The *fine* line between sleep and wakefulness

Understanding hypersomnolence disorders and sleep in the Intensive Care environment

A.E. van der Hoeven

The *fine* line between sleep and wakefulness

Understanding hypersomnolence disorders
and sleep in the Intensive Care environment

A.E. van der Hoeven

Financial support for the printing of this thesis was provided by:

Slaapgeneeskunde Vereniging Nederland (SVNL)

Nederlandse Vereniging voor Slaap- en Waakonderzoek (NSWO)

ISBN	978-94-6506-302-7
Lay-out	Adrienne van der Hoeven and Anton Tuluk
Cover design	Adrienne van der Hoeven
Printed by	Ridderprint | www.ridderprint.nl

The *fine* line between sleep and wakefulness

Understanding hypersomnolence disorders and sleep in the Intensive Care environment

Proefschrift

ter verkrijging van
de graad van doctor aan de Universiteit Leiden,
op gezag van rector magnificus prof. dr. ir. H. Bijl,
volgens besluit van het college voor promoties
te verdedigen op donderdag 10 oktober 2024
klokke 16:00 uur

door

Adrienne Elisabeth van der Hoeven
geboren te Haarlem
in 1992

Promotor	prof. dr. G.J. Lammers
Copromotoren	dr. R. Fronczek dr. D. Bijlenga
Promotiecommissie	prof. dr. C.M. Cobbaert prof. dr. S. Overeem (Technische Universiteit Eindhoven en Kempenhaeghe, Heeze) prof. dr. R. Khatami (Universiteit van Bern) prof. dr. J.E. Tulleken (Rijksuniversiteit Groningen)

Contents

Chapter 1	General introduction	8

Part I: Aspects of Sleep in Central Disorders of Hypersomnolence

Chapter 2	Intermediate hypocretin-1 cerebrospinal fluid levels and typical cataplexy: their significance in the diagnosis of narcolepsy type 1	28
Chapter 3	Hypocretin-1 measurements in cerebrospinal fluid using radioimmunoassay: within and between assay reliability and limit of quantification	50
Chapter 4	Applicability of the Sustained Attention to Response Task in hypersomnolence: experience and results from a Tertiary Referral Center	70

Part II: Sleep in the Intensive Care Environment

Chapter 5	Sleep in the intensive and intermediate care units: exploring related factors of delirium, benzodiazepine use and mortality	96
Chapter 6	Dream experiences during ICU stay: prevalence, content, vividness and associated factors	120
Chapter 7	Summary, discussion and future perspectives	148

Appendices	A Samenvatting, discussie en toekomstige perspectieven	166
	B Curriculum Vitae	179
	C List of publications and articles	180
	D Dankwoord	181

1

Introduction

Chapter 1 General introduction

The Importance of Sleep

Sleep is an essential aspect to life and well-being seen throughout the animal kingdom. During sleep, the ability to drink, eat or react to outside stimuli is temporarily suspended. In spite of this, sleep is a state conserved across animal species, indicating an evolutionary value [1]. Thus, the benefits of sleep appear to outweigh its limitations [2]. Various functions of sleep were discovered by studying sleep deprivation. Poor sleep negatively affects memory consolidation, emotional processing, cognitive function, the immune system, hormone secretion and physical performance [2-8]. Sleep disruption and a sleep duration that deviates from the norm are associated with increased mortality and morbidity [9-11]. These negative effects underline the importance of researching sleep: to help those with poor sleep by improving our understanding of the various risk factors involved in sleep disruption and by developing better diagnostic methods.

How is Sleep defined?

Having a clear definition of sleep is beneficial as such a definition provides a foundation and common framework when conducting research. Moreover, by establishing clear criteria for normal sleep, abnormal sleep can also more easily be identified, investigated and addressed [12]. Generally speaking, sleep can be defined as a recurring, reversible neuro-behavioral state characterized by a reduced awareness and responsiveness to the surroundings, often (in humans) accompanied by lying down, decreased activity, and closed eyes [13]. The *AASM manual for the scoring of sleep and associated events: rules, terminology and technical specifications* defines broadly two distinct forms of sleep: rapid eye movement (REM) sleep and non-REM sleep [14, 15]. Normal sleep architecture is characterized by alternations between REM and non-REM sleep throughout the night in cycles of approximately 90 minutes [14, 16]. Non-REM sleep is further divided into N1, N2 (both light sleep) and N3 (deep sleep or slow wave sleep, i.e. SWS) [16].

SWS is considered to be the most restorative sleep, and is associated with human growth hormone secretion and metabolite clearance [17, 18]. Dreams mostly occur during REM sleep, but can also take place during non-REM sleep [16, 19]. However, the dream content differs:

dreams during REM sleep are generally longer, tend to follow a narrative, and are more vivid, story-like and bizarre [19]. Research suggests that emotional processing mostly occurs during REM sleep [7, 20] and that both REM and non-REM sleep play a role in memory consolidation [21, 22].

Intrinsic and Extrinsic Factors

Sleep is impacted by an interplay of intrinsic and extrinsic factors. Intrinsic factors, such as age, sex, illnesses, and underlying disorders, are inherent to the individual [23-28]. Extrinsic factors refer to environmental factors such as lifestyle, ambient temperature, noise and light exposure [29, 30]. Sleep difficulties are common in the general population but tend to be more pronounced in hospitalized patients due to combinations of both intrinsic and extrinsic factors [31]. This thesis examines different aspects of sleep and sleep disorders in two different contexts: in central disorders of hypersomnolence (Part I) and in the intensive care environment (Part II).

Summary

- Sleep is important for health and well-being.
- Normal sleep architecture involves cycles of REM and non-REM sleep. The latter is further divided into N1, N2, and N3 stages. N3 is also named deep sleep or slow wave sleep.
- Vivid, story-like and bizarre dreams mostly occur during REM sleep.
- Sleep is influenced by intrinsic (e.g. genetics, age, illness, health) and extrinsic factors (e.g. lifestyle, environment).
- This thesis explores aspects of sleep in central disorders of hypersomnolence and intensive care settings.

Part I: Aspects of Sleep in Central Disorders of Hypersomnolence

1

Central disorders of hypersomnolence (CDH) are rare but can be debilitating. They are characterized by excessive daytime sleepiness (EDS) [32]. In the third edition of the International Classification of Sleep Disorders (ICSD3) EDS is defined as an irrepressible need to sleep, or daytime lapses into sleep, occurring almost daily for at least three months [33]. The three most common central disorders of hypersomnolence are narcolepsy type 1 (NT1), narcolepsy type 2 (NT2), and idiopathic hypersomnia (IH) [33].

Narcolepsy Type 1

NT1 is a chronic sleep-wake disorder with a prevalence of 25 to 50 per 100,000 people [34], that typically starts in adolescence or young adulthood [35, 36]. NT1 is associated with a deficiency of the neuropeptide hypocretin-1 in cerebrospinal fluid (CSF). Due to associations with a specific HLA subtype, the H1N1 influenza pandemic, and a variety of other infections, it has been hypothesized that the condition is caused by an autoimmune response targeting hypocretin-producing neurons in the hypothalamus [37, 38]. Determining hypcretin-1 deficiency in CSF is considered the gold standard of the NT1 diagnosis, as it can be sufficient for a diagnosis when combined with EDS complaints [33].

In essence, those with NT1 experience a bidirectional intrusion of aspects of the sleep and wake states, blurring the boundaries between sleep and wakefulness. EDS, cataplexy, sleep paralysis, and hypnogogic hallucinations are main features of NT1 [32]. EDS frequently results in vigilance complaints during the day [39]. In addition to these classic symptoms, individuals with NT1 commonly report disrupted nighttime sleep [40]. The impact of these symptoms can be substantial, affecting individuals in their educational, professional and social lives. It affects driving and increases financial burden due to health care expenses. It is also associated with higher rates of mental health issues such as depression, and impacts work and school performances [41]. Generally, with a few exceptions, NT1 tends to start gradually. Only a small portion (<10%) of NT1 patients have all the main characteristics at disease onset. The gradual start and low prevalence of the disorder contribute to a median diagnostic delay of 10.5 years as the disorder often goes unnoticed or is misdiagnosed [42] (see also figure 1.1 in which both the spread in age of onset and age of diagnosis of a large group with narcolepsy are shown [43]).

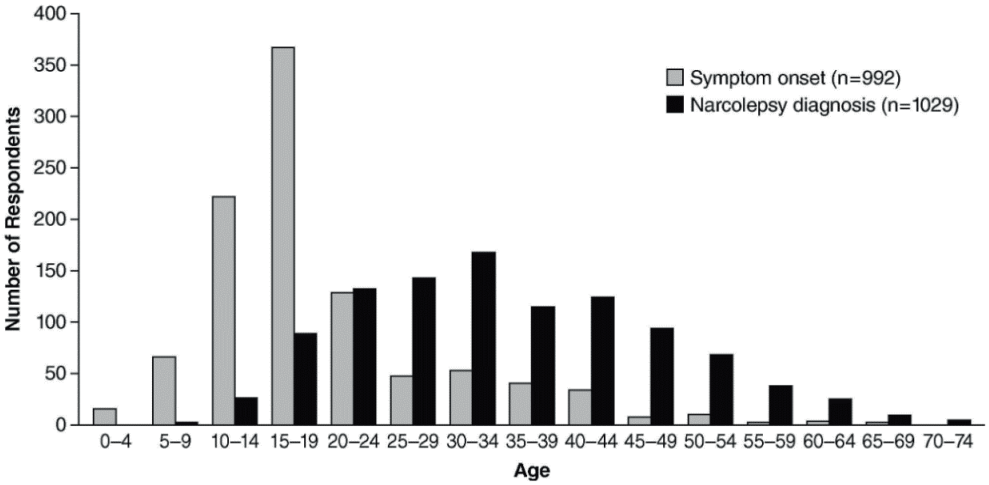


Figure 1.1. Narcolepsy onset and diagnosis. Figure adapted from Thorpy MJ, Krieger AC. Delayed diagnosis of narcolepsy: characterization and impact. *Sleep Medicine*. 2014;15(5):502-507.

Cataplexy is the most specific symptom of NT1, and its early recognition can reduce diagnostic delay [44]. It is defined as a sudden, brief, generally bilateral (partial or complete) loss of muscle tone triggered by strong (usually positive) emotions with retained consciousness (see figure 1.2). Return of muscle activity is typically abrupt [33]. Cataplexy can be conceptualized as a pathological manifestation of a REM sleep feature (lack of muscle tone) during wakefulness [32, 45, 46]. The presentation of cataplexy varies, and a distinction has been made between typical and atypical cataplexy, with typical cataplexy being most specific for NT1. While typical cataplexy is characterized by the presence of the aforementioned characteristics, atypical cataplexy has not yet been clearly defined [47, 48]. **Chapter 2** delves deeper in the clinical significance of the differentiation between typical and atypical cataplexy. In this chapter we also re-examine the currently used diagnostic hypocretin-1 cutoff value in CSF, which was established over 20 years ago, before the current classification of NT1 and NT2 was in place [49].

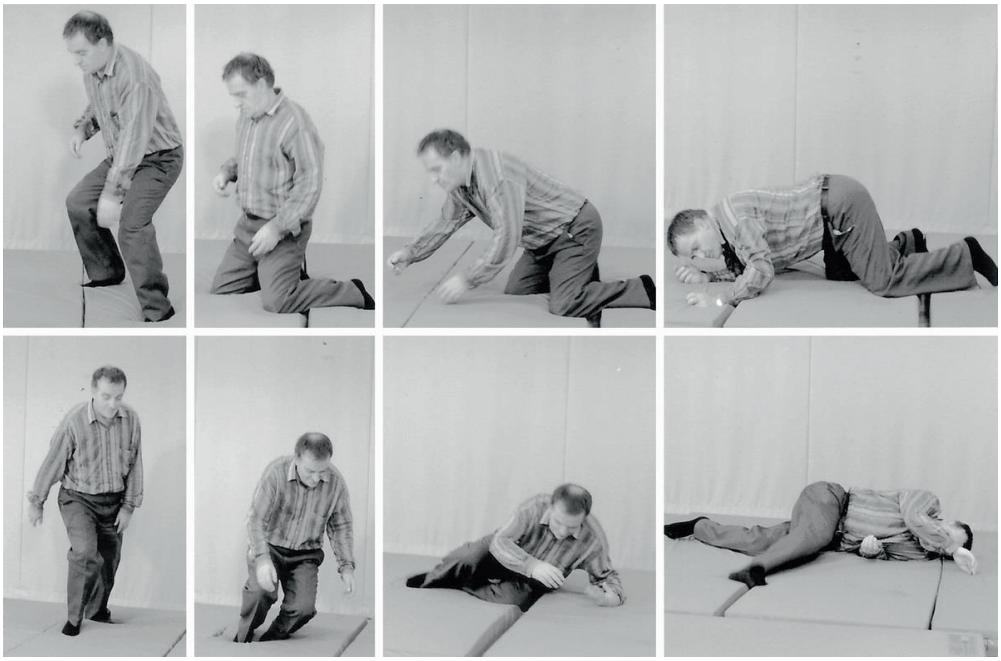


Figure 1.2. A cataplectic episode. Figure adapted from Overeem S, Mignot E, van Dijk JB, Lammers GJ. Narcolepsy: clinical features, new pathophysiologic insights, and future perspectives. *J Clin Neurophysiol.* 2001;18(2):78–105.

Diagnosing Narcolepsy Type 1

The diagnosis of NT1 is based upon clinical evaluation and diagnostic tests. The diagnostic criteria of the third edition of the International classification of sleep disorders (ICSD3) can be found in the first column of table 1.1. Mid 2023 a text revision was published (ICSD3-TR), with some minor changes to the criteria. However, the ICSD3 criteria are still presented in the table because they were used for the studies included in this thesis. The changes in the TR edition are listed in the notes below the table.

Important outcome parameters are provided by the combination of an overnight polysomnography (PSG) followed by a multiple sleep latency test (MSLT) the next day [33]. The PSG is a sleep study during which different body functions are monitored, including brain and muscle activity, breathing, eye movements and heart rhythm [50]. In this context, it is primarily used to rule out other sleep disorders and to provide information regarding the sleep architecture and presence of abnormal REM sleep patterns [51]. In NT1, sleep is typically fragmented, with frequent awakenings and sleep-onset REM (SOREM) sleep (defined as REM sleep within 15 minutes after sleep onset) [51]. The MSLT measures sleep propensity during daytime naps. Short mean sleep latencies (i.e. $\leq$ 8 minutes) and the presence of SOREM

periods, which are indicative of narcolepsy [33, 52]. The outcomes of the MSLT can resemble those of a narcoleptic patient due to factors like insufficient sleep syndrome and circadian rhythm misalignment. To address this, actigraphy is required to assess sleep patterns in the weeks leading up to the MSLT [53].

As mentioned before, the determination of a hypocretin-1 deficiency in CSF using radioimmunoassay (RIA) is the gold standard for diagnosing NT1. However, despite its status as the gold standard, a comprehensive evaluation of the variability in results when applying this method has not been conducted. In **Chapter 3**, we report on the reliability of this method. Despite its importance in the diagnosis of NT1, hypocretin-1 levels are not always measured as it requires a lumbar puncture which is an invasive and uncomfortable procedure for patients [33, 49].

Genetic testing for specific human leukocyte antigen (HLA) alleles, such as the HLA DQB1 06:02 allele, can also be helpful when diagnosing NT1. HLA DQB1 06:02 positivity is regarded as a perquisite for the development of NT1, with over 90% of NT1 patients being HLA DQB1 06:02 positive. It is not specific however, as 12-40% of the general population, depending on the ethnicity, is also HLA DQB1 06:02 positive [54, 55].

Table 1.1. Diagnostic criteria for narcolepsy type 1, type 2 and idiopathic hypersomnia following the ICSD3

Narcolepsy type 1	Narcolepsy type 2	Idiopathic hypersomnia
1. Daily EDS for ≥ 3 months	1. Daily EDS for ≥ 3 months	1. Daily EDS for ≥ 3 months
2. One or both of the following: a. Cataplexy* and mean SL ≤ 8 minutes and ≥ 2 SOREMPS on MSLT (1 SOREMP may be on preceding PSG)** b. Hypocretin-1 ≤ 110 pg/mL or < 1/3 of mean values with the same assay in healthy patients	2. Mean SL ≤ 8 minutes and ≥ 2 SOREMPS on MSLT (1 SOREMP may be on preceding PSG) 3. Absence of cataplexy* 4. One of the following: a. Hypocretin-1 not measured b. Hypocretin-1 > 110 pg/mL or > 1/3 of mean values with the same assay in healthy patients 5. EDS and/or MSLT results not better explained by other causes	2. Insufficient sleep syndrome is confirmed to be absent 3. MSLT and PSG show < 2 SOREMPS 4. One or both of the following: a. Mean SL ≤ 8 minutes on MSLT b. Total 24-hours sleep time ≥ 660 minutes (long sleep subtype) 5. Absence of cataplexy* 6. EDS and/or MSLT results not better explained by other causes

Abbreviations: ICSD3, 3rd edition of the International Classification of Sleep Disorders; EDS, excessive daytime sleepiness; SL, sleep latency; SOREMP, sleep-onset rapid eye movement period; MSLT, multiple sleep latency test; PSG, polysomnography

* This criterion was altered in the International Classification of Sleep Disorders, 3rd edition text revision (ICSD-3-TR). Cataplexy must now meet the criteria for typical cataplexy.

**This criterion was also altered in the ICSD-3-TR. Now cataplexy and only an overnight SOREMP can fulfill criterion a.

Narcolepsy Type 2 and Idiopathic Hypersomnia

NT2 presents with similar symptoms and characteristics as NT1, except that in NT2 typical cataplexy is absent and there is no hypocretin-1 deficiency (see table 1.1). Atypical cataplexy or cataplexy-like symptoms may be present [56]. Because cataplexy is not always present at the earlier stages of NT1, some individuals may be initially misdiagnosed as NT2. The diagnosis is later changed to NT1 after cataplexy symptoms emerge [57]. The prevalence is estimated at 20 to 34 per 100,000 people [35, 36], although this may be an overestimation based on more current research [58]. The cause of NT2 is unknown, and the lack of hypocretin-1 deficiency suggests a different pathophysiology than NT1. Moreover, the percentage of HLA DQB1 06:02 positivity is much lower in NT2 compared to NT1 (only 40-50%) [57]. There are several hypotheses, including a less severe reduction in hypocretin-producing neurons, compromised signaling through hypocretin receptors, or an entirely distinct mechanism [59]. It is also possible that a portion of those diagnosed with NT2 have a false-positive diagnosis based on a reversible behavioral component, particularly if reversible causes such as insufficient sleep syndrome were not thoroughly eliminated during the diagnostic process [58].

IH is characterized by EDS that is not explained by other factors or conditions. The prevalence is estimated at approximately 2 to 5 cases per 100,000 based on sleep center referrals [60-62]. Symptoms typically start at adolescence or young adulthood [60]. Cataplexy, sleep paralysis, or hypnogogic hallucinations are not typically present in individuals with IH [33, 60, 62]. Although, it is worth noting that, with the exception of cataplexy, these symptoms may occasionally be present in some cases of IH. Individuals with IH have difficulty awakening in the morning and frequently have long naps, which are often unrefreshing. This contrasts with narcolepsy patients, who report refreshing effects of short naps [60]. There are two subtypes of IH: IH with a long sleep duration (at least 660 minutes of sleep time measured over 24 hours of PSG monitoring) and IH with a normal sleep duration [56]. IH can affect many aspects of a person's life, including interpersonal relationships, work and school obligations, and safety, for example while driving [63]. A PSG and MSLT need to be performed to rule out other causes and confirm an IH diagnosis. The cause of IH is unknown, although in one study the frequency of autoinflammatory disorders was significantly higher in those with IH as compared to controls [64]. Unlike NT1 and NT2, HLA DQB1 06:02 positivity is similar to that in the general population [60, 65].

Differentiating between NT2 and IH with a normal sleep duration can be challenging, as both lack the consistent presence of distinctive and pathognomonic symptoms and biomarkers. The MSLT shows low test-retest reliability in NT2 and IH, as opposed to NT1 [66, 67]. Because of

this, it is questioned whether NT2 and IH represent distinct entities or if they simply are broad, somewhat overlapping, categories that include various underlying sleep disorders with different pathophysiologies [68, 69]. At the 7th International Symposium on Narcolepsy in 2018, several clinician-scientist even suggested combining NT2 and the IH with a normal sleep duration subtype into a single disease entity, resulting in the publication of a consensus review on this topic by this group of scientists [70].

Vigilance in Central Disorders of Hypersomnolence

Individuals with CDH experience EDS, making it challenging to maintain sustained attention throughout the day. Vigilance is a rarely assessed aspect of EDS in hypersomnolence disorders, but has a profound impact on daily functioning and poses safety risks [39]. The Sustained Attention to Response Task (SART) can be used to assess the presence and severity of vigilance complaints in CDH. This task measures the ability to sustain attention over a short period of time during multiple assessments over the course of the day [71]. **Chapter 4** focusses on the associations between SART results and MSTL and PSG outcomes, and investigates its value in the assessment of CDH.

Summary

- Central disorders of hypersomnolence (CDH), including narcolepsy type 1 (NT1), type 2 (NT2), and idiopathic hypersomnia (IH), are characterized by excessive daytime sleepiness (EDS).
- To diagnose CDH polysomnography (PSG), the multiple sleep latency test (MSLT), actigraphy, HLA typing, and hypocretin-1 measurements in CSF using radioimmunoassay (RIA) can/should be used in various combinations.
- NT1 is associated with typical cataplexy and hypocretin-1 deficiency in the brain, which can be measured in cerebrospinal fluid (CSF). The diagnostic criteria for NT2 are similar to NT1, with the exceptions of these associations.
- Cataplexy can be typical (meeting all the official criteria), or less commonly, atypical.
- Diagnosing and differentiating NT2 and IH with a normal sleep duration can be challenging due to a lack of reliable distinctive biomarkers and frequently overlapping symptoms.
- Vigilance complaints greatly impact daily functioning in CDH, but are rarely assessed.
- The susta10ined attention to response task (SART) can be used to assess vigilance.

Part II: Sleep in the Intensive Care Environment

Moving from the exploration of aspects of hypersomnolence disorders, this part of the thesis considers the impact of critical care settings on sleep quality and quantity of hospitalized patients. Part II of this thesis focuses on sleep at the Intensive Care Unit (ICU) and Intermediate Care Unit (IMCU), which provide specialized and intensive medical care for patients with severe and/or critical conditions. ICU patients generally require close monitoring and stabilizing measures due to a variety of conditions such as severe trauma, (acute) organ failure, cardiac emergencies, respiratory distress, or post-operative recovery. Treatment options such as mechanical ventilation, hemodynamic support and the maintenance of anesthesia are available in the ICU [72, 73]. An IMCU can provide care for patients who require a level of monitoring and treatment between that of the general ward and ICU [74]. While sleep difficulties in the ICU have been extensively studied and reported on, there is minimal literature available on sleep in the IMCU.

Sleep Disruption in the ICU

Disrupted sleep and circadian rhythm abnormalities are often observed in ICU patients. Sleep in the ICU tends to be highly fragmented, with a decrease in the proportion of time spent in deep sleep and REM sleep stages. Additionally, over half of the proportion of total sleep can occur during the day [75].

Various factors contribute to disturbed sleep in the ICU, including medication use, mechanical ventilation, severity of illness, stress, the frequency, timing and intensity of required care, and relatively high noise and light levels at night [76, 77]. In the ICU, nighttime sleep disruption is associated with the onset of delirium and psychological complaints after discharge. Moreover, in the general population sleep disruption has been shown to negatively affect the immune system, pulmonary muscle endurance and hormone levels [76]. Although the extent is unknown, these effects have a negative impact on patient recovery and wellbeing [76-78].

It is a challenge to both treat the critical illness properly and optimize sleep in critically ill patients. Multiple sleep-promoting interventions have been applied and studied, with varying results. Currently, there is a lack of adequate treatment options [79]. Pharmacologic treatment alone is often insufficient, and could result in both polypharmacy and increased risk of developing delirium, instead of actually promoting sleep. Nonetheless, such medications, mostly benzodiazepine hypnotics, are frequently prescribed despite a lack of evidence

regarding positive outcomes [75]. Based on existing data, a multicomponent protocolled approach that favors non-pharmacologic measures, such as improving sleep hygiene and lowering noise and light levels, is considered to be the best option [75]. All in all, better treatment strategies are needed to prevent and treat sleep disturbances in critically ill individuals.

In **Chapter 5** we report on the prevalence, causes, and negative impact of sleep disturbances in patients admitted to the ICU and IMCU, as well as the use and effect of often prescribed benzodiazepine hypnotics, such as temazepam, oxazepam, and lorazepam.

Dream Experiences in the ICU

Vivid and often distressing dreams are frequently reported in the ICU [80, 81]. These dreams can be bizarre and frightening and can impact patients' mental wellbeing and quality of life even after discharge [82-84]. A potential cause could be the physical and emotional stress ICU patients experience, which could affect their dream content [85, 86]. Other possible causes are the use of sedatives and analgesics, which are known to alter sleep architecture and can impact dream recall [87, 88], or the presence of delirium, which occurs frequently in critically ill individuals [89]. The impact of vivid and negative dream experiences in ICU patients is a complex and under-explored research area. This phenomenon has not been studied in hospitalized patients with a longer length of stay. As a result, the exact factors that contribute to their occurrence are unknown. In **Chapter 6** we describe the frequency, content and associated factors of dream experiences in the ICU. Recognizing, studying, and addressing such dream experiences could contribute to patients' overall well-being by taking measures to prevent or minimize the impact of the dreams on their mental health. Moreover, research in this area can further improve understanding of the negative impact of sleep and dreaming in the ICU.

Summary

- The Intensive and Intermediate Care Units (ICU and IMCU) provide care for patients with severe or critical medical conditions.
- Sleep disruption is common in the ICU, due to sickness, pain, discomfort, mechanical ventilation, noise and light disturbances, care activities, medication and stress. Sleep in the IMCU is under-investigated.
- Current treatment options for sleep disturbances at the ICU/IMCU are limited, and pharmacologic treatment alone may not be effective and can increase the risk of delirium.
- Dream experiences in the ICU are frequently reported, and can be frightening, vivid and distressing.
- Further research on sleep disturbances and dream experiencers in the ICU and IMCU is needed to understand the contributing factors and to develop strategies to address the impact on patient well-being.

Aims and Thesis Outline

This thesis explores two distinct scenarios in which sleep disturbances occur: in the context of central disorders of hypersomnolence (Part I) and in the ICU and IMCU environment (Part II). Although the causes of disrupted sleep differ in these two settings, the resulting sleep patterns exhibit similar and overlapping characteristics. Enhancing our understanding of sleep-related aspects in one domain could enhance our understanding of sleep in the other domain as well.

Part I: the diagnostic methods for hypersomnolence disorders are investigated, with particular focus on NT1.

In **Chapter 2** the clinical significance of differentiating between typical and atypical cataplexy when diagnosing NT1 is explored. This includes a re-assessment of the current hypocretin-1 cutoff value in CSF and its relevance given the new classification of NT1 and NT2. Additionally, the clinical and diagnostic characteristics of patients with intermediate hypocretin-1 levels were determined. The aim is to improve the diagnostic process for individuals with suspected narcolepsy.

Chapter 3 provides a systematic evaluation of the reliability of RIAs in determining CSF hypocretin-1 levels. The aim was to assess intra- and inter-assay variability and report the lowest concentration that can be measured with a stated amount of variability (the lower limit of quantification). This is relevant in clinical and scientific practice, since hypocretin-1 CSF values below the detection level are unjustly still often reported. It is recommended that if

below the lower limit of quantification, these low values are reported as *undetectable*. The outcomes of this evaluation, based on 20 years of experience, have implications for both previously published studies and future research involving CSF hypocretin-1 analysis.

Chapter 4 aims to assess the practical application of the SART as a tool to evaluate vigilance, an often overlooked aspect of EDS in hypersomnolence disorders. By examining differences in SART outcomes between hypersomnolence disorders and associations between the SART and MSLT and PSG outcomes, value of the SART in the assessment of CDH is investigated.

Part II: sleep disruptions and dream experiences in the ICU and IMCU settings.

Moving on to Part II, **Chapter 5** focusses on sleep disturbances in patients admitted to the ICU and IMCU. We explored the prevalence, causes, and negative impact of these sleep disruptions, as well as the use and effects of benzodiazepine hypnotics. This chapter highlights the common occurrence of sleep difficulties in this population and the lack of efficacy of benzodiazepine hypnotics to alleviate these complaints. Additionally, we emphasized the need for consistent monitoring of sleep in critical care settings. This could aid in gaining a better understanding of the relationship between sleep and clinical outcomes and could potentially help in detecting delirium at an earlier time point or prevent it altogether.

Chapter 6 addresses the under-explored problem of vivid and often frightening dreams in critical care settings. The frequency, content, and associated factors are investigated. Recognizing, studying, and looking for possible ways to address these dream experiences, could improve patient's overall well-being in the ICU and IMCU and also enhance our understanding of sleep in these settings as well.

Finally, the conclusions of this thesis are discussed in **Chapter 7**. The clinical implications and recommendations are outlined here as well.

References

1. Keene, A.C. and E.R. Duboue, *The origins and evolution of sleep.* J Exp Biol, 2018. **221**(Pt 11).
2. Krueger, J.M., et al., *Sleep function: Toward elucidating an enigma.* Sleep Med Rev, 2016. **28**: p. 46-54.
3. Irwin, M.R., *Why sleep is important for health: a psychoneuroimmunology perspective.* Annu Rev Psychol, 2015. **66**: p. 143-72.
4. Depner, C.M., E.R. Stothard, and K.P. Wright, Jr., *Metabolic consequences of sleep and circadian disorders.* Curr Diab Rep, 2014. **14**(7): p. 507.
5. Leproult, R. and E. Van Cauter, *Role of sleep and sleep loss in hormonal release and metabolism.* Endocr Dev, 2010. **17**: p. 11-21.
6. Cousins, J.N. and G. Fernández, *The impact of sleep deprivation on declarative memory.* Prog Brain Res, 2019. **246**: p. 27-53.
7. Tempesta, D., et al., *Sleep and emotional processing.* Sleep Med Rev, 2018. **40**: p. 183-195.
8. Copenhaver, E.A. and A.B. Diamond, *The Value of Sleep on Athletic Performance, Injury, and Recovery in the Young Athlete.* Pediatr Ann, 2017. **46**(3): p. e106-e111.
9. da Silva, A.A., et al., *Sleep duration and mortality in the elderly: a systematic review with meta-analysis.* BMJ Open, 2016. **6**(2): p. e008119.
10. Li, J., et al., *Sleep duration and health outcomes: an umbrella review.* Sleep Breath, 2022. **26**(3): p. 1479-1501.
11. Medic, G., M. Wille, and M.E. Hemels, *Short- and long-term health consequences of sleep disruption.* Nat Sci Sleep, 2017. **9**: p. 151-161.
12. Hirshkowitz, M., *Normal human sleep: an overview.* Med Clin North Am, 2004. **88**(3): p. 551-65, vii.
13. Carskadon, M. and W. Dement, *Principles and Practice of Sleep Medicine 5th edn, ed M H Kryger et al*. 2011, Amsterdam: Elsevier.
14. Carley, D.W. and S.S. Farabi, *Physiology of Sleep.* Diabetes Spectr, 2016. **29**(1): p. 5-9.
15. Iber, C., et al., *The AASM Manual for the Scoring of Sleep and Associated Events: Rules, Terminology and Technical Specifications.* Westchester, IL: American Academy of Sleep Medicine, 2007.
16. Patel, A.K., et al., *Physiology, Sleep Stages*, in *StatPearls*. 2023, StatPearls Publishing
17. Copyright © 2023, StatPearls Publishing LLC.: Treasure Island (FL).
18. Sassin, J.F., et al., *Human growth hormone release: relation to slow-wave sleep and sleep-walking cycles.* Science, 1969. **165**(3892): p. 513-5.
19. Xie, L., et al., *Sleep drives metabolite clearance from the adult brain.* Science, 2013. **342**(6156): p. 373-7.
20. Tsunematsu, T., *What are the neural mechanisms and physiological functions of dreams?* Neurosci Res, 2023. **189**: p. 54-59.
21. Werner, G.G., et al., *Differential Effects of REM Sleep on Emotional Processing: Initial Evidence for Increased Short-term Emotional Responses and Reduced Long-term Intrusive Memories.* Behavioral Sleep Medicine, 2021. **19**(1): p. 83-98.
22. Boyce, R., S. Williams, and A. Adamantidis, *REM sleep and memory.* Current Opinion in Neurobiology, 2017. **44**: p. 167-177.
23. Dijk, D.J., *Regulation and functional correlates of slow wave sleep.* J Clin Sleep Med, 2009. **5**(2 Suppl): p. S6-15.
24. Redline, S., et al., *The effects of age, sex, ethnicity, and sleep-disordered breathing on sleep architecture.* Arch Intern Med, 2004. **164**(4): p. 406-18.
25. Carrier, J., et al., *Sex differences in age-related changes in the sleep-wake cycle.* Front Neuroendocrinol, 2017. **47**: p. 66-85.
26. Singareddy, R.K. and R. Balon, *Sleep in posttraumatic stress disorder.* Ann Clin Psychiatry, 2002. **14**(3): p. 183-90.
27. Zhang, Y., et al., *Sleep in Parkinson's disease: A systematic review and meta-analysis of polysomnographic findings.* Sleep Med Rev, 2020. **51**: p. 101281.
28. Zhang, Y., et al., *Polysomnographic nighttime features of narcolepsy: A systematic review and meta-analysis.* Sleep Med Rev, 2021. **58**: p. 101488.
29. Watson, N.F., et al., *Transcriptional Signatures of Sleep Duration Discordance in Monozygotic Twins.* Sleep, 2017. **40**(1).
30. Ikeda, Y., et al., *Relationships between sleep efficiency and lifestyle evaluated by objective sleep assessment: SLeep Epidemiology Project at University of Tsukuba.* Nagoya J Med Sci, 2022. **84**(3): p. 554-569.
31. Caddick, Z.A., et al., *A review of the environmental parameters necessary for an optimal sleep environment.* Building and Environment, 2018. **132**: p. 11-20.

32. Bano, M., et al., *The Influence of Environmental Factors on Sleep Quality in Hospitalized Medical Patients.* Frontiers in Neurology, 2014. **5**.
33. Khan, Z. and L.M. Trotti, *Central Disorders of Hypersomnolence: Focus on the Narcolepsies and Idiopathic Hypersomnia.* Chest, 2015. **148**(1): p. 262-273.
34. American Academy of Sleep, M., *The international classification of sleep disorders*. 2014.
35. Longstreth, W.T., Jr., et al., *The epidemiology of narcolepsy.* Sleep, 2007. **30**(1): p. 13-26.
36. Silber, M.H., et al., *The epidemiology of narcolepsy in Olmsted County, Minnesota: a population-based study.* Sleep, 2002. **25**(2): p. 197-202.
37. Shin, Y.K., et al., *Prevalence of narcolepsy-cataplexy in Korean adolescents.* Acta Neurol Scand, 2008. **117**(4): p. 273-8.
38. Liblau, R.S., et al., *Hypocretin (orexin) biology and the pathophysiology of narcolepsy with cataplexy.* Lancet Neurol, 2015. **14**(3): p. 318-28.
39. Bonvalet, M., et al., *Autoimmunity in narcolepsy.* Curr Opin Pulm Med, 2017. **23**(6): p. 522-529.
40. van Schie, M.K.M., et al., *Vigilance: discussion of related concepts and proposal for a definition.* Sleep Med, 2021. **83**: p. 175-181.
41. Roth, T., et al., *Disrupted nighttime sleep in narcolepsy.* J Clin Sleep Med, 2013. **9**(9): p. 955-65.
42. Jennum, P., et al., *Health, social, and economic consequences of narcolepsy: a controlled national study evaluating the societal effect on patients and their partners.* Sleep medicine, 2012. **13**(8): p. 1086-1093.
43. Morrish, E., et al., *Factors associated with a delay in the diagnosis of narcolepsy.* Sleep Med, 2004. **5**(1): p. 37-41.
44. Thorpy, M. and S. Cronin. *Age of onset and time to diagnosis of narcolepsy*. in *Neurology*. 1999. LIPPINCOTT WILLIAMS & WILKINS 530 WALNUT ST, PHILADELPHIA, PA 19106-3621 USA.
45. Taddei, R.N., et al., *Diagnostic delay in narcolepsy type 1: combining the patients' and the doctors' perspectives.* Journal of Sleep Research, 2016. **25**(6): p. 709-715.
46. Mahoney, C.E., et al., *The neurobiological basis of narcolepsy.* Nat Rev Neurosci, 2019. **20**(2): p. 83-93.
47. Seifinejad, A., A. Vassalli, and M. Tafti, *Neurobiology of cataplexy.* Sleep Med Rev, 2021. **60**: p. 101546.
48. Lammers, G.J., et al., *Diagnosis of central disorders of hypersomnolence: A reappraisal by European experts.* Sleep Medicine Reviews, 2020. **52**: p. 101306.
49. Anic-Labat, S., et al., *Validation of a Cataplexy Questionnaire in 983 Sleep-disorders Patients.* Sleep, 1999. **22**(1): p. 77-87.
50. Mignot, E., et al., *The Role of Cerebrospinal Fluid Hypocretin Measurement in the Diagnosis of Narcolepsy and Other Hypersomnias.* Archives of Neurology, 2002. **59**(10): p. 1553-1562.
51. Markun, L.C. and A. Sampat, *Clinician-Focused Overview and Developments in Polysomnography.* Current Sleep Medicine Reports, 2020. **6**(4): p. 309-321.
52. Pizza, F., et al., *Nocturnal Sleep Dynamics Identify Narcolepsy Type 1.* Sleep, 2015. **38**(8): p. 1277-84.
53. Carskadon, M.A., et al., *Guidelines for the multiple sleep latency test (MSLT): a standard measure of sleepiness.* Sleep, 1986. **9**(4): p. 519-24.
54. Torstensen, E.W., et al., *Diagnostic value of actigraphy in hypersomnolence disorders.* Sleep Med, 2021. **85**: p. 1-7.
55. Capittini, C., et al., *Correlation between HLA-DQB1*06:02 and narcolepsy with and without cataplexy: approving a safe and sensitive genetic test in four major ethnic groups. A systematic meta-analysis.* Sleep Med, 2018. **52**: p. 150-157.
56. Tafti, M., et al., *DQB1 locus alone explains most of the risk and protection in narcolepsy with cataplexy in Europe.* Sleep, 2014. **37**(1): p. 19-25.
57. Medicine, A.A.o.S., *International classification of sleep disorders—third edition (ICSD-3).* AASM Resour Libr, 2014. **281**: p. 2313.
58. Andlauer, O., et al., *Predictors of hypocretin (orexin) deficiency in narcolepsy without cataplexy.* Sleep, 2012. **35**(9): p. 1247-55F.
59. Baumann-Vogel, H., et al., *Narcolepsy type 2: A rare, yet existing entity.* J Sleep Res, 2021. **30**(3): p. e13203.
60. Thannickal, T.C., R. Nienhuis, and J.M. Siegel, *Localized loss of hypocretin (orexin) cells in narcolepsy without cataplexy.* Sleep, 2009. **32**(8): p. 993-8.
61. Anderson, K.N., et al., *Idiopathic hypersomnia: a study of 77 cases.* Sleep, 2007. **30**(10): p. 1274-81.
62. Bassetti, C. and M.S. Aldrich, *Idiopathic hypersomnia. A series of 42 patients.* Brain, 1997. **120 (Pt 8)**: p. 1423-35.
63. Billiard, M., *Idiopathic hypersomnia.* Neurol Clin, 1996. **14**(3): p. 573-82.
64. Susta, M., et al., *Idiopathic Hypersomnia—A Dynamic Simulation Model.* Frontiers in Neurology, 2022. **13**.

65. Barateau, L., et al., *Comorbidity between central disorders of hypersomnolence and immune-based disorders.* Neurology, 2017. **88**(1): p. 93-100.
66. Miyagawa, T., et al., *An association analysis of HLA-DQB1 with narcolepsy without cataplexy and idiopathic hypersomnia with/without long sleep time in a Japanese population.* Hum Genome Var, 2015. **2**: p. 15031.
67. Ruoff, C., et al., *The MSLT is Repeatable in Narcolepsy Type 1 But Not Narcolepsy Type 2: A Retrospective Patient Study.* J Clin Sleep Med, 2018. **14**(1): p. 65-74.
68. Trotti, L.M., B.A. Staab, and D.B. Rye, *Test-retest reliability of the multiple sleep latency test in narcolepsy without cataplexy and idiopathic hypersomnia.* J Clin Sleep Med, 2013. **9**(8): p. 789-95.
69. Šonka, K., M. Šusta, and M. Billiard, *Narcolepsy with and without cataplexy, idiopathic hypersomnia with and without long sleep time: a cluster analysis.* Sleep Med, 2015. **16**(2): p. 225-31.
70. Gool, J.K., et al., *Data-Driven Phenotyping of Central Disorders of Hypersomnolence With Unsupervised Clustering.* Neurology, 2022. **98**(23): p. e2387-e2400.
71. Fronczek, R., et al., *To split or to lump? Classifying the central disorders of hypersomnolence.* Sleep, 2020. **43**(8).
72. Robertson, I.H., et al., *`Oops!': Performance correlates of everyday attentional failures in traumatic brain injured and normal subjects.* Neuropsychologia, 1997. **35**(6): p. 747-758.
73. Wendon, J. and J. Coltart, *Intensive care medicine--a review.* Postgrad Med J, 1989. **65**(770): p. 875-80.
74. Hall, J.B., G.A. Schmidt, and L.D.H. Wood, *Principles of Critical Care*. 1998: McGraw-Hill, Health Professions Division.
75. Plate, J.D.J., et al., *Utilisation of Intermediate Care Units: A Systematic Review.* Crit Care Res Pract, 2017. **2017**: p. 8038460.
76. Devlin, J.W., et al., *Clinical Practice Guidelines for the Prevention and Management of Pain, Agitation/Sedation, Delirium, Immobility, and Sleep Disruption in Adult Patients in the ICU.* Crit Care Med, 2018. **46**(9): p. e825-e873.
77. Pisani, M.A., et al., *Sleep in the intensive care unit.* Am J Respir Crit Care Med, 2015. **191**(7): p. 731-8.
78. Pulak, L.M. and L. Jensen, *Sleep in the Intensive Care Unit:A Review.* Journal of Intensive Care Medicine, 2016. **31**(1): p. 14-23.
79. Lewandowska, K., et al., *Sleep Deprivation from the Perspective of a Patient Hospitalized in the Intensive Care Unit—Qualitative Study.* Healthcare, 2020. **8**(3): p. 351.
80. Nilius, G., M. Richter, and M. Schroeder, *Updated Perspectives on the Management of Sleep Disorders in the Intensive Care Unit.* Nat Sci Sleep, 2021. **13**: p. 751-762.
81. Roberts, B. and W. Chaboyer, *Patients' dreams and unreal experiences following intensive care unit admission.* Nursing in Critical Care, 2004. **9**(4): p. 173-180.
82. Rundshagen, I., et al., *Incidence of recall, nightmares, and hallucinations during analgosedation in intensive care.* Intensive care medicine, 2002. **28**: p. 38-43.
83. Laitinen, H., *Patients' experience of confusion in the intensive care unit following cardiac surgery.* Intensive Crit Care Nurs, 1996. **12**(2): p. 79-83.
84. Granja, C., et al., *Patients' recollections of experiences in the intensive care unit may affect their quality of life.* Critical Care, 2005. **9**(2): p. R96.
85. Doig, L. and K. Solverson, *Wanting to Forget: Intrusive and Delusional Memories from Critical Illness.* Case Rep Crit Care, 2020. **2020**: p. 7324185.
86. Miller, T.W. and D.T. Miller, *Dreaming: the impact of life stress events.* Psychiatr Dev, 1989. **7**(4): p. 367-73.
87. Garcia, O., et al., *What goes around comes around: Nightmares and daily stress are bidirectionally associated in nurses.* Stress Health, 2021. **37**(5): p. 1035-1042.
88. Date, A., et al., *Differences between natural sleep and the anesthetic state*. Future Sci OA. 2020 Nov 30:FSO664. doi: 10.2144/fsoa-2020-0149. eCollection 2020 Dec.
89. Stait, M.L., K. Leslie, and R. Bailey, *Dreaming and recall during sedation for colonoscopy.* Anaesth Intensive Care, 2008. **36**(5): p. 685-90.
90. Stollings, J.L., et al., *Delirium in critical illness: clinical manifestations, outcomes, and management.* Intensive Care Medicine, 2021. **47**(10): p. 1089-1103.

Part I

Aspects of Sleep in Central Disorders of Hypersomnolence

2

Intermediate hypocretin-1 cerebrospinal fluid levels and typical cataplexy

their significance in the diagnosis of narcolepsy type 1

Adrienne Elisabeth van der Hoeven[1,2], Rolf Fronczek[1,2], Mink Sebastian Schinkelshoek[1,2], Frederik Willem Cornelis Roelandse[3], Jaap Adriaan Bakker[3], Sebastiaan Overeem[4], Denise Bijlenga[1,2], Gert Jan Lammers[1,2]

[1]Department of Neurology, Leiden University Medical Center, Leiden, the Netherlands, [2]Sleep-Wake Center, Stichting Epilepsie Instellingen Nederland (SEIN), Heemstede, the Netherlands, [3]Department of Clinical Chemistry and Laboratory Medicine Leiden University Medical Center, Leiden, the Netherlands, [4]Sleep Medicine Center, Kempenhaeghe, Heeze, the Netherlands

Sleep. 2022 May

Abstract

Study Objectives

The diagnosis of narcolepsy type 1 (NT1) is based upon the presence of cataplexy and/or a cerebrospinal fluid (CSF) hypocretin-1/orexin-A level ≤110 pg/mL. We determined the clinical and diagnostic characteristics of patients with intermediate hypocretin-1 levels (111-200 pg/mL) and the diagnostic value of cataplexy characteristics in individuals with central disorders of hypersomnolence.

Methods

Retrospective cross-sectional study of 355 people with known CSF hypocretin-1 levels who visited specialized Sleep-Wake Centers in the Netherlands. For n=271, we had full data on cataplexy type ('typical' or 'atypical' cataplexy).

Results

Compared to those with normal hypocretin-1 levels (>200 pg/mL), a higher percentage of individuals with intermediate hypocretin-1 levels had typical cataplexy (75% or 12/16 vs 9% or 8/88, p<.05), and/or met the diagnostic polysomnographic (PSG) and Multiple Sleep Latency Test (MSLT) criteria for narcolepsy (50 vs 6%, p<.001).

Of those with typical cataplexy, 88% had low, 7% intermediate, and 5% normal hypocretin-1 levels (p<.001). Atypical cataplexy was also associated with hypocretin deficiency but to a lesser extent.

A hypocretin-1 cutoff of 150 pg/mL best predicted the presence of typical cataplexy and/or positive PSG and MSLT findings.

Conclusions

Individuals with intermediate hypocretin-1 levels or typical cataplexy more often have outcomes fitting the PSG and MSLT criteria for narcolepsy than those with normal levels or atypical cataplexy. In addition, typical cataplexy has a much stronger association with hypocretin-1 deficiency than atypical cataplexy. We suggest increasing the NT1 diagnostic hypocretin-1 cutoff and adding the presence of clearly-defined typical cataplexy to the diagnostic criteria of NT1.

Statement of significance

The diagnostic value of intermediate CSF hypocretin-1 levels, 111 - 200 pg/mL, has not been established. Neither has the relevance of distinguishing typical and atypical cataplexy been evaluated. Additionally, the currently used cutoff value to diagnose narcolepsy was determined when the current classification (narcolepsy type 1 and 2) was not in place. A (re-)examination of this cutoff value and more information regarding the diagnostic implications of typical cataplexy and clinical characteristics of individuals with intermediate CSF hypocretin-1 levels will improve the diagnostic process of people with suspected narcolepsy.

Introduction

Narcolepsy type 1 (NT1) is a debilitating sleep-wake disorder with a prevalence of 0.02 to 0.06% (1, 2). The clinical presentation involves excessive daytime sleepiness (EDS), disrupted nighttime sleep, and cataplexy. Cataplexy is currently defined in the International Classification of Sleep Disorders (3rd edition, ICSD3) as more than one episode of generally brief (<2 min), usually bilateral symmetric, sudden loss of muscle tone with retained consciousness. The episode is provoked by strong emotion (mainly of a positive nature), and ends with an abrupt return of muscle activity (3). The clinical symptoms of narcolepsy are presumed to be caused by insufficient cerebral hypocretin (also named orexin) transmission. In idiopathic narcolepsy (about 95% of cases in humans) (4), the probable cause is an autoimmune induced loss of hypocretin-producing neurons in the lateral hypothalamus due to a combination of genetic and environmental factors. The loss of these cells leads to low or undetectable hypocretin-1 levels in the CSF (5-7). A CSF hypocretin-1 deficiency (≤ 110 pg/mL) combined with EDS is sufficient for the diagnosis of NT1. As an alternative to the assessment of CSF hypocretin-1 concentration, narcolepsy can be diagnosed by polysomnography (PSG) during the night, followed by a Multiple Sleep Latency Test (MSLT) during the day (3). According to the ICSD3, NT1 is diagnosed when EDS is present for ≥ 3 months, and when the MSLT yields a mean sleep latency of ≤ 8 minutes and ≥ 2 sleep-onset rapid eye movement periods (SOREMPs). A nocturnal SOREMP during the PSG may count as one of the required SOREMPs. When applying these diagnostic criteria, cataplexy must be present.

Mignot et al. (2002) established the currently used CSF hypocretin-1 cutoff value of ≤ 110 pg/mL within a subject group of 274 people with sleep disorders, including 157 with narcolepsy, and 296 controls (healthy controls and individuals with other neurological disorders). This threshold had a sensitivity of 60% and specificity of 98% for diagnosis of ICSD2-defined narcolepsy, while a threshold of > 200 pg/mL had the best sensitivity/specificity ratio for healthy controls versus all other subject samples (8). Subsequent studies have implemented these cutoffs, (9, 10) as has the ICSD3 (3, 11). Consequently, an intermediate CSF hypocretin-1 range of 111-200 pg/mL was created, resulting in a patient subgroup that is more difficult to diagnose, unless there are clear features of typical cataplexy and positive MSLT and PSG findings (12). Moreover, at the time these hypocretin-1 cutoff values were determined, the criteria for the differential diagnosis of NT1 and narcolepsy type 2 (NT2) were not yet established, as narcolepsy was then classified as narcolepsy with or without cataplexy (8) We need to characterize individuals with intermediate CSF hypocretin-1 levels, by re-examining the current hypocretin-1 cutoff value.

The presence of cataplexy is considered pathognomonic for NT1, and early recognition can help to avoid misdiagnoses and reduce diagnostic delay (13, 14). The determination of cataplexy is usually solely based on patient reports, which can be quite difficult to interpret (15, 16). For example, some symptoms seen in syncope, epileptic and psychogenic attacks, may resemble cataplexy (16). Generally, loss of consciousness from the start of the attack indicates that the attacks are not cataplectic. Also, cataplexy attacks sometimes change in frequency and intensity over time (17) and can develop years after EDS onset (18). There is also diversity in the expression of cataplexy, which led to the introduction of the terms 'typical' and 'atypical' cataplexy, which are however not clearly defined (19). Typical cataplexy has been reported to be associated with CSF hypocretin-1 levels ≤ 110 pg/mL in 90-95% of people. Thus, typical cataplexy is hypothesised to be more specific for NT1 than its atypical form (20). The diagnostic significance of intermediate hypocretin-1 levels and of clearly defined typical and atypical cataplexy in the diagnosis of NT1 is yet unresolved. We aimed to: (1) assess the prevalence of intermediate hypocretin-1 levels in people with hypersomnolence complaints; (2) evaluate how clinical aspects (the results of auxiliary investigations as well as cataplexy presence and characteristics) relate to low, intermediate and normal hypocretin-1 levels; (3) examine the diagnostic value of defined typical versus atypical cataplexy; and (4) evaluate the current CSF hypocretin-1 cutoff point for NT1. The results of this study may improve the diagnostic accuracy of NT1.

Methods

In this retrospective cross-sectional study, we analyzed data of individuals who were referred with complaints of hypersomnolence to any of the following Dutch sleep-wake clinics, between October 2001 and December 2019: Leiden University Medical Center (LUMC, n=116), Stichting Epilepsie Instellingen Nederland (SEIN, n=158), and/or Kempenhaeghe (n=81). Individuals were included if their hypocretin-1 CSF level was assessed and if their electronic health record was available.

Hypocretin-1 CSF levels were determined at the LUMC department of Clinical Chemistry and Laboratory Medicine using the radioimmunoassay (RIA) kit of Phoenix Pharmaceuticals (Phoenix Pharmaceuticals Inc, Burlingame, CA, USA) and harmonized using a Stanford reference sample with a known concentration (used to correct for inter-assay variation between RIAs). Hypocretin-1 levels below 75 pg/mL are deemed "undetectable" in this article, as there is currently local consensus at the LUMC that levels below this value cannot be measured reliably.

Individual characteristics (gender and age at time of CSF assessment) and diagnoses and results of auxiliary investigations (HLA DQB1*06:02 positivity, and PSG and MSLT results at time of diagnostic evaluation) were extracted from the electronic health records. In 77% of patients, PSG reports were available, and the parameters Total Sleep Time (TST), Time In Bed (TIB), Sleep Efficiency (SE) and the presence of a nighttime sleep onset-REM period (SOREMP) were extracted. When reports were not available, data could sometimes be extracted from physician notes, referral letters and letters to the family doctor, resulting in night-time SOREMP data being available in 86% of cases. From the MSLT reports, the presence of SOREMPs and the sleep latency (SL) were extracted. Data on cataplexy presence and type were available for people from the LUMC and SEIN clinics (n=271, for 3 individuals from these centers cataplexy data were not available). At SEIN and the LUMC, the MSLT was performed according to Littner et al (21) with 5 nap opportunities starting at 09:00 AM, at least an hour and a half after the termination of nocturnal sleep. The MSLT was preceded by a PSG, usually in an ambulant setting. Individuals were instructed to try to sleep for at least 6 hours. However, this procedure was not followed in all instances, as a substantial subgroup of individuals was referred after the MSLT and PSG were already performed at the referring sleep centers. These registrations were not always repeated, particularly not if pharmacotherapy was started and the earlier registrations were performed in certified centers (according to ESRS

guidelines). Most of these referring centers applied 4 nap opportunities when performing the MSLT. A study by Kawai et al. (2015) (22) indicated that the clinical presentation of narcolepsy differed between ethnicities. Specific data regarding ethnicity were not available; the majority were Caucasians.

The final diagnoses were made by experienced physicians (GJL, RF & SO) based on the ICSD3 criteria. In rare cases of familial narcolepsy (defined as more than one first degree relative with EDS and cataplexy) the diagnosis was sometimes made despite not fulfilling all polysomnographic criteria or in the absence of typical cataplexy. Diagnoses made before the introduction of the ICSD3 were adjusted by applying the ICSD3 criteria.

All individuals were categorized into the following clinical diagnostic categories: NT1, NT2, familial narcolepsy, idiopathic hypersomnia or "other". For the individuals with available information regarding the presence of typical cataplexy, a categorization was also made according to the official ICSD3 criteria for NT1 and NT2.

Table 2.1. Characteristics of typical cataplexy and atypical cataplexy as defined by Lammers et al (2020)

Typical cataplexy	Atypical cataplexy
Meets all of the following ICSD3 criteria for cataplexy:	*Meets one of the following criteria, in addition or contradiction to all other typical cataplexy criteria:*
Bilaterally symmetrical (some asymmetry may be experienced)	Purely unilateral episodes
Generally brief (< 2 min)	Prolonged duration (e.g. > 3 minutes) without remaining precipitant or recent discontinuation of anti-cataplectic medication
Provoked by strong emotion, particularly of positive nature (occasional spontaneous attacks may occur)	No identifiable trigger or only negative emotions as trigger
≥ 1 episode of loss of muscle tone	Hyperacute generalized muscle weakness without build-up over seconds, leading to falls or injuries
Abrupt return of muscle activity after episode	Prolonged recovery (several minutes or longer)
Retained consciousness	Exclusively generalized attacks without history of partial episodes

*Abbreviations: ICSD3, International Classification of Sleep Disorders third edition

Typical and atypical cataplexy

The definitions of typical and atypical attacks are shown in table 2.1. 'Typical' cataplexy was defined as the presence of all phenomena in the current ICSD3 definition as noted in the left column of table 2.1 and in addition the *absence* of atypical characteristics (23) as noted in the right column of table 2.1. If only one of these atypical characteristics was present, cataplexy was defined as 'atypical', as defined by Lammers et al (2020) (23) mainly based on expert opinion and the results from Overeem et al (2011) (19). If more than two atypical characteristics we defined the attacks as non-cataplexic.

Ethics statement

The study was conducted per the Helsinki Declaration as revised in 2013. Due to the retrospective design of this study, a waiver of the requirement for informed consent was obtained from the Medical Ethical Committee of Leiden-Den Haag-Delft (registration number G20.139).

Data availability statement

The datasets generated during and/or analyzed during the current study are not publicly available to protect participant confidentiality but are available from the corresponding author on reasonable request.

Statistical analysis

Frequencies (in numbers or percentages) were used to describe categorical variables, and continuous data were presented using means and standard deviations (SDs), or medians and interquartile range (IQR) depending on the distribution of the data.

Differences in continuous variables between two groups were analyzed using T-tests when normally distributed and the Mann-Whitney-U test when non-normally distributed. Differences between more than two groups were computed using a one-way ANOVA or the Kruskal Wallis test, depending on the distribution of the data. To analyze differences between categorical variables, a Chi-Square test was used: if one or more cells had an expected frequency <5, Fisher's exact test was used instead.

To establish the optimal CSF hypocretin-1 threshold for diagnosis of NT1, we performed multiple receiver operating characteristic (ROC) analyses, using various outcome parameters for NT1. The following four sets of parameters were used: (i) positive MSLT and PSG findings according to the ICSD3 criteria for narcolepsy, (ii) typical cataplexy, (iii) any cataplexy (typical or atypical), (iv) typical cataplexy and/or positive PSG and MSLT findings, and (v) typical cataplexy and positive PSG and MSLT findings. To determine the optimal cutoff for hypocretin-1, we chose the point of the ROC curve as defined by the Index of Union (IU) method, where (1) the sensitivity and specificity are simultaneously close to the AUC value and (2) the difference between sensitivity and specificity is minimal (24).

The data were analyzed using SPSS version 25.0. Pairwise deletion was applied in case of missing data, as automatically performed by SPSS. Graphs were made using GraphPad Prism 8.4.2. A statistical significance level of $\alpha = 0.05$ (2-tailed) was used.

Results

Patients

CSF hypocretin-1 levels and electronic health records were available for 355 individuals. Cataplexy characteristics were available for 271 of them. As shown in figure 2.1, 342 out of 355 individuals (96.3%) were categorized with the clinical diagnosis NT1, NT2, familial narcolepsy, idiopathic hypersomnia or "other". The remaining 13 individuals had no final clinical diagnosis. Two thirds (n=235) were clinically diagnosed with NT1, 4.4% (n=15) with NT2, 1.5% (n=5) with familial narcolepsy, 6.1% (n=21) with idiopathic hypersomnia, and 19.3% (n=66) received a different (or no) sleep diagnosis. Low (≤ 110 pg/mL) hypocretin-1 level was present in 58.9% of people, intermediate level (111-200 pg/mL) in 5.3% and normal hypocretin-1 level (> 200 pg/mL) in 35.8%. The medians and IQRs of the CSF hypocretin-1 levels of the whole group and the diagnostic groups (based on clinical diagnoses) are displayed in figure 2.1.

In the group with information on the presence of typical cataplexy a categorization according to the ICSD3 criteria was also made. Out of 271, 175 were classified with NT1 and 5 as with NT2 according to official criteria. Respectively 19 and 20 cases could not be classified due to missing data.

As shown in table 2.2, 53% were male and the median age was 28 years.

Hypocretin-1 level groups

An overview of the diagnostic outcomes of patients with low, intermediate and normal hypocretin-1 levels is shown in table 2.2.

In the group with an intermediate hypocretin-1 level there was a significantly higher percentage of HLA positivity for DQB1*0602 compared to the normal hypocretin-1 group (94.7 vs 36.8%, p=.001). The MLST results of the intermediate hypocretin-1 group showed that their sleep latency was longer than the low hypocretin-1 group (6.2 vs 3.3 minutes; p<.005), and the frequency of SOREMPs was higher as compared to the normal hypocretin-1 group. The PSG results showed that nighttime SOREMPs occurred significantly more often in the intermediate group as compared to the normal hypocretin-1 group (46.7 vs 3.7%, p<.001).

In the group with intermediate hypocretin-1 levels there were significantly more NT1 diagnoses according to the ICSD3 criteria than in those with normal hypocretin-1 levels (41.7 vs 0%, p<.001), while the percentage was not significantly different from those with low

hypocretin-1 levels (64.8%). This means that only 83 out of 128 individuals with low hypocretin-1 levels and available MSTL and PSG findings received the diagnosis of NT1 based solely on their MSLT and PSG results and cataplexy presence. Non-NT1 diagnoses based on ICSD3 criteria were adjusted to NT1 as a result of an established CSF hypocretin-1 deficiency in 5 out of 10 individuals thought to have NT2, and 45 out of 121 (35.2%) individuals not meeting the ICSD3 criteria for NT1, 36 (80.0%) of whom did have typical cataplexy.

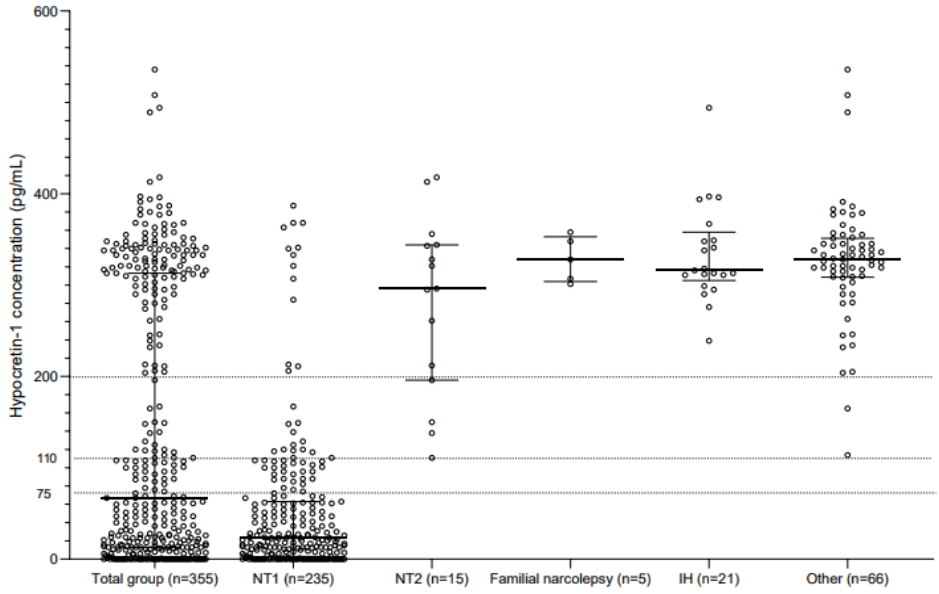


Figure 2.1. CSF hypocretin-1 levels of the total group and the clinical diagnostic groups, N=355. The bars represent median and interquartile ranges. The dashed lines represent the lower reliable hypocretin-1 detection limit (75 pg/mL), and the 'intermediate' hypocretin-1 range (111 - 200 pg/mL).
*Abbreviations: NT1, narcolepsy type 1; NT2, narcolepsy type 2; IH, idiopathic hypersomnia

In addition to the analyses using diagnoses according to the ICSD-3 criteria that can be seen in table 2.2, an analysis of the clinical diagnoses of individuals with intermediate hypocretin-1 levels was performed. More clinical diagnoses than ICSD3 diagnoses were available as some of the MSLT and PSG results, needed to diagnose according to the ICSD3 criteria, were missing. Of the 19 patients with intermediate hypocretin-1 levels, 13 patients received the clinical (i.e. not necessarily ICSD3 compliant) diagnosis NT1 (of whom one had a secondary narcolepsy diagnosis as it was likely caused by angiitis), four had the clinical diagnosis NT2 and two were not diagnosed with a disorder of hypersomnolence (EDS e.c.i., i.e. EDS complaints without explanatory diagnosis).

Table 2.2. Demographic characteristics, HLA, MSLT, PSG, and diagnostic characteristics of the total group and of the groups with low, intermediate and normal CSF hypcretin-1 levels (N=355)

	Total (N=355)	Low: ≤ 110 pg/mL (n=209)	Intermediate: 111-200 pg/mL (n=19)	Normal: > 200 pg/mL (n=127)	Test statistic	P-value
Age at lumbar puncture, n	28.0 (20-40), 312	27.3 (18-39), 184	34.0 (23-44), 18	29.7 (21-41), 110	H=5.538	.063
Gender, count/n (% male)	188/355 (53.0)	114/209 (54.5)	13/19 (68.4)	61/127 (48.0)	χ^2=3.272	.204
HLA positive, count/n (%)	225/285 (78.9)	175/179 (97.8)	18/19 (94.7)	32/87 (36.8)	FET	<.001[a]
MSLT results:	n=302	n=180	n=15	n=107		
- Sleep latency in minutes, n	5.0 (2.5-8.6), 280	3.3 (1.9-5.4), 161	6.2 (4.1-10.3), 15	8.5 (5.1-13.9), 104	H=83.513	<.001[b]
- Number of SOREMPs, n	2 (0-3), 300	3 (1-4), 180	2 (0-3), 14	0 (0-0), 106	H=98.485	<.001[a]
- ≥2 SOREMPs, count/n (%)	158/302 (52.3)	134/180 (74.4)	9/15 (60.0)	15/107 (14.0)	χ^2=98.596	<.001[a]
PSG results:	n=302	n=179	n=15	n=108		
- TIB in minutes (mean ±SD), n	504 ±95, 274	505 ±100, 155	501 ±72, 15	505 ±90, 104	F=.009	.991
- TST in minutes (mean ±SD), n	437 ±86, 280	437 ±92, 159	441 ±89, 15	435 ±78, 106	F=.028	.973
- SE %, n	90.1 (84-94), 284	89.8 (84-94), 162	88.7 (84-95), 15	91.0 (83-95), 107	H=.385	.825
- SOREMP present, count/n (%)	91/302 (30.1)	80/179 (44.7)	7/15 (46.7)	4/108 (3.7)	χ^2=98.596	<.001[a]
Diagnosis:	n=252	n=170	n=12	n=70		
Narcolepsy diagnoses (ICSD3) per group, count/n (%)	180/251 (71.7)	170/170 (100.0)	6/12 (50.0)	4/69 (5.8)	FET	<.001[c]
Of whom the diagnosis is based on:						
- NT1 (ICSD3 criteria excl. hypocretin-1 measurement), count/n (%)	88/128 (42.1)	83/128 (64.8)	5/12 (41.7)	0/69 (0.0)	χ^2=77.332	<.001[a]
- NT1 (ICSD3), count/n (%)	175/252 (69.4)	170/170 (100.0)	5/12 (41.7)	0/70 (0.0)	FET	<.001[c]
- NT2 (ICSD-3 criteria excl. hypocretin-1 measurement), count/n (%)	10/209 (4.8)	5/128 (3.9)	1/12 (8.3)	4/69 (5.8)	FET	0.450
- NT2 (ICSD3), count/n (%)	5/251 (2.0)	0/170 (0.0)	1/12 (8.3)	4/69 (5.8)	FET	.003[d]

*Median (IQR) is used unless specified otherwise

*Abbreviations: FET, Fisher's exact test; MSLT, Multiple Sleep Latency Test; SOREMP, sleep-onset rapid eye movement period; PSG, polysomnography; TIB, time in bed; TST, total sleep time; SE, sleep efficiency; NT1, narcolepsy type 1; NT2, narcolepsy type 2; ICSD3, International Classification of Sleep Disorders third edition, patients meet the following criteria: sleepiness >3 months, ≥2 SOREMPs, sleep latency ≤8 minutes and/or cerebrospinal fluid hypocretin-1 ≤110 pg/mL and/or cataplexy (in the case of NT1)"

[a] significant differences between the normal hypocretin-1 group and the low and intermediate groups; [b] the difference is significant between the low hypocretin-1 group and intermediate and normal groups; [c] the difference is significant between all groups; [d] the difference is significant between the low and normal groups

The five adults with familial narcolepsy (4 females, 1 male) all had normal hypocretin-1 levels. As familial narcolepsy has an atypical presentation, these individuals did not meet all formal PSG and MSLT criteria for narcolepsy type 1 or 2.

Cataplexy presence

The presence and type of cataplexy per hypocretin-1 level group is shown in table 2.3. Of the five people with familial narcolepsy, four had cataplexy (two typical and two atypical). The prevalence of cataplexy (typical and atypical) differed significantly between all groups, with the highest prevalence in those with low, and lowest prevalence in those with normal hypocretin-1 levels (p<.001). The prevalence of typical cataplexy was significantly higher in those with low hypocretin-1 levels (88.6%) as well as those with intermediate hypocretin-1 levels (75%), compared with those with normal hypocretin-1 levels (9.1%, p<.001). When excluding familial narcolepsy cases the percentage of patients with normal hypocretin-1 levels and typical cataplexy dropped to 7.0%.

Table 2.3. Comparison of cataplexy among individuals with low, intermediate and normal CSF hypcretin-1 levels (n=271)

	Total group (n=271)	Low: ≤ 110 pg/mL (n=167)	Intermediate: 111-200 pg/mL n=(16)	Normal: > 200 pg/mL (n=88)	Test statistic	P-value
Cataplexy, count/n (%)	184/271 (67.9)	157/167 (94.0)	12/16 (75.0)	15/88 (17.0) *4 familial narcolepsy*	χ^2=157.020	<.001[a]
- Typical, count/n (%)	167/271 (61.6)	148/167 (88.6)	12/16 (75.0)	8/88 (9.1) *2 familial narcolepsy*	χ^2=155.936	<.001[b]
- Atypical, count/n (%)	17/271 (6.3)	9/167 (5.4)	0/16 (0.0)	7/88 (8.0) *2 familial narcolepsy*	FET	.634

*[a] the difference is significant between all groups; [b] significant differences between the normal hypocretin-1 group and low and intermediate groups exist

The comparisons among people with no cataplexy, typical cataplexy and atypical cataplexy are shown in table 2.4. Most (88.1%) of the people with typical cataplexy had low hypocretin-1 levels compared to 56.3% of those with atypical cataplexy and 11.5% of those without cataplexy (p<.01). Figure 2.2A also shows the distribution of hypocretin-1 levels per cataplexy type, clearly showing a relationship between typical cataplexy and lower hypocretin-1 levels. The distributions of hypocretin-1 levels per cataplexy type of only HLA DQB1*0602 negative patients are displayed in figure 2.2B. None of the HLA negative individuals without cataplexy (n=21) had a hypocretin-1 level below 200 pg/mL.

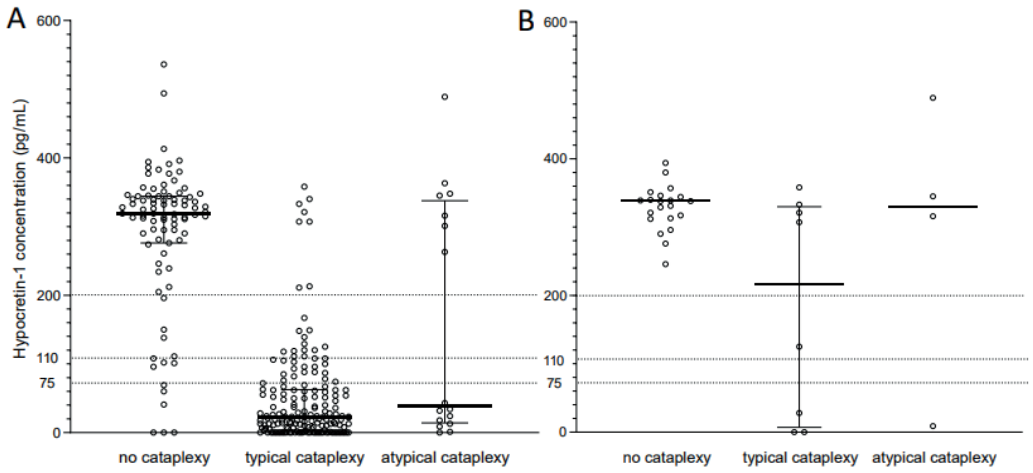


Figure 2.2. Cerebrospinal fluid hypocretin-1 concentration categorized by cataplexy category of: (A) all included individuals (n=271), (B) only HLA- individuals (n=33). Median and interquartile range are displayed for each category. The dashed lines represent the lower reliable hypocretin-1 detection limit (75 pg/mL), and the 'intermediate' hypocretin-1 range (111-200 pg/mL).

The presence of HLA DQB1*0602 positivity was significantly higher in those with typical cataplexy (94%) as compared to those without cataplexy (61%) and those with atypical cataplexy (73%; p<.05).

Those with typical cataplexy had a significantly shorter sleep latency and higher number of SOREMPs during the MSLT than those without cataplexy (p<.001). The prevalence of nighttime SOREMPs during the PSG was significantly higher in those with typical cataplexy (47%) than in those without cataplexy (9.6%, p<.001). Other PSG parameters did not differ between groups.

The percentage of people with a NT1 diagnosis according to the ICSD3 criteria differed significantly between all groups, with 93.9% of people with typical cataplexy meeting the criteria for NT1, 64.3% of people with atypical cataplexy and 13.9% of those without cataplexy meeting the criteria for NT1 (p<.001). When not taking hypocretin-1 measurements into account when making diagnoses, the percentage of people meeting the NT1 ICSD3 criteria without cataplexy became 0%, differing significantly from those with typical or atypical cataplexy (p<.001). Moreover, in this instance, the percentage of people meeting the NT1 ICSD3 criteria was lower in those with typical and those with atypical cataplexy.

Table 2.4. Comparison of characteristics in people with no cataplexy, cataplexy and atypical cataplexy (n=271)

n=271	No cataplexy (n=87)	Typical cataplexy (n=168)	Atypical cataplexy (n=16)	Test statistic	P-value
Age at lumbar puncture, n	29.0 (21-42), 75	27.8 (18-40), 142	26.6 (15-49), 14	H=.995	.608
Gender, count/n (% male)	46/87 (52.9)	88/168 (52.4)	8/16 (50.0)	χ^2=.045	1.000
Hypocretin-1 level:				FET	<.001[a]
- Low, count/n (%)	10/87 (11.5)	148/168 (88.1)	9/16 (56.3)		
- Intermediate, count/n (%)	4/87 (4.6)	12/168 (7.1)	0/16 (0.0)		
- Normal, count/n (%)	73/87 (83.7)	8/168 (4.8)	7/16 (43.8)		
HLA+, count/n (%)	33/54 (61.1)	134/142 (94.4)	11/15 (73.3)	FET	<.001[b]
MSLT (n=222)	n=71	n=138	n=13		
Sleep latency in minutes, n	8.2 (4.9-13.2), 69	4.0 (2.0-6.4), 121	5.9 (4.0-8.7), 13	H=38.763	<.001[c]
Number of SOREMPs, n	0.0 (0-1), 70	3.0 (1-4), 137	1.0 (0-4), 13	H=53.636	<.001[c]
≥2 SOREMPs, count/n (%)	15/71 (21.1)	96/138 (69.6)	6/13 (46.2)	χ^2=44.364	<.001[c]
PSG (n=224)	n=73	n=137	n=14		
TIB in minutes (mean ±SD) , n	503 ±109, 71	490 ±94, 117	509 ±78, 10	F=.474	.623
TST in minutes (mean ±SD) , n	426 ±82, 70	423 ±92, 120	474 ±111, 13	F=1.856	.159
SE %, n	87.0 (82-94), 73	89.7 (83-94), 124	89.9 (81-95), 11	H=.475	.789
SOREMP, count/n (% present)	7/73 (9.6)	65/137 (47.4)	4/14 (28.6)	FET	<.001[c]
Diagnosis (n=249)	n=72	n=163	n=14		
NT1 (ICSD3 criteria excl. hypocretin-1 measurement), count/n (%)	0/68 (0.0)	82/1278 (64.1)	6/13 (46.2)	χ^2=74.858	<.001[d]
NT1 (ICSD3), count/n (%)	10/72 (13.9)	153/163 (93.9)	9/14 (64.3)	χ^2=149.698	<.001[e]
NT2 (ICSD3 criteria excl. hypocretin-1 measurement), count/n (%)	10/68 (14.7)	0/128 (0.0)	0/13 (0.0)	FET	<.001[c]
NT2 (ICSD3), count/n (%)	5/71 (7.0)	0/163 (0.0)	0/14 (0.0)	FET	.004[c]

*Median (IQR) is used unless specified otherwise

*Abbreviations: FET, Fisher's exact test; MSLT, Multiple Sleep Latency Test; SOREMP, sleep-onset rapid eye movement period; PSG, polysomnography; TIB, time in bed; TST, total sleep time; SE, sleep efficiency; NT1, narcolepsy type 1; NT2, narcolepsy type 2; ICSD3, International Classification of Sleep Disorders third edition, patients meet the following criteria: sleepiness >3 months, ≥2 SOREMPs, sleep latency ≤8 minutes and/or cerebrospinal fluid hypocretin-1 ≤110 pg/mL and/or cataplexy (in the case of NT1)"

* [a] significant between all groups excluding the intermediate hypocretin-1 level row; [b] significant between the typical cataplexy and the no and atypical cataplexy groups; [c] significant between the typical cataplexy group and no cataplexy group; [d] significant between the no cataplexy and the typical and atypical cataplexy groups; [e] significant between all groups

Diagnostic value of typical cataplexy

The diagnostic value of typical cataplexy versus atypical and no cataplexy to determine a CSF hypocretin-1 level ≤ 110 pg/mL has a sensitivity of 88.6% with a specificity of 80.8% (n=271). The presence of typical cataplexy results in a positive predictive value (PPV) of 88.1% and a negative predictive value (NPV) of 81.6% when predicting the presence of a low hypocretin-1 level (≤ 110 pg/mL).

The sensitivity increases (94.0%) while the specificity decreases (74.0%) when using cataplexy in general (typical and atypical) to test for hypocretin-1 deficiency (PPV is 85.3% and NPV is 88.5%).

In contrast, when positivity for the ICSD3 PSG and MSLT criteria is used to predict hypocretin-1 deficiency, the sensitivity and specificity are 67.9% and 86.6% (n=213), and the PPV is 89.0% and NPV is 62.8%.

Characteristics of atypical cataplexy

The criterion of atypical cataplexy most often seen was the lack of an identifiable trigger or only negative emotions being the trigger for a cataplexy attack. 7 out of the 16 individuals suspected of having atypical cataplexy displayed this criterion, with 5 of those having no identifiable trigger and 2 having only negative emotions as a trigger.

Atypical features did not significantly differ between those without eventual diagnosis of narcolepsy (n=4) and those with NT1 or familial narcolepsy (n=12, p=.608), nor did they significantly differ between hypocretin-1 range groups (p=.927).

Establishing the optimal hypocretin-1 level threshold

A CSF hypocretin-1 threshold of 55.0 pg/mL was optimal to determine whether patients had positive PSG and MLST findings (i), with a sensitivity of 78.0% and a specificity of 77.9%. The area under the curve (AUC) was .837 (fig. 2.3A, p<.001). When using the presence of typical cataplexy versus atypical or no cataplexy as diagnostic for narcolepsy (ii), a threshold of 101.5 pg/mL was found with a specificity of 83.5% and a sensitivity of 85.1%, with an AUC of .894 (fig. 2.3B, p<.001). A hypocretin-1 threshold of 119.5 pg/mL was optimal to determine cataplexy (typical and atypical) presence (iii), with a sensitivity of 87.5% and specificity of 87.4% and an AUC of .915 (fig. 2.3C, n=271, p<.001). Lastly, the outcome parameter perhaps closest to the clinical practice of diagnosing narcolepsy – typical cataplexy and/or positive PSG and MLST findings (iv) – resulted in an optimal cutoff value of 149.4 pg/mL, with a sensitivity of 93.0% and a specificity of 92.3%. In this case, the AUC was .952 (fig. 2.3D, n=250, p<.001). Notably, when applying the outcome parameter typical cataplexy and positive PSG and MSLT findings (v) the found threshold was far lower, namely 40.5 pg/mL with both a sensitivity and a specificity of 78.0% and an AUC of .846.

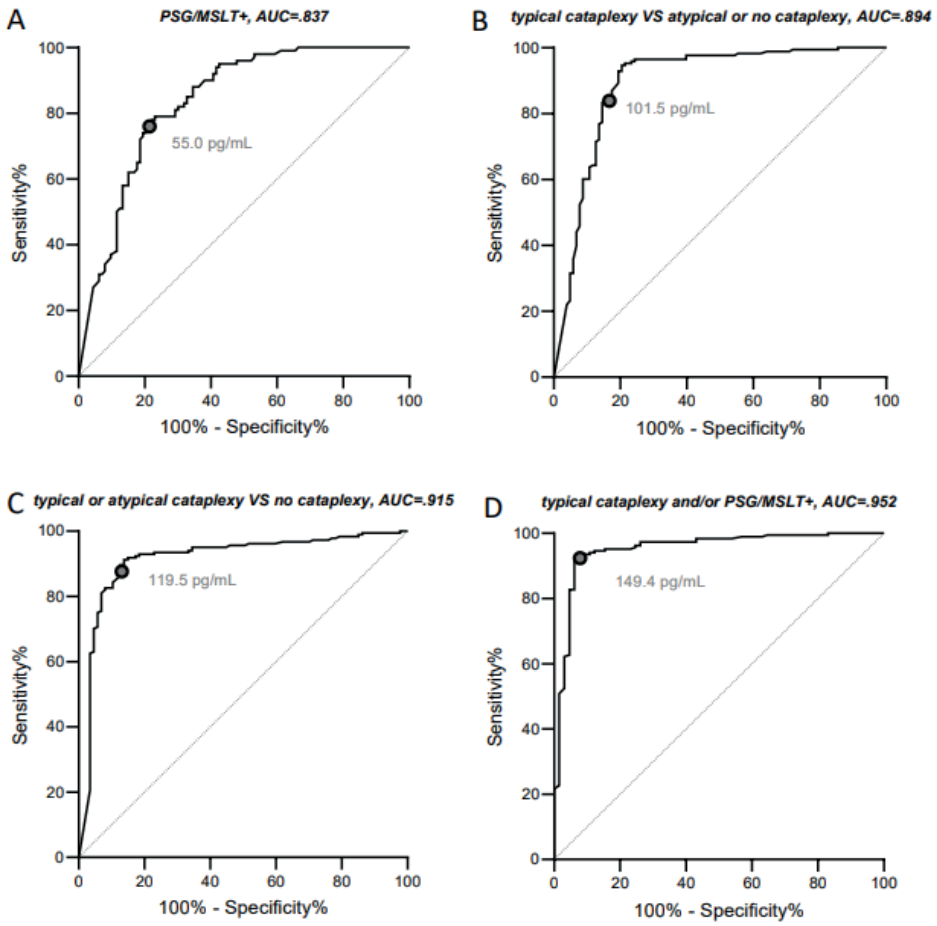


Figure 2.3. ROC curves of CSF hypocretin-1 concentrations for the presence of: (A) n=223; positive polysomnography (PSG) and Multiple Sleep Latency Test (MSLT); (B) n=271; typical cataplexy (vs atypical and no cataplexy); (C) n=271; cataplexy (atypical and typical); (D) n=250; the combination of typical cataplexy and/or positive PSG and MSLT findings. Optimal cut-off values are shown in grey.

Discussion

We examined the significance of intermediate CSF hypocretin-1 values and presence of typical cataplexy, using
historic data of people with EDS from highly specialized sleep-wake centers. Our results show that: (1) only 5.3% had hypocretin-1 levels in the intermediate range; (2) In general, individuals with intermediate hypocretin-1 levels had more features (both of cataplexy presence and auxiliary findings) associated with NT1 than those with normal hypocretin-1 levels (> 200 pg/mL); (3) When categorizing groups based on cataplexy type, we found that those with typical cataplexy had more positive diagnostic findings for NT1 than those with atypical or no cataplexy. Compared to cataplexy in general (atypical and typical cataplexy), typical cataplexy has a higher specificity, but a lower sensitivity for NT1 (as determined by hypocretin-1 ≤ 110 pg/mL). (4) We found that a higher cutoff value for hypocretin-1 is needed to increase diagnostic accuracy for NT1. These results suggest a re-appraisal of both the diagnostic value of typical cataplexy and of the current hypocretin-1 threshold values in narcolepsy.

The diagnostic value of typical cataplexy

People with typical cataplexy more often fulfill the diagnostic criteria for NT1 and more often have low hypocretin-1 levels than those with atypical or no cataplexy. When using typical cataplexy as a predictor for hypocretin-1 deficiency, a sensitivity of 88% and a specificity of 81% was found. In short, typical cataplexy is a good predictor of NT1 and should therefore have more weight when diagnosing NT1.

While the term “typical cataplexy” (as opposed to cataplexy in general) has been used in scientific research, including in the article from Mignot et al (2002) (8) on which the current hypocretin-1 cutoff is based, it generally has not been defined clearly. We clearly define both typical and atypical cataplexy and at the same time show the usefulness of typical cataplexy to diagnose NT1.

Individuals with intermediate CSF hypocretin-1 levels

There were only 19 people (5.3%) with an intermediate CSF hypocretin-1 level. Of these, 17 were diagnosed with either NT1 or NT2 based on clinical characteristics (i.e. they not necessarily met all ICSD3 criteria for NT1 or NT2, only 40% did). These individuals had a median MSLT sleep latency ≤ 8 minutes and significantly more SOREMPs (during PSG and

MSLT) than those with normal hypocretin-1 levels. This suggests that the current diagnostic criteria, the MSLT and PSG findings and the hypocretin-1 cutoff of 110 pg/mL, are insufficient accurately to diagnose individuals who have hypocretin-1 levels in the intermediate range. A higher CSF hypocretin-1 cutoff may mitigate this problem. Visual inspection of the range of hypocretin-1 levels in our sample (see figure 2.1) would lend support to this option. Two distinct clusters can be distinguished, with a lower cluster that ranges well beyond the current cutoff of 110 pg/mL. Other studies have also suggested a higher cutoff. Andlauer et al. (2012) (9) found an optimal cutoff CSF hypocretin-1 level for narcolepsy without cataplexy of 200 pg/mL rather than 110 pg/mL, with a high specificity of 99% but a low sensitivity of 33%. A similar conclusion, that a higher cutoff value may be feasible to determine hypocretin-1 deficiency, was drawn by Heier et al (2007) (25).

Diagnosing patients with intermediate hypocretin-1 levels is further complicated by variability in the determination of hypocretin-1 concentrations using Phoenix Pharmaceuticals radioimmunoassay kits. Inter-assay variability is quite high when not corrected using a reference sample. This could cause a patient to fall within a different hypocretin-1 range category (i.e. intermediate instead of low). Thus, when evaluating hypocretin-1 levels it should be made sure the concentration has been corrected.

The optimal CSF hypocretin-1 cutoff

The results of our own ROC analyses varied depending on the chosen outcome parameter used as substitute or "gold standard" for the NT1 diagnosis. At the moment, the closest to a gold standard is a CSF hypocretin-1 level ≤ 110 pg/mL. As this outcome measure cannot be used to determine the optimal CSF hypocretin-1 threshold value, other outcome parameters were chosen based on the ICSD3 criteria.

We found that using positive PSG and MSLT findings results in a cutoff value of 55 pg/mL (sensitivity and specificity 78%), far lower than the currently used cutoff value. This, however, does not mean that a 55 pg/mL threshold is suitable to diagnose narcolepsy. In our study 29% of individuals with a CSF hypocretin-1 level ≤ 110 pg/mL did not meet the PSG and MSLT criteria for narcolepsy. It seems that PSG and MSTL findings tend to be especially positive in people with a more severely decreased CSF hypocretin-1 level, resulting in a lower cutoff value, while we are also interested in those with intermediate levels. Thus, positive PSG and MSLT findings alone may not be the best measure of narcolepsy.

When we use cataplexy (typical or both typical and atypical) as outcome measure, the optimal cutoff value (101 pg/mL) comes closer to the one currently in use. However, just as is the case with using positive PSG and MSLT findings as outcome measure, typical cataplexy is not the gold standard for NT1. In spite of cataplexy being highly pathognomonic for NT1, not all people with NT1 have cataplexy (6% of people in our study with CSF hypocretin-1 ≤110 pg/mL had no cataplexy). It is not uncommon for the people to develop cataplexy a couple of years after their diagnosis (18). We had follow up data of 10 out of 14 patients who had a low hypocretin-1 level despite the absence of cataplexy. Of these, 6 developed typical cataplexy at a later stage, emphasizing the strong correlation of typical cataplexy and hypocretin-1 deficiency. In addition, 12% of patients with typical cataplexy had a CSF hypocretin-1 level > 110 pg/mL. It should however be mentioned that the area under the curve is considerably higher when typical cataplexy or cataplexy in general is used compared to when PSG and MSLT findings are used as outcome measure (see figure 2.3). This would suggest that measured hypocretin-1 level is a better predictor of the presence of (typical) cataplexy than of positive PSG and MSLT findings.

The diagnosis of NT1 is complicated and a perfect gold standard does not exist. Given that no single option is a perfect substitute which can be used to predict hypocretin-1 deficiency, we consider the combination of typical cataplexy and/or positive PSG and MSLT findings the best approximation of a NT1 diagnosis to be used for this purpose. We found that, with an area under the curve of 0.952, a hypocretin-1 cutoff value of <150 pg/mL best predicts this combination.

Some limitations should be mentioned. First of all, we used data from people who visited a sleep-wake clinic because of complaints of EDS and with suspected narcolepsy. As such, the results are not generalizable to the general population. Moreover in some people, the CSF hypocretin-1 levels were determined as part of scientific research. These individuals already had a narcolepsy diagnosis and would normally not have undergone a lumbar puncture. Thus, it is possible that our analysis is influenced by inclusion bias given that the a priori probability of a narcolepsy diagnosis was higher than would be the case in clinical practice. Secondly, because the number of individuals with NT1 is far higher than those with NT2 in our population, this may impact the hypocretin-1 cut-off we found. However, this reflects the prevalence of these disorders in Europe. Thus, we believe the result to be relevant in clinical practice. Lastly, as this is a retrospective study, we were not able to use the newly defined hypocretin-1 cutoff value and use of typical cataplexy in a test sample to determine actual

positive and negative predictive values of NT1. Future studies with a new sample should evaluate this in prospectively.

Conclusion

We come to several conclusions that could improve the diagnostic process of NT1, especially where it concerns individuals with narcolepsy symptoms and intermediate hypocretin-1 levels. Firstly, given that all non-cataplexic HLA negative patients in our sample had CSF hypocretin-1 levels above 200 pg/mL, we conclude that determining hypocretin-1 levels in patients with these characteristics is largely redundant.

Secondly, the current diagnostic process is mainly focused on the hypocretin-1 CSF concentration combined with PSG and MSLT results. The presence of typical cataplexy has a higher diagnostic than PSG and MSLT findings. Correct identification of typical cataplexy thus improves diagnostic accuracy. Lastly, modification of the currently used hypocretin-1 cutoff should be considered. We suggest a new cutoff of < 150 pg/mL. In conclusion, adding typical cataplexy to the diagnostic criteria, preventing unnecessary lumbar punctures and altering the hypocretin-1 cutoff value would enhance the diagnostic accuracy and patient care in narcolepsy.

Disclosure statement

Financial disclosure: S. Overeem consulted for Bioprojet, Jazz Pharmaceuticals, UCB Pharma and Takeda, all paid to institution, and not related to present work. R. Fronczek consulted for and/or received lecture fees from Bioprojet, Takeda, Lundbeck, TEVA, Lilly, Novartis and Allergan, and grant support from Jazz Pharmaceuticals and Bioprojet, all not related to the present work. D. Bijlenga received a reimbursement for a course from Bioprojet, not related to the present work. G.J. Lammers consulted for Bioprojet, Jazz Pharmaceuticals, UCB Pharma, NLS and Takeda, all paid to institution, and not related to present work. He also served as member of advisory boards on narcolepsy: Bioprojet, Jazz Pharmaceuticals, UCB Pharma, NLS. The other authors have indicated no potential financial conflicts of interest.

Non-financial disclosure: The authors declare that there is no non-financial conflict of interest.

References

1. Ohayon MM, Priest RG, Zulley J, Smirne S, Paiva T. Prevalence of narcolepsy symptomatology and diagnosis in the European general population. Neurology. 2002;58(12):1826-33.
2. Longstreth WT, Jr., Koepsell TD, Ton TG, Hendrickson AF, van Belle G. The Epidemiology of Narcolepsy. Sleep. 2007;30(1):13-26.
3. American Academy of Sleep M. The international classification of sleep disorders2014.
4. Hartwig G, Harsh J, Ripley B, Nishino S, Mignot E. Low cerebrospinal fluid hypocretin levels found in familial narcolepsy. Sleep medicine. 2001;2(5):451-3.
5. Liblau RS, Vassalli A, Seifinejad A, Tafti M. Hypocretin (orexin) biology and the pathophysiology of narcolepsy with cataplexy. The Lancet Neurology. 2015;14(3):318-28.
6. Mignot E. Genetic and familial aspects of narcolepsy. Neurology. 1998;50(2 Suppl 1):S16.
7. Mahlios J, De la Herrán Arita AK, Mignot E. The Autoimmune Basis of Narcolepsy. Current opinion in neurobiology. 2013;23.
8. Mignot E, Lammers GJ, Ripley B, Okun M, Nevsimalova S, Overeem S, et al. The Role of Cerebrospinal Fluid Hypocretin Measurement in the Diagnosis of Narcolepsy and Other Hypersomnias. Archives of Neurology. 2002;59(10):1553-62.
9. Andlauer O, Moore Ht, Hong S-C, Dauvilliers Y, Kanbayashi T, Nishino S, et al. Predictors of hypocretin (orexin) deficiency in narcolepsy without cataplexy. Sleep. 2012;35(9):1247-55F.
10. Dauvilliers Y, Baumann C, Carlander B, Bischof M, Blatter T, Lecendreux M, et al. CSF hypocretin-1 levels in narcolepsy, Kleine-Levin syndrome, and other hypersomnias and neurological conditions. Journal of neurology, neurosurgery, and psychiatry. 2004;74:1667-73.
11. Ruoff C, Rye D. The ICSD-3 and DSM-5 guidelines for diagnosing narcolepsy: clinical relevance and practicality. Current Medical Research and Opinion. 2016;32(10):1611-22.
12. Fernandes G, Cremaschi R, Tufik S, Coelho F. Grey zone between narcolepsy type 1 and type 2. Arquivos de Neuro-Psiquiatria. 2017;75:682-3.
13. Taddei RN, Werth E, Poryazova R, Baumann CR, Valko PO. Diagnostic delay in narcolepsy type 1: combining the patients' and the doctors' perspectives. Journal of sleep research. 2016;25(6):709-15.
14. Bhattarai J, Sumerall S. Diagnostic Delay of Narcolepsy: Contributing Factors and Implications for Clinicians. Sleep and Vigilance. 2018;2.
15. Dauvilliers Y, Siegel JM, Lopez R, Torontali ZA, Peever JH. Cataplexy—clinical aspects, pathophysiology and management strategy. Nature Reviews Neurology. 2014;10(7):386-95.
16. Pillen S, Pizza F, Dhondt K, Scammell TE, Overeem S. Cataplexy and Its Mimics: Clinical Recognition and Management. Current Treatment Options in Neurology. 2017;19(6):23.
17. Lopez R, Barateau L, Evangelista E, Chenini S, Robert P, Jaussent I, et al. Temporal Changes in the Cerebrospinal Fluid Level of Hypocretin-1 and Histamine in Narcolepsy. Sleep. 2017;40(1).
18. Pérez-Carbonell L, Leschziner G. Clinical update on central hypersomnias. Journal of thoracic disease. 2018;10(Suppl 1):S112-S23.
19. Overeem S, van Nues SJ, van der Zande WL, Donjacour CE, van Mierlo P, Lammers GJ. The clinical features of cataplexy: A questionnaire study in narcolepsy patients with and without hypocretin-1 deficiency. Sleep medicine. 2011;12(1):12-8.
20. Bourgin P, Zeitzer JM, Mignot E. CSF hypocretin-1 assessment in sleep and neurological disorders. The Lancet Neurology. 2008;7(7):649-62.
21. Littner MR, Kushida C, Wise M, Davila DG, Morgenthaler T, Lee-Chiong T, et al. Practice parameters for clinical use of the multiple sleep latency test and the maintenance of wakefulness test. Sleep. 2005;28(1):113-21.
22. Kawai M, O'Hara R, Einen M, Lin L, Mignot E. Narcolepsy in African Americans. Sleep. 2015;38(11):1673-81.
23. Lammers GJ, Bassetti CLA, Dolenc-Groselj L, Jennum PJ, Kallweit U, Khatami R, et al. Diagnosis of central disorders of hypersomnolence: A reappraisal by European experts. Sleep medicine reviews. 2020;52:101306.
24. Unal I. Defining an Optimal Cut-Point Value in ROC Analysis: An Alternative Approach. Computational and Mathematical Methods in Medicine. 2017;2017:3762651.
25. Heier MS, Evsiukova T, Vilming S, Gjerstad MD, Schrader H, Gautvik K. CSF hypocretin-1 levels and clinical profiles in narcolepsy and idiopathic CNS hypersomnia in Norway. Sleep. 2007;30(8):969-73.

3

Hypocretin-1 measurements in cerebrospinal fluid using radioimmunoassay

within and between assay reliability and limit of quantification

Adrienne Elisabeth van der Hoeven[1,2], Kevin van Waaij[1], Denise Bijlenga[1,2], Frederik Willem Cornelis Roelandse[3], Sebastiaan Overeem[4], Jaap Adriaan Bakker[3], Rolf Fronczek[1,2], Gert Jan Lammers[1,2]

[1]Department of Neurology Leiden University Medical Center, Leiden, the Netherlands, [2]Sleep-Wake Center, Stichting Epilepsie Instellingen Nederlands (SEIN), Heemstede, the Netherlands, [3]Department of Clinical Chemistry and Laboratory Medicine Leiden University Medical Center, Leiden, the Netherlands, [4]Sleep Medicine Center, Kempenhaeghe, Heeze, the Netherlands

Sleep. 2022 July

Abstract

Study objectives

The most sensitive and specific investigative method for the diagnosis of narcolepsy type 1 (NT1) is the determination of hypocretin-1 (orexin A) deficiency (≤110 pg/mL) in cerebrospinal fluid using a radioimmunoassay (RIA). We aimed to assess the reliability of the Phoenix Pharmaceuticals hypocretin-1 RIA, by determining the lower limit of quantification (LLOQ), the variability around the cut-off of 110 pg/mL and the inter-and intra-assay variability.

Methods

Raw data of 80 consecutive hypocretin-1 RIAs were used to estimate the intra- and inter-assay coefficient of variation (CV). The LLOQ was established, defined as the lowest converted concentration with a CV<25%; the conversion is performed using a harmonization sample which is internationally used to minimize variation between RIAs.

Results

The mean intra-assay CV was 4.7%, while the unconverted inter-assay CV was 28.3% (18.5% excluding 2 outliers) and 7.5% when converted to international values. The LLOQ was determined as 27.9 pg/mL. The intra-assay CV of RIAs with lower specific radioactive activity showed a median of 5.6% (n=41, range 1.6-17.0%), which was significantly higher than in RIAs with higher specific activity (n=36; median 3.2%, range 0.4-11.6%, p=.013). The CV around the 110 pg/mL cut-off was <7%.

Conclusions

Hypocretin-1 RIAs should always be harmonized using standard reference material. The specific activity of an RIA has a significant impact on its reliability, because of the decay of ^{125}I radioactivity. Values around the hypocretin-1 cutt-off can reliably be measured. Hypocretin-1 concentrations below 28 pg/mL should be reported as “undetectable” when measured with the Phoenix Pharmaceuticals RIA.

Statement of significance

The reliability of radioimmunoassay to determine cerebrospinal fluid hypocretin-1 levels has not systematically been evaluated. We found high intra-assay reliability and low inter-assay reliability when concentrations were not converted to international standard values. Harmonization using a commonly used reference sample from Stanford greatly improved inter-assay reliability. Additionally, there is a clear lower limit of quantification. In spite of this, numerous previous studies have reported very low hypocretin-1 values, which were used in various analyses. These reported values lie far below the lower limit of quantification and should be presented as such. The outcomes of this evaluation of 20 years' experience have implications for both previously published and future research.

Introduction

Narcolepsy type 1 (NT1) is a rare chronic neurological sleep-wake disorder, characterized by excessive daytime sleepiness, sleep fragmentation, sleep paralysis, hypnagogic hallucinations, and cataplexy (1). In addition, a decreased or undetectable concentration of hypocretin-1 (also known as orexin A) in cerebrospinal fluid (CSF) can be used to confirm a NT1 diagnosis. This hypocretin-1 deficiency is hypothesized to be caused by the autoimmune destruction of hypocretin-producing neurons (2). Hypocretin-1 deficiency is found in few other disorders (3, 4) and is considered the gold standard for diagnosing NT1 (5).

The cut-off value for hypocretin deficiency is 110 pg/mL (adjusted for Stanford values) (6, 7), and is commonly determined using the radioimmunoassay (RIA) kit from Phoenix Pharmaceuticals Inc (Burlingame, CA, USA). To determine the hypocretin-1 concentration, the RIA uses a known quantity of radioactive (I^{125}) labelled hypocretin-1, competing for a known quantity of hypocretin-1 antibody binding sites with the unlabelled hypocretin-1 in the sample. Then, the unbound antigens are washed away and the radioactivity of the precipitate is measured (8). More unlabelled antigens results in less specific activity (i.e. the activity per mass of radionuclide (9), expressed as counts per minute (CPM)). Thus, the hypocretin concentration is determined using a calibration curve based on standards with known concentrations.

However, RIAs have limitations. As radioactive materials are involved, precautions have to be taken concerning the use and disposal of the materials . Furthermore, antibody batches can vary greatly in their binding potential (8). Other factors to consider include differences in experience using RIAs to measure hypocretin-1 between laboratories, and possible cross reactivity with matrix constituents (10, 11). The fast decay of I^{125} radioactivity is another limitation, leading to short expiration dates (12). Differences in shelf lifes thus cause variability in specific radioactivity between RIAs, potentially impacting their reliability. The extent of the variability in outcomes caused by these factors has not been systematically evaluated.

Test duplicate samples are used to estimate for each pair the coefficient of variation (CV, i.e. the ratio of the standard deviation to the mean value of the measured concentrations). The CV gives an indication of the variability and can be used to determine the *intra*-assay variability of the RIA (13). The *inter*-assay variation (i.e. inter-assay CV) can be estimated by analyzing aliquots of the same sample in different assay runs (14). Harmonization samples from Stanford are used to verify the reliability of a RIA and to correct the measured values (15). These reference samples have a known concentration of hypocretin-1 and can be used for many RIAs due to the fact that hypocretin-1 concentrations in CSF remain very stable over time even after freezing and thawing (16). With the use of these reference samples, a conversion factor can be estimated for the harmonization of the individual RIAs (15).

The reliability of measurements around the cut-off point of 110 pg/mL is of clinical relevance but has not been determined yet. Also, even though most studies report CSF hypocretin-1 concentrations in absolute numbers, the the lowest hypocretin-1 concentration that can be reliably determined (i.e. lower limit of quantification, LLOQ) is yet unknown (17, 18). It is internationally agreed that the LLOQ should be within 25% of the nominal value, while the measurement of quality control (QC) samples (or standards used to create a calibration curve) need to have a repeatability of 20% or less (19, 20).

To further validate the Phoenix Pharmaceuticals hypocretin-1 RIA kit, we established the inter-assay variability and the intra-assay variability in general and around the cut-off of 110 pg/mL, and the LLOQ. We also assessed the impact of differences in specific radioactivity between different antibody batches on the reliability of the measured hypocretin-1 levels.

Methods

Data collection

The Leiden University Medical Center (LUMC) Department of Neurology has an international reference center and provides a service for hypocretin-1 measurement. The LUMC Department of Clinical Chemistry has over 20 years of experience of measuring hypocretin-1 in CSF, using the Phoenix Pharmaceutical RIA kit for the quantification of hypocretin-1. Between October 2001 and December 2021, 80 RIAs were performed to determine the CSF hypocretin-1 concentrations of people with suspected NT1. We analyzed raw data from the 80 consecutive RIAs.

During validation or verification of new analytical methods determination, the intra- and inter-assay CV is required. The Guideline on bioanalytical method validation of the European Medicines Agency (EMA), like others, (20) recommends measuring multiple QC samples at 5 concentration levels covering the calibration curve range: around the anticipated LLOQ, a low, medium and high concentration and around the anticipated ULOQ (upper limit of quantification). To determine the inter-assay CV, these QC samples need to be measured on different assay runs. We could not apply this preferred procedure due to the retrospective nature of our study. We therefore mimicked this using available data. To enhance feasibility and to reduce selection bias we choose to limit the collection of raw data (duplicate CPM measurements) to:

- The RIA calibration curve standards, which have the following standard hypocretin-1 concentrations: 10, 20, 40, 80, 160, 320, 640 and 1280 pg/mL.
- QC samples, which include:
    - Stanford reference sample measurements per RIA (harmonization samples, mostly measured in quadruplicate or, in a few cases, in duplicate). These reference samples have a known concentration of 280 pg/mL (16 RIAs) or 329 pg/mL (64 RIAs) and are used to correct for variation between RIAs. Two samples with different concentrations were used because after performing RIA 16, the 280 pg/mL reference sample was used up, and a new sample with a different concentration (329 pg/mL) was taken into use.
    - In line with updated quality requirements, internal control sample measurements per RIA (measured in duplicate) were introduced to monitor long time stability of the assay. These QC samples were created by the department

of Clinical Chemistry of the LUMC by pooling individual samples, resulting in an internal control sample with an intermediate hypocretin-1 concentration, which is used to internally verify reliability of RIAs over a longer period. Internal control sample measurements of 49 RIAs were available.

- Kit control sample measurements per RIA (measured in duplicate); an extra QC sample was provided by the supplier of the kit. It should be noted that, unlike the Stanford harmonization and internal control samples, the kit control samples differed for each RIA. As such, they cannot be used to determine the inter-assay CV. Kit control duplicate measurements of 57 RIAs were available.

- Individuals samples with a hypocretin-1 concentration below 200 pg/mL. These were added to determine the LLOQ and the CV in the lower range.

Outliers due to analytical errors were excluded. If the CV of a CPM duplicate was ≥20%, the sample was not collected, since this is likely due to an error during the RIA's execution.

Sample collection, storage and preparation

The CSF was obtained by lumbar puncture during regular clinical practise of local centers. Centers centrifuged the samples at 4000 rpm for 5 minutes, as part of the general practice to remove cell debris, before transport and sent the supernatant liquid. Internal control CSF was stored at -70°C, while regular patient samples were stored at -20°C. CSF hypocretin-1 levels were determined using RIA every three months. The total process of preparation and measurement takes 3 days. The third day, the RIA buffer is added to the samples (in brosilicate glass tubes) and the samples are centrifuged at 3600 rpm for 25 minutes. Then, the specific radioactive activity is measured. The RIAs were performed under the conditions recommended by the supplier of the kit.

Analysis

Intra-assay CV

The intra-assay CV was determined for each RIA and the average intra-assay CV for all RIAs. The measurements of all three QC samples (Stanford harmonization, internal control and kit control) were used for this calculation when available (the internal and kit control samples were not measured for every RIA and two RIAs lacked QC sample measurements). The following method was used to estimate the intra-assay CV (see figure 3.1 for an overview of this method). The calibration curve for each RIA was plotted and used to convert the CPM measurements of the QC samples to concentrations. Using these concentrations, 3 sub intra-assay CVs could be

estimated per RIA with the following formula: (SD quadruplicates or duplicates) / (mean quadruplicates or duplicates) * 100%. The sub intra-assay CVs were then used to estimate the total intra-assay CV for each RIA with another formula: $\sqrt{[\Sigma CV_i^2 /N]}$ (21). If only one sub CV was available, that one was used as the total intra-assay CV. Subsequently, the average of all the total intra-assays was determined. 77 of 80 RIAs could be used in this calculation: the Stanford harmonization, internal control and kit sample measurements of two RIAs were missing, while the results for the calibration curve were missing for a third RIA.

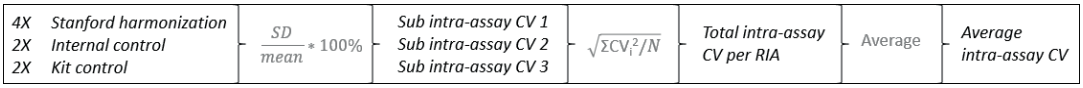


Figure 3.1. Overview of method used to determine the intra-assay CV for each RIA and the average intra-assay CV for all RIAs (n=77).

Inter-assay CV

The Stanford harmonization and internal control samples were used to determine the inter-assay CV. As with the intra-assay CV, the converted hypocretin-1 concentrations were used rather than the direct quadruplicate or duplicate CPM measurements. The variability between RIAs was estimated twice, using Stanford converted concentrations, and using unconverted concentrations. This way, we assessed the necessity and impact of the current process of harmonizing RIA outcomes using reference samples from Stanford.

Stanford harmonization measurements (those with the same known concentration of 329 pg/mL) were only used in the calculation of the unconverted inter-assay CV. The internal control measurements were used for the unconverted and Stanford converted inter-assay CV, giving a total of three inter-assay CVs.=:

1. The inter-assay CV using the new Stanford harmonization measurements not converted to Stanford-values was determined as follows: (SD mean Stanford harmonization sample concentrations)/(average mean Stanford harmonization sample concentration) * 100%. 58 RIAs were available for this calculation. The RIAs that could not be used consisted of four RIAs with Stanford harmonization outliers, one RIA with missing Stanford harmonization measurements and one RIA lacking the data for the calibration curve.
2. The inter-assay CV estimated using the internal control samples, converted and unconverted for Stanford, was obtained using the following formula: (SD mean Stanford (un)converted internal control sample concentration/average mean Stanford (un)converted internal control sample concentration) * 100%. 40 out of 74 RIAs were

available for this calculation. Excluded were: two RIAs with internal control sample outliers, two RIAs for which it was not possible to use internal control sample concentrations due to faulty measurements, one RIA missing a calibration curve and 29 RIAs lacking internal control sample measurements.

3. To correct the internal control duplicate, the conversion factor was estimated for each RIA by dividing 329 pg/mL by the mean of the measured Stanford harmonization concentrations derived from the calibration curves. Afterwards the conversion factor was applied to the measured internal control concentrations, resulting in Stanford converted internal control concentrations.

Determining the Lower Limit of Quantification

The LLOQ was set as the concentration with a CV of ≤25%, following international guidelines (19, 20).

To determine the LLOQ we used the raw data of all samples with an average concentration below 200 pg/mL (unconverted as well as converted to Stanford values). Duplicate measurements were excluded when the CV of CPM measurements was more than 20%, as this suggests an analytical error as opposed to normal variation. Using these data, the intra-assay CVs of the concentration measurements were estimated. By plotting samples with the CV per sample on the y-axis and average concentration per sample on the x-axis, logarithmic modelling was applied and the LLOQ was interpolated.

To estimate the CV at a 110 pg/mL concentration, we used the equations based on logarithmic modelling using the above-mentioned duplicate individual sample measurements.

Determining the impact of ^{125}I radioactivity decay on RIA reliability

To assess the impact of shelf life and amount of specific activity on RIA reliability, we performed an additional analysis in which we divided the RIAs into 2 separate groups: a group with < 3000 CPM and a group with ≥3000 CPM. Due to radioactive decay, we assumed that the RIAs with <3000 CPM were nearer the expiration date. Due to non-normal distribution, we used the Mann-Whitney U test to examine whether the intra-assay CVs differed significantly between groups. We also assessed whether radioactive decay affects the LLOQ, we measured the LLOQ in the group with lower and in the group with higher specific activity.

Data were analyzed using Microsoft Excel 2016 and SPSS version 25.0. For all statistical analyses, a statistical significance level of $\alpha = .05$ (2-tailed) was used. Figures were created using the GraphPad Prism 8 software.

Data availability statement

The data are available from the corresponding author upon reasonable request.

Results

Calibration curves

79 sigmoidal calibration curves were obtained from which concentrations were estimated (see figure 3.2 for two examples). Calibration curve standards were missing for one RIA, excluding this RIA from the analyses. The CPM range of the calibration curves is plotted in figure 3.3.

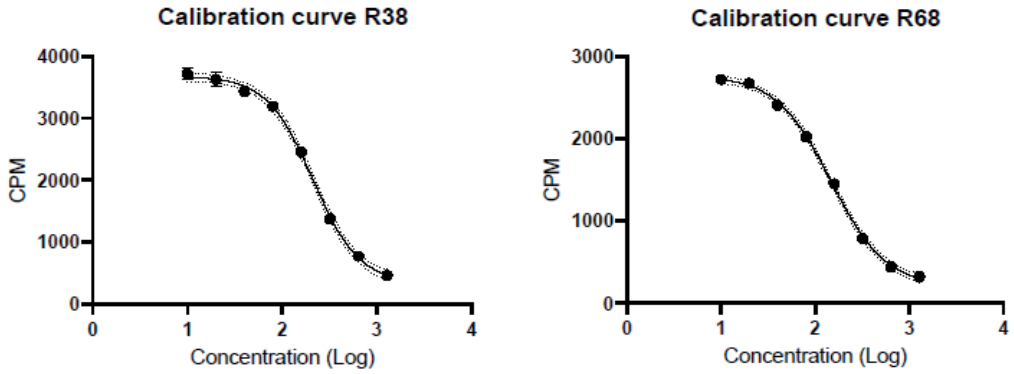


Figure 3.2. Two examples of RIA calibration curves. RIA R38 left (A), R68 right (B) On the X-axis is the concentration transformed to a log scale and on the Y-axis the CPM.

Variability within RIAs: the Intra-assay CV

The mean intra-assay CV was 4.73% (SD ±2.82%) when using all available RIAs. The mean intra-assay CV was 5.04% (SD ±2.85%) when excluding the 29 RIAs that missed one or two (of three) sub intra-assay CVs (those estimated with either the Stanford harmonization sample alone or with the kit or internal control sample).

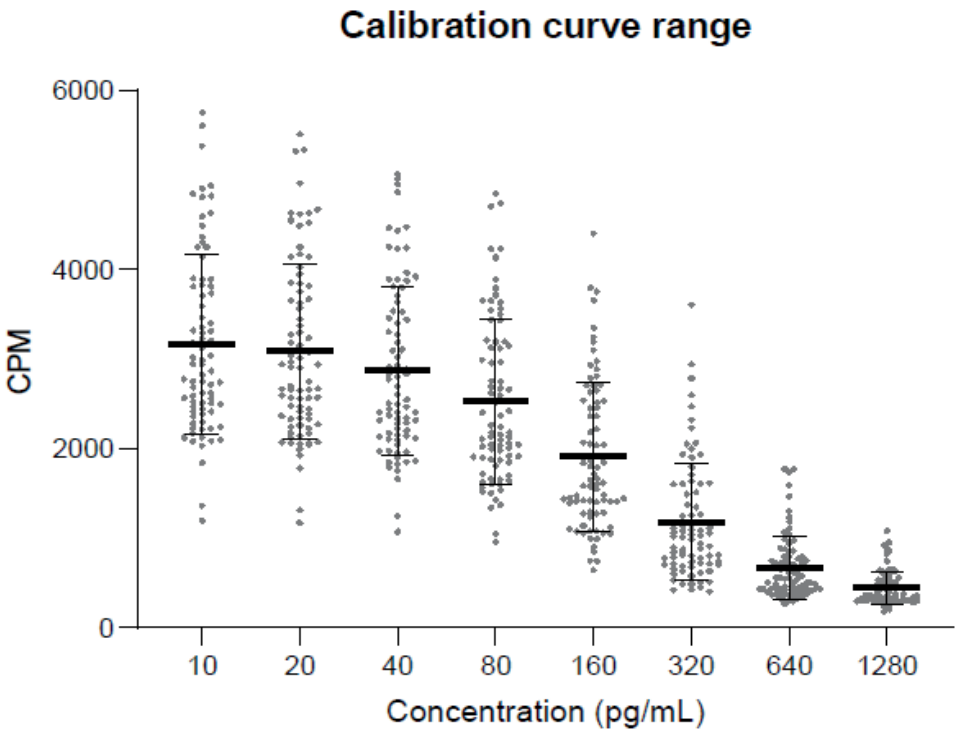


Figure 3.3. RIA calibration curve range comparison, the bars represent the means and SDs of each standard concentration. Shown here are all standards from 79 RIAs plotted against their CPMs, visualizing the wide spread of the CPM range across the RIAs. As the concentration increases, the CPM spread decreases.

Variability between RIAs: the Inter-assay CV

The inter-assay CV was computed three times: twice using the internal controls (unconverted and as converted for Stanford) and once using the new Stanford harmonization samples (those with the known concentration of 329 pg/mL).

Using the internal controls, the inter-assay CV was 28.28% (n=48) when not converted using the Stanford harmonization sample (18.49% without the two outliers, n=46) and 7.46% when converted (7.49% without the two outliers). The average of the applied conversion factors was 1.23, causing the interpolated semilog line to shift to the right (as can be seen figure 3.4).

The inter-assay CV estimated using the new Stanford harmonization samples (n=62) was 33.17% when outliers were included, and 23.14% when four outliers were excluded (n=58).

Lower Limit of Quantification

The LLOQ graphs were based on 809 unconverted samples and 635 Stanford converted samples. The resulting graphs gave the following LLOQs (with a CV of <25%): 22.9 pg/mL

for the unconverted <200 pg/mL samples and 27.9 pg/mL for the Stanford converted <200 pg/mL samples, see figure 3.4.

The concentration with a CV of 20% was 34.5 pg/mL when unconverted and 40.7 pg/mL when converted.

Reliability around 110 pg/mL

The CV with a unconverted concentration of 110 pg/mL was 5.89% and 6.79% with a converted concentration.

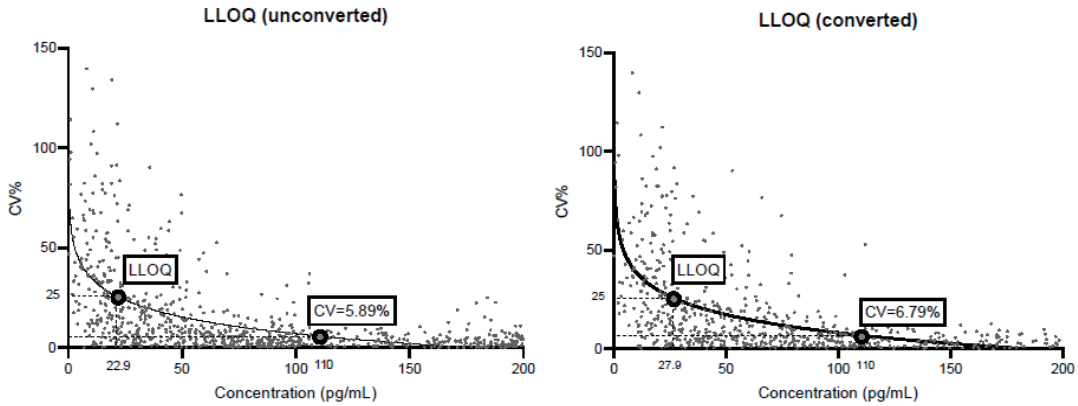


Figure 3.4. Lower Limits of Quantification (LLOQ). Shown here are the graphs from which the LLOQs were determined (at a concentration CV of 25%). For the 809 unconverted <200 pg/mL samples (A), the LLOQ is 22.9 pg/mL (formula: y=-12.17*ln(x)+63.091). For the 635 converted samples (B), the LLOQ is 27.9 pg/mL (formula: y=-13.29*ln(x)+69.255).

Impact of ^{125}I radioactivity decay on RIA reliability

The intra-assay CV of RIAs with lower specific activity (CPM range<3000, n=41) was significantly higher than the intra-assay CV of RIAs with higher specific activity (CPM range ≥3000, n=36), with a median of 5.58% (min 1.6%, max 17.0%) for the lower activity and 3.24% (min 0.4%, max 11.6%, p=.013) for the higher activity.

When assessing whether radioactive decay impacts the LLOQ determined using converted hypocretin-1 concentration measurements, a lower LLOQ was found in the group with lower specific activity (22.6 pg/mL, y=-12.08*ln(x)+62.683) than in the group with higher specific activity (35.1 pg/mL, y=-14.22*ln(x)+75.603). When assessing the impact on unconverted concentration measurements, however, the LLOQ was almost the same (20.2 pg/mL, y=-11.39*ln(x)+59.218 and 24.5 pg/mL, y=-12.55*ln(x)+65.149.

Discussion

We evaluated technical aspects of CSF hypocretin-1 measurements using the RIA kit from Phoenix Pharmaceuticals Inc which is currently used worldwide. We found (1) a high intra-assay reliability, including around the cut-off point of 110 pg/mL; (2) a low inter-assay reliability without conversion to Stanford values, which improved substantially after conversion; (3) an LLOQ (with a CV ≤ 25%) obtained from the converted samples of 27.9 pg/mL; and (4) that the CPM range differences between RIAs varies widely. These results have important implications for the clinical use of RIAs for hypocretin-1 measurements in CSF for the diagnosis of NT1.

Intra- and inter-assay CV

Specifications of inter-individual biologic variation and what is considered an acceptable level of imprecision for laboratory testing are listed for many analytes (22, 23), but this is not the case for hypocretin-1. While guidelines generally recommend an acceptance criterion of a ≤15% deviation within and between assays, when it comes to ligand-binding assays such as the RIA, a higher deviation of ≤20% is accepted (19, 20, 24, 25).

The intra-assay CV we estimated using our data is acceptable, with even the intra-assay variability around the cut-off point of 110 pg/mL being far below the accepted deviation of 20%, and similar to intra-assay variations found in earlier studies (3, 26-28). Specific methods on how these variations were estimated in these other studies were not provided. To illustrate the importance of the method used, Keating et al (12) found an average intra-assay hypocretin-1 variation of 9.4% (higher than the present study and other studies). In order to obtain an intra-assay variation, the authors divided the absolute value of the difference in a duplicate measurement by the higher of the two values, which tends to give a higher intra-assay variation than with the CV calculation used in this study. More than one method is available for the calculation of the intra- and inter-assay CV. This makes it all the more important that the methods used are clearly reported, to facilitate the comparison between outcomes of past and present studies. It is known that the hypocretin-1 protein, among other proteins is very hydrophobic (29). Not using the proper collection or storage tubes can result in loss of hypocretin-1 due to absorption to the tube wall. No studies have been performed to determine the type of tube best suited to transport and store CSF samples for hypocretin-1 measurements. However, based on studies (30-33) regarding hydrophobic proteins in CSF, polypropylene

tubes are recommended for storage and collection. For the analysis borosilicate glass tubes are used (34).

We performed two assessments of inter-assay variability without conversion for Stanford, based on the Stanford harmonization and internal control measurements, and found high inter-assay CVs. Our results show a low inter-assay reliability when measurements are not corrected using a harmonization sample. Previous studies have also shown large inter-assay variations although, as with the intra-assay variations, methods of calculation are not mentioned (27, 35). In contrast, certain other applications of the RIA tend to yield lower inter-assay variations (13, 36-39). One reason for this is the difference in the unit of measurement used in different applications of the RIA. If the sample contains a large quantity of the substance measured, the concentrations of interfering compounds are relatively lower, reducing the impact of cross-reactivity. Another possible explanation may be that the antibodies in these other RIAs cross-react less with matrix proteins. For example, Sakai et al (28) found inter-assay CVs between 2.5% and 6.4% in their hypocretin-1 RIAs, using self-created polyclonal anti-hypocretin-1 antibodies, which were about 20 times more sensitive than the antibodies provided by Phoenix Pharmaceuticals.

After converting the mean internal control concentrations using the Stanford conversion factor, the internal control inter-assay CV was reduced to a more acceptable level. The exact concentration of the internal control sample was unknown in this study, but the average internal control mean after conversion (314 pg/mL) was quite close to the known concentration of the new Stanford harmonization samples (329 pg/mL), which were used in calculating the conversion factors that corrected the internal control sample values. It may be more interesting to see what would happen to the inter-assay CV when converting sample concentrations that are substantially different from the Stanford harmonization sample concentrations used. No repeated sample data were available to conduct such an assessment during this retrospective study.

The effect of outliers was also illustrated by our results, in that two or four outliers can increase the inter-assay CV enormously. The consistent deviation of these RIAs suggests that a systematic human error may have taken place in these cases.

Lower Limit of quantification

The LLOQ before Stanford conversion of the hypocretin-1 concentrations was lower than the LLOQ after conversion, which is as expected since most of the samples used were corrected

with a conversion factor above 1.0 (average of 1.23). This increases the concentration of most samples, causing the interpolated semilog line to shift to the right.

Previous studies mostly used the limit of detection (LOD) instead of the LLOQ (3, 26, 27, 40). Our study used the LLOQ, which as mentioned previously, shows the point from which a concentration can be reliably determined. This is clinically more relevant than a detection limit, as the LOD only indicates the lowest concentration distinguishable from background noise. Thus, the presence of an analyte can reliably be detected, but the reported concentration is not necessarily reliable (17). Past studies have reported hypocretin-1 concentrations below 27.9 pg/mL without the caveat that these concentrations may not be accurate. We suggest that hypocretin-1 concentrations measured using RIA should be reported as undetectable when below 27.9 pg/mL.

Variation in RIA calibration curve range and the impact on RIA reliability

The calibration curves used in our study varied widely (see Figure 3.3), a variation that can be explained by the fact that, as the RIA kit gets older, the radioactivity of the I^{131} labelled hypocretin-1 decreases. This causes narrowing and lowering of the CPM range, thus flattening its calibration curve.

The effects of this flattening were noticeable in the comparison between the intra-assay CVs from RIAs with an average CPM range above 3000 and RIAs with an average CPM range below 3000, with the latter assumed to be RIAs nearing the supplier-defined expiration date of the RIA. The intra-assay CV was substantially and significantly higher for the older RIAs (CPM range <3000) than for the newer RIAs (CPM range ≥3000). As mentioned above, this is likely due to differences in calibration curve spread. When the calibration curve spread decreases small differences in CPM measurements will have more impact when converting the measured CPM to a hypocretin-1 concentration, causing more variation and less intra-reliability. This effect is also seen when we look at the difference in the LLOQ measured using raw data from RIAs with high and RIAs with low specific activity. While the difference is minimal when using unconverted concentrations, it becomes more evident when using converted concentrations, with the RIAs with higher specific activity having a higher LLOQ, indicating more variability.

Keating et al (12) also found that RIAs with a longer shelf time have higher intra-assay variations when determining hypocretin-1 concentrations. While presently, per operating instructions, RIAs need to be used within a period of six weeks, these results further suggest

that expeditious use of RIAs is desirable. Even more so as we found that the radioactivity of Phoenix Pharmaceuticals RIA kits has, on average, been lower in recent years (see figure S3.1).

Limitations

Some limitations should be mentioned. First of all, for multiple calculations we could only use the samples that were measured with consecutive RIAs as part of the standard protocol (Stanford harmonization, internal control and kit samples). Thus while assessment of the accuracy of RIAs at low, intermediate and high concentration levels is recommended by the EMA (and other) guidelines for the validation of analytical methods (6), this was not possible in this retrospective study. The Stanford harmonization, internal control and kit samples we used to determine variability all had intermediate concentration levels and, given the sigmoidal shape of calibration curves, it is also of interest to assess the inter-assay variability of low and high concentration samples. Furthermore, these measurements were performed in a single laboratory, which potentially decreases the generalizability of the results. Prospective inter-laboratory research adhering to the recommendations of the FDA and EMA (or other relevant) guidelines would be preferred. We are aware that recently more sophisticated mass spectrometry based methods for the determination of hypocretin-1 were published, which is a very promising evolution (10, 11, 41). At the moment the field of quantitative clinical chemistry proteomics is developing fast (42) and measurement of CSF hypocretin-1 by this technique can overcome some of the limitations of the RIAs. Still, the RIA remains the only test that is sufficiently validated to assess hypocretin deficiency to diagnose narcolepsy. The enzyme-linked immunosorbent assay (ELISA) cannot be used as a reliable alternative method to measure hypocretin-1 in CSF (41, 43).

Conclusion

In conclusion, the intra-assay variation when measuring hypocretin-1 using the commonly used RIA kit from Phoenix Pharmaceuticals is of an acceptable level. The low intra-assay variability around the current cut-off point of 110 pg/mL is particularly of interest in clinical practise. In contrast, the inter-assay variation, and thus the inter-assay reliability, is too high without conversion using a harmonization sample. After conversion, the inter-assay variation becomes acceptable, though this correction should be tested on samples with different concentrations in the future, preferably including concentrations on more extreme ends of the calibration curve. In addition, the shelf life of the RIA kits should be kept in mind, as the amount of specific activity has a significant impact on intra-assay reliability. A hypocretin-1 concentration below

28 pg/mL should be reported as undetectable. In addition, we find that concentrations around the currently used cut-off value to diagnose NT1 can be determined reliably. Lastly, we recommend standardization of RIAs following our findings. This is achieved by using polypropylene plastic tubes for collection and storage of the CSF, using RIA kits immediately after receiving them from Phoenix Pharmaceuticals preventing a lower specific activity of the RIA, and by applying a harmonization sample. This way, results of RIAs for hypocretin-1 measurements will be more reliable and comparable across studies.

Disclosure statement

S. Overeem consulted for Bioprojet, Jazz Pharmaceuticals, UCB Pharma and Takeda, all paid to the institution, and not related to the present work. R. Fronczek consulted for and/or received lecture fees from Bioprojet, Takeda, Lundbeck, TEVA, Lilly, Novartis and Allergan, and grant support from Jazz Pharmaceuticals and Bioprojet, all not related to the present work. D. Bijlenga received a reimbursement for a course from Bioprojet, not related to the present work. G.J. Lammers consulted for Bioprojet, Jazz Pharmaceuticals, UCB Pharma, NLS and Takeda, all paid to the institution, and not related to the present work. He also served as a member of advisory boards on narcolepsy: Bioprojet, Jazz Pharmaceuticals, UCB Pharma, NLS. The other authors have indicated no potential financial conflicts of interest.

Supplementary material

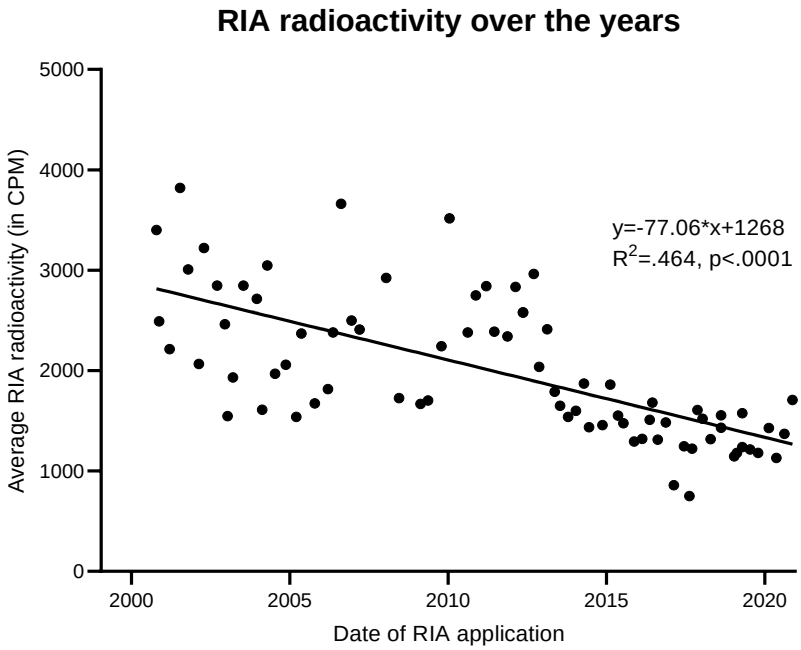


Figure S3.1. A significant decrease in radioactivity (expressed in count per minute, i.e. CPM) of radioimmunoassays (RIAs), measured using the average of calibration curves, can be seen in recent years. The line represents a linear regression analysis (n=79).

References

1. Longstreth WT, Jr., Koepsell TD, Ton TG, Hendrickson AF, van Belle G. The Epidemiology of Narcolepsy. Sleep. 2007;30(1):13-26.
2. Liblau RS, Vassalli A, Seifinejad A, Tafti M. Hypocretin (orexin) biology and the pathophysiology of narcolepsy with cataplexy. The Lancet Neurology. 2015;14(3):318-28.
3. Ripley B, Overeem S, Fujiki N, Nevsimalova S, Uchino M, Yesavage J, et al. CSF hypocretin/orexin levels in narcolepsy and other neurological conditions. Neurology. 2001;57(12):2253-8.
4. Nishino S, Kanbayashi T, Fujiki N, Uchino M, Ripley B, Watanabe M, et al. CSF hypocretin levels in Guillain-Barré syndrome and other inflammatory neuropathies. Neurology. 2003;61(6):823-5.
5. American Academy of Sleep M. The international classification of sleep disorders2014.
6. Mignot E, Lammers GJ, Ripley B, Okun M, Nevsimalova S, Overeem S, et al. The Role of Cerebrospinal Fluid Hypocretin Measurement in the Diagnosis of Narcolepsy and Other Hypersomnias. Archives of Neurology. 2002;59(10):1553-62.
7. Postiglione E, Barateau L, Pizza F, Lopez R, Antelmi E, Rassu AL, et al. Narcolepsy with intermediate cerebrospinal level of hypocretin-1. Sleep. 2022;45(2).
8. Goldsmith SJ. Radioimmunoassay: Review of basic principles. Seminars in Nuclear Medicine. 1975;5(2):125-52.
9. Groch MW. Radioactive decay. RadioGraphics. 1998;18(5):1247-56.
10. Bardsen K, Gjerstad MD, Partinen M, Kvivik I, Tjensvoll AB, Ruoff P, et al. Considerably Lower Levels of Hypocretin-1 in Cerebrospinal Fluid Is Revealed by a Novel Mass Spectrometry Method Compared with Standard Radioimmunoassay. Analytical chemistry. 2019;91(14):9323-9.
11. Hirtz C, Vialaret J, Gabelle A, Nowak N, Dauvilliers Y, Lehmann S. From radioimmunoassay to mass spectrometry: a new method to quantify orexin-A (hypocretin-1) in cerebrospinal fluid. Scientific reports. 2016;6:25162.
12. Keating G, Bliwise DL, Saini P, Rye DB, Trotti LM. Hypocretin measurement: shelf age of radioimmunoassay kit, but not freezer time, influences assay variability. Scand J Clin Lab Invest. 2017;77(5):390-3.
13. Hanneman SK, Cox CD, Green KE, Kang D-H. Estimating Intra- and Inter-Assay Variability in Salivary Cortisol. Biological Research For Nursing. 2011;13(3):243-50.
14. Rodbard D. Statistical Quality Control and Routine Data Processing for Radioimmunoassays and Immunoradiometric Assays. Clinical Chemistry. 1974;20(10):1255-70.
15. Sudhansu C, Luigi F-S. Oxford Textbook of Sleep Disorders. Oxford, UK: Oxford University Press; 2017-06.
16. Nishino S, Ripley B, Overeem S, Nevsimalova S, Lammers GJ, Vankova J, et al. Low cerebrospinal fluid hypocretin (orexin) and altered energy homeostasis in human narcolepsy. Annals of Neurology. 2001;50(3):381-8.
17. Armbruster DA, Pry T. Limit of blank, limit of detection and limit of quantitation. Clin Biochem Rev. 2008;29 Suppl 1(Suppl 1):S49-S52.
18. Wolfinger RD, Beedanagari S, Boitier E, Chen T, Couttet P, Ellinger-Ziegelbauer H, et al. Two approaches for estimating the lower limit of quantitation (LLOQ) of microRNA levels assayed as exploratory biomarkers by RT-qPCR. BMC Biotechnol. 2018;18(1):6-.
19. Han F, Lin L, Schormair B, Pizza F, Plazzi G, Ollila HM, et al. HLA DQB1*06:02 negative narcolepsy with hypocretin/orexin deficiency. Sleep. 2014;37(10):1601-8.
20. European Medicines Agency. Guideline on Bioanalytical Method Validation (2011). www.ema.europa.eu/docs/en GB/document library/Scientific guideline/2011/08/WC500109686.pdf.
21. Influentialpoints. Coefficient of variation [Available from: https://influentialpoints.com/Training/coefficient_of_variation-principles-properties-assumptions.htm.
22. Luraschi P, Brambilla S, Pagani A, Infusino I, Franzini C. Updating biological variation databases and selecting reasonably correct published values. Scandinavian Journal of Clinical and Laboratory Investigation. 2000;60(8):733-5.
23. Fraser CG. Biological variation in clinical chemistry. An update: collated data, 1988-1991. Archives of Pathology & Laboratory Medicine. 1992;116(9):916-23.
24. Administration F. USFDA. Guidance for Industry: Bioanalytical Method Validation. US Department of Health and Human Services, Food and Drug Administration. 2001;66.
25. Viswanathan CT, Bansal S, Booth B, DeStefano AJ, Rose MJ, Sailstad J, et al. Quantitative Bioanalytical Methods Validation and Implementation: Best Practices for Chromatographic and Ligand Binding Assays. Pharmaceutical Research. 2007;24(10):1962-73.
26. Nishino S, Ripley B, Overeem S, Lammers GJ, Mignot E. Hypocretin (orexin) deficiency in human narcolepsy. Lancet. 2000;355(9197):39-40.

27. Heier MS, Evsiukova T, Vilming S, Gjerstad MD, Schrader H, Gautvik K. CSF hypocretin-1 levels and clinical profiles in narcolepsy and idiopathic CNS hypersomnia in Norway. Sleep. 2007;30(8):969-73.
28. Sakai N, Matsumura M, Lin L, Mignot E, Nishino S. HPLC analysis of CSF hypocretin-1 in type 1 and 2 narcolepsy. Scientific reports. 2019;9(1):477.
29. Hungs M, Mignot E. Hypocretin/orexin, sleep and narcolepsy. Bioessays. 2001;23(5):397-408.
30. Hansson O, Mikulskis A, Fagan AM, Teunissen C, Zetterberg H, Vanderstichele H, et al. The impact of preanalytical variables on measuring cerebrospinal fluid biomarkers for Alzheimer's disease diagnosis: A review. Alzheimers Dement. 2018;14(10):1313-33.
31. Toombs J, Foiani MS, Wellington H, Paterson RW, Arber C, Heslegrave A, et al. Amyloid β peptides are differentially vulnerable to preanalytical surface exposure, an effect incompletely mitigated by the use of ratios. Alzheimers Dement (Amst). 2018;10:311-21.
32. Vanderstichele HM, Janelidze S, Demeyer L, Coart E, Stoops E, Herbst V, et al. Optimized Standard Operating Procedures for the Analysis of Cerebrospinal Fluid Aβ42 and the Ratios of Aβ Isoforms Using Low Protein Binding Tubes. J Alzheimers Dis. 2016;53(3):1121-32.
33. Simonsen AH, Bahl JM, Danborg PB, Lindstrom V, Larsen SO, Grubb A, et al. Pre-analytical factors influencing the stability of cerebrospinal fluid proteins. J Neurosci Methods. 2013;215(2):234-40.
34. Goebel-Stengel M, Stengel A, Taché Y, Reeve JR, Jr. The importance of using the optimal plasticware and glassware in studies involving peptides. Anal Biochem. 2011;414(1):38-46.
35. Lopez R, Barateau L, Evangelista E, Chenini S, Robert P, Jaussent I, et al. Temporal Changes in the Cerebrospinal Fluid Level of Hypocretin-1 and Histamine in Narcolepsy. Sleep. 2017;40(1).
36. Heilmann RM, Paddock CG, Ruhnke I, Berghoff N, Suchodolski JS, Steiner JM. Development and analytical validation of a radioimmunoassay for the measurement of alpha1-proteinase inhibitor concentrations in feces from healthy puppies and adult dogs. Journal of Veterinary Diagnostic Investigation. 2011;23(3):476-85.
37. Buono A, Lidbury JA, Wood C, Wilson-Robles H, Dangott LJ, Allenspach K, et al. Development, analytical validation, and initial clinical evaluation of a radioimmunoassay for the measurement of soluble CD25 concentrations in canine serum. Vet Immunol Immunopathol. 2019;215:109904.
38. Benabdelaziz A, Boudjemai S, Khelili R, Besbaci M, Kaidi R. Development of in-house RIA kit for progesterone in cow's skim milk. J Immunoassay Immunochem. 2020;41(2):195-207.
39. Deberg M, Houssa P, Frank BH, Sodoyez-Goffaux F, Sodoyez JC. Highly specific radioimmunoassay for human insulin based on immune exclusion of all insulin precursors. Clin Chem. 1998;44(7):1504-13.
40. Knudsen S, Jennum PJ, Alving J, Sheikh SP, Gammeltoft S. Validation of the ICSD-2 criteria for CSF hypocretin-1 measurements in the diagnosis of narcolepsy in the Danish population. Sleep. 2010;33(2):169-76.
41. Lindström M, Schinkelshoek M, Tienari PJ, Kukkonen JP, Renkonen R, Fronczek R, et al. Orexin-A measurement in narcolepsy: A stability study and a comparison of LC-MS/MS and immunoassays. Clin Biochem. 2021;90:34-9.
42. Smit NPM, Ruhaak LR, Romijn F, Pieterse MM, van der Burgt YEM, Cobbaert CM. The Time Has Come for Quantitative Protein Mass Spectrometry Tests That Target Unmet Clinical Needs. J Am Soc Mass Spectrom. 2021;32(3):636-47.
43. Ono T, Kanbayashi T, Yoshizawa K, Nishino S, Shimizu T. Measurement of cerebrospinal fluid orexin-A (hypocretin-1) by enzyme-linked immunosorbent assay: A comparison with radioimmunoassay. Psychiatry and clinical neurosciences. 2018;72(11):849-50.

4

Applicability of the Sustained Attention to Response Task (SART) in hypersomnolence

experience and results from a tertiary referral center

Adrienne Elisabeth van der Hoeven[a,b], Denise Bijlenga[a,b], Puck Bouhuijs[b], Mojca Kristina Maria van Schie[a,b], Gert Jan Lammers[a,b], Rolf Fronczek[a,b]

[a]Department of Neurology Leiden University Medical Center, Leiden, the Netherlands, [b]Sleep-Wake Center, Stichting Epilepsie Instellingen Nederlands (SEIN), Heemstede, the Netherlands

***S**leep Medicine*. 2023 August

Abstract

Objective/Background

Evaluation of hypersomnolence disorders ideally includes an assessment of vigilance using the short Sustained Attention to Response Task (SART). We evaluated whether this task can differentiate between hypersomnolence disorders, whether it correlates with subjective and objective sleepiness, whether it is affected by the time of day, and symptoms of anxiety and depression.

Patients/Methods

We analyzed diagnostic data of 306 individuals with hypersomnolence complaints diagnosed with narcolepsy type 1 (n=100), narcolepsy type 2 (n=20), idiopathic hypersomnia (n=49), obstructive sleep apnea (n=27) and other causes or without explanatory diagnosis (n=110). We included the Multiple Sleep Latency Test (MSLT), polysomnography, Epworth Sleepiness Scale (ESS), Hospital Anxiety and Depression Scale and SART, which were administered five times during the day (outcomes: reaction time, total, commission and omission errors).

Results

The SART outcomes did not differ between groups when adjusted for relevant covariates. Higher ESS scores were associated with longer reaction times and more commission errors ($p<.01$). The main outcome, total errors, did not differ between times of the day. Reaction times and omission errors were impacted ($p<.05$).

Conclusions

The SART quantifies disturbed vigilance, an important dimension of disorders of hypersomnolence. Results do not suggest that depressive symptoms influence SART outcomes. A practice session is advised. Testing time should be taken into account when interpreting results. We conclude that the SART does not differentiate between central disorders of hypersomnolence. It may be a helpful addition to the standard diagnostic workup and monitoring of these disorders.

Statement of significance

The Sustained Attention to Response Task (SART) is suitable for assessing vigilance in individuals with hypersomnolence. It is easy to implement and requires little time and resources when combined with the Multiple Sleep Latency Test (MSLT) in the routine workup. Our study assessed the association between SART and polysomnography (PSG) outcomes. We also evaluated the implementation of the SART in a large sample in clinical practice. This sample showed that the SART measures a specific aspect, namely vigilance, that is often overlooked despite having a significant impact on people's daily lives.

Introduction

Individuals suffering from disorders of hypersomnolence often complain of disturbed vigilance, resulting in disturbed sustained attention [1-3], which has the potential to have a profound, negative impact on daily functioning [4-7]. The most commonly used tests to evaluate these disorders focus solely on sleep and wakefulness. These tests are either subjective, e.g. the Epworth Sleepiness Scale (ESS), or objective, using polysomnography (PSG) and/or the Multiple Sleep Latency Test (MSLT) [8-10]. Daytime vigilance is still rarely assessed in clinical practice of hypersomnolence disorders. Vigilance tests such as the Sustained Attention to Response Task (SART) generally take little time and provide important insight into an often neglected disease aspect [11].

The SART is a short (<5 minutes), and inexpensive computer task in which the subject should click a button when a target appears on the screen and inhibit to a non-target. It can be easily used in combination with the MSLT, as it is assessed five times over the day, e.g. just before each MSLT session [12, 13]. It was initially developed in 1997 to measure vigilance in individuals with traumatic brain injury and has since been used in various disorders, including central disorders of hypersomnolence: narcolepsy types 1 and 2 (NT1, NT2) and idiopathic hypersomnia (IH) [12-15]. The Psychomotor Vigilance Task (PVT) [16] and Oxford Sleep Resistance test (OSLER) [17] have also been used to measure vigilance impairment in sleep-wake disorders. However the SART is more sensitive to measure treatment effects in individuals with narcolepsy type 1 than the PVT [18], and while the OSLER is also suitable to measure vigilance, it takes considerably more time to perform [17].

We previously showed that the SART could differentiate between healthy controls and individuals with NT1 [12], but not between NT1, NT2 and IH [13]. The study also showed that the NT1 group was relatively slower in their reaction to the presented stimuli and that they made more mistakes on the SART in morning sessions than in the afternoon sessions [13]. This was either due to a learning or a time-of-the-day effect. Follow-up research among healthy subjects [15] showed that this was presumably due to a learning effect, but that applied to the specific test instruction that accuracy was more important than speed, whereas in our earlier research among patients with sleep disorders a different instruction was used, namely that accuracy and speed were equally important. No correlations were found between the SART scores and the MSLT or ESS outcomes. Thus, SART scores reflected a different aspect of disorders of hypersomnolence [12, 13]. However, the relationship between SART and PSG results has not yet been assessed.

This report provides clinical data from a tertiary sleep center, collected over a period when the SART was routinely assessed in the diagnostic workup of suspected hypersomnolence. In addition to providing data regarding the daily clinical practice of implementing the SART, we explored whether the SART (1) can be used to distinguish between different disorders of hypersomnolence, (2) measures a different disease aspect than the PSG, MSLT or ESS, (3) is affected by testing time in clinical practice, and (4) is affected by anxiety or depression (as measured using the Hospital Anxiety and Depression Scale; HADS).

Methods

Data collection

Data were collected between March 2014 and October 2021 for clinical purposes in a tertiary Sleep-Wake center (SEIN, Heemstede, the Netherlands). Included were data of people (≥16 years old) who completed the SART, MSLT, PSG and/or ESS as part of routine diagnostic workup (the diagnostic sample), or completed the SART and ESS for their driver's license evaluation (the driver's license sample). Individuals with multiple sleep diagnoses were excluded, except for mild obstructive sleep apnea (OSA; i.e. apnea-hypopnea index <15). This secondary sleep diagnosis is often an incidental finding after a diagnostic polysomnography [19, 20].

Experienced neurologists-somnologists made the clinical diagnoses of people in the diagnostic sample. The diagnoses of people in the driver's license sample were verified based upon clinical information obtained from the electronic health records from the various sleep-wake clinics or (if this was not possible) extracted from referral letters. Individuals were classified into the following primary diagnostic groups: NT1, NT2, IH, OSA or complaints of excessive daytime sleepiness without explanatory diagnosis (CEDS). The CEDS group consisted of the following categories: insomnia (20.0%), no primary sleep diagnosis (18.2%), unclear diagnosis but no central hypersomnolence disorder (16.4%), behaviorally induced insufficient sleep syndrome (10.0%), restless legs syndrome (6.4%), suboptimal sleep hygiene (6.4%), mood issues (3.6%), circadian rhythm disorder (3.6%), psychogenic non-epileptic seizures (2.7%), hypersomnolence due to a medical disorder (2.7%), and 10.0% other. A secondary analysis was performed on the individuals strictly meeting the ICSD3 criteria (taking into account the results of any previous diagnostic testing performed) and the CEDS group and can be found in the supplementary material.

Diagnostic workup

The diagnostic workup consisted of the ESS, SART, and MSLT performed in a clinical setting on a single day. A PSG was performed either in the person's home or a clinical setting on a preceding night (see figure 4.1). Before the first SART session, individuals had the opportunity to practice the SART for 30 seconds. A SART session preceded every MSLT nap opportunity.

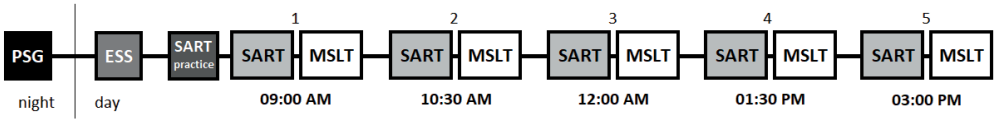


Figure 4.1. Representation of our routine diagnostic work-up in a clinical setting. During the night a polysomnography (PSG; black) is performed, followed by the Epworth Sleepiness Scale (ESS; dark grey), five sessions of the Sustained Attention to Response Task (SART; light grey) the first preceded by a SART practice round and the Multiple Sleep Latency Test (MSLT; white).

Materials

Sustained Attention to Response Task (SART)

The SART [21] is a short (4 minutes and 20 seconds) computerized go/no-go task and was administered on a laptop model HP ProBook 6570b with a monitor refresh of 60 Hz and utilizing the operating system Windows 7 Professional. To improve the reliability of measured reaction times, the SART computer program was executed in priority mode, with minimal

background programs running. It is performed in a quiet room with dimmed lights, which is the same room where the MSLT is also conducted. Numbers 1 to 9 are consecutively presented in a random sequence on a computer screen. Participants are instructed to press a button whenever a number (the target) appears on the screen, except for number 3 (the non-target). Each number is shown 25 times (225 in total), for 250 milliseconds each, as white numbers on a black screen, followed by a black screen for 900 milliseconds. Individuals have to respond before the following number appears. Participants are instructed to aim for accuracy over speed [13]. The main outcome is the mean total error score over the five SART sessions, consisting of the sum of commission errors (pressing a key after a non-target) and omission errors (not pressing a key after a target). Another SART outcome was the mean reaction time (SART RT) which is estimated per session over correct responses.

Multiple Sleep Latency Test (MSLT)

The MSLT is an objective test that can provide EDS or daytime sleep pressure information. The MSLT is performed under standard conditions with dimmed light. Individuals undergoing this test lie in bed and are instructed to nap for up to 20 minutes on five occasions on a single day, with electro-encephalographic recording of sleep [22]. The outcome parameters of the MSLT are sleep latency (SL, with a cut-off of ≤8 minutes), and the presence of sleep-onset REM periods (SOREMP, with a cut-off of ≥2 SOREMPs), which are part of the diagnostic criteria of NT1 and NT2 [9].

Polysomnography (PSG)

From the PSG performed the preceding night, the parameters relevant to the diagnostic process based on the official ICSD-3 criteria were used: sleep latency (PSG SL), duration of the time in bed (TIB), total sleep time (TST), sleep efficiency (SE, i.e. TST/TIB*100%) and the REM-sleep latency, which was used to determine whether a SOREMP was present [23]. Furthermore, the apnea hypopnea index was collected, representing the average number of apneas and hypopneas per hour individuals experience during the night. The hypopnea definition used in this study was a decrease of at least 30% in airflow, coupled with a desaturation of at least 3%, as is recommended by the American Academy of Sleep Medicine (AASM) [24].

Epworth Sleepiness Scale (ESS)

The ESS [10] is a self-report questionnaire developed to measure subjective daytime sleepiness. Individuals are asked to estimate their likelihood of falling asleep in certain

situations. Total scores range from 0 to 24; the higher the score, the higher the subjective sleepiness during the day. A cut-off of ≥10 is used to indicate EDS [12, 25].

Hospital Anxiety and Depression Scale (HADS)

The HADS is a questionnaire screening tool for anxiety and depression [26]. A cut-off of ≥8 (out of 21) indicates increased anxiety or depression symptoms [27].

Ethics statement

The study was conducted following the Helsinki Declaration as revised in 2013. Due to the historical nature of data, the Medical Ethical Committee of Leiden-Den Haag-Delft (registration number: G20.044) allowed a waiver of the requirement for informed consent.

Data availability statement

The data are available from the corresponding author upon reasonable request.

Data analyses

Data were analyzed using SPSS version 24 (Chicago, IL). Unless mentioned otherwise, a statistical significance level of α=0.05 (2-tailed) was used.

Prevalences (frequencies and percentages) were used to describe categorical variables. Continuous data were presented using means and standard deviations or median and interquartile range (IQR) depending on the distribution. Pearson's Chi-Square test compared categorical data except where an expected count was below 5, then Fisher's exact test was used. Depending on distributions, one-way ANOVA or the Kruskal-Wallis test was used to compare multiple groups. In case of significant differences across three or more groups, post hoc analyses were performed to examine which groups differed significantly, using pairwise comparisons (Dunn's test) in the case of the Kruskal-Wallis test and multiple comparisons (Tukey's test) in the case of one-way ANOVA analysis. Pairwise Chi-Square tests were performed in case of categorical variables.

We first used an ANOVA or Kruskal-Wallis test, depending on the distribution, to determine whether SART outcomes differed significantly between groups (aim 1). We used multiple linear regression models for the significant outcomes of the SART, corrected for age, BMI and sex.

Regression analyses were used to determine whether SART outcomes were related to MSLT, ESS, PSG (aim 2) or HADS (aim 4) outcomes after verifying that the relevant key assumptions

were met: linearity, multivariate normality, no multi-collinearity, homoscedasticity and independence of observations. Univariate (with reaction time, commission and omission errors as independent variables) and multivariate linear regression analyses (with additional variables age, sex and BMI) were performed. The dependent variables used in separate models were: MSLT SL, PSG TST, ESS, HADS total, anxiety and depression scores. The variables were entered simultaneously. One outlier (defined as a standardized residual more prominent than 3 in absolute value) was excluded in the association between SART outcomes and the MSLT SL and the PSG TST. As the distributions of the PSG SL, AHI and SE residuals were skewed, Spearman's Rank correlation was used. Where there were significant associations, log transformations were applied to the PSG SL, AHI and SE outcomes and further analyzed in a regression analysis with correction for age, BMI and sex.

The Friedman test and post hoc testing Wilcoxon Signed Ranks Test were used to determine any impact of the five SART session times on SART outcomes (aim 3).

A secondary analysis was performed exclusively on the outcomes of individuals strictly meeting the ICSD3 criteria and those of the CEDS group [28]. The outcomes of this secondary analysis were similar to the primary analysis and can be found in the accompanying supplementary materials.

Results

Study population

Data of 306 individuals were collected, of whom 233 (76%) were tested as part of the normal diagnostic procedure and 73 (24%) were tested as part of their driver's license medical evaluation. Individuals were categorized according to their final clinical diagnosis: NT1 (diagnostic sample: n=43, driver's license sample: n=57), NT2 (diagnostic sample: n=10, driver's license sample=10), IH (diagnostic sample: n=43, driver's license sample 6), OSA (n=27) or CEDS (n=110). Of these, 77.0% of individuals with NT1, 70.0% with NT2, 42.9% with IH and 100% with OSA strictly met the ICSD3 criteria. 3 out of 27 individuals from the OSA group had received CPAP during the diagnostic procedures, which they underwent due to remaining hypersomnolence complaints in spite of treatment. Descriptive statistics are presented in Table 4.1 and 4.2. The outcomes of the secondary analysis were congruent with

those of the primary analysis and can be seen in the supplementary material (tables S4.1, S4.2 and S4.4).

Table 4.2. Descriptive statistics and diagnostic outcomes for all subgroups of the driver's license sample

Driver's license sample (n=73)	1 NT1 (n=57)	2 NT2 (n=10)	3 IH (n=6)	Test statistic	p-value
Age in years	37 (25-48)	41 (30-58)	29 (27-44)	H=1.469	.480
Sex, count (% male)	32 (56.1)	7 (70.0)	2 (33.3)	FET	.373
BMI, n	26.8 (23-30), 53	25.1 (24-33), 10	21.7 (19-26), 6	H=4.722	.094
Medication:	n=57	n=10	n=6		
Use of stimulants, SXB and/or AD: count/n (%)	52 (91.2)	9 (90.0)	4 (66.7)	FET	.195
Stimulants: count (%)	42 (73.7)	8 (80.0)	4 (66.7)	FET	.802
Sodium oxybate: count (%)	27 (47.4) *2,3	1 (10.0) *1	0 (0.0) *1	FET	.008*
Antidepressant: count (%)	14 (24.6)	0 (0.0)	0 (0.0)	FET	.088
Questionnaires:	n=56	n=10	n=6		
ESS	7.0 (6-9) *3	8.5 (7-9) *3	5.5 (3-7) *1,2	H=6.253	.044*
SART outcomes:	n=57	n=10	n=6		
Commission errors	4.8 (2-7)	3.5 (2-5)	2.0 (2-4)	H=3.154	.207
Omission errors	0.3 (0-1)	0.6 (0-2)	0.4 (0-1)	H=.514	.774
Total errors	5.8 (3-9)	3.9 (2-7)	2.8 (2-5)	H=3.663	.160
Reaction time, ms	413 (359-468)	384 (350-434)	434 (402-515)	H=3.417	.181

Abbreviations: NT1, narcolepsy type 1; NT2, narcolepsy type 2; IH, idiopathic hypersomnia; Driver's license sample, individuals with treated sleep disorders as concluded by the Dutch central driving license office.
Notes: Median (IQR) was used unless specified otherwise. * p<.05. Significant pairwise difference with: (1) NT1, (2) NT2, (3) IH.

Table 4.1. Descriptive statistics of characteristics and diagnostic outcomes for all subgroups of the diagnostic sample

Diagnostic sample (n=233)	1 NT1 (n=43)	2 NT2 (n=10)	3 IH (n=43)	4 OSA (n=27)	5 CEDS (n=110)	Test statistic	p-value
Age in years	27 (21-44) *2,4,5	24 (19-27) *1,4,5	29 (24-41) *4,5	58 (47-65) *1,2,3,5	47 (29-54) *1,2,3,4	H=64.942	<.001*
Sex: count (% male)	21 (48.8) *2,4	0 (0.0) *1,3,4,5	16 (37.2) *2,4	24 (88.9) *1,2,3,5	48 (43.6) *2,4	FET	<.001*
BMI	25.0 (23-28) *4	23.1 (21-27) *4	24.6 (21-28) *4	29.9 (26-33) *1,2,3,5	25.2 (22-29) *4	H=18.899	.001*
Use of antidepressants: count (%)	0 (0.0) *3,5	0 (0.0)	7 (16.3) *1	3 (11.1)	23 (20.9) *1	FET	.003*
PSG:	n=42	n=10	n=41	n=27	n=108		
SL in minutes	5.4 (2-9) *3,4,5	4.7 (2-10) *3,4,5	10.4 (7-15) *1,2,5	10.8 (7-16) *1,2	15.4 (7-26) *1,2,3	H=33.400	<.001*
TST in hours: mean ±SD	6.4 ±1.1	7.3 ±1.4	7.1 ±0.7 *4,5	6.2 ±1.2 *3	6.4 ±1.2 *3	F=5.038	.001*
TIB, hours, mean ±SD	7.6 ±1.0	7.8 ±1.4	7.9 ±0.8	7.5 ±1.0	7.5 ±1.1	F=1.167	.326
SE	88.3 (80-92) *2,3	94.7 (91-96) *1,4,5	92.5 (89-94) *1,4,5	84.2 (79-93) *2,3	87.5 (82-91) *2,3	H=28.808	<.001*
SOREMP present, count (%)	20 (47.6) *3,4,5	5 (50.0) *3,4,5	1 (2.4) *1,2	0 (0.0) *1,2	1 (0.9) *1,2	FET	<.001*
AHI	1.4 (0.1-4.6) *3,4	0.4 (0.2-1.8) *4	0.4 (0.1-1.3) *1,4,5	18.4 (13.7-31.8) *1,2,3,5	1.6 (0.2-4.2) *3,4	H=77.891	<.001*
Location: count (% ambulatory)	38 (90.5)	10 (100.0)	34 (82.9)	24 (88.9)	99 (91.7)	FET	.501
MSLT:	n=42	n=10	n=43	n=26	n=109		
SL, minutes	4.6 (3-7) *3,4,5	3.7 (2-5) *3,4,5	8.5 (6-11) *1,2,5	10.0 (6-13) *1,2,5	13.5 (11-16) *1,2,3,4	H=91.164	<.001*
Number of SOREMPs	3 (2-4) *3,4,5	3 (1-4) *3,4,5	0 (0-0) *1,2	0 (0-0) *1,2	0 (0-0) *1,2	H=144.557	<.001*
SOREMP present, count (%)	36 (85.7) *3,4,5	8 (80.0) *3,4,5	5 (11.6) *1,2	3 (11.5) *1,2	5 (4.6) *1,2	FET	<.001*
Questionnaires:	n=36	n=7	n=34	n=21	n=81		
ESS, n	16 (14-19), 36 *5	16 (15-18), 7	15 (12-17), 34	15 (10-19), 21	13 (10-16), 76 *1	H=18.315	.001*
HADS, n	11.5 (7-14), 32	5.5 (3-11), 6	11.5 (8-15), 28	12.0 (5-21), 13	11.0 (7-15), 81	H=6.263	.180
-Anxiety	7.0 (4-9)	2.5 (1-4)	6.0 (3-8)	7.0 (3-11)	6.0 (4-9)	H=8.518	.074
-Anxiety score ≥8, count (% ≥8)	11 (34.4)	0 (0.0)	9 (32.1)	4 (30.8)	27 (33.3)	FET	.603
-Depression	5.0 (3-7)	3.0 (2-6)	6.0 (3-9)	5.0 (2-10)	5.0 (3-9)	H=3.010	.556
-Depression score ≥8, count (% ≥8)	7 (21.9)	1 (16.7)	9 (32.1)	6 (46.2)	24 (29.6)	FET	.554

Abbreviations: NT1, narcolepsy type 1; NT2, narcolepsy type 2; IH, idiopathic hypersomnia; OSA, obstructive sleep apnea; CEDS, complaints of excessive daytime sleepiness; FET, Fisher's exact test; BMI, body mass index; PSG, polysomnography; SL, sleep latency; TST, total sleep time; TIB, time in bed; SE, sleep efficiency; SOREMp, sleep-onset rapid eye movement period; AHI, apnea-hypopnea index; MSLT, Multiple Sleep Latency Test; ESS, Epworth Sleepiness Scale; HADS, Hospital Anxiety and Depression Scale.
Notes: Median (IQR) was used unless specified otherwise. * p<.05. Significant pairwise difference with: (1) NT1, (2) NT2, (3) IH, (4) OSA, (5) CEDS.

Individual characteristics and diagnostic outcomes

In the diagnostic sample, the OSA group was significantly older and significantly higher BMI than the other groups. The OSA group also had substantially more males ($p<.05$). The NT2 group was younger and had a lower proportion of males (0%) than the NT1, CEDS and OSA groups ($p<.05$).

The NT1 and NT2 groups had a shorter MSLT and PSG sleep latency and more SOREMPs than the other groups (all $p<.05$). SOREMPs were not limited to narcolepsy patient groups: 5 out of 109 individuals with CEDS, 3 out of 26 individuals with OSA and 6 out of 43 individuals with IH had at least one SOREMP during the MSLT and PSG. ESS scores were higher in the NT1 group than CEDS group ($p<.05$). The percentage of antidepressant use was lower in the narcolepsy type 1 group than the IH and CEDS groups ($p<.05$). Sodium oxybate and stimulants were not used in any of the diagnostic sample subgroups. HADS scores did not differ significantly between groups.

The characteristics of the driver's license group (see table 4.2) did not differ between diagnostic groups (NT1, NT2 and IH).

Vigilance across diagnostic groups

Differences in SART outcomes were analyzed between diagnostic groups (see table S4.3). Figure 4.2 illustrates the median outcomes, the IQR, the minimum and maximum for each group and significance levels corrected for age, sex and BMI. No significant differences were found. Additional correction for antidepressant use did not substantially impact the outcomes.

Confounders

Age was negatively associated to SART commission errors ($B=-.058$, $p=.013$, $R^2=.024$), and positively associated to SART reaction time ($B=1.468$, $p<.001$, $R^2=.092$). Age was not associated to other SART outcomes. Correction for diagnosis did not change these outcomes. Sex and BMI were not associated with any SART outcomes.

Associations between SART and PSG, MSLT and ESS

There were no associations between MSLT sleep latency and any SART outcomes in both the univariate and multivariate regression analysis. PSG TST, SE and SL were also not significantly associated with SART RT, commission or omission errors. A higher AHI was associated with a longer reaction time when not adjusted for confounders ($p<.001$), however after adjusting for age, sex and BMI the association was no longer significant ($p=.259$).

There were significant positive associations between the ESS score and SART RT and commission errors, in the univariate and multivariate regression analyses. I.e. subjects with higher sleepiness scores tend to react slower and make more commission errors (see table 4.3). These significant associations remained when only including individuals from the diagnostic sample (see table S4.5) and adding antidepressant use as independent variable in the multivariate analysis.

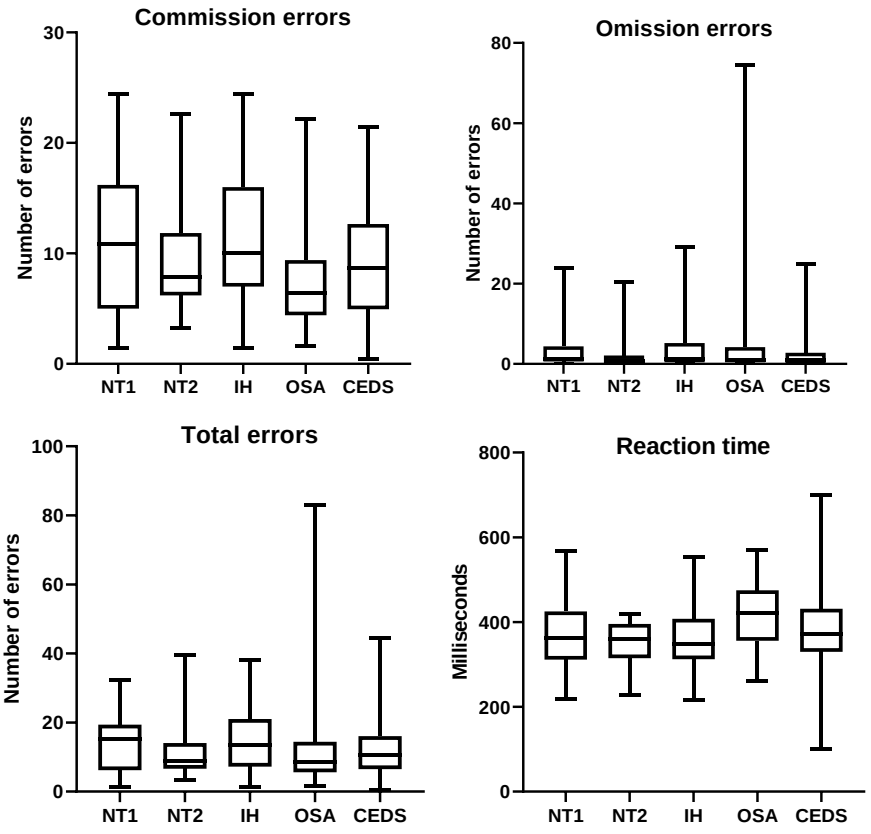


Figure 4.2. Differences SART outcomes between diagnostic sample groups: narcolepsy type 1 (NT1, n=43), narcolepsy type 2 (NT2, n=10), idiopathic hypersomnia (IH, n=43), obstructive sleep apnea (OSA, n=27), individuals with complaints of excessive daytime sleepiness (CEDS, n=110). Boxplots represent median, IQR, minimum and maximum.

Table 4.3. Associations between ESS as dependent variable, and: SART outcomes (univariate) or SART outcomes and general characteristics (multivariate)

	Dependent variable: ESS score			
	Univariate (n=245)		Multivariate (n=241)	
	Estimate	p-value	B	p-value
SART Reaction time	.018	.006*	.021	.002*
SART Commission errors	.448	<.001*	.458	<.001*
SART Omission errors	-.089	.295	-.072	.417
Age	[a]	[a]	-.028	.228
Sex	[a]	[a]	1.023	.110
BMI	[a]	[a]	.071	.262
Model information	R^2= .093, p<.001*		R^2=.108, p<.001*	

Abbreviations: ESS, Epworth Sleepiness Scale.
Notes: * p<.05. [a] Unused variables in univariate regression analysis.

Relationship between depression, anxiety, and vigilance

There was no significant association between HADS total score (or the anxiety and depression sub-scores) and any SART outcome.

Comparisons of SART outcomes between individuals from the diagnostic sample group based on HADS depression and anxiety cut-off scores are shown in table 4.4. People with a higher anxiety score (≥8) made near-significantly more omission errors than those with a lower score (<8, p=.055).

Table 4.4. Comparisons in SART outcomes based on HADS depression and anxiety cut-off scores

Diagnostic sample (n=160)	Depression<8 (n=113)	Depression≥8 (n=47)	Test statistic	p-value
Commission errors	8.6 (5-14)	9.4 (5-14)	H=2862	.440
Omission errors	0.8 (0-3)	1.0 (0-3)	H=2779	.644
Total errors	9.8 (6-18)	11.0 (6-21)	H=2895	.371
Reaction time, ms	372 (326-427)	346 (316-446)	H=2570	.749
	Anxiety<8 (n=109)	Anxiety≥8 (n=51)		
Commission errors	8.8 (5-14)	9.0 (5-14)	H=2879	.591
Omission errors	0.8 (0-3)	1.4 (0-3)	H=3302	.055
Total errors	9.8 (6-18)	11.0 (7-19)	H=3136	.192
Reaction time, ms	370 (323-423)	367 (316-455)	H=2817	.892

Abbreviations: HADS, Hospital Anxiety and Depression Scale.
Notes: Median (IQR) was used unless specified otherwise. * p<.05.

Time of day effects on vigilance outcomes

Commission and total errors did not differ significantly between SART sessions (see figure 4.3).

Reaction time was significantly longer during the first 3 sessions than in the last 2 (median and IQR: session 1, 382 (343-439); session 2, 381 (332-446); session 3, 380 (328-435); session 4, 373 (326-422); session 5, 355 (314-414), in all comparisons p<.05).

Omission errors, although scarce, differed significantly between sessions 1 and 2 (median and IQR 1 (0-3) vs 0 (0-2), p<.005), 1 and 3 (median and IQR 0 (0-2), p<.001), 2 and 3 (0 (0-2)), 3 and 4 (0 (0-2) vs 1 (0-2), p<.05) and 3 and 5 (1 (0-3), p<.05).

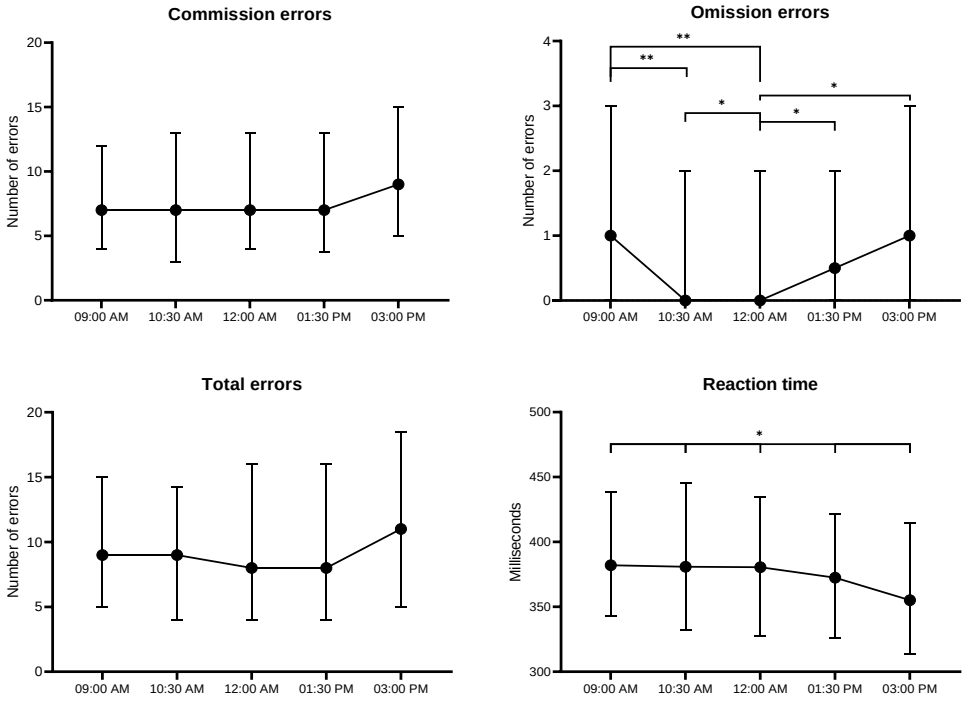


Figure 4.3. Differences in Sustained Attention to Response Task outcomes between five sessions starting at 09:00 AM, 10:30 AM, 12:00 AM, 01:30 PM and 03:00 PM. Median and IQR are shown. * p<.05; ** p<.005. Reaction time differed significantly between all sessions.

Driver's license subgroup

The number of SART errors (omission, commission and total errors) was lower in the driver's license group than in the diagnostic sample (in the NT1 group 0.3 vs 1.1, 4.8 vs 10.6 and 5.8 vs 13.8, see supplemental table S4.2). On the other hand, reaction times were longer in the driver's license sample (in the NT1 group 413 vs 361 ms).

Discussion

We evaluated the applicability of the SART in the diagnostic workup and driver's license evaluations of a tertiary referral center for disorders of hypersomnolence. We assessed the association between SART and PSG outcomes. Additionally, while associations with ESS and MSLT outcomes were previously reported, we examined them in larger samples relevant to clinical practice, including a sample consisting of people with EDS complaints without a diagnosis of a disorder of hypersomnolence. While associations between ESS and MSLT and vigilance outcomes have been assessed before in individuals without explanatory diagnosis [29, 30], the SART was not explicitly evaluated. We conclude that the SART (1) can detect disturbed vigilance in hypersomnolence disorders but cannot differentiate between different conditions, (2) measures a different EDS disease aspect than the MSLT and PSG, (3) is influenced by the time of day and (4) is not influenced by depression. Notably, individuals with increased anxiety (sub-score ≥8) did make near-significantly more omission errors. Additionally, (5) total, commission and omission errors are lower in individuals undergoing a driver's license evaluation, probably because of treatment and/or motivational factors.

It is worth noting that there is currently no universally accepted cut-off for the SART total error score in the context of hypersomnolence disorders. While a cut-off of 5 total errors was established by Fronczek et al (2006) in a small study with 15 untreated narcoleptics and 15 matched controls, it has not been further validated in larger or more diverse populations [12]. Despite this, our assessment of different hypersomnolence disorders demonstrated impaired vigilance, as evidenced by all groups exhibiting a median total error score above this previously determined cut-off. Previous research [15], conducted with healthy participants, showed that this cut-off was only applicable when using the SART with the instruction that accuracy is more important than response speed. Therefore, the finding of our previous descriptive study [13] that individuals with hypersomnolence disorders have abnormal SART outcome measures became questionable. However, abnormal SART results have now also been reproduced in individuals who were instructed to prioritize accuracy, thereby providing further support to previous findings. Nonetheless, to arrive at definitive conclusions, further research is required, which should involve both individuals with hypersomnolence disorders and a healthy control group, as well as testing the effects of differential instructions.

In line with previous research [13], we conclude that the SART is not a diagnostic tool for a specific sleep disorder. However, it may have potential as a monitoring tool during follow-up.

The SART has already been proven capable of measuring treatment effects in individuals with narcolepsy [31, 32].

There was an overlap between subjective EDS as measured by the ESS and vigilance impairment as measured by the SART. To rule out a confounding effect of treatment on these associations, a post hoc analysis was performed using only the diagnostic sample and adjusting for antidepressant use. This association was not found between the SART on the one hand, and PSG and MSLT outcomes on the other hand. Vigilance impairments are thus associated with subjective sleepiness but not with objective sleepiness. This suggests that the SART measures a different disease aspect, which is supported by multiple, previously conducted studies [33-35].

We found a clear association between SART outcomes and age, where older individuals tend to have longer reaction times. Increased age was associated with longer reaction times and fewer commission errors. This may be related to the speed-accuracy trade-off [15], where longer reaction times result in fewer errors. These results are in line with expectations, given the known relationship between decreasing reaction times with increasing age [36, 37]. It is possible that age needs to be taken into account when interpreting SART results. Notably, the primary SART outcome measure, total errors, was not significantly associated with age, making this association less relevant in clinical practice.

The number of total errors also did not differ significantly between different times of the day. The time of the day significantly impacted reaction times and the number of omission errors, with reaction time being higher in the morning and fewer omission errors around noon. The differences in omission errors were marginal and not clinically relevant. Based on previous research we suspect these differences are more likely to be caused by a time of day effect than a learning effect [15]. Unlike two previous studies by our research group, we did not find that total error scores were significantly higher in the first session [13, 15]. The longer reaction times in the early morning confirmed past results [15]. We found no evidence of a clinically relevant time-of-day effect on total errors scores. However the time of the day should be taken into account when interpreting the reaction times.

There were no associations between subjective depression symptoms and any SART outcomes. This may seem counterintuitive as multiple studies have shown attention deficits in people with depression [38, 39]. In our study population only two individuals had a cut-off score ≥16 on the HADS (both from the CEDS group), indicating severe depression symptoms [27]. Therefore, we cannot reliably examine the effect of depression on SART outcomes. It was, however, found that individuals with increased anxiety (sub-score ≥8) made near-significantly

more omission errors. This finding aligns with a previous report that anxiety is associated with response inhibition [40]. As with depression, there were few people in this study with severe anxiety complaints (the maximum anxiety sub-score in the sample was 13/21).

The impact of motivation on SART performance is an interesting topic that needs further examination. The driver's license group made fewer errors during the SART than the diagnostic group. There were substantial differences between these groups: the driver's license group may have been more motivated to stay awake and had received treatment. Despite treatment, individuals with severe complaints would not be considered for a driver's license testing, resulting in selection bias. These factors make it difficult to determine the cause of differences between the groups.

Strengths and limitations

Our study population reflected the daily clinical practice of a highly specialized Sleep-Wake center. Therefore, our OSA group is not generalizable to the general population with OSA, as not all people with OSA will experience complaints of hypersomnolence severe enough to result in referral to a Sleep-Wake center.

Our diagnostic groups were relatively large, but the NT2 group consisted of only 10 individuals, thus making it difficult to draw definite conclusions regarding this sub-group.

In addition, it should be noted that the SART program used in our study was run in priority mode and was not performed on a calibrated platform, which may have introduced some variability in the results. However, the impact on the results is expected to be small given the large sample size used.

Another limitation is related to the setup of the diagnostic procedures. The PSG is always followed by tests during the following day starting at 09:00 AM, this has probably shortened the natural sleep duration of at least some of those included. This means that the PSG TST and TIB results are probably underestimations and need to be interpreted with care.

Finally, no healthy control group was included in this study with which the sub-groups could be compared.

Conclusion

Impaired vigilance can significantly affect quality of life, as vigilance is needed for everyday tasks at work, school, parenting tasks, household, and social interactions such as a conversation [41, 42]. Thus, not only EDS complaints but also vigilance impairment should be regularly monitored in individuals with central disorders of hypersomnolence. The SART can be used to assess vigilance in these people. Unlike the MSLT and PSG, the SART is a short and easy to administer task and a good addition to the standard diagnostic workup and monitoring of central disorders of hypersomnolence.

Supplementary material

Table S4.1. Descriptive statistics and diagnostic outcomes for all subgroups of the diagnostic sample (ICSD3 and CEDS)

Diagnostic sample (n=198)	1 NT1 (n=34)	2 NT2 (n=9)	3 IH (n=18)	4 OSA (n=27)	5 CEDS (n=110)	Test statistic	p-value
Age in years	27 (21-41) *4,5	24 (19-27) *4,5	31 (23-43) *4,5	58 (47-65) *1,2,3,5	47 (29-54) *1,2,3,4	H=56.988	<.001*
Sex, count (% male)	16 (47.1) *2,4	0 (0.0) *1,4,5	6 (33.3) *4	24 (88.9) *1,2,3,5	48 (43.6) *2,4	FET	.001*
BMI	25.0 (23-27), *4	23.2 (21-27) *4	25.6 (23-29) *4	29.9 (26-33) *1,2,3,5	25.2 (22-29) *4	H=14.234	.007*
Use of antidepressants: count (%)	0 (0.0) *3,5	0 (0.0)	4 (22.2) *1	3 (11.1)	23 (20.9) *1	FET	.007*
PSG:	n=33	n=9	n=17	n=27	n=108		
SL, minutes	4.8 (2-8) *4,5	4.8 (2-11) *5	9.4 (5-13) *5	10.8 (7-16) *1	15.4 (7-26) *1,2,3	H=32.674	<.001*
TST, hours, mean ±SD	6.3 ±1.2 *2	7.6 ±1.0 *1,4,5	7.0 ±0.7	6.2 ±1.2 *3	6.4 ±1.2 *3	F=3.882	.005*
TIB, hours, mean ±SD	7.5 ±1.1	8.1 ±1.1	7.9 ±0.8	7.5 ±1.0	7.5 ±1.1	F=.925	.450
SE	87.2 (79-92) *2,3	95.1 (92-96) *1,4,5	92.3 (89-95) *1,4,5	84.2 (79-93) *2,3	87.5 (82-91) *2,3	H=23.403	<.001*
SOREMP present, count (%)	17 (51.5) *3,4,5	5 (55.6) *3,4,5	1 (5.9) *1,2	0 (0.0) *1,2	1 (0.9) *1,2	FET	<.001*
AHI	1.5 (0.1-5.0) *4	0.5 (0.2-1.8) *4	0.5 (0.2-2.1) *4	18.4 (13.7-31.8) *1,2,3,5	1.6 (0.2-4.2) *4	H=70.092	<.001*
Location: count (% ambulatory)	30 (90.9)	9 (100.0)	15 (88.2)	24 (88.9)	99 (91.7)	FET	.874
MSLT:	n=34	n=9	n=18	n=26	n=109		
SL, minutes	3.8 (2-6) *4,5	3.5 (2-4) *4,5	5.7 (5-7) *4,5	10.0 (6-13) *1,2,3,5	13.5 (11-16) *1,2,3,4	H=106.411	<.001*
Number of SOREMPs	3 (3-4) *3,4,5	3 (1-5) *3,4,5	0 (0-0) *1,2	0 (0-0) *1,2	0 (0-0) *1,2	H=149.447	<.001*
SOREMP present, count (%)	33 (97.1) *3,4,5	7 (77.8) *3,4,5	1 (5.6) *1,2	3 (11.5) *1,2	5 (4.6) *1,2	FET	<.001*
Questionnaire:	n=28	n=6	n=14	n=21	n=81		
ESS, n	16 (14-19), 28 *5	16 (14-19), 6	15 (12-17), 14	15 (10-19), 21	13 (10-16), 76 *1	H=17.599	.001*
HADS, n	12.0 (8-18), 25	5.5 (3-11), 6	9.0 (4-16), 10	12.0 (5-21), 13	11.0 (7-15), 81	H=4.827	.306
-Anxiety	7.0 (4-10)	2.5 (1-4)	6.5 (2-8)	7.0 (3-11)	6.0 (4-9)	H=7.404	.116
-Anxiety score ≥8, count (% ≥8)	10 (40.0)	0 (0.0)	4 (40.0)	4 (30.8)	27 (33.3)	FET	.459
-Depression	5.0 (4-9)	3.0 (2-6)	4.5 (2-8)	5.0 (2-10)	5.0 (3-9)	H=1.819	.769
-Depression score ≥8, count (% ≥8)	7 (28.0)	1 (16.7)	2 (20.0)	6 (46.2)	24 (29.6)	FET	.684

Abbreviations: ICSD3, International Classification of Sleep Disorders third edition; NT1, narcolepsy type 1; NT2, narcolepsy type 2; IH, idiopathic hypersomnia; OSA, obstructive sleep apnea; CEDS, complaints of excessive

daytime sleepiness; FET, Fisher's exact test; PSG, polysomnography; ESS, Epworth Sleepiness Scale; HADS, Hospital Anxiety and Depression Scale.
Notes: Only individuals with a diagnosis meeting the ICSD3 criteria were included. Median (IQR) was used unless specified otherwise. * p<.05. Significant pairwise difference with: (1) NT1, (2) NT2, (3) IH, (4) OSA, (5) CEDS.

Table S4.2. Descriptive statistics and diagnostic outcomes for all subgroups of the driver's license sample (ICSD3)

Driver's license sample (n=48)	NT1 (n=40)	NT2 (n=5)	IH (n=3)	Test statistic	p-value
Age in years	32 (25-48)	32 (24-56)	29	H=.258	.879
Gender, count (% male)	22 (55.0)	4 (80.0)	1 (33.3)	FET	.436
BMI	27.8 (25-32)	30.4 (22-38)	20.9	H=1.639	.441
Questionnaires:	n=39	n=5	n=3		
ESS	7 (6-9)	8 (6-10)	6	H=1.784	.410
SART outcomes:	n=40	n=5	n=3		
Commission errors	4.6 (1-7)	2.8 (2-4)	1.8	H=1.484	.476
Omission errors	0.3 (0-1)	1.5 (0-2)	0.0	H=1.265	.531
Total errors	5.6 (2-9)	2.8 (2-6)	2.0	H=1.907	.385
Reaction time, ms	412 (346-469)	388 (368-458)	437	H=1.805	.406

Abbreviations: ICSD3, International Classification of Sleep Disorders third edition; NT1, narcolepsy type 1; NT2, narcolepsy type 2; IH, idiopathic hypersomnia; Driver's license sample, individuals with treated sleep disorders as concluded by the Dutch central driving license office.
Notes: Only individuals with a diagnosis meeting the ICSD3 criteria were included. Median (IQR) was used unless specified otherwise.

Table S4.3. Differences SART outcomes between subgroups of the diagnostic sample

Total sample (n=233)	1 NT1 (n=43)	2 NT2 (n=10)	3 IH (n=43)	4 OSA (n=27)	5 CEDS (n=110)	Test statistic	p-value
Commission errors	10.8 (5-16)	7.9 (6-12)	10.0 (7-16)	6.4 (4-9)	8.7 (5-13)	H=7.897	.095
Omission errors	1.2 (1-4)	0.6 (0-2)	1.2 (0-5)	1.0 (0-4)	0.9 (0-3)	H=2.279	.685
Total errors	15.2 (6-19)	8.8 (7-14)	13.6 (7-21)	8.6 (6-14)	10.6 (6-16)	H=5.801	.214
Reaction time, ms	363 (311-425) *4	360 (315-395)	348 (312-408) *4	421 (356-475) *1,3,5	372 (330-431) *4	H=10.933	.027*

Abbreviations: NT1, narcolepsy type 1; NT2, narcolepsy type 2; IH, idiopathic hypersomnia; OSA, obstructive sleep apnea; CEDS, complaints of excessive daytime sleepiness
Notes: Only individuals with either CEDS or a diagnosis meeting the ICSD3 criteria were included. Median (IQR) was used. * p<.05. Significant pairwise difference with: (1) NT1, (2) NT2, (3) IH, (4) OSA, (5) CEDS.

Table S4.4. Differences SART outcomes between subgroups of the diagnostic sample (ICSD3 and CEDS)

Total sample (n=198)	1 NT1 (n=34)	2 NT2 (n=9)	3 IH (n=18)	4 OSA (n=27)	5 CEDS (n=110)	Test statistic	p-value
Commission errors	10.1 (5-16)	8.0 (6-12)	10.6 (8-18)	6.4 (4-9)	8.7 (5-13)	H=8.014	.091
Omission errors	1.0 (0-4)	0.6 (0-3)	1.1 (0-8)	1.0 (0-4)	0.9 (0-3)	H=1.145	.887
Total errors	11.0 (6-19)	9.6 (6-18)	15.9 (8-23)	8.6 (6-14)	10.6 (6-16)	H=5.826	.214
Reaction time, ms	375 (315-430) *4	357 (302-399) *4	345 (319-430) *4	421 (356-475) *1,2,3,5	372 (330-431) *4	H=9.304	.054

Abbreviations: ICSD3, International Classification of Sleep Disorders third edition; NT1, narcolepsy type 1; NT2, narcolepsy type 2; IH, idiopathic hypersomnia; OSA, obstructive sleep apnea; CEDS, complaints of excessive daytime sleepiness

Notes: Only individuals with either CEDS or a diagnosis meeting the ICSD3 criteria were included. Median (IQR) was used. * p<.05. Significant pairwise difference with: (1) NT1, (2) NT2, (3) IH, (4) OSA, (5) CEDS.

Table S4.5. Associations between ESS as dependent variable, and: SART outcomes (univariate) or SART outcomes and general characteristics (multivariate) in the diagnostic sample

	Dependent variable: ESS score Univariate (n=172)		Multivariate (n=172)	
	Estimate	p-value	B	p-value
SART Reaction time	.014	.070	.019	.020*
SART Commission errors	.192	.078	.219	.045*
SART Omission errors	-.049	.596	-.046	.622
Age	[a]	[a]	-.050	.047*
Sex	[a]	[a]	.860	.209
BMI	[a]	[a]	.106	.149
Model information	R^2= .022, p=.295		R^2=.059, p=.116	

Abbreviations: ESS, Epworth Sleepiness Scale. Notes: * p<.05. [a] Unused variables in univariate regression analysis.

References

1. Davies, D.R. and R. Parasuraman, *The psychology of vigilance*. 1982: Academic Press.
2. Oken, B.S., M.C. Salinsky, and S.M. Elsas, *Vigilance, alertness, or sustained attention: physiological basis and measurement.* Clinical neurophysiology : official journal of the International Federation of Clinical Neurophysiology, 2006. **117**(9): p. 1885-1901.
3. van Schie, M.K.M., et al., *Vigilance: discussion of related concepts and proposal for a definition.* Sleep Med, 2021. **83**: p. 175-181.
4. Khan, Z. and L.M. Trotti, *Central Disorders of Hypersomnolence: Focus on the Narcolepsies and Idiopathic Hypersomnia.* Chest, 2015. **148**(1): p. 262-273.
5. Al-Shargie, F., et al., *Vigilance Decrement and Enhancement Techniques: A Review.* Brain sciences, 2019. **9**(8): p. 178.
6. Mathis, J. and C.W. Hess, *Sleepiness and vigilance tests.* Swiss Med Wkly, 2009. **139**(15-16): p. 214-9.
7. Lyznicki, J.M., et al., *Sleepiness, driving, and motor vehicle crashes. Council on Scientific Affairs, American Medical Association.* Jama, 1998. **279**(23): p. 1908-13.
8. Komada, Y., et al., *Difference in the characteristics of subjective and objective sleepiness between narcolepsy and essential hypersomnia.* Psychiatry Clin Neurosci, 2005. **59**(2): p. 194-9.
9. Carskadon, M.A., et al., *Guidelines for the multiple sleep latency test (MSLT): a standard measure of sleepiness.* Sleep, 1986. **9**(4): p. 519-24.
10. Johns, M.W., *Reliability and factor analysis of the Epworth Sleepiness Scale.* Sleep, 1992. **15**(4): p. 376-81.
11. Jung, C.M., et al., *Comparison of sustained attention assessed by auditory and visual psychomotor vigilance tasks prior to and during sleep deprivation.* Journal of sleep research, 2011. **20**(2): p. 348-355.
12. Fronczek, R., et al., *Focusing on vigilance instead of sleepiness in the assessment of narcolepsy: high sensitivity of the Sustained Attention to Response Task (SART).* Sleep, 2006. **29**(2): p. 187-91.
13. Van Schie, M.K., et al., *Sustained attention to response task (SART) shows impaired vigilance in a spectrum of disorders of excessive daytime sleepiness.* J Sleep Res, 2012. **21**(4): p. 390-5.
14. Whyte, J., et al., *Measuring sustained attention after traumatic brain injury: Differences in key findings from the sustained attention to response task (SART).* Neuropsychologia, 2006. **44**(10): p. 2007-2014.
15. van Schie, M.K., et al., *The influences of task repetition, napping, time of day, and instruction on the Sustained Attention to Response Task.* J Clin Exp Neuropsychol, 2014. **36**(10): p. 1055-65.
16. Wilkinson, R.T. and D. Houghton, *Field test of arousal: a portable reaction timer with data storage.* Hum Factors, 1982. **24**(4): p. 487-93.
17. Bennett, L.S., J.R. Stradling, and R.J. Davies, *A behavioural test to assess daytime sleepiness in obstructive sleep apnoea.* J Sleep Res, 1997. **6**(2): p. 142-5.
18. van Schie, M.K., et al., *Improved vigilance after sodium oxybate treatment in narcolepsy: a comparison between in-field and in-laboratory measurements.* J Sleep Res, 2016. **25**(4): p. 486-96.
19. Chokroverty, S., *Sleep apnea in narcolepsy.* Sleep, 1986. **9**(1 Pt 2): p. 250-3.
20. Pataka, A.D., et al., *Narcolepsy and sleep-disordered breathing.* Eur J Neurol, 2012. **19**(5): p. 696-702.
21. Robertson, I.H., et al., *`Oops!': Performance correlates of everyday attentional failures in traumatic brain injured and normal subjects.* Neuropsychologia, 1997. **35**(6): p. 747-758.
22. Baiardi, S. and S. Mondini, *Inside the clinical evaluation of sleepiness: subjective and objective tools.* Sleep Breath, 2020. **24**(1): p. 369-377.
23. Ibanez, V., J. Silva, and O. Cauli, *A survey on sleep assessment methods.* PeerJ, 2018. **6**: p. e4849.
24. Iber, C., et al., *The AASM Manual for the Scoring of Sleep and Associated Events: Rules, Terminology and Technical Specifications.* Westchester, IL: American Academy of Sleep Medicine, 2007.
25. Johns, M.W., *A new method for measuring daytime sleepiness: the Epworth sleepiness scale.* Sleep, 1991. **14**(6): p. 540-5.
26. Brennan, C., et al., *The Hospital Anxiety and Depression Scale: A diagnostic meta-analysis of case-finding ability.* Journal of Psychosomatic Research, 2010. **69**(4): p. 371-378.
27. Bjelland, I., et al., *The validity of the Hospital Anxiety and Depression Scale: An updated literature review.* Journal of Psychosomatic Research, 2002. **52**(2): p. 69-77.
28. American Academy of Sleep, M., *The international classification of sleep disorders*. 2014.
29. Ramm, M., et al., *Disease-specific attention impairment in disorders of chronic excessive daytime sleepiness.* Sleep Medicine, 2019. **53**: p. 133-140.
30. Mathis, J., et al., *The diagnostic value of sleep and vigilance tests in central disorders of hypersomnolence.* Sleep, 2022. **45**(3).
31. Dauvilliers, Y., et al., *Pitolisant versus placebo or modafinil in patients with narcolepsy: a double-blind, randomised trial.* The Lancet Neurology, 2013. **12**(11): p. 1068-1075.

32. van der Heide, A., et al., *Comparing Treatment Effect Measurements in Narcolepsy: The Sustained Attention to Response Task, Epworth Sleepiness Scale and Maintenance of Wakefulness Test.* Sleep, 2015. **38**(7): p. 1051-8.
33. Chervin, R.D., et al., *Comparison of the results of the Epworth Sleepiness Scale and the Multiple Sleep Latency Test.* Journal of Psychosomatic Research, 1997. **42**(2): p. 145-155.
34. Sangal, R.B., M.M. Mitler, and J.M. Sangal, *Subjective sleepiness ratings (Epworth sleepiness scale) do not reflect the same parameter of sleepiness as objective sleepiness (maintenance of wakefulness test) in patients with narcolepsy.* Clinical Neurophysiology, 1999. **110**(12): p. 2131-2135.
35. Furuta, H., et al., *Epworth Sleepiness Scale and sleep studies in patients with obstructive sleep apnea syndrome.* Psychiatry and Clinical Neurosciences, 1999. **53**(2): p. 301-302.
36. Fozard, J.L., et al., *Age Differences and Changes in Reaction Time: The Baltimore Longitudinal Study of Aging.* Journal of Gerontology, 1994. **49**(4): p. P179-P189.
37. Bento-Torres, N.V., et al., *Influence of schooling and age on cognitive performance in healthy older adults.* Braz J Med Biol Res, 2017. **50**(4): p. e5892.
38. Li, X., et al., *Study on the executive function of attention in depression patients based on SPECT technology.* International journal of clinical and experimental medicine, 2014. **7**(4): p. 1110-1115.
39. Keller, A.S., et al., *Paying attention to attention in depression.* Translational Psychiatry, 2019. **9**(1): p. 279.
40. Grillon, C., et al., *Clinical anxiety promotes excessive response inhibition.* Psychological medicine, 2017. **47**(3): p. 484-494.
41. Valley, V. and R. Broughton, *Daytime performance deficits and physiological vigilance in untreated patients with narcolepsy-cataplexy compared to controls.* Rev Electroencephalogr Neurophysiol Clin, 1981. **11**(1): p. 133-9.
42. Lee, I.-S., et al., *The Relationship between psychomotor vigilance performance and quality of life in obstructive sleep apnea.* Journal of clinical sleep medicine : JCSM : official publication of the American Academy of Sleep Medicine, 2011. **7**(3): p. 254-260.

Part II

Sleep in the Intensive Care Environment

5

Sleep in the intensive and intermediate care units

exploring related factors of delirium, benzodiazepine use and mortality

Adrienne E. van der Hoeven[1,2], Denise Bijlenga[1,2], Ernst van der Hoeven, Mink S. Schinkelshoek[1,2], Floor W. Hiemstra[3,4], Laura Kervezee[4], David J. van Westerloo[3], Rolf Fronczek[1,2], Gert Jan Lammers[1,2]

[1]Department of Neurology, Leiden University Medical Center, Leiden, the Netherlands, [2]Stichting Epilepsie Instellingen Nederland (SEIN), Sleep-Wake Center, Heemstede, the Netherlands, [3]Department of Intensive Care, Leiden University Medical Center, Leiden, the Netherlands, [4]Group of Neurophysiology, Department of Cell and Chemical Biology, Leiden University Medical Center, Leiden, the Netherlands

Intensive and Critical Care Nursing. 2024 January

Abstract

Aim of the study

The primary purpose was to examine sleep difficulties and delirium in the Intensive and Intermediate Care Unit. Secondarily, factors impacting night-time sleep duration and quality, mortality, and the impact of benzodiazepine use on sleep outcomes were investigated.

Materials and Methods

This retrospective study encompassed data from 323 Intensive and Intermediate Care Unit admissions collected in the Netherlands, spanning from November 2018 to May 2020. Sleep quality was measured using the Richards-Campbell Sleep Questionnaire. Night-time sleep duration was nurse-reported. We investigated associations of these sleep outcomes with age, sex, length-of-stay, natural daylight, disease severity, mechanical ventilation, benzodiazepine use, and delirium using Generalized Estimating Equations models. Associations with 1-year post-discharge mortality were analyzed using Cox regression.

Results

Night-time sleep duration was short (median 4.5h) and sleep quality poor (mean score 4.9/10). Benzodiazepine use was common (24% of included nights) and was negatively associated with night-time sleep duration and quality (B=-.558 and -.533, $p<.001$). Delirium and overnight transfers were negatively associated with sleep quality (B=-.716 and -1.831, $p<.05$). The day-to-night sleep ratio was higher in the three days before delirium onset than in non-delirious individuals ($p<.05$). Age, disease severity and female sex were associated with increased 1-year mortality. Sleep quality was negatively, but not-significantly, associated with mortality ($p=.070$).

Conclusions

Night-time sleep in the critical care environment has a short duration and poor quality. Benzodiazepine use was not associated with improved sleep. Sleep patterns change ahead of delirium onset.

Implications for Clinical Practice

Consistent sleep monitoring should be part of routine nursing practice, using a validated instrument like the Richards-Campbell Sleep Questionnaire. Given the lack of proven efficacy of benzodiazepines in promoting sleep in critical care settings, it is vital to develop more effective sleep treatments that include non-benzodiazepine medication and sleep hygiene strategies.

Statement of significance

This study highlights the prevalence and negative impact of sleep disturbances in critically ill individuals admitted to the Intensive and Intermediate Care Units. Despite the well-documented negative effects of disrupted sleep on immune function, cognition, and wound healing in the general population, there is still a lack of interventions to address this issue. We show that sleep difficulties are common and our study suggests that benzodiazepines, a common hypnotic treatment to improve sleep, are negatively associated to sleep outcomes. One should be cognizant of sleep in this vulnerable population to better understand the relationship with overall wellbeing and clinical outcomes. In addition, more research is required to develop effective treatment strategies to prevent sleep disorders in the Intensive Care Unit.

Introduction

Sleep problems are common in individuals admitted to the Intensive Care Unit (ICU), including disrupted 24-hour sleep-wake cycles, fragmented night-time sleep, and increased use of sleep medication [1, 2], resulting in sleep deprivation and poor sleep quality during the night [3]. In the ICU, persistent sleep disturbances have a variety of negative effects, including an association with the onset of delirium [4], which in turn is associated with increased morbidity and mortality [5, 6], and psychological complaints after discharge [7]. A meta-analysis of population-based studies showed that deviation from 7-8 hours of sleep during the night is associated with an increased risk of cardiovascular events and mortality [8]. Animal studies also show causational effect of prolonged sleep deprivation on mortality [9, 10]. An ICU study by Boyko et al. shows an association between increased 90-day mortality and atypical sleep patterns observed with polysomnography. More specifically, this was linked to the absence of K-complexes or sleep spindles, which are integral to the normal sleep cycle [11]. Currently, the exact associations between sleep disruptions in the ICU, and morbidity and mortality are unknown.

Multiple factors contribute to disturbed nocturnal sleep in an ICU setting, including medication use, mechanical ventilation, severity of illness, and the intensity of required care. Relatively high noise and light levels characterize the critical healthcare environment at night, adding to disrupted natural sleep patterns [1, 2, 12]. Although sleep difficulties in the ICU are well-

documented, there is limited literature on sleep difficulties in the Intermediate Care Unit (IMCU), probably due to its more recent establishment and the relatively low number of IMCUs worldwide. The IMCU commonly provides care for individuals who are too ill for a general ward but not ill enough for ICU-level care [13].

Objective sleep measurements are not part of routine critical care. Polysomnography, the gold standard [14, 15], can be difficult to implement in the ICU/IMCU due to the labor-intensive nature of the test, the burden it places on critically ill individuals, and the difficulties in scoring the results in this population due to disruption of the signals [15, 16]. Actigraphy, while easier to apply, determines sleep primarily based on movement, which is also less ideal for critically ill patients who tend to have reduced mobility [15]. Alternatively, subjective sleep measurements, such as sleep quality or duration reported by patients or nurses, have the benefit of easy integration into routine clinical practice. The Richards-Campbell Sleep Questionnaire (RCSQ) is a brief sleep assessment tool that is frequently used in ICUs to measure sleep quality during the night [17-20].

We aimed to investigate (1) the sleep characteristics and sleep-related outcomes of individuals admitted to the ICU and IMCU; (2) the factors that are associated with decreased sleep quality and duration; (3) whether delirium is preceded by a deteriorated sleep quality and/or a shift in day-to-night distribution of sleep; (4) if the use of benzodiazepines provided as hypnotic to improve nocturnal sleep was associated with improved sleep outcomes; and (5) the associations of various (sleep-related) factors to the 1 year post-ICU discharge mortality.

Material and methods

We conducted a retrospective longitudinal observational study using data from electronic health records (EHR) of all individuals admitted to the ICU and/or IMCU of the Leiden University Medical Center (LUMC) in the Netherlands from November 2018 until May 2020. The design allowed us to investigate potential associations between the different sleep and delirium- related outcome measures, and study factors associated with increased mortality. The period was chosen because of a systems change in 2018 and the start of data collection in 2020. We analyzed data on individuals aged 18 years or older, who were admitted for at least one night to the ICU and/or IMCU, and from whose admission at least one RCSQ score was available.

The ICU of our medical center is located adjacent to the IMCU as part of the same ward. While the IMCU typically caters to patients who are not as critically ill as those in the ICU, interventions such as non-invasive ventilation, specialized wound care, or intravenous medications are available. IMCU patients often include those stepping down from the ICU or those requiring intensive monitoring and specialized treatments not provided in general wards.

Ethics statement

The study was conducted following the Helsinki Declaration as revised in 2013. The study does not fall under the laws of the Medical Research Involving Human Subjects Act. The medical ethical committee of the LUMC (registration number: 22-3043) granted a waiver for the requirement of informed consent.

Assessments

All data were collected as part of routine clinical care, during which they were entered into the EHRs (Metavision). Data were exported from the EHR into a dataset for data-analysis which included age, sex, medication use, subjective sleep parameters, disease severity, presence of delirium, and if the individual received invasive mechanical ventilation during the night. Varying time windows were used to aggregate variables for each included night in order to most effectively address the research objectives (see figure 5.1). We chose to focus on night-time ventilation as it best enabled us to address our research objective concerning the relation between ventilation and sleep outcomes. Moreover, for each admission, admitting specialty, date of admittance and discharge, and whether an overnight transfer took place were added to the dataset. Additionally, the availability of direct natural daylight in the patient's room was included in the dataset. Rooms were either located externally, receiving direct natural daylight, or internally, receiving daylight in a more indirect manner through roof windows located above an internal plaza. The variable availability of direct natural daylight was based solely on the location: can there be direct daylight in the patient's room or not. Missing data were noted if scores were unavailable.

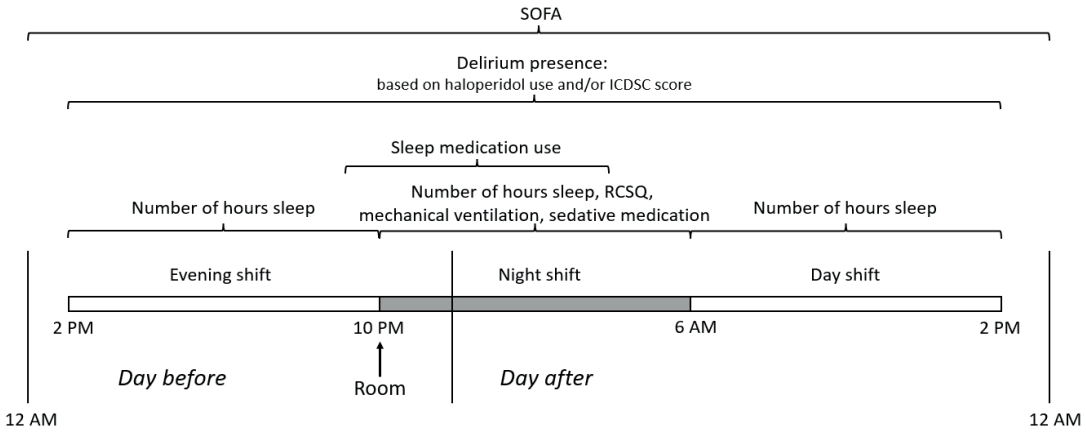


Figure 5.1. The scores, room number and medication use data were aggregated for each included night using different time windows. Collected data were: the average Sequential Organ Failure Assessment (SOFA) score of the date before and after the included night, delirium presence (based on haloperidol use and the Intensive Care Delirium Screening Checklist, i.e. ICDSC) between 2 PM on the day before and after the night shift, sleep medication use between 9 PM on day 1 and 4 AM on day 2, the Richards-Campbell Sleep Questionnaire (RCSQ) score taken at 6 AM, the number of hours sleep during each shift, minutes of sedative medication use (propofol and/or continuous midazolam), whether invasive mechanical ventilation took place during the night shift, and the room number of the individuals at 10 PM.

Concerning medication, the use of non-continuous benzodiazepines and melatonin between 9:00 PM and 4:00 AM were extracted from the EHR. This time window was chosen to capture the period during which the administration of sleep promoting medication was anticipated to have a substantial impact on the sleep outcomes measured at 6 AM. Temazepam, lorazepam, oxazepam, midazolam, and diazepam were registered. Zopiclone was also registered due to its similar sleep-promoting properties. Medication usage was entered into EHRs by nurses who thereby confirmed that the medication had been dispensed. A distinction was made between continuous and non-continuous use of midazolam, as the former is used for sedative purposes. Use of continuous midazolam and propofol between 10:00 PM and 6:00 AM was collected in the dataset. A continuous sedation-free night was defined as a night with less than 60 minutes of sedation with continuous midazolam and/or propofol. This threshold was chosen to differentiate between minor and more significant sedation effects on sleep, as opposed to no sedation or sedation at any time point.

When information about an individual's death has reached the health care provider/LUMC (via family or other health care providers), the date of death is registered in the EHR. Mortality up to 1 year after discharge was also extracted by the first author from the EHR used in the general

ward (HiX 6.3). If survival status after 1 year after discharge was unknown, we recorded the last contact date.

The recorded subjective sleep parameters involved sleep duration and sleep quality. The sleep duration was estimated based on observations and reported in hours for each shift by the attending nurse as part of daily routine, aligning with the rigorous sleep observation practices emphasized in a consistent manner in both the ICU and IMCU. Sleep quality was assessed with the Richards-Campbell Sleep Questionnaire (RCSQ), which is a brief assessment that is commonly used in the ICU to measure sleep quality during the night. It consists of five questions that can be answered by either the hospitalized individual or the attending nurse and has been validated for use in the ICU [19]. The questions pertain to sleep quality, sleep depth, number of awakenings, sleep-onset latency and sleep latency after an awakening [17-20]. The RCSQ is scored from 0 to 10 for each question with a higher score indicating better sleep quality. In this study the total score is the average of the subscores. A cut-off of <5 is commonly used to indicate poor sleep [19]. The RCSQ has been found to be a reliable measure of sleep in critically ill populations and is routinely used in the LUMC hospital to monitor sleep quality since 2016 [21, 22]. The RCSQ score is assessed only during nights free of continuous sedation. The attending nurse must confirm this status to record questionnaire responses and scores in the healthcare record. In addition, the day-night sleep ratio was calculated, which was determined by dividing the number of hours slept during the day and evening shifts (6:00 AM - 10:00 PM) by the number of hours slept during the night shift (10:00 PM - 6:00 AM). All sleep outcomes were retrieved from the EHRs. The sole exceptions were average sleep outcomes before and after benzodiazepine initiation, which were manually included in a dataset aggregated per admission by the first author.

Disease severity was assessed daily using the Sequential Organ Failure Assessment (SOFA). This is a widely used tool for the assessment of organ dysfunction severity in critically ill individuals. It consists of six organ systems: respiratory, cardiovascular, hepatic, renal, neurological, and coagulation. Each system is scored on a scale of 0 to 4, with the total score ranging from 0 to 24, with higher scores indicating greater illness severity [23]. The SOFA score has good validity and reliability to predict ICU mortality in various populations [24-26]. Presence of delirium was identified using the Intensive Care Delirium Screening Checklist (ICDSC) score, a tool validated in the ICU environment [27], which indicates delirium with a score of ≥4 [28] and/or prescription of haloperidol, since haloperidol is prescribed specifically for the treatment of delirium in the LUMC.

At our institution, the ICU and IMCU are located adjacent to each other as part of the same ward. For the purpose of this study, people who were transferred between the ICU and IMCU during their admission were counted as a single admission. Furthermore, if a person was readmitted to the ICU/IMCU within 24 hours after discharge, this was also counted as part of a single admission.

Analyses

All statistical analyses were performed using SPSS version 25, with a two-tailed significance level of α=0.05.

Numerical data are expressed as mean and the standard deviation (written as mean ±SD) or median and the interquartile range (written as median (IQR)) when data follow a normal or non-normal distribution, respectively. To determine the significance of differences of continuous variables between two groups the unpaired T-test was performed in case of normal distribution, otherwise the rank sum Mann-Whitney U test was performed. The t-test and U test statistics were reported, respectively. Categorical data are expressed as percentages of the total. Any differences between the ICU and IMCU for the averages of the RCSQ scores, daytime sleep durations, and night-time sleep durations were examined using Mann-Whitney U tests.

Two generalized Estimating Equation (GEE) models were used to determine the association of certain independent variables with night-time sleep duration and sleep quality [29]. A first-order autoregressive (AR[1]) correlation matrix was used to account for potential correlation among repeated measures within subjects. The following independent variables were included: age in years, sex, length of stay at the ICU/IMCU in days, SOFA score, invasive mechanical ventilation (yes/no), benzodiazepine use (yes/no), the presence of direct natural daylight (yes/no), delirium presence (yes/no) and overnight transfer (yes/no).

The paired Wilcoxon signed-ranks test was used to assess the association between benzodiazepine use and sleep outcomes, by comparing average sleep parameters (RCSQ score, night-time sleep duration and day-to-night sleep ratio) between two days prior to the benzodiazepine use to the first two days after start of benzodiazepine use. Cases in which the individual used another benzodiazepine in the two prior days were excluded from this analysis. Of note, only the night-time sleep duration analysis included two patients who had two admissions. They were each treated as unique admissions for analysis; the other analyses did not include multiple admissions of single patients.

Finally, a Cox regression analysis was used to determine the associations between variables (minimum RCSQ score, >1 night invasive mechanical ventilation, mean SOFA score, delirium presence on >1 days, length of stay, age, and sex) and mortality at 1 year after discharge [30]. The odds ratio (OR) and confidence interval (CI) were reported. For binary variables the odds ratio signifies the odds of the outcome for the category specified between brackets. This analysis excluded individuals who died prior to their discharge from the ICU. In some cases, censoring (a common occurrence in survival analysis) took place due to a lack of follow up opportunities. These observations were retained in the analysis until the point at which they were last observed.

Results

Study population

Data of 285 individuals with 323 admissions meeting the inclusion criteria (representative of a population with a ICU/IMCU stay of 4 or more days) consisting of 3,756 nights between November 2018 and May 2020 were analyzed. We analyzed data of 1,481 nights with an available RCSQ score and 1,570 nights with available night-time sleep duration (respectively 23.0% and 24.3% of all continuous sedation-free nights in the IMCU/ICU during this period). The mortality status of 230 out of 285 individuals (80.7%) could be confirmed with certainty, either because the individual had died or because they were verified as being alive 1 year post-ICU discharge.

An overview of outcomes and characteristics of the admitted individuals is shown in table 5.1. The median age was 66.5 years and 63.2% of individuals were male. The departments for which the individuals were admitted can also be found in table 5.1. The majority (59.9%) of admissions were postoperative.

Table 5.1. Characteristics of ICU and/or IMCU admissions with ≥ 1 RCSQ score, N=323 admissions

	N=323 admissions
Age in years, median (IQR)	66.5 (54.6-71.9)
Sex, count (% male)	204 (63.2)
Length of stay, median (IQR)	6.9 (3.8-13.8)
≥1 overnight transfer during admission (%)	38 (11.8)
Disease severity	
SOFA score, mean ±SD	7.5 ±2.8
Received invasive mechanical ventilation overnight, count (%)	244 (75.5)
Sleep	
Night-time sleep duration in hours (10 PM-6 AM), median (IQR)	4.6 (3.8-5.7), n=305
Daytime sleep duration in hours (6 AM-10 PM), median (IQR)	3.2 (2.0-5.0), n=239
24 hour sleep duration, median (IQR)	7.7 (6.0-10.0), n=239
RCSQ score total, mean ±SD	4.9 ±1.8
Subscale Sleep depth, mean ±SD	4.5 ±1.9
Subscale Sleep latency, mean ±SD	5.0 ±2.1
Subscale Awakenings, mean ±SD	5.0 ±1.8
Subscale Sleep latency after awakening, mean ±SD	5.1 ±1.9
Subscale Quality, mean ±SD	5.0 ±1.9
Delirium	
Delirium, count (%)	148 (45.8)
Number of days until start delirium after ICU admission, median (IQR)	3.0 (1.5-7.0)
Medication (used one or more times during admission)	
Haloperidol, count (%)	118 (36.5)
Benzodiazepines, count (%)	198 (61.3)
Of which non-continuous benzodiazepine, count (%)	180 (55.7)
Melatonin, count (%)	28 (8.7)
Propofol, count (%)	210 (65.0)
Sedation >60 minutes, count (%)	215 (66.6)
Admission department	
Cardiothoracic surgery, count (%)	87 (26.9)
Surgery, count (%)	76 (23.5)
Gastroenterology, count (%)	34 (10.5)
Internal medicine, count (%)	29 (9.0)
Cardiology, count (%)	22 (6.8)
Neurosurgery, count (%)	17 (5.3)
Neurology, count (%)	16 (5.0)
Pulmology, count (%)	12 (3.7)
Other, count (%)	30 (9.3)

*Abbreviations: IQR; interquartile range, ICU; Intensive Care Unit, IMCU; Intermediate Care Unit, RCSQ; Richards-Campbell Sleep Questionnaire, SOFA; sequential organ failure assessment, SD; standard deviation
*n is only mentioned if not 323.

Sleep in the ICU/IMCU

On average, individuals admitted to the ICU and/or IMCU had a low RCSQ total score (4.9 ±1.8, out of 10, see table 5.1). The median sleep duration during the night was 4.6 (IQR: 3.8-5.7) hours. Median daytime sleep duration was 3.2 hours. One or more overnight transfers occurred in 38 admissions (11.8%) which resulted in a total of 40 transfers. Of these transfers, 15% took place within the ICU, 15% within the IMCU, 43% from the IMCU to ICU, and 28% from the ICU to IMCU.

The median RCSQ score during included nights was 4.9 (IQR: 3.6-5.9) in the ICU (n=251 ICU specific admissions) and 5.2 (IQR: 3.8-6.5) in the IMCU (n=155 IMCU specific admissions, U=17517.5, p=.092). Median night-time sleep duration did not differ significantly between the ICU and IMCU (4.6 vs 5.0, n=232 and 149 admissions, respectively, U=16718.5, p=.589). Median daytime sleep duration was longer in the ICU than the IMCU (3.5 vs 2.0 hours, n=173 and 119 admissions respectively, U=7324.5, p<.001).

Delirium in the ICU/IMCU

Delirium was experienced during 148 out of 323 (45.8%) admissions. The delirium started at a median of three days after admission to the ICU or IMCU. The average RCSQ score in the three days preceding the onset of (detected) delirium was 4.7 ±1.8 (n=43), which was not significantly different from the average RCSQ score of those without delirium (5.2 ±1.7, n=175, t= 1.509; p=.149). The median day-to-night sleep ratio (hours slept during the day/hours slept during the night) in the three days preceding delirium was significantly higher than in individuals without delirium: 1.13 (IQR: 0.5-1.7, n=24) vs 0.57 (IQR: 0.3-1.0, n=105; U=719, p=.001).

Table 5.2. Use of sleep medication in the ICU and IMCU (N=3,754 included nights)

Use of sleep medication, count (%) of total included nights	
Melatonin	306 (8.1)
Benzodiazepine	895 (23.8)
Of which:	
Temazepam	473 (12.6)
Oxazepam	271 (7.3)
Lorazepam	180 (4.8)
Midazolam (excl. continuous)	39 (1.1)
Diazepam	29 (0.7)
Zopiclone	12 (0.3)

*Abbreviations: ICU; Intensive Care Unit, IMCU; Median Care Unit,

Medication use

Benzodiazepines were frequently administered, with non-continuous benzodiazepines being prescribed as hypnotics (i.e. not administered to induce continuous sedation) between 9 PM and 4 AM at least once in 55.7% of the admissions. Benzodiazepines were used 23.8% of the total number of included nights (N=3,754, see table 5.2). The medication was started at a median of 4.1 days after admission (IQR: 1.7-8.5). The sleep outcomes (RCSQ score, night-time sleep duration and day-night sleep ratio) of the two days prior to benzodiazepine initiation were averaged and compared to the same sleep outcomes of the first two days after starting benzodiazepine use (not including the initiation night). There was no association between benzodiazepine use and improvement in recorded sleep outcomes (see table 5.3).

Table 5.3. Paired analyses of average sleep outcomes 2 days prior vs 2 days after benzodiazepine initiation

	n	Prior initiation	After initiation	Test statistic	p-value
RCSQ score	23	4.2 (3.0-6.2)	3.2 (2.2-5.6)	Z=-1.060	.300
Sleep duration night (10 PM-6 AM)	26	4.0 (3.0-5.5)	4.0 (2.9-6.0)	Z=-.127	.905
Day-night sleep ratio	11	0.6 (0.3-1.3)	0.5 (0.4-1.1)	Z=-.890	.398

*Median (IQR) is used

*Abbreviations: RCSQ; Richards-Campbell Sleep Questionnaire

Table 5.4. Associations with sleep outcomes in the ICU and IMCU

Dependent variable:	RCSQ score, n=1481 included nights of 319 admissions				Sleep duration night (10 PM-6 AM), n=1570 included nights of 299 admissions			
Independent variables	B	Lower 95% CI	Upper 95% CI	p-value	B	Lower 95% CI	Upper 95% CI	p-value
Age (in years)	.013	.000	.027	.048*	.008	-.005	.020	.223
Sex (male)	.263	-.054	.580	.104	.117	-.151	.384	.393
Length of stay (in days)	.010	.000	.020	.056	.011	.003	.019	.007*
SOFA score	-.018	-.058	.023	.394	-.008	-.046	.029	.671
Invasive mechanical ventilation	.061	-.265	.386	.715	.154	-.132	.439	.292
Benzodiazepine use	-.558	-.840	-.275	<.001*	-.533	-.780	-.285	<.001*
Delirium	-.716	-1.047	-.386	<.001*	-.697	-.993	-.402	<.001*
Overnight transfer	-1.831	-2.917	-.746	.001*	-1.250	-2.876	.375	.132
External daylight present	.069	-.241	.379	.663	.056	-.200	.312	.667

*Abbreviations: ICU; Intensive Care Unit, IMCU; Intermediate Care Unit, RCSQ; Richards-Campbell Sleep Questionnaire, CI; confidence interval, SOFA; sequential organ failure assessment

B=Unstandardized coefficient estimate

Factors associated with sleep in the ICU

The outcomes of two GEE-models using RCSQ score and sleep duration during included nights as dependent variables are shown in table 5.4. The higher the age , the better the RCSQ score (p=.048). A longer length of stay was associated with a longer night-time sleep duration (p=.007). Overnight transfers were associated with a worse RCSQ score (p=.001). Benzodiazepine use and delirium presence were associated with negative sleep outcomes (a lower RCSQ score and a shorter night-time sleep duration, both p<.001).

Mortality

28 patients died during their ICU and/or IMCU admission. The mortality rate within the first month after discharge was 16.4% (this includes patients who died during their ICU/IMCU admission). At 6 months after discharge, the mortality rate was 26.6% and at 1 year 29.7%. RCSQ score was not significantly associated with 1 year post-discharge mortality (OR [95% CI]: 0.904 [0.810-1.008]; p=.070, see table 5.5). Increased mortality was associated with older age, a higher SOFA score, and female sex (all p<.05).

Table 5.5. Associations with mortality at 1 year after discharge from the ICU or IMCU (n=289 [a])

Dependent variable:	1 year mortality			
Independent variables	OR	Lower 95% CI	Upper 95% CI	p-value
Age (in years)	1.045	1.021	1.069	<.001*
Sex (male)	.562	.342	.924	.023*
Length of stay (in days)	.992	.973	1.013	.464
SOFA score	1.266	1.153	1.391	<.001*
Invasive mechanical ventilation (>1x)	.654	.372	1.151	.141
Delirium (>1x)	1.151	.661	2.004	.619
RCSQ score	.904	.810	1.008	.070

*Abbreviations: ICU; Intensive Care Unit, IMCU; Intermediate Care Unit, CI; confidence interval, OR; odds ratio, SOFA; sequential organ failure assessment, RCSQ; Richards-Campbell Sleep Questionnaire

[a] 28 out of 323 cases were excluded from the analysis as the individual died during the relevant ICU/IMCU admission. 6 additional cases were excluded, as they were censored before the earliest event in a stratum.

Discussion

We evaluated the quality and quantity of sleep as reported by nurses in both the ICU and IMCU. We investigated various factors that may affect these sleep outcomes and their association with mortality up to 1 year after discharge. Results indicated poor sleep quality and short sleep duration in the ICU/IMCU environment. The use of benzodiazepines, presence of delirium and overnight transfers were associated with negative sleep outcomes. Surprisingly, a longer length of stay was associated with a longer night-time sleep duration and an older age was associated with better sleep quality. Delirium was preceded by changes in the day-night sleep ratio. Furthermore, initiation of benzodiazepines as hypnotic did not lead to improved sleep quality or a longer night-time sleep duration. Greater disease severity, older age, and the female sex were associated with an increased post discharge mortality rate. There was a non-significant association, adjusted for relevant variables such as disease severity, between decreased sleep quality and increased mortality 1 year post-discharge.

Sleep quantity and quality

The median night-time sleep duration (between 10 PM and 6 AM) in the ICU/IMCU was short (4.6 hours). In contrast, the median 24-hour sleep duration was 7.7 hours. Other studies using polysomnography to monitor sleep in the ICU [31-34] reported a wide range of 24 hour sleep durations, depending on the study populations. Consistent with our findings, the average sleep duration tends to be between 6 and 9 hours according to these aforementioned studies. Sleep in these studies (in which heavily sedated individuals were excluded) was generally very fragmented, with roughly half of the total sleep time occurring during the day. Our study, using subjective instead of objective outcome measures, also finds a high portion of sleep during the day.

The quality of night-time sleep in the ICU/IMCU was below the commonly used cut-off of 5, indicating poor sleep quality [19]. In seven previous studies where the RCSQ was assessed in ICU or ICU-like settings, with study populations with inclusions numbers ranging from 31 to 222 (median 60), the RCSQ score ranged between 4.6 and 6.0 [19, 22, 32, 35-39]. Our results are consistent with these findings. However, a direct comparison is difficult as the studied populations differ.

Factors impacting sleep in the ICU

Benzodiazepine use and delirium were associated with negative sleep outcomes. Disrupted nocturnal sleep is a well-known symptom of delirium [40]. This explains the association between delirium, and both night-time sleep duration and quality. Previous studies, conducted in non-ICU environments, have shown that benzodiazepines negatively affect sleep quality, despite decreasing sleep latency and improving total sleep time [41]. In our study population benzodiazepines were commonly used (in 56% at least once during an admission, not including continuous midazolam use). The observed negative association between benzodiazepine use and sleep quality and duration, when studying this relationship longitudinally using a GEE model, may partly be attributed to the fact that benzodiazepines are more commonly prescribed for individuals experiencing sleep difficulties. However, we looked not only at associations, but also sleep in the days before and after initiation of benzodiazepine use. Despite very frequent prescribing, initiation was not associated with improvements in sleep quality, duration of night-time sleep or day-night sleep ratio. On top of that, benzodiazepine use is not recommended in critically ill individuals, as it may increase the risk of delirium and adverse clinical outcome [42, 43].

To adequately treat sleep difficulties in the ICU, guidelines recommend use of a sleep promoting protocol that favors non-pharmacologic measures such as improving sleep hygiene (i.e. behavioral and environmental factors that support healthy sleep patterns [44]) and lowering noise and light levels. Efforts have been made to create sleep-enhancing protocols, but a widely adopted standardized protocol has not yet been established [45, 46]. Complementary and alternative methods have also been suggested, including various relaxation techniques, which showed some promising results [47-49]. However, further research with adequate sample sizes is required to strengthen the evidence for such interventions [50-53]. We also encourage research into the effects of other classes of medication previously shown to improve night-time sleep, such as sodium oxybate [54] or dexmedetomidine [55, 56]. Considering these factors, health care providers should exercise caution and consider alternative treatments to improve outcomes. The development of a standardized sleep promoting protocol could improve both patient care and scientific research in this field, as it would enable more meaningful comparisons to be made.

As expected, transfers during the night were negatively associated with sleep quality. Previous studies have found associations between out-of-hours ICU discharge and increased readmission rates and mortality [57, 58]. However, the relation with sleep has not been investigated yet. In many cases it was not medically necessary for the individual to be

transferred overnight. The reason for a transfer was not always reported in the EHR, but was likely related to increasing the availability of ICU beds. The number of available ICU beds has, in general, a substantial impact on the management and outcomes of critically ill individuals experiencing sudden clinical deterioration [59]. It should also be noted that in 43% of cases the overnight transfer was from the IMCU to the ICU, which does indicate medical necessity, and may be a further explanation for the association between transfers and reduced sleep quality.
Length of stay was significantly associated with a longer night-time sleep duration, but not with sleep quality. Other studies also did not find an association between length of ICU stay and sleep quality [60, 61]. A potential relationship between length of stay and night-time sleep duration has not been thoroughly investigated. Factors such as pain or anxiety, which have been demonstrated to be significant factors influencing sleep, may be more prevalent at the onset of an ICU admission, providing a potential explanation for the found association [61, 62].
A significant association between older age and better sleep quality was found, which is counterintuitive given the negative association between age and sleep quality in the general population [63]. Prior ICU-based research yielded mixed results regarding this association. Studies performed in the ICU suggest that increased age is associated with negative sleep outcomes [64, 65]. On the other hand, a study by Bihari et al. (2012) found a positive association between age and sleep outcomes in the ICU, but only in women [60]. Further research is needed to elucidate this association and any underlying mechanisms.

Delirium

The average day-night sleep ratio in the three days preceding delirium differed significantly from individuals without delirium, indicating that the day-night rhythm changes prior to (noticeable) delirium onset. The average RCSQ score was also slightly but non-significantly lower just before delirium onset. Previous research of Kamdar et al. (2015) supports this finding [66], as no significant association was found between RCSQ score and the transition to delirium within 24 hours. While the bi-directional nature of the relation between delirium and sleep disturbance is well-known [67-69], studies have mainly focused on sleep after the onset of delirium, not preceding it. Only a limited number of studies have specifically investigated sleep changes that precede the onset of delirium. One study showed a loss of melatonin rhythmicity and a drop of average melatonin levels as early as three days before the diagnosis of delirium, indicating a relationship between circadian rhythm disruption and delirium onset [70]. Another study [71] showed that critically ill individuals with sleep deprivation were more likely to develop mental status changes. Additionally, ICU nurses and physicians believe that

poor sleep is associated with the development of delirium [72]. Currently, a shift in day-night sleep rhythm is not mentioned as precipitating risk factor associated with delirium occurrence in the ICU in the Clinical Practice Guidelines for the Management of Pain, Agitation, and Delirium in Adult Patients in the ICU of 2018 [42]. However, non-ICU specific guidelines have acknowledged sleep disturbance as a risk factor for delirium [73], and a recent publication proposed revising the guidelines to incorporate it as a risk factor for delirium as well [74]. Despite the fact that the exact relationship is still unclear, monitoring sleep patterns could potentially aid in the early detection and prevention of delirium.

Mortality

No statistically significant association was found between reduced sleep quality and increased mortality (p=.070). However, given the low p-value, further research in a larger sample powered to detect mortality differences is warranted to investigate a potential direct relation between sleep and mortality. Additionally, there were significant associations between 1 year mortality post-discharge and older age and increased disease severity, which are known predictors for increased mortality post-ICU discharge [75, 76], as well as female sex. Previous studies report conflicting results regarding sex and mortality post-ICU discharge [75, 77-79]. Additional research with large, representative populations and appropriate confounder adjustments is necessary to clarify these findings.

Strengths and limitations

The longitudinal design and consistent completion of RCSQ scores by nurses as opposed to patients are strengths of this study. This makes a possible potential bias towards exclusion of specific subgroups, such as patients unable to provide self-assessment scores due to various circumstances, less likely.

However, sleep-related outcomes were subjective and also missing in many cases, presenting a limitation. Nurses often rate sleep quality higher than patients on the RCSQ [35]. It would be best to include both patient and nurse ratings in future research where possible. Determining whether a patient is asleep or simply has their eyes closed can be challenging, as a result, nurses tend to overestimate total sleep time [80]. Additional limitations include a single-site design with mainly post-surgical patients, retrospective data, and unexamined variables like reason for admittance, pain, comorbidities, and environmental noise levels. The study did not use the Confusion Assessment Method for the ICU (CAM-ICU) score which is used in many other ICUs. Moreover, it primarily focused on common sedative medications like propofol and

midazolam, omitting other relevant sedatives and their potential impact on delirium risk such as opioids [81]. Due to missing data, we opted to determine delirium presence using both the ICDSC score and haloperidol use, which may have been continued after the delirium has remitted.

The availability of recorded RCSQ scores was limited. Only 16.0% of admissions with >1 complete night shift had RCSQ scores, and only 23.0% of available sedation-free nights had RCSQ scores recorded. This is partly due to a high discharge rate within 24 hours (38.2%), with many patients emerging from deep sedation post-surgery, reducing the likelihood of sleep scoring being conducted. The data of those with a longer length of stay were more often available. Although the median length of stay was longer in the studied group than in the total population (6.9 vs 1.7 days), it equalled 10.0 days for those with admissions lasting at least 4 days. This suggests that the studied group, which accounts for 42.5% of admissions with a length of stay ≥4 days, is representative of this population for whom sleep difficulties may also pose a greater issue as exposure to ICU/IMCU-related factors linked to adverse sleep outcomes would be of a longer duration.

Conclusions

Poor sleep quality and a short night-time sleep duration are common in the ICU and IMCU setting. Night-time transfers are associated with worse sleep quality and delirium is associated with both worse sleep quality and a shorter night-time sleep duration. Delirium is preceded by a shift in day-to-night distribution of sleep. The use of benzodiazepine hypnotics is not associated with an improvement of night-time sleep. Based on our findings, we recommend that efforts are undertaken to prevent medically unnecessary night-time transfers. Caution should also be undertaken when prescribing benzodiazepines, considering their potential risks and the lack of proven efficacy in the critical care environment. It is vital to develop more effective sleep treatments that include non-benzodiazepine medication and sleep hygiene strategies. More focus on sleep promotion protocols is required in clinical practice and future research. Lastly, we recommend a consist monitoring of sleep as part of delirium prevention strategies.

Acknowledgements

We acknowledge the ICU and IMCU nurses who recorded the scores we used for this study. We also thank dr. Erik van Zwet for his helpful advice regarding the statistical analyses.

References

1. Pulak, L.M. and L. Jensen, *Sleep in the Intensive Care Unit:A Review.* Journal of Intensive Care Medicine, 2016. **31**(1): p. 14-23.
2. Pisani, M.A., et al., *Sleep in the intensive care unit.* Am J Respir Crit Care Med, 2015. **191**(7): p. 731-8.
3. Little, A., et al., *A patient survey of sleep quality in the Intensive Care Unit.* Minerva Anestesiol, 2012. **78**(4): p. 406-14.
4. Stuck, A., M.J. Clark, and C.D. Connelly, *Preventing intensive care unit delirium: a patient-centered approach to reducing sleep disruption.* Dimens Crit Care Nurs, 2011. **30**(6): p. 315-20.
5. Aung Thein, M.Z., et al., *A call to action for delirium research: Meta-analysis and regression of delirium associated mortality.* BMC Geriatrics, 2020. **20**(1): p. 325.
6. Kiely, D.K., et al., *Persistent delirium predicts greater mortality.* J Am Geriatr Soc, 2009. **57**(1): p. 55-61.
7. Telias, I. and M.E. Wilcox, *Sleep and Circadian Rhythm in Critical Illness.* Critical Care, 2019. **23**(1): p. 82.
8. Kwok, C.S., et al., *Self-Reported Sleep Duration and Quality and Cardiovascular Disease and Mortality: A Dose-Response Meta-Analysis.* J Am Heart Assoc, 2018. **7**(15): p. e008552.
9. Maragno-Correa, J.M., et al., *Sleep deprivation increases mortality in female mice bearing Ehrlich ascitic tumor.* Neuroimmunomodulation, 2013. **20**(3): p. 134-40.
10. Friese, R.S., B. Bruns, and C.M. Sinton, *Sleep deprivation after septic insult increases mortality independent of age.* J Trauma, 2009. **66**(1): p. 50-4.
11. Boyko, Y., et al., *Atypical sleep in critically ill patients on mechanical ventilation is associated with increased mortality.* Sleep Breath, 2019. **23**(1): p. 379-388.
12. Aaron, J.N., et al., *Environmental noise as a cause of sleep disruption in an intermediate respiratory care unit.* Sleep, 1996. **19**(9): p. 707-10.
13. Plate, J.D.J., et al., *Utilisation of Intermediate Care Units: A Systematic Review.* Crit Care Res Pract, 2017. **2017**: p. 8038460.
14. Elías, M.N., *Assessment and Monitoring of Sleep in the Intensive Care Unit.* Crit Care Nurs Clin North Am, 2021. **33**(2): p. 109-119.
15. Richards, K.C., et al., *A Systematic Review of Sleep Measurement in Critically Ill Patients.* Front Neurol, 2020. **11**: p. 542529.
16. Drouot, X., et al., *A new classification for sleep analysis in critically ill patients.* Sleep Med, 2012. **13**(1): p. 7-14.
17. Locihová, H., et al., *Sleep quality assessment in intensive care: actigraphy vs. Richards-Campbell sleep questionnaire.* Sleep Sci, 2020. **13**(4): p. 235-241.
18. Menear, A., et al., *Repeated sleep-quality assessment and use of sleep-promoting interventions in ICU.* Nurs Crit Care, 2017. **22**(6): p. 348-354.
19. Richards, K.C., P.S. O'Sullivan, and R.L. Phillips, *Measurement of sleep in critically ill patients.* J Nurs Meas, 2000. **8**(2): p. 131-44.
20. Ritmala-Castren, M., et al., *Investigating the construct and concurrent validity of the Richards-Campbell Sleep Questionnaire with intensive care unit patients and home sleepers.* Australian Critical Care, 2022. **35**(2): p. 130-135.
21. Varella, N.C., et al., *Cross-cultural adaptation of the Richards-Campbell Sleep Questionnaire for intensive care unit inpatients in Brazil: internal consistency, test-retest reliability, and measurement error.* Sleep Med, 2021. **85**: p. 38-44.
22. Frisk, U. and G. Nordström, *Patients' sleep in an intensive care unit--patients' and nurses' perception.* Intensive Crit Care Nurs, 2003. **19**(6): p. 342-9.
23. Singer, M., et al., *The Third International Consensus Definitions for Sepsis and Septic Shock (Sepsis-3).* Jama, 2016. **315**(8): p. 801-10.
24. Ferreira, F.L., et al., *Serial evaluation of the SOFA score to predict outcome in critically ill patients.* Jama, 2001. **286**(14): p. 1754-8.
25. Vincent, J.L., et al., *The SOFA (Sepsis-related Organ Failure Assessment) score to describe organ dysfunction/failure. On behalf of the Working Group on Sepsis-Related Problems of the European Society of Intensive Care Medicine.* Intensive Care Med, 1996. **22**(7): p. 707-10.
26. Minne, L., A. Abu-Hanna, and E. de Jonge, *Evaluation of SOFA-based models for predicting mortality in the ICU: A systematic review.* Crit Care, 2008. **12**(6): p. R161.
27. Devlin, J.W., et al., *Delirium assessment in the critically ill.* Intensive Care Med, 2007. **33**(6): p. 929-40.
28. Bergeron, N., et al., *Intensive Care Delirium Screening Checklist: evaluation of a new screening tool.* Intensive Care Med, 2001. **27**(5): p. 859-64.

29. Zeger, S.L., K.-Y. Liang, and P.S. Albert, *Models for Longitudinal Data: A Generalized Estimating Equation Approach.* Biometrics, 1988. **44**(4): p. 1049-1060.
30. Abd ElHafeez, S., et al., *Methods to Analyze Time-to-Event Data: The Cox Regression Analysis.* Oxidative Medicine and Cellular Longevity, 2021. **2021**: p. 1302811.
31. Friese, R.S., et al., *Quantity and Quality of Sleep in the Surgical Intensive Care Unit: Are Our Patients Sleeping?* Journal of Trauma and Acute Care Surgery, 2007. **63**(6): p. 1210-1214.
32. Elliott, R., et al., *Characterisation of sleep in intensive care using 24-hour polysomnography: an observational study.* Crit Care, 2013. **17**(2): p. R46.
33. Gabor, J.Y., et al., *Contribution of the Intensive Care Unit Environment to Sleep Disruption in Mechanically Ventilated Patients and Healthy Subjects.* American Journal of Respiratory and Critical Care Medicine, 2003. **167**(5): p. 708-715.
34. Freedman, N.S., et al., *Abnormal sleep/wake cycles and the effect of environmental noise on sleep disruption in the intensive care unit.* Am J Respir Crit Care Med, 2001. **163**(2): p. 451-7.
35. Kamdar, B.B., et al., *Patient-nurse interrater reliability and agreement of the Richards-Campbell sleep questionnaire.* Am J Crit Care, 2012. **21**(4): p. 261-9.
36. Li, S.Y., et al., *Efficacy of controlling night-time noise and activities to improve patients' sleep quality in a surgical intensive care unit.* J Clin Nurs, 2011. **20**(3-4): p. 396-407.
37. McKinley, S., et al., *Sleep and psychological health during early recovery from critical illness: An observational study.* Journal of Psychosomatic Research, 2013. **75**(6): p. 539-545.
38. Nicolás, A., et al., *Perception of night-time sleep by surgical patients in an intensive care unit.* Nurs Crit Care, 2008. **13**(1): p. 25-33.
39. Kamdar, B.B., D.M. Needham, and N.A. Collop, *Sleep deprivation in critical illness: its role in physical and psychological recovery.* J Intensive Care Med, 2012. **27**(2): p. 97-111.
40. Inouye, S.K., *Delirium in older persons.* N Engl J Med, 2006. **354**(11): p. 1157-65.
41. Achermann, P. and A.A. Borbély, *Dynamics of EEG slow wave activity during physiological sleep and after administration of benzodiazepine hypnotics.* Hum Neurobiol, 1987. **6**(3): p. 203-10.
42. Devlin, J.W., et al., *Clinical Practice Guidelines for the Prevention and Management of Pain, Agitation/Sedation, Delirium, Immobility, and Sleep Disruption in Adult Patients in the ICU.* Crit Care Med, 2018. **46**(9): p. e825-e873.
43. Dorsch, J.J., et al., *Sleep in the Intensive Care Unit: Strategies for Improvement.* Semin Respir Crit Care Med, 2019. **40**(5): p. 614-628.
44. Alanazi, E.M., et al., *Sleep Hygiene Practices and Its Impact on Mental Health and Functional Performance Among Adults in Tabuk City: A Cross-Sectional Study.* Cureus, 2023. **15**(3): p. e36221.
45. Elliott, R. and S. McKinley, *The development of a clinical practice guideline to improve sleep in intensive care patients: a solution focused approach.* Intensive Crit Care Nurs, 2014. **30**(5): p. 246-56.
46. Topcu, N. and Z. Tosun, *Efforts to improve sleep quality in a medical intensive care unit: effect of a protocol of non-pharmacological interventions.* Sleep Breath, 2022. **26**(2): p. 803-810.
47. Lee, S.Y. and J. Kang, *Effect of virtual reality meditation on sleep quality of intensive care unit patients: A randomised controlled trial.* Intensive Crit Care Nurs, 2020. **59**: p. 102849.
48. Mermer, E. and S. Arslan, *The effect of audiobooks on sleep quality and vital signs in intensive care patients.* Intensive Crit Care Nurs, 2023. **80**: p. 103552.
49. Tan, J.X.J., J.S. Cai, and J. Ignacio, *Effectiveness of aromatherapy on anxiety and sleep quality among adult patients admitted into intensive care units: A systematic review.* Intensive Crit Care Nurs, 2023. **76**: p. 103396.
50. Shih, C.Y., et al., *Comparative efficacy of nonpharmacological interventions on sleep quality in people who are critically ill: A systematic review and network meta-analysis.* Int J Nurs Stud, 2022. **130**: p. 104220.
51. Jun, J., M.C. Kapella, and P.E. Hershberger, *Non-pharmacological sleep interventions for adult patients in intensive care Units: A systematic review.* Intensive Crit Care Nurs, 2021. **67**: p. 103124.
52. Hu, R.F., et al., *Non-pharmacological interventions for sleep promotion in the intensive care unit.* Cochrane Database Syst Rev, 2015. **2015**(10): p. Cd008808.
53. Cooke, M., et al., *Effectiveness of complementary and alternative medicine interventions for sleep quality in adult intensive care patients: A systematic review.* Int J Nurs Stud, 2020. **107**: p. 103582.
54. Roth, T., et al., *Effect of sodium oxybate on disrupted nighttime sleep in patients with narcolepsy.* J Sleep Res, 2017. **26**(4): p. 407-414.
55. Alexopoulou, C., et al., *Effects of dexmedetomidine on sleep quality in critically ill patients: a pilot study.* Anesthesiology, 2014. **121**(4): p. 801-7.
56. Skrobik, Y., et al., *Low-Dose Nocturnal Dexmedetomidine Prevents ICU Delirium. A Randomized, Placebo-controlled Trial.* Am J Respir Crit Care Med, 2018. **197**(9): p. 1147-1156.

57. Vollam, S., et al., *Out-of-hours discharge from intensive care, in-hospital mortality and intensive care readmission rates: a systematic review and meta-analysis.* Intensive Care Med, 2018. **44**(7): p. 1115-1129.
58. Laupland, K.B., et al., *Mortality associated with timing of admission to and discharge from ICU: a retrospective cohort study.* BMC Health Serv Res, 2011. **11**: p. 321.
59. Stelfox, H.T., et al., *Intensive Care Unit Bed Availability and Outcomes for Hospitalized Patients With Sudden Clinical Deterioration.* Archives of Internal Medicine, 2012. **172**(6): p. 467-474.
60. Bihari, S., et al., *Factors affecting sleep quality of patients in intensive care unit.* J Clin Sleep Med, 2012. **8**(3): p. 301-7.
61. Miranda-Ackerman, R.C., et al., *Associations between stressors and difficulty sleeping in critically ill patients admitted to the intensive care unit: a cohort study.* BMC Health Services Research, 2020. **20**(1): p. 631.
62. Carrera-Hernández, L., et al., *Patients' perceptions of sleep in a Critical Care Unit.* Enfermería Intensiva (English Edition), 2018. **29**(2): p. 53-63.
63. Hublin, C., et al., *Changes in sleep quality with age–a 36-year follow-up study of Finnish working-aged adults.* Journal of Sleep Research, 2018. **27**(4): p. e12623.
64. Naik, R.D., et al., *Sleep Quality and Quantity in Intensive Care Unit Patients: A Cross-sectional Study.* Indian J Crit Care Med, 2018. **22**(6): p. 408-414.
65. Bernat Adell, M.D., et al., *Factors affecting sleep quality in Intensive Care Units.* Medicina Intensiva (English Edition), 2021. **45**(8): p. 470-476.
66. Kamdar, B.B., et al., *Delirium transitions in the medical ICU: exploring the role of sleep quality and other factors.* Crit Care Med, 2015. **43**(1): p. 135-141.
67. Figueroa-Ramos, M.I., et al., *Sleep and delirium in ICU patients: a review of mechanisms and manifestations.* Intensive Care Med, 2009. **35**(5): p. 781-95.
68. Mistraletti, G., et al., *Sleep and delirium in the intensive care unit.* Minerva Anestesiol, 2008. **74**(6): p. 329-33.
69. Weinhouse, G.L., et al., *Bench-to-bedside review: Delirium in ICU patients - importance of sleep deprivation.* Critical Care, 2009. **13**(6): p. 234.
70. Ángeles-Castellanos, M., et al., *Loss of melatonin daily rhythmicity is asociated with delirium development in hospitalized older adults.* Sleep Sci, 2016. **9**(4): p. 285-288.
71. Helton, M.C., S.H. Gordon, and S.L. Nunnery, *The correlation between sleep deprivation and the intensive care unit syndrome.* Heart Lung, 1980. **9**(3): p. 464-8.
72. Kamdar, B.B., et al., *Perceptions and Practices Regarding Sleep in the Intensive Care Unit. A Survey of 1,223 Critical Care Providers.* Ann Am Thorac Soc, 2016. **13**(8): p. 1370-7.
73. *National Institute for Health and Care Excellence: Guidelines*, in *Delirium: prevention, diagnosis and management in hospital and long-term care*. 2023, National Institute for Health and Care Excellence (NICE)
74. Seo, Y., et al., *2021 KSCCM clinical practice guidelines for pain, agitation, delirium, immobility, and sleep disturbance in the intensive care unit.* Acute Crit Care, 2022. **37**(1): p. 1-25.
75. Ma, J.-G., et al., *Gender- and age-based differences in outcomes of mechanically ventilated ICU patients: a Chinese multicentre retrospective study.* BMC Anesthesiology, 2022. **22**(1): p. 18.
76. Soares Pinheiro, F.G.d.M., et al., *Mortality Predictors and Associated Factors in Patients in the Intensive Care Unit: A Cross-Sectional Study.* Critical Care Research and Practice, 2020. **2020**: p. 1483827.
77. Schoeneberg, C., et al., *Gender-specific differences in severely injured patients between 2002 and 2011: data analysis with matched-pair analysis.* Crit Care, 2013. **17**(6): p. R277.
78. Fowler, R.A., et al., *Sex-and age-based differences in the delivery and outcomes of critical care.* Cmaj, 2007. **177**(12): p. 1513-9.
79. Samuelsson, C., et al., *Gender differences in outcome and use of resources do exist in Swedish intensive care, but to no advantage for women of premenopausal age.* Crit Care, 2015. **19**(1): p. 129.
80. Beecroft, J.M., et al., *Sleep monitoring in the intensive care unit: comparison of nurse assessment, actigraphy and polysomnography.* Intensive Care Med, 2008. **34**(11): p. 2076-83.
81. Clegg, A. and J.B. Young, *Which medications to avoid in people at risk of delirium: a systematic review.* Age Ageing, 2011. **40**(1): p. 23-9.

6

Dream experiences during intensive care unit stay

prevalence, content, vividness, and associated factors

Adrienne E van der Hoeven[1,2], Rolf Fronczek[1,2], Denise Bijlenga[1,2], Sarah Hadra[1,2], Caro Ridder[1], Marlies Henneman[3,4], Jessica A Maas[3,4], Suzanna A Goedemans-de Graaf[3,4], Gert Jan Lammers[1,2], David J van Westerloo[3], Mink S Schinkelshoek[1,2]

[1]Department of Neurology, Leiden University Medical Center, Leiden, the Netherlands, [2]Stichting Epilepsie Instellingen Nederland, Sleep-Wake Center, Heemstede, the Netherlands, [3]Department of Intensive Care, Leiden University Medical Center, Leiden, the Netherlands, [4]Aftercare working group, department of Intensive Care, Leiden University Medical Center, Leiden, the Netherlands

Submitted

Abstract

Objectives

Vivid dream experiences in the Intensive Care Unit (ICU) are common, but poorly understood. We investigated the prevalence, vividness, content, emotional impact, and associated factors of dream experiences in the ICU.

Research Methodology/Design

Retrospective qualitative study with subjects ≥18 years, previously admitted to the ICU for ≥4 days and/or due to COVID-19, who were not sedated for ≥24 hours during their stay (n=80). Participants were interviewed by telephone. Clinical data were collected from patient files.

Results

Interviews were conducted a median of 9 months post-discharge. At the time of the interview, dream experiences were reported by 79%. Of patients with dream experiences, 73% reported "life-like" dreams, and 49% associated their dreams with negative emotions. Some participants (28.6%) continued to have similar dreams at home. The dream content was often related to the ICU admission. Younger age and longer length of stay were related to vivid dream experiences. Of participants with dream experiences 62.5% had experienced delirium during their ICU stay. Perceptual disturbances were also frequently reported by patients (50%) and only 45% could clearly distinguish them from dream experiences. There was an overlap between participants reporting perceptual disturbances and confirmed delirium (70%).

Conclusion

Life-like dream experiences are common in patients in the ICU and often have a negative emotional impact. To mitigate this impact, some participants suggest receiving information during their hospital stay about the potential for vivid dream experiences could be beneficial. Future studies should dive deeper into effective ways to distinguish dreams, delirium, and perceptual disturbances and how to reduce their impact.

Implications for clinical practice

Because of the psychological consequences of an ICU stay it is important to provide appropriate support during and after discharge.

Statement of significance

This study investigates prevalence and content of (vivid) dream experiences in the Intensive Care Unit (ICU), as well as associated factors. A majority of patients report having had such experiences, often with a negative and vivid content. This is a poorly understood and under-investigated aspect of ICU admission. There are indications of long-term psychological consequences, given that a substantial number of patients report having these experiences even months post-ICU discharge. Age, the use of anaesthetics and admission duration were associated with the report of more vivid dream experiences. These findings highlight the need to further investigate the psychological aspects and consequences of an ICU stay and the importance of providing appropriate support and guidance during and after discharge.

Introduction

In critically ill patients, ensuring restorative sleep is very important, as disrupted sleep is associated with impaired cognition, delirium, immune dysfunction and neuroendocrine stress system changes [1]. Sleep disruption and alterations of the circadian rhythm are prevalent in the Intensive Care Unit (ICU) due to a variety of factors, including patient care activities, severity of illness, pain, discomfort, stress, medications and mechanical ventilation [2]. Despite substantial research on sleep in the ICU [3-5], the investigation into dream experiences in this setting remains insufficient.

This study was initiated by an internal report from nurses providing ICU aftercare, revealing that many former ICU patients described experiencing strange, vivid, and often negative dreams. Our aim was to determine the prevalence of such experiences in the ICU and to identify associated factors.

Literature suggests that vivid and often unsettling dreams are frequent in critically ill patients [6, 7], causing stress even after leaving the ICU [8]. Factors such as length of stay, mechanical ventilation and certain medications are thought to be associated with these dream experiences [6, 9-12].

Previous studies have shown inconsistencies in the frequency and content of dreams and hallucinations, and often conflate the two phenomena [6, 9, 13]. Reported occurrence of delusional memories varies across studies [14], partly due to differing interview times post-

hospitalization. Existing studies tend to focus on certain patient groups within the ICU, such as those that were delirious, mechanically ventilated, or sedated [13, 15]. Lastly, the sample sizes of each of these studies were relatively small [6-10, 16-20]. Larger data sets and broader inclusion criteria are needed for a more comprehensive understanding of dream experiences in the overall ICU population.

We used broad inclusion criteria in a general ICU population to assess the extent, content, impact, and related factors of dream experiences in the ICU in general, and life-like dream experiences in particular. We also explored the overlap between delirium, dream experiences and perceptual disturbances. The findings of this study can provide further insights for the improvement of care provided to ICU patients during and after their hospitalization.

Method

Participants

From June 2021 to November 2022 we reached out to individuals discharged from the Leiden University Medical Center (LUMC) ICU. Inclusion criteria mirrored those guiding ICU aftercare participation, and consisted of age ≥ 18 years, a length of stay (LOS) of ≥4 days and/or admittance due to COVID-19, not sedated (as measured using a Richmond Agitation-Sedation Scale, RASS, score of -1 or higher) for at least 24 hours during the admission period, as well as the ability to provide informed consent and perform an interview via telephone. Individuals were not approached if they were discharged from the ICU over a year ago to reduce recollection bias.

The LUMC ICU has an adjacent Intermediate Care Unit (IMCU) as part of the same ward. For the purpose of this study, patient who were transferred between the ICU and IMCU during their admission were regarded as a single ICU admission. In cases of multiple admissions meeting the inclusion criteria, only the latest admission applied, unless the patient was re-admitted within a week for the same reason, in which case the admittance days were combined.

Ethics statement

The study was conducted per the Helsinki Declaration as revised in 2013 and in accordance with local statutory requirements (registration number N21.058). All participants provided written informed consent for participation and publication.

Study procedures

Individuals who met the inclusion criteria were identified by nurses who performed the post-ICU admission aftercare follow-up calls (JAM, MH, SAG). Normally this happens three to four months after ICU discharge. However, due to COVID-19, calls were sometimes made months later. Individuals were asked if they consented to being additionally approached regarding a study on dream experiences in the ICU. Individuals were explicitly informed that participation was possible for both those who had dream experiences and those who did not. After showing interest in participating, participant information and consent forms were provided. After a week, individuals were called to further discuss study participation and included if they gave informed consent.

Interview process

Participants were interviewed by telephone and were asked about dream experiences during their ICU admission using a self-designed questionnaire comprised of both closed and open-ended questions (for the English translation of interview questions, see Appendix A). Interviews took between 5 and 60 minutes depending on the content and were not recorded. Instead, notes were taken by the interviewer during and directly after the interview. After an interview, a copy was sent by email to the participant to allow for corrections to be made. The interviewers were aware of the participant's name, age, sex, LOS, and contact information but did not know any clinical details. Participants were approached and interviewed by AEH (a PhD student in medicine), SH (a medicine bachelor student), and CR (a health sciences bachelor student), or by one of three other students, including a psychology master student and two medicine master students.

Data collection

In addition to the interview, patient group data were retrieved from electronic health records (see table S6.1 for an overview of collected variables) to describe the study population and determine potential factors related with the occurrence of (vivid) dream experiences. For group comparisons, only factors that did not have to be aggregated over the admission were included:

age, sex, LOS, time since ICU discharge, SOFA at admission, COVID-19 positivity at admission, elective/non-elective admission, intubation during admission (yes/no), and overall subjective sleep quality.

Sleep quality was measured using the Richards-Campbell Sleep Questionnaire (RCSQ) score, which is validated in critically ill patients [21]. The RCSQ score consists of five questions pertaining to sleep which are scored from 0 to 10, with a higher score indicating better sleep. The total score is the sum of all sub-scores. Sleep duration was recorded as the total hours of sleep during the night shift (between 10 PM and 6 AM), as reported by the attending nurse.

The presence of delirium was determined using the Intensive Care Delirium Screening Checklist (ICDSC) score, which indicates delirium with a score of ≥4 [22] and/or prescription of haloperidol, as haloperidol is specifically prescribed for the treatment of delirium in the LUMC. Occurrence of perceptual disturbances was ascertained using the following question: "Based on the fact that others did not share your perceptions or for other reasons, do you think that you experienced hallucinations (sensory experiences that were not real) during your admission?".

Disease severity was recorded using the Sequential Organ Failure Assessment (SOFA) score at ICU admission. This is a widely used score, ranging from 0 to 24, that measures the severity of organ dysfunction in critically ill patients [23]. Pain was assessed using the Numeric Rating Scale (NRS) pain score (range 0-10). Respectively, a higher score indicates greater disease or pain severity [23, 24]. Participants were asked to what extent ambient noise bothered them when awake during their admission (options: not, hardly, somewhat, quite a lot, and a lot). The duration of invasive mechanical ventilation was calculated by summing up the number of days.

Ventilation was considered to have ended if weaning had started and there was no ventilation for ≥12 hours, unless it resumed and exceeded the previous day's hours. Restarting ventilation after ≥48 hours was considered a separate period.

Medication use was documented for opioids, anaesthetics, benzodiazepines, corticosteroids, antipsychotics, antidepressants, and antiepileptic drugs. Whether it was used at all and the percentage of the LOS (number of days prescribed/LOS) it was used were recorded. No distinction was made between benzodiazepine use for hypnotic or sedative purposes.

Data analysis

Quantitative data

To determine the primary objective, the prevalence of reported dream experiences in the ICU at the time of the interview, we used the interview question pertaining to the ICU admission: "Can you describe a dream that you remember the most?". If the participant was able to describe dream content, they were classified as having had dream experiences during their ICU stay. The following methods were applied to determine the secondary objectives regarding dream content and associated factors. First of all, descriptions of dream experiences are reported qualitatively. Recurring descriptions are reported as a percentage of all included participants and of the total number of dreams of the participants. Second, to identify potential associated factors, differences in variable outcomes were assessed between participants with and without reported dream experiences, and between those with and without vivid dream experiences. The latter group was determined based on responses to the interview question "How life-like were the dream experiences?", where participants could choose from six options, including "Like being awake" for vivid dreams and "no clear content".

All statistical analyses were performed using SPSS version 25, with a two-tailed significance level of $\alpha=0.05$. Normally distributed continuous data are presented as mean and standard deviation (SD). Non-normally distributed continuous data are presented as median and interquartile range (IQR). Depending on the distribution, the unpaired T-test or the Mann-Whitney U test were used to analyse the significance of differences in continuous variables. Categorical data are presented as counts and percentages of the total. The Chi-square test was applied to test for differences when comparing categorical data. In case of <5 observations within a contingency table cell, Fisher's Exact Test (FET) was used.

Qualitative data

A qualitative content analysis was performed based on the method suggested by Graneheim and Landman (2004) in which text units are categorised into themes, categories and sub-categories [25]. Two investigators (AEH and CR) independently reviewed interview notes focusing on participants' most notable dream content, based on the following interview question: "Can you describe a dream that you remember the most?". The dream content was then organized into distinct themes, not defined beforehand. Initially, an overall understanding of the participants' dream content was obtained by examining the text overview of the notes of all participant dream experience content descriptions. Subsequently, a thorough re-reading was

performed, during which text units were condensed into an underlining meaning and categorized into themes. Additionally, the investigators separately assessed the overall emotional tone of each dream, categorizing them as positive, neutral, or negative. Following individual data review, the two investigators collaborated to determine the final tone classification for each dream and to establish the final theme, categories and sub-categories based on consensus after comparing their respective classifications.
When referring to a specific participant, their subject number (1-80) is mentioned.

Results

Study population

Of the 233 former ICU patients contacted during regular post-discharge follow-up procedures, 134 consented to being contacted. Of those, 80 could be included (see figure 6.1).
The median age of the participants was 61 (53-68), 71.3% was male (see table 6.1). The median LOS and the median time after ICU discharge were 10 and 273 days respectively. Only one participant was admitted less than 4 days. 53.8% was COVID-19 positive during their admission.

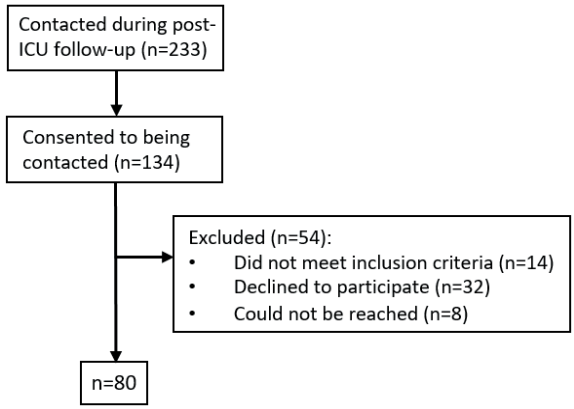


Figure 6.1. Inclusions.

Medication use during ICU admission

An overview of medication used by participants in the ICU is shown in table S6.2. Opioids, anesthetics, benzodiazepines, corticosteroids, and antipsychotics were prescribed in the majority of cases (all in more than 50% of participants).

Table 6.1. Participant characteristics (n=80)

Demographics	
Age in years	61 (53-68)
Sex, count (% male)	57 (71.3)
Admittance characteristics	
LOS in days	10 (7-19)
Time in days after discharge	273 (218-310)
Elective admission, count (%)	14 (17.5)
SOFA at admission, mean ±SD	7.5 ±3.1
Average NRS pain score	0.25 (0.0-1.0), n=77
Maximum NRS pain score	2.0 (.0-5.5), n=77
COVID-19 positivity, count (%)	43 (53.8)
Mechanical ventilation, count (%)	72 (90.0)
Mechanical ventilation in days	7 (2-10)
Ambient noise[a], count (%)	24 (30.8), n=78

*Notes: if not specified median and interquartile range (IQR) are reported
*Abbreviations: LOS; length of stay, SOFA; sequential organ failure assessment, NRS; numerical rating scale, COVID-19; coronavirus disease 2019
[a] Count and percentage of patients bothered quite a lot or a lot by ambient noise

Sleep in the ICU

The average RCSQ score was 22.9 ±9.1 out of a maximum of 50 (n=56). The average night-time (between 10 PM and 6 AM) sleep duration was 4.0 ±1.4 hours (n=58). Normal/regular self-reported sleep was generally better than the sleep during the ICU admission (see figure 6.2).

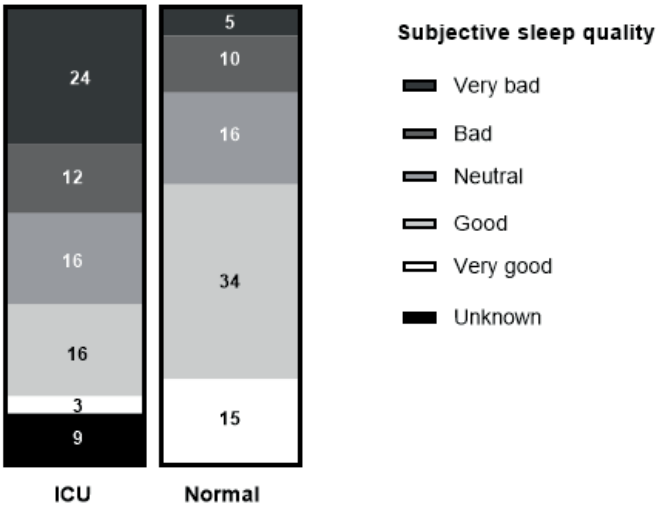


Figure 6.2. Self-reported sleep quality during and outside of the Intensive Care Unit (ICU) admission.

Dream experiences in the ICU

78.8% (63/80) reported dream experiences (see figure 6.3A). Of these 63 participants, 73.0% described their dream experiences as life-like (see figure 6.3B) and 73.0% said they had auditory sensations during dreams. Somatic sensations during dreams were experienced by 49.2% of the participants with dream experiences.

Within the group of participants with dream experiences, 49.2% (31/63) said they experienced negative feelings about them (see figure 6.4). A respective 19.4% (50/63) and 90.5% (57/63) indicated that both the frequency and content of their dreams in the ICU differed from their dreams outside of the ICU. Additionally, 28.6% (18/63) reported having similar dream experiences post-ICU discharge, while only 6.3% (4/63) reported having had such experiences before their ICU admission.

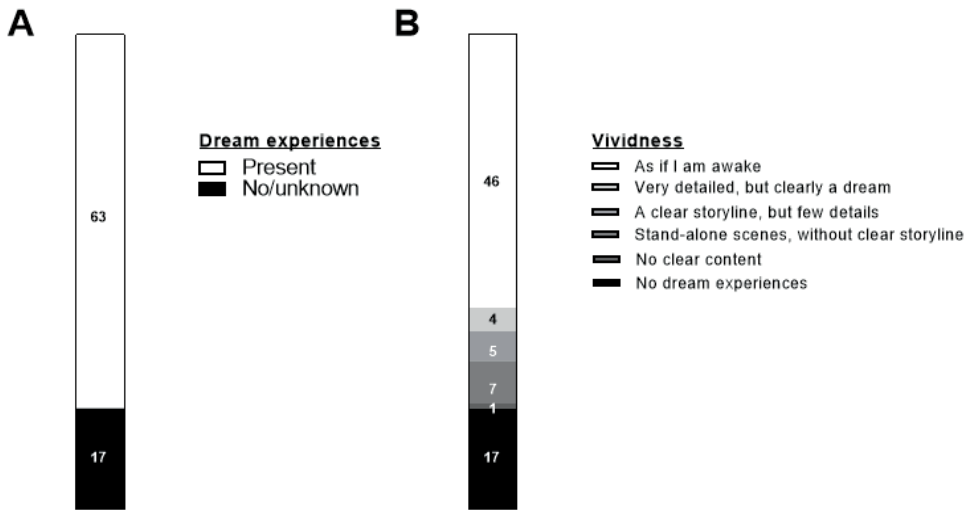


Figure 6.3. Intensive care unit stay dream presence (3A) and vividness (3B) percentages (n=80).

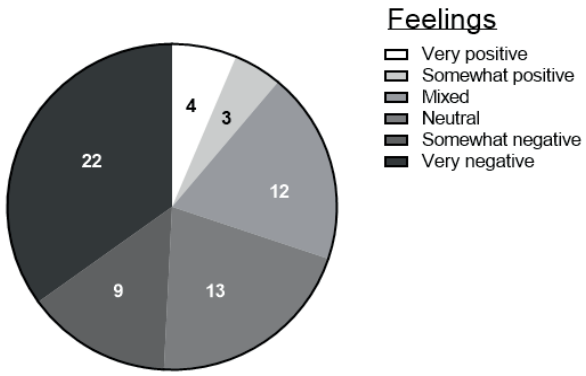


Figure 6.4. Feelings regarding dream experiences (n=63).

Factors associated with life-like dream experiences

Participants with life-like dream experiences (n=46) were younger and had a longer length of stay than those who did not report life-like dream experiences (n=34, all p<.05, see table S6.3). While non-significant, those with life-like dream experiences had been intubated more to a greater extent (p=.066). Sleep outcomes, SOFA and the percentage of elective admissions score did not differ significantly between those with and without life-like dream experiences.

Participants who reported having had dream experiences in general (n=63) had a longer admittance duration than those who did not report this (n=17, median of 13.0 versus 7.0 days, p<.001, see table S6.4) and had been intubated more often (95.2 versus 70.6%, FET, p=.009).

Partial overlap between reported perceptual disturbances, dream experiences, and delirium

The majority (50/80) of participants had a positive delirium score and/or used haloperidol at any time during their admission (see table 6.2). In addition to that, a substantial number of participants (40/80) reported experiencing perceptual disturbances. There was a large overlap between those with delirium and those experiencing perceptual disturbances, as 70.0% of the perceptual disturbances subgroup had a confirmed delirium. 55.0% could not say with certainty if they had had a dream or perceptual disturbance. Notably, the occurrence of perceptual disturbances was not noticed during the ICU/IMCU stay by the attending physicians or nurses. The percentage of reported dream experiences was higher in the perceptual disturbances subgroup than in all participants. Dream experiences in general were reported by 85.0%, and life-like dream experiences were reported by 77.5%. This compared to 80.0% and 57.5% respectively in all participants.

Content of dream experiences

The participants with dream experiences (n=63) described one or more experiences that were most notable/memorable to them (median of 2 dreams, maximum 5, total n=127). Regarding the dream experiences overall, 73.0% of participants described negative, 22.2% positive and 42.9% neutral dream experiences. 5 participants (7.9%) described having had a “continuous dream” (2, 25, 26, 55, 68). Notably, these participants were all delirious at some point during their ICU stay. Dream content was mostly related to the hospital stay (49.2%), helplessness (47.6%), family/friends/acquaintances (44.4%), and death (36.5%). Participants also frequently dreamt about lying down in a bed, on a platform or in a vehicle (25.4%) or about various

conflict situations (25.4%) (see table S6.5 for a complete overview of dream themes and sub-themes).

The most common dream themes were often present within the same dream. In some cases, participants described that they were abandoned by hospital staff and/or family and friends in their dreams, increasing their feeling of helplessness. One participant described dreaming about an interaction with a nurse during which the nurse said, “I’m not helping you anymore, you can sort it out with my colleague”, abandoning the patient in a dark, cold basement (27). Most participants (69.8%) believe that their dream experiences were substantively related to experiences during their ICU admission. For example, three participants reported dreams in which they were underwater or just below the surface and unable to escape (11, 29, 57). One of them suggested that this could be linked to his breathing difficulties during the ICU stay. Two out of three were admitted to the ICU due to cardiopulmonary-related problems, requiring supplemental oxygen.

Impact of ICU dream experiences

Some participants found their dreams to be quite pleasant or believed that they mitigated the intensity of their experience in the ICU (8, 41, 65). However, multiple participants reported a negative effect from the dreams even after waking up. 54.0% stated the dreams continued to impact them after awakening. To illustrate this, one participant dreamed that someone was throwing his bones into a large fire. Upon waking he said to his wife, "This dream has prophetic value. I am going to die!". Another participant mentioned difficulties sleeping after being back at home, due to the traumatic nature of his dreams during his ICU stay (47). Others remained convinced that the events they had dreamt about had really happened, only realizing later that they dreamt it (11, 21, 42, 49, 54, 63). A participant dreamt that her husband was angry with hospital staff because they would not allow her to be discharged. Upon waking she was under the impression that her husband was still angry (54). Another participant had a dream in which a small man wearing a yellow vest groped him. It took over a week to reconcile with the fact that this event could not have taken place in reality (80). Another example is of a participant who said he dreamed that a nurse was mean to him, after he woke up he tried to kick her (63).

Table 6.2. Delirium related outcomes in the total group (n=80) and within a subgroup reporting perceptual disturbances during their intensive care unit stay (n=40).

All participants (n=80)	
Delirium	
ICDSC score ≥ 4, count (%)	27 (39.7), n=68
ICDSC score ≥ 4 and/or haloperidol use, count (%)	50 (62.5)
Patient reported perceptual disturbances, count (%)	40 (50.0)
Participants with reported perceptual disturbances (n=40)	
Of whom reported :	
- Any dream experience, count (%)	34 (85.0)
- Life-like dream experiences, count (%)	31 (77.5)
- Perceptual disturbances distinguishable from dream experiences, count (%)	18 (45.0)
Of whom reported type of perceptual disturbance:	
- Visual, count (%)	36 (90)
- Auditory, count (%)	20 (50.0)
- Somatic, count (%)	11 (27.5)
- Unknown, count (%)	2 (5.0)
Of whom reported the time of perceptual disturbance:	
- During waking or falling asleep, count (%)	12 (30.0)
- When awake, count (%)	10 (25.0)
- When awake and when falling asleep, count (%)	11 (27.5)
- Unknown, count (%)	7 (17.5)
Of whom had a delirium:	
ICDSC score ≥ 4, count (%)	15 (44.1), n=34
ICDSC score ≥ 4 and/or haloperidol use, count (%)	28 (70.0)

*Abbreviations: ICDSC; Intensive Care Delirium Screening Checklist
*The n is mentioned when not 80 or 40 respectively.

Participant perspectives on coping with ICU dream experiences

When concluding the dream experience part of the interview, participants were asked if they had any additional comments regarding their ICU admission in general. 11 of them did. Two participants said they had been relieved to hear that other ICU patients had similar experiences during their admission (44, 49). Three participants said that having someone (familiar) present to support them or explain what was happening with them and around them would have been beneficial (16, 69, 79). Four expressed interest in speaking to the nurses and/or physicians who had cared for them in the ICU or returning to the ICU (6, 13, 14, 74). Two others mentioned that they had talked about their dream experiences with a psychiatrist after discharge, and that they believe this has helped them (11, 39).

Discussion

Dream experiences are frequently reported and often called life-like among participants. Even after several months, most participants had memories of very specific, mostly negative, dream experiences they had during their IC stay, reflecting the impact of these experiences. A number of factors may be related to the vividness of the reported dream experiences, including age, length of stay and use of anaesthetics. Half of the participants reported perceptual disturbances, but many of them could not distinguish these from dream experiences. There is a large overlap between those with perceptual disturbances and those who, with certainty, had a delirium. The most common dream themes were related to the ICU admission itself.

Prevalence and associated factors

The prevalence of self-reported dream experiences was high (79% after a median of 9 months). Roberts et al. (2004) [6] performed interviews 12-18 months after ICU discharge and reported 74% of patients who had ICU stays of at least 3 days experienced dreams. Another study with 41 participants who had ICU stays of at least 36 hours found a dream experience prevalence of 44% after up to two years post-discharge [9].

Vividness of dream experiences in the ICU has not been specifically investigated before, although it was also reported in a previous study [6]. While not specifically studied in connection to vivid dreams, previous research does support an association between reported dream experiences and length of stay [6, 9, 10], as well as an association with invasive mechanical ventilation [6, 9], supporting our findings. A connection between age and the occurrence of dream experiences was not found in prior studies. More, preferably prospective and longitudinal, research is needed to gain further insights into the various factors affecting the occurrence of dream experiences in the ICU.

The fragmented sleep often observed in hospitalized patients and the traumatic nature of an ICU stay could play a role in the high prevalence and frequency of vivid dream experiences in the ICU [26]. Psychological trauma is related to a higher frequency of dreams, often nightmares [27, 28], while sleep fragmentation can cause awakenings during REM sleep, thereby increasing dream recall [29].

Dream content

While dream content in the ICU generally differs from the general population [30, 31], some typical themes like being chased, attack, trapped, or attempting something repeatedly were mentioned. For instance, a participant described recurring dreams of battling a "Corona [virus] monster", failing twice, and finally defeating it with a pan, then consuming the monster (15). This exemplifies a common theme observed in reported dreams, specifically their connection to the ICU stay.

Differences in dream content between our study population and the general population could be caused by the different study population and methodology, as we did not use the Typical Dream Questionnaire [30-32], but conducted a semi structured interview focused on the specific situation of the participants. Participants described their most memorable dreams after an extended period of time, potentially impacting recall and introducing selection bias, and skewing the descriptions away from more "typical" dream themes [13].

Many participants considered the content of their dreams to be substantially related to the ICU admission itself. These dreams often had a negative content (helplessness, darkness, not being heard or seen, or being lost). Similarly, another study found that 60.9% of participating ICU patients experienced "scary" dreams [6] or dreams associated with negative emotions [6, 9, 10]. Several studies hypothesize that dreams could have an adaptive function in emotion regulation by preparing the individual to react to and resolve emotional conflicts, achieving emotional mastery [33, 34]. The heightened amygdala activity during REM sleep, regulating responses to stress, might explain negative or scary dream content [34, 35]. Although the proportion of time spent in REM sleep decreases in the ICU, the increased sleep fragmentation people experience in the ICU could result in more frequent awakenings during REM sleep, resulting in a greater recall of negative dreams [29, 36]. Additionally, dreams tend to be a continuation of experiences that people are likely to have in the waking world [37], perhaps explaining the high prevalence of ICU-related dream content. Overall, our findings suggest that ICU patients experience dreams that are often related to their admission and frequently have a negative content. These results contribute to our knowledge regarding the emotional and psychological experiences of ICU patients and could provide a starting point for further research in this area.

Relationship between dream experiences and perceptual disturbances

We found a large overlap between reported perceptual disturbances, dream experiences and delirium. A multicenter cohort study found a (non-significant) trend towards increased

prevalence of dreaming among those with a delirium compared to non-delirious patients [9]. Similarities have previously been noted between characteristics of dreams and delirium [38], hallucinations [39] and even psychosis [40]. Possible explanations for why many patients reported perceptual disturbances, even though a substantial portion did not have an established delirium, could be unnoticed delirium, medication side-effects and/or the effects of prolonged sleep deprivation without delirium being present (yet) [41-43]. The presence of hypnagogic and/or hypnopompic hallucinations (which occur during transitional states between wakefulness and sleep) could be an additional explanation [44, 45]. Furthermore, many ICU patients also sleep during the day, which could make it more difficult to distinguish between experiences during wakefulness and sleep. The findings of the current study do not clarify the exact relationship between dream experiences, perceptual disturbances and delirium. These findings do illustrate that the distinction can be difficult to make for both the people having these experiences as well as the researchers studying them. In future studies, the differentiation between dreams, delirium, and perceptual disturbances should preferably be made as they occur instead of retrospectively.

The post-discharge consequences of dream experiences

Many participants mentioned having had strange, often negative, dream experiences during their ICU stay. Some patients even report experiencing similar dreams after discharge from the ICU. This is consistent with findings from Roberts et al. (2004) who reported a recurrence of dream experiences in 22% [6]. Previous research has also shown that fear associated with such experiences can be partially mitigated by care actions and support by relatives [7]. However, patients usually do not inform others of unreal experiences [46]. In this study, several participants indicated that knowing others have similar experiences, receiving support from familiar individuals, or discussing their experiences with their health care providers would be beneficial. These findings emphasize the importance of monitoring and addressing patients' dream experiences during and after ICU admission.

Strengths and limitations

The study design, use of both closed and open-ended questions and blinding of interviewers for patients' medical backgrounds during the interview to reduce potential bias in data collection are strengths of this study.

Limitations include a relatively small sample size, which could limit generalizability. There is a potential for selection bias, as patients with more dream experiences may have been more

likely to agree to participate. The data is self-reported and the follow up time was around 9 months, which introduced the possibility of recall bias. Furthermore, by asking participants to describe notable dream(s) this could have prompted them to recall extremely negative or positive experiences. The framing of this question should be taken into account when interpreting the results. Lastly, the reported variables were aggregated across the ICU stay (i.e. means and medians of the whole admission were reported instead of per day). A longitudinal prospective design may be necessary to examine direct impacts of factors such as medication use and delirium on dream experiences.

Conclusion

Dream experiences were common and frequently associated with a negative emotional tone among ICU patients eligible for ICU aftercare. The majority of the reported dream experiences were vivid, with a substantial number of participants reporting auditory and sensory sensations. Younger age and longer length of stay are potentially related to more occurrence of life-like dream experiences. Dream content often pertained to the ICU admission, with the hospital environment and feelings of helplessness being common themes. Half of the participants reported perceptual disturbances, with a substantial overlap between these experiences and delirium. Many could not distinguish perceptual disturbances and dream experiences. These findings highlight the need for further research on dream experiences in ICU patients, including the causes, the extent of their impact during the ICU admission, potential impact on post-ICU mental well-being, and ways to minimize negative experiences. For example, by letting patients know these experiences are common during longer ICU admissions.

Acknowledgments

We thank the ICU aftercare working group for bringing this matter to our attention and for their valuable contribution to participant recruitment. We also thank the students who performed interviews and helped with data collection.

Supplementary material

Table S6.1. Collected variables

Category	Items
Demographics	Age
	Sex
General medical information	Date of ICU admission and ICU discharge
	Duration of ICU stay
	Relevant comorbidities (such as sleep disorders)
	SOFA score at admission
	Reason for admission
	Pain scale scores (Numerical Rating Scale)
	COVID-19 positivity
	Mechanical ventilation details during ICU stay
	Subjective sound intensity during ICU stay
	Medication use during ICU stay
	Sleep scores (RCSQ score and sleep duration during the night shift)
	Presence of delirium during ICU stay

*Abbreviations: ICU; Intensive Care Unit, COVID-19; Coronavirus disease 2019, SOFA; Sequential Organ Failure Assessment, NRS; numerical rating scale, RCSQ; Richards-Campbell Sleep Questionnaire

Table S6.2. Medication use

n=80	
Medication use	
Opioids, count (%)	75 (93.8)
- % of time used	70.0 (50-81)
Anesthetics, count (%)	73 (91.3)
- % of time used	57.1 (31-74)
Benzodiazepines, count (%)	72 (90.0)
- % of time used	43.9 (25-76)
Corticosteroids, count (%)	62 (77.5)
- % of time used	87.9 (51-100)
Antipsychotics, count (%)	44 (55.0)
- % of time used	21.6 (11-37)
Antidepressants, count (%)	4 (5.0)
- % of time used	85.7 (46-100)
Antiepileptic drugs, count (%)	4 (5.0)
- % of time used	81.3 (40-93)

*Notes: if not specified median and interquartile range (IQR) are reported

Table S6.3. Differences between participants who report life-like dream experience content and those who do not

n=80	Life-like dream experience, n=46	Unknown/no life-like dream experience, n=34	Test statistic	P-value
Demographics				
Age, mean ±SD	58 (51-63)	64 (59-71)	H=533.000	.015*
Sex, count (% male)	34 (73.9)	23 (67.6)	χ^2=.375	.620
Admittance characteristics				
Length of stay in days	12.5 (9-20)	8 (6-15)	U=1017.000	.022*
Time in days after discharge	274 (221-310)	262 (207-313)	U=836.000	.599
Elective admission, count (%)	5 (10.9)	9 (26.5)	χ^2=3.296	.082
SOFA at admission	7.5 (5-9)	7.0 (6-9)	U=727.000	.590
COVID-19 positivity at admission, count (%)	28 (60.9)	15 (44.1)	χ^2=2.207	.175
Mechanical ventilation, count (%)	44 (95.7)	28 (82.4)	FET	.066
During ICU admission - self reported sleep quality, count (%)				
Very bad	17 (37.0)	7 (20.6)	FET	.368
Bad	5 (10.9)	7 (20.6)	-	-
Neutral	10 (21.7)	6 (17.6)	-	-
Good	9 (19.6)	7 (20.6)	-	-
Very good	2 (4.3)	1 (2.9)	-	-
Unknown	3 (6.5)	6 (17.6)	-	-
Outside of ICU admission - self reported sleep quality, count (%)				
Very bad	5 (10.9)	0 (0.0)	FET	.233
Bad	5 (10.9)	5 (14.7)	-	-
Neutral	11 (23.9)	5 (14.7)	-	-
Good	17 (37.0)	17 (50.0)	-	-
Very good	8 (17.4)	7 (20.6)	-	-

*Abbreviations: SOFA; Sequential Organ Failure Assessment, COVID-19; Coronavirus disease 2019

*Notes: if not specified median and interquartile range (IQR) are reported

Table S6.4. Differences between participants who report dream experience content and those who do not

n=80	Dream experience, n=63	Unknown/no dream experience, n=17	Test statistic	P-value
Demographics				
Age	60 (53-64)	67 (54-72)	U=388.500	.084
Sex, count (% male)	46 (73.0)	11 (64.7)	FET	.552
Admittance characteristics				
Length of stay in days	13.0 (9.0-19.0)	7.0 (4.5-8.5)	U=850.500	<.001*
Time in days after discharge	276 (222-311)	238 (179-293)	U=677.500	.095
Elective admission, count (%)	10 (15.9)	4 (23.5)	FET	.482
SOFA at admission	8.0 (6.0-10.0)	7.0 (5.5-8.5)	U=600.000	.445
COVID-19 positivity at admission, count (%)	36 (57.1)	7 (41.2)	χ^2=1.373	.281
Mechanical ventilation, count (%)	60 (95.2)	12 (70.6)	FET	.009*
During ICU admission - self reported sleep quality, count (%)				
Very bad	20 (31.7)	4 (23.5)	FET	.839
Bad	9 (14.3)	3 (17.6)	-	-
Neutral	13 (20.6)	3 (17.6)	-	-
Good	13 (20.6)	3 (17.6)	-	-
Very good	2 (3.2)	1 (5.9)	-	-
Unknown	6 (9.5)	3 (17.6)	-	-
Outside of ICU admission - self reported sleep quality, count (%)				
Very bad	5 (7.9)	0 (0.0)	FET	.596
Bad	8 (12.7)	2 (11.8)	-	-
Neutral	14 (22.2)	2 (11.8)	-	-
Good	24 (38.1)	10 (58.8)	-	-
Very good	12 (19.0)	3 (17.6)	-	-

*Abbreviations: SOFA; Sequential Organ Failure Assessment, COVID-19; Coronavirus disease 2019

*Notes: if not specified median and interquartile range (IQR) are reported

Table S6.5. Dream content frequency in the patient group (n=63)

Theme	**Sub-theme**		**Dreams aggregated per patient (n=63)**		**All dreams (n=127)**	
			Theme	**Sub-theme**	**Theme**	**Sub-theme**
Death/dying	Physical passage to death		23 (36.5)	5 (7.9)	27 (21.3)	5 (3.9)
	Seeing someone who is deceased alive			3 (4.8)		4 (3.1)
	Participants' own death	Imminent/potential death		14 (22.2)	7 (11.1)	14 (11.0)	7 (5.5)
		Actual death			7 (11.1)		7 (5.5)
	Someone else's death	Imminent/potential death		8 (12.7)	2 (3.2)	8 (6.3)	2 (1.6)
		Actual death			6 (9.5)		6 (4.7)
Travel	In vehicle		20 (31.7)	15 (23.8)	22 (17.3)	16 (12.6)	
	Going from one location to another			8 (12.7)		8 (6.3)	
	Time travel			3 (4.8)		3 (2.4)	
	Being lost			3 (4.8)		3 (2.4)	

Helplessness	Being lost		30 (47.6)	3 (4.8)		44 (34.6)	3 (2.4)	
	Being stuck somewhere			25 (39.7)			30 (23.6)	
	Not being seen or heard			10 (15.9)			12 (9.4)	
	Facing an obstacle/barrier that cannot be overcome			3 (4.8)			3 (2.4)	
	No control over one's body or actions			8 (12.7)			8 (6.3)	
Conflict	Own involvement	Escaping from someone	16 (25.4)	10 (15.9)	4 (6.3)	18 (14.2)	12 (9.4)	4 (3.1)
		Physical abuse			6 (9.5)			6 (4.7)
		Physically fighting			2 (3.2)			2 (1.6)
		Argument			5 (7.9)			5 (3.9)
	No involvement from participant	Physical abuse		4 (6.3)	2 (3.2)		5 (3.9)	1 (0.8)
		Argument			2 (3.2)			3 (2.4)
		Riots			1 (1.6)			1 (0.8)
Participants' social circle	Partner		28 (44.4)	11 (17.5)		40 (31.5)	13 (10.2)	
	Family			18 (28.6)			22 (17.3)	
	Friends			5 (7.9)			6 (4.7)	
	Acquaintances			12 (19.0)			12 (9.4)	
Hospital-related	Hospital personnel		31 (49.2)	21 (33.3)		38 (29.9)	23 (18.1)	
	Caregiving/examination activities			10 (15.9)			10 (7.9)	
	In a hospital			20 (31.7)			25 (19.7)	
Location	In a hospital		36 (57.1)	20 (31.7)		48 (37.8)	25 (19.7)	
	At home			8 (12.7)			8 (6.3)	
	Strange/unknown house			6 (9.5)			7 (5.5)	
	Foreign country			10 (15.9)			11 (8.7)	
Employment	Participants' own job	Current	5 (7.9)	4 (6.3)	1 (1.6)	5 (3.9)	4 (3.1)	1 (0.8)
		Previous			3 (4.8)			3 (2.4)
	New/other job			1 (1.6)			1 (0.8)	
Recreational activities	Festivities/partying		10 (15.9)	5 (7.9)		12 (9.4)	5 (3.9)	
	Outing			2 (3.2)			3 (2.5)	
	Television/theater show	Participation		4 (6.3)	3 (4.8)		4 (3.1)	3 (2.4)
		Observation			1 (1.6)			1 (0.8)
Creatures/unreal persons	Fantasy figures	Humanlike	13 (20.6)	8 (12.7)	3 (4.8)	18 (14.2)	12 (9.4)	6 (4.7)
		Monster			2 (3.2)			3 (2.4)
		Other			3 (4.8)			3 (2.4)
	Animals			8 (12.7)			8 (6.3)	
Position	Lying down		23 (36.5)	16 (25.4)		28 (22.0)	20 (15.7)	
	Other specifically mentioned position			8 (12.7)			8 (6.3)	
Feeling physically unwell	Sick/pain		8 (12.7)	3 (4.8)		9 (7.1)	3 (2.4)	
	Feeling suffocated			3 (4.8)			4 (3.1)	
	Cold			2 (3.2)			2 (1.6)	
	Tired			1 (1.6)			1 (0.8)	
Ambient noises	People talking	To participant	26 (41.3)	15 (23.8)	7 (11.1)	27 (21.3)	15 (11.8)	7 (5.5)
		Around participant			8 (12.7)			8 (6.3)
	Music			4 (6.3)			4 (3.1)	
	Loud noise			4 (6.3)			4 (3.1)	
	Other			4 (6.3)			4 (3.1)	
Emotional support	Being checked on/being seen		7 (11.1)	4 (6.3)		7 (5.5)	4 (3.1)	
	Being reassured			3 (4.8)			3 (2.4)	
	Being protected			1 (1.6)			1 (0.8)	

Appendix A: Patient participant interview

_____	_____	_____
Name of participant	Date	Subject code

Semi-structured interview Dream experiences in the ICU

Conducted by:

Dream Experiences

1. Did you have any dream experiences during your ICU admission?
o Yes
o No
Explanation:

If no, see question 19

2. When did these dream experiences take place?
□ While awake
□ Around falling asleep or waking up
□ During sleep

3. Can you remember the content of the dream experiences?
o Yes
o No
Explanation:

4. Can you describe a dream that you remember the most?

5. Were there any dream experiences that recurred several times during the ICU admission?
o Yes, all times the same dream experience
o Yes, several times the same dream experience
o Very occasionally a recurring dream experience
o No

6. Did the dream experiences in the ICU differ in content from the dream experiences usual for you outside the ICU environment?
o Yes
o No
Explanation:

7. Did the dream experiences in the ICU differ in terms of frequency (how often they occur) from the dream experiences outside the ICU environment that were usual for you?
o Yes
o No
Explanation:

8. Have you had these dream experiences before (for example, during a previous hospitalization)?
o Yes
o No
Explanation:

9. Did you continue to have these dream experiences after you were discharged from the ICU?
o Yes
o No
Explanation:

10. Does the content of your dream experiences relate to experiences you had in the ICU?
o Yes
o No
Explanation:

11. What feeling did the dream experiences give you when looking at the entire ICU admission?
o A very negative feeling
o Somewhat negative feeling
o A neutral feeling
o A little positive feeling
o A very positive feeling
o Alternately positive and negative feeling

12. How life-like were the dream experiences?
o Like being awake
o Very detailed, but clearly a dream
o A clear story line, but few details
o Self-contained scenes with no clear line
o Mainly sounds, shapes, smells, tastes or other sensory perceptions
o No clear content

13. Were you aware that you were dreaming during the dream experience?
o Yes
o No

14. Did you feel that you could direct the events in the dream experience as you wished?
o Yes
o No
Explanation:

15. Did the dream experiences affect your functioning during the time you were awake?
o Yes
o No
Explanation:

16. Did you sometimes have to reason to know it was a dream?
o Yes
o No
Explanation:

17. During your dreams, did you only see things or did you also hear and feel things?
o Only seeing
o Also hearing
Also feeling
o Also hearing and feeling

18. Did any acquaintances or family members appear in the dream experience(s)?
o Yes
o No
Explanation:

19. To what extent do dream experiences interest you in daily life?
o Not
o Hardly
o A little
o Quite a lot
o A lot
Explanation:

20. Did ambient noise bother you when you were awake?
o Not
o Hardly
o Somewhat
o Quite a lot
o A lot
Explanation:

21. How would you rate your normal sleep quality?
o Very poor
o Poor
o Neutral
o Good
o Very good
Explanation:

22. How would you rate your sleep quality during the intensive care unit stay?
o Very poor
o Poor
o Neutral
o Good
o Very good
Explanation:

23. Is there anything left unsaid that you would like to share?

Hallucinations

1. Because others did not share your perceptions or by otherwise reasoning, do you think you hallucinated (had sensory experiences that were not real) during your admission?
o Yes
o No
o Don't know
Explanation:

If yes:

2. What kind of hallucinations were they?
□ Visual (you saw something that was not real)
□ Auditory (you heard something that was not real)
□ Tactile (multiple options possible) (you felt something that was not real)
Explanation:

3. Were the hallucinations and real-life dream experiences distinguishable from each other?
o Yes
o No
o Don't know
Explanation:

4. When did the hallucinations take place?
o Around falling asleep or waking up
o While awake
o Both around falling asleep and while awake
Explanation:

5. What feeling did the hallucinations give you when looking at the entire ICU admission?
o A very negative feeling
o Somewhat negative feeling
o A neutral feeling
o A little positive feeling
o A very positive feeling
o Alternately positive and negative
Explanation:

References

1. Pulak, L.M. and L. Jensen, *Sleep in the Intensive Care Unit: A Review.* Journal of Intensive Care Medicine, 2014. 31(1): p. 14-23.
2. Pisani, M.A., et al., *Sleep in the intensive care unit.* American journal of respiratory and critical care medicine, 2015. 191(7): p. 731-738.
3. Pisani, M.A., et al., *Sleep in the intensive care unit.* Am J Respir Crit Care Med, 2015. 191(7): p. 731-8.
4. Pulak, L.M. and L. Jensen, *Sleep in the Intensive Care Unit:A Review.* Journal of Intensive Care Medicine, 2016. 31(1): p. 14-23.
5. Kakar, E., et al., *Sleep assessment in critically ill adults: A systematic review and meta-analysis.* Journal of Critical Care, 2022. 71: p. 154102.
6. Roberts, B. and W. Chaboyer, *Patients' dreams and unreal experiences following intensive care unit admission.* Nursing in Critical Care, 2004. 9(4): p. 173-180.
7. Granberg, A., I. Bergborn Engberg, and D. Lundberg, *Acute confusion and unreal experiences in intensive care patients in relation to the ICU syndrome. Part II.* Intensive and Critical Care Nursing, 1999. 15(1): p. 19-33.
8. Granja, C., et al., *Patients' recollections of experiences in the intensive care unit may affect their quality of life.* Critical care (London, England), 2005. 9(2): p. R96-R109.
9. Roberts, B., et al., *Patients' dreams in ICU: Recall at two years post discharge and comparison to delirium status during ICU admission.* Intensive & critical care nursing : the official journal of the British Association of Critical Care Nurses, 2006. 22: p. 264-73.
10. Rundshagen, I., et al., *Incidence of recall, nightmares, and hallucinations during analgosedation in intensive care.* Intensive care medicine, 2002. 28: p. 38-43.
11. Leslie, K., et al., *Dreaming during anesthesia and anesthetic depth in elective surgery patients: a prospective cohort study.* Anesthesiology, 2007. 106(1): p. 33-42.
12. Cascella, M., et al., *Anesthetic dreaming, anesthesia awareness and patient satisfaction after deep sedation with propofol target controlled infusion: A prospective cohort study of patients undergoing day case breast surgery.* Oncotarget, 2017. 8(45): p. 79248-79256.
13. Löf, L., L. Berggren, and G. Ahlström, *Severely ill ICU patients recall of factual events and unreal experiences of hospital admission and ICU stay—3 and 12 months after discharge.* Intensive and Critical Care Nursing, 2006. 22(3): p. 154-166.
14. Kiekkas, P., et al., *Psychological distress and delusional memories after critical care: a literature review.* International Nursing Review, 2010. 57(3): p. 288-296.
15. Cochen, V., et al., *Vivid dreams, hallucinations, psychosis and REM sleep in Guillain–Barré syndrome.* Brain, 2005. 128(11): p. 2535-2545.
16. Hough, C.L. and J.R. Curtis, *Long-term sequelae of critical illness: memories and health-related quality of life.* Critical care (London, England), 2005. 9(2): p. 145-146.
17. Akintomide, G.S. and H. Rickards, *Narcolepsy: a review.* Neuropsychiatric disease and treatment, 2011. 7: p. 507-518.
18. Hall-Smith, J., C. Ball, and J. Coakley, *Follow-up services and the development of a clinical nurse specialist in intensive care.* Intensive and Critical Care Nursing, 1997. 13(5): p. 243-248.
19. Swaiss, I.G., *Discomfort, awareness and recall of patients in the intensive care: still a problem?* Critical Care, 1999. 3(Suppl 1): p. P231-P231.
20. Green, A., *An exploratory study of patients' memory recall of their stay in an adult intensive therapy unit.* Intensive and Critical Care Nursing, 1996. 12(3): p. 131-137.
21. Richards, K.C., P.S. O'Sullivan, and R.L. Phillips, *Measurement of sleep in critically ill patients.* J Nurs Meas, 2000. 8(2): p. 131-44.
22. Bergeron, N., et al., *Intensive Care Delirium Screening Checklist: evaluation of a new screening tool.* Intensive Care Med, 2001. 27(5): p. 859-64.
23. Ferreira, F.L., et al., *Serial evaluation of the SOFA score to predict outcome in critically ill patients.* Jama, 2001. 286(14): p. 1754-8.
24. Breivik, H., et al., *Assessment of pain.* Br J Anaesth, 2008. 101(1): p. 17-24.
25. Graneheim, U.H. and B. Lundman, *Qualitative content analysis in nursing research: concepts, procedures and measures to achieve trustworthiness.* Nurse Educ Today, 2004. 24(2): p. 105-12.
26. Jackson, J.C., et al., *Post-traumatic stress disorder and post-traumatic stress symptoms following critical illness in medical intensive care unit patients: assessing the magnitude of the problem.* Critical Care, 2007. 11(1): p. R27.
27. Valli, K., et al., *The effect of trauma on dream content - A field study of Palestinian children.* Dreaming, 2006. 16: p. 63-87.

28. Duval, M. and A. Zadra, *Frequency and Content of Dreams Associated with Trauma.* Sleep Medicine Clinics, 2010. 5: p. 249-260.
29. Carskadon, M. and W. Dement, *Principles and Practice of Sleep Medicine 5th edn, ed M H Kryger et al*. 2011, Amsterdam: Elsevier.
30. Maggiolini, A., *Typical dreams across the life cycle.* International Journal of Dream Research, 2020.
31. Schredl, M., et al., *Typical dreams: stability and gender differences.* J Psychol, 2004. 138(6): p. 485-94.
32. Zadra, A., et al., *The Typical Dreams of Canadian University Students.* Dreaming, 2003. 13: p. 211-235.
33. Scarpelli, S., et al., *The Functional Role of Dreaming in Emotional Processes.* Front Psychol, 2019. 10: p. 459.
34. Desseilles, M., et al., *Cognitive and emotional processes during dreaming: a neuroimaging view.* Conscious Cogn, 2011. 20(4): p. 998-1008.
35. Nir, Y. and G. Tononi, *Dreaming and the brain: from phenomenology to neurophysiology.* Trends Cogn Sci, 2010. 14(2): p. 88-100.
36. Devlin, J.W., et al., *Clinical Practice Guidelines for the Prevention and Management of Pain, Agitation/Sedation, Delirium, Immobility, and Sleep Disruption in Adult Patients in the ICU.* Crit Care Med, 2018. 46(9): p. e825-e873.
37. Fogli, A., L. Maria Aiello, and D. Quercia, *Our dreams, our selves: automatic analysis of dream reports.* Royal Society Open Science, 2020. 7(8): p. 192080.
38. Hobson, J.A., *Dreaming as delirium: a mental status analysis of our nightly madness.* Semin Neurol, 1997. 17(2): p. 121-8.
39. Waters, F., et al., *What Is the Link Between Hallucinations, Dreams, and Hypnagogic-Hypnopompic Experiences?* Schizophr Bull, 2016. 42(5): p. 1098-109.
40. Fosse, R. and F. Larøi, *Quantifying auditory impressions in dreams in order to assess the relevance of dreaming as a model for psychosis.* PLoS One, 2020. 15(3): p. e0230212.
41. Rolland, B., et al., *Pharmacology of hallucinations: several mechanisms for one single symptom?* Biomed Res Int, 2014. 2014: p. 307106.
42. Waters, F., et al., *Severe Sleep Deprivation Causes Hallucinations and a Gradual Progression Toward Psychosis With Increasing Time Awake.* Front Psychiatry, 2018. 9: p. 303.
43. Babkoff, H., et al., *Perceptual distortions and hallucinations reported during the course of sleep deprivation.* Percept Mot Skills, 1989. 68(3 Pt 1): p. 787-98.
44. Ohayon, M.M., *Prevalence of hallucinations and their pathological associations in the general population.* Psychiatry Res, 2000. 97(2-3): p. 153-64.
45. Bless, J.J., et al., *In the twilight zone: An epidemiological study of sleep-related hallucinations.* Compr Psychiatry, 2021. 108: p. 152247.
46. Granberg, A., I.B. Engberg, and D. Lundberg, *Patients' experience of being critically ill or severely injured and cared for in an intensive care unit in relation to the ICU syndrome. Part I.* Intensive and Critical Care Nursing, 1998. 14(6): p. 294-307.

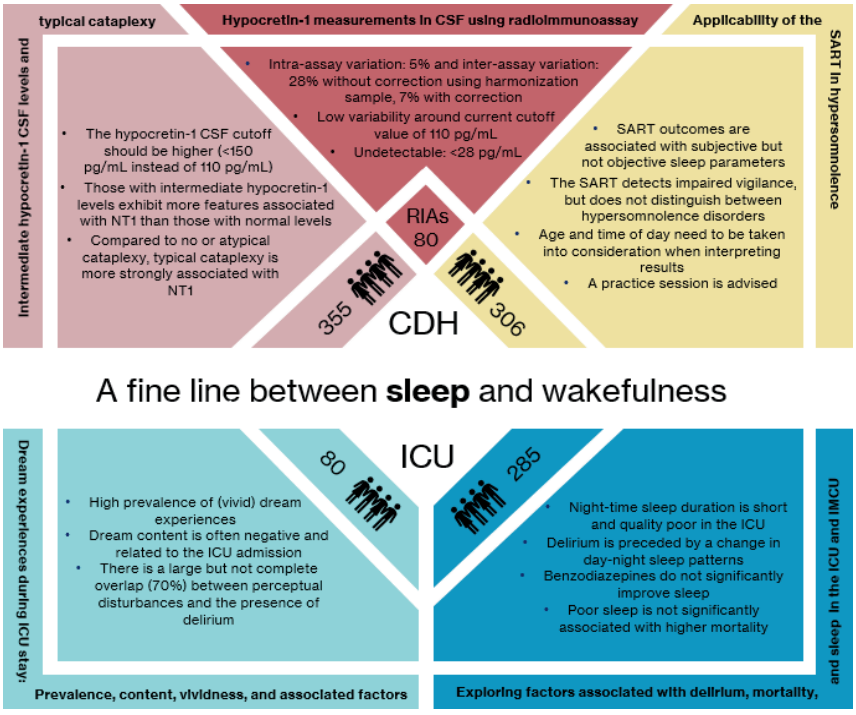
7

Discussion

Chapter 7 Summary, discussion and future perspectives

Summary

This thesis examines aspects of sleep disturbances in central disorders of hypersomnolence (CDH) and in critical care settings. It investigates diagnostic methods for hypersomnolence disorders in part I (**Chapters 2, 3 and 4**). Specifically by re-establishing the most optimal hypocretin-1 cerebrospinal fluid (CSF) cut-off value, assessing the value of typical cataplexy in the narcolepsy type 1 (NT1) diagnostic process, evaluating the reliability of CSF hypocretin-1 measurement using radioimmunoassay (RIA), and by determining the value of the sustained attention to response task (SART) when assessing CDH. Part II (**Chapter 5 and 6**) concerns sleep in the intensive care environment. This part includes a discussion of the sleep difficulties that occur in the Intensive and Intermediate Care Units (ICU and IMCU), the factors that influence sleep, the relationship between sleep and mortality, and the occurrence and factors involved in vivid and negative dream experiences in the ICU. Below, the findings of the chapters are summarized and discussed. Potential future directions are discussed at the end of both Parts I and II.



Part I: Aspects of Sleep in Central Disorders of Hypersomnolence

Intermediate hypocretin-1 cerebrospinal fluid levels and typical cataplexy: their significance in the diagnosis of narcolepsy type 1

Measurement of hypocretin-1 in CSF is the gold standard in the diagnosis of NT1. The currently used cutoff values for deficient and normal levels (110 and 200 pg/mL, respectively) were established over two decades ago [1]. Despite changes in the diagnostic criteria for narcolepsy, where individuals are now classified type 1 or type 2 (NT2) instead of with or without cataplexy, no reassessment of these cutoff values has taken place. This also poses a challenge for diagnosing individuals with intermediate hypocretin-1 levels in the range of 111-200 pg/mL, as they are in the "grey area" between deficiency and what is considered normal. In addition, although not the gold standard, determining typical cataplexy is very important in the diagnostic process of NT1 due to its high specificity for this disorder. The expression of cataplexy can vary, and atypical variations exist. However, the diagnostic value of typical compared to atypical cataplexy as not been comprehensively researched until now.

In **Chapter 2**, we found that those with intermediate hypocretin-1 CSF levels exhibited more features associated with NT1 than those with normal levels. We reported that intermediate hypocretin-1 CSF levels are relatively rare, about 5%, in line with recent research [2, 3]. Additionally, when categorizing groups based on cataplexy type, individuals with typical cataplexy had more positive diagnostic findings for NT1 than those with atypical or no cataplexy. Combining both typical and atypical cataplexy into a single category yielded higher sensitivity but lower specificity for diagnosing NT1 compared to using typical cataplexy alone. All non-cataplectic HLA-negative patients in the sample had normal hypocretin-1 CSF levels. Thus it could be inferred that determining hypocretin-1 in these patients is largely redundant as the concentration will likely be normal. However, a recent case series did report seven cases of HLA-negative patients with CSF hypocretin-1 deficiency. These patients were positive for other high-risk HLA-DQB1 alleles such as DQB1*03:01, DQB1*03:02 and DQB1*02:01 [4]. Studying the likelihood of hypocretin-1 deficiency in individuals without these high-risk alleles would be an intriguing avenue for future research.

The Optimal Hypocretin-1 Threshold

Regarding the cutoff value for hypocretin-1 in CSF, our findings indicate that a higher threshold of 150 pg/mL would optimize the diagnostic accuracy for NT1. Recently, another study has also highlighted the need for reconsideration of the current hypocretin-1 cutoff [5, 6]. Torstensen et al. (2023) replicated our results in a Danish and an Italian population. Respectively, the ideal cutoffs were 179 and 129 pg/mL [5]. The currently used cutoff is 110 pg/mL, but now that the results of two studies have shown higher optimal cutoffs, a modification should be considered.

The Clinical Relevance of MSLT and PSG findings

We found a hypocretin-1 cutoff value of 55 pg/mL when using only positive multiple sleep latency test (MSLT) and polysomnography (PSG) findings to predict hypocretin-1 deficiency, with relatively more false positive and false negative outcomes compared to options that include cataplexy in addition to and/or instead of PSG/MSLT findings. This is a value far below the currently used threshold of 110 pg/mL, indicating that MSLT/PSG findings tend show a higher rate of positivity in individuals with a more pronounced hypocretin-1 deficiency. It also suggest that the currently used MSLT and PSG markers are insufficient to adequately diagnose NT1. This sentiment has been supported in recent literature [3]. The MSLT has a relatively low sensitivity, specificity and re-test reliability, although this issue is less pronounced when diagnosing in NT1 as opposed to NT2 and idiopathic hypersomnia (IH) [7, 8]. Moreover, its validation specifically for CDH patients (aside from NT1 patients) is lacking [9]. Several recent studies have suggested alternative MSLT and PSG criteria than the ones currently in use [3, 10, 11]. Very recently, the ICSD-3 has been revised to address some of these concerns. Now, in addition to the previous criteria, the combination of cataplexy and an overnight PSG SOREMP is sufficient for a NT1 diagnosis. Unfortunately, this addition is not likely to solve all the problems with the current criteria. Although the presence of a PSG SOREMP is very specific, it lacks sensitivity as a diagnostic marker for NT1 [6, 12]. Thus, not only the weight of typical cataplexy presence as diagnostic criterion and the hypocretin-1 CSF cutoff value need to be re-assessed, but also the currently used MSLT and PSG criteria.

Summary

- Two important aspects when diagnosing narcolepsy type 1 are hypocretin-1 measurements in CSF and typical cataplexy.
- Those with intermediate hypocretin-1 CSF levels exhibit more features associated with NT1 than those with normal levels.
- Typical cataplexy is more strongly associated with positive diagnostic findings for NT1 than no or atypical cataplexy.
- Combining both typical and atypical cataplexy improves sensitivity but reduces specificity for the NT1 diagnosis compared to typical cataplexy alone.
- A cutoff value of 150 pg/mL for hypocretin-1 improves diagnostic accuracy for NT1.
- Current MSLT and PSG markers alone are insufficient to accurately diagnose NT1.
- The diagnostic criteria for NT1 should be re-evaluated, including the weight of typical cataplexy, the hypocretin-1 cutoff value, and MSTL/PSG criteria.

Hypocretin-1 measurements in cerebrospinal fluid using radioimmunoassay: within and between assay reliability and limit of quantification

Given the importance of hypocretin-1 CSF measurements to the NT1 diagnosis, we aimed to assess the reliability of the currently used measurement techniques. **Chapter 3** evaluates the technical aspects of CSF hypocretin-1 measurements using the RIA kit from Phoenix Pharmaceuticals Inc., which is commonly used for this purpose worldwide [13]. Several important findings are put forward in this chapter. Firstly, the intra-assay reliability was high, even around the currently used cutoff value of 110 pg/mL, which is critical for an accurate diagnosis. However, the inter-assay variability was too low without the commonly applied conversion to Stanford values using harmonization samples, but substantially improved after conversion. Further assessment is needed to determine the impact of converting substantially different sample concentrations. The lowest concentration that could reliably be measured was determined to be 28 pg/mL, providing an applicable threshold for accurate measurement for future studies. Concentrations below this threshold should be reported as 'undetectable'. Additionally, it was found that the different ranges of calibration curves varied widely between RIAs. The impact on intra-assay reliability was significant, but not clinically relevant as the reliability was high in any case. These findings have important implications for the clinical use of RIAs in measuring hypocretin-1 for diagnosing NT1.

Alternative methods: RIA versus mass spectrometry

However, RIAs have downsides. In addition to high inter-assay variability that needs to be compensated by using harmonization samples, RIAs necessitate specialized equipment and

trained personnel. Performing a RIA is technically complex and requires training and adherence to safety protocols to handle the radioactive isotopes that are used in the process. Consequently, the technique is expensive and its accessibility is limited. For example, the Leiden University Medical Center (LUMC) laboratory serves as central facility for hypocretin-1 RIAs for the Netherlands and Belgium.

Due to these drawbacks, alternative nonradioactive techniques have been considered. One commonly used technique, enzyme-linked immunosorbent assay (ELISA), has been demonstrated to be unreliable for hypocretin-1 measurement in CSF [14-16]. However, liquid chromatography combined with mass spectrometry (LC-MS), shows more promise [17, 18]. A study by Bårdsen et al. (2019) measured the same samples with LC-MS and RIA and found a high correlation between the two. Interestingly, hypocretin-1 levels obtained by LC-MS measurements were far lower and had a wider distribution than the levels obtained using RIA [17]. This is likely due to the very specific nature of LC-MS. In LC-MS only hypocretin-1 is quantified, while antibodies used in RIA may cross-react with other molecules. Variability within and between assays was 8% and 15%, respectively. Given that high inter-assay variability in RIA is commonly cited as a reason to seek for alternative methods [17, 18], it is worth noting that the RIAs investigated in our study performed better when measurements were converted to Stanford values (7.5% variability between RIAs after conversion).

Recently, sharp diurnal hypocretin-1 fluctuations were reported in a study applying LC-MS to CSF of cynomolgus monkeys, with the minimum concentration being 83% lower than the maximum concentration [19]. This is far more variability than the slight diurnal variation that was previously found with RIA in healthy human subjects, with only a 4% difference between 5 AM and 5 PM [20] and 10% variation over a period of 24 hours found in a different study [21]. Two other animal studies which did use RIA and showed fluctuations of 31% and 27% in squirrel monkeys and rats respectively [22, 23]. Caution should be taken when comparing the results, as the investigated species and methods differ. However, given the high level of specificity of the LC-MS technique, it would not be surprising if this method is more suitable for the detection of diurnal variations, while the RIA is more likely to provide a relatively stable measurement by also measuring inactive hypocretin-1 fragments [24].

It is important to acknowledge that the presence of substantial diurnal fluctuations of hypocretin-1 levels could complicate the diagnostic process of NT1, which is one of the reasons why this phenomenon requires further investigation before implementation in clinical practice is advisable. A comparative study applying both RIA and LC-MS in humans would provide a clearer picture of their respective capabilities to measure diurnal fluctuations of hypocretin-1

in CSF. Moreover, this study would provide an opportunity to determine the extent and timing of these fluctuations (as measured using different techniques) in a larger sample, which could lead to further insight into the role and function of hypocretin-1.

Before considering the application of LC-MS in clinical practice, all the specific factors that influence its variability would need to be fully understood. This understanding will enable fair interpretation of results and/or the establishment of standardized measurement practices. Additionally, a suitable LC-MS hypocretin-1 CSF threshold needs to be defined before clinical application is possible. Therefore, although LC-MS has certain benefits, more research is necessary before it can be implemented clinically. Additionally, LC-MS is relatively expensive compared to RIA, which could hinder its large scale clinical implementation. Despite its limitations, the RIA remains the most reliable method to determine hypocretin-1 CSF deficiency in the context of diagnosing NT1.

In light of this, we also suggest that assessment of the inter-assay variability of hypocretin-1 CSF measurements at low and high concentration levels using RIA should take place, which has not been done as of now.

Summary

- Radioimmunoassay (RIA) is the commonly used method to measure hypocretin-1 in CSF, the gold standard for the NT1 diagnosis. Thus, it is important to assess the reliability of this method.
- RIA intra-assay reliability is high, inter-assay reliability is low, unless converted to Stanford values.
- CSF hypocretin-1 values below 28 pg/mL should be reported as undetectable in future studies.
- Calibration curves varied considerably between RIAs, the impact on reliability was significant but low.
- RIAs require specialized equipment and trained personnel.
- Alternative nonradioactive techniques are considered: enzyme-linked immunosorbent assay (ELISA) is unreliable, but liquid chromatography combined with mass spectrometry (LC-MS) shows promise.
- LC-MS may be better suited for detection of hypocretin-1 variability in CSF, while RIA provides more stable measurements with less diurnal variation.
- Clinical application of LC-MS requires more research. Currently, RIA remains the most reliable method to assess hypocretin-1 in CSF in the diagnostic process of narcolepsy.

Applicability of the Sustained Attention to Response Task (SART) in hypersomnolence: experience and results from a tertiary referral center

We searched for additional tools to better understand and monitor CDH and focused on an under-investigated aspect of these disorders: vigilance. In **Chapter 4,** we assessed the applicability of the Sustained Attention to Response Task (SART) in the clinical practice of a tertiary referral center for hypersomnolence disorders. Both individuals undergoing regular diagnostic procedures and those undergoing a driver's license evaluation were included in this evaluation. Associations between SART outcomes and PSG, MSLT and ESS results were assessed. There is no universally accepted cutoff value for the SART, nevertheless we demonstrate that the SART can detect impaired vigilance based on a previously established cutoff value from Fronczek et al. (2006) [25]. However, the SART cannot differentiate between different hypersomnolence disorders after adjustments for relevant confounders had taken place. It measures a different aspect of EDS compared to the MSLT and PSG, is influenced by the time of day, and does not seem to be influenced by symptoms of depression. Individuals with increased anxiety showed a tendency to make more omission errors. Additionally, it was found that individuals undergoing a driver's license evaluation had lower error rates, possibly due to treatment and/or motivational factors. Of course, the reason for the better performance of the driver's license evaluation group is difficult to interpret because there are multiple differences between the two groups and there may have been a selection bias (i.e. more severely affected individuals may have been less likely to undergo such an evaluation). Based on the results, age and time of day need to be considered when interpreting SART results.

Barriers to the assessment of vigilance in clinical practice

The results provide valuable insights into the application of the SART in the clinical practice of a tertiary sleep-wake center. However, further research is needed to draw definitive conclusions on the applicability of the SART as a monitoring tool. This research should include a more diverse study population consisting of both healthy individuals and those diagnosed with hypersomnolence disorders. Moreover, it is crucial to determine a definitive SART cutoff value in a large diverse populations that includes both healthy individuals and those with different hypersomnolence disorders. Impact of age and time of day need to be further explored, also when it comes to cutoff values, given that our results indicate that age is an associated factor to the SART results.

It is evident that vigilance impairments in the context of CDH are under-investigated. But it is still unclear what the best measurement method is. For example, many studies use the Psychomotor Vigilance Task (PVT) instead [26-29]. Comparing studies becomes more challenging when there are substantial differences in the methods used. It would be beneficial to conduct comparative studies, evaluating alternative assessment methods with the SART and determining their respective efficacy and reliability. While some reviews have made such comparisons based on studies assessing a single method, there is a notable lack of direct comparisons in clinical trials [30]. Therefore, further investigation in this area is needed.

Summary

- Vigilance is an important but underexplored aspect of hypersomnolence disorders.
- The SART can detect and quantify impaired vigilance, but cannot differentiate between different causes of hypersomnolence.
- The SART measures a different aspect of EDS compared to the MSLT and PSG, is influenced by the time of day, and is not affected by depression symptoms.
- Age and time of day should be considered when interpreting SART results.
- A SART cutoff value needs to be established, perhaps for different age and gender categories.
- Future research should compare the SART directly with alternative tools used to measure vigilance to determine the most appropriate methods and procedures and to increase standardization across different sleep-wake centers.

Part II: Sleep in the Intensive Care Environment

Exploring sleep quality, duration, and related factors in the intensive and intermediate care units

Chapter 5 highlights the prevalence and negative impact of sleep disturbances in individuals admitted to the ICU and IMCU. The detrimental effects of poor sleep on patients' recovery and its association with the onset of delirium are well established [31-33]. Delirium is a condition that often prolongs hospital stays and leads to increased morbidity and mortality [31]. Furthermore, animal studies have even shown a direct causal relationship between sleep deprivation and mortality [34, 35]. This raises the question of whether sleep disruption could independently impact mortality in the ICU, even when adjusted for the presence of delirium. Given the known negative physical and mental effects of sleep disruption, identifying factors associated with sleep disruption could help prevent negative effects by addressing those specific factors.

We found a short sleep duration during the night (from 10 PM to 6 AM) of only 4.6 out of a maximum of 8 hours. Sleep quality was generally poor, with an average nurse-reported Richards-Campbell Sleep Questionnaire (RSCQ) score of 4.9 out of 10. Poor sleep quality or duration were associated with the use of benzodiazepines, delirium, and overnight transfers.

Sleep and delirium

The average day-night sleep ratio during the three days before the onset of delirium was notably different (with more sleep during the day) compared to the average of the total ICU stay of individuals who did not experience delirium. The RCSQ score was also lower, but not significantly so.

Sleep disturbances are currently not mentioned as a risk factor for the development of delirium in the Clinical Practice Guidelines for the Management of Pain, Agitation, and Delirium [36]. In spite of this, there are indications suggesting that pre-existing sleep disturbances are likely associated with the onset of delirium [37-39]. Moreover, delirium and sleep deprivation share similar symptoms, risk factors and are thought to be associated with malfunction of the same central nervus system regions (prefrontal cortex, thalamus, and posterior parietal cortex) [40], further suggesting a connection (see figure 7.1).

So far, these connections have primarily been explored in post-operative and non-critical care settings. Additionally, studies tend to focus on sleep after the start of delirium, not on sleep

preceding delirium. Only a few studies have been designed with the intent of studying the potential of sleep disturbances as risk factor for the development of delirium in the ICU. Helton et al (1980) found that sleep deprivation were more likely to develop mental status changes and Ángeles-Castellanos et al (2016) found average melatonin levels decreasing and a loss of melatonin rhythmicity three days before a delirium was diagnosed [41]. Moreover, two studies implementing ICU-wide interventions to improve sleep demonstrated a significant increase in delirium-free days [42, 43]. These studies support our findings, although additional research is necessary to better understand the bi-directional link between disrupted sleep and delirium.

A prospective trial utilizing a combination of 24-hours PSG, the RCSQ, and delirium screening would be ideal to determine the nature of the link between sleep disruption and delirium. Employing PSG in a large sample is generally not feasible due to the significant burden it places on patients and the substantial labor that is involved, but could be done in a smaller subset.

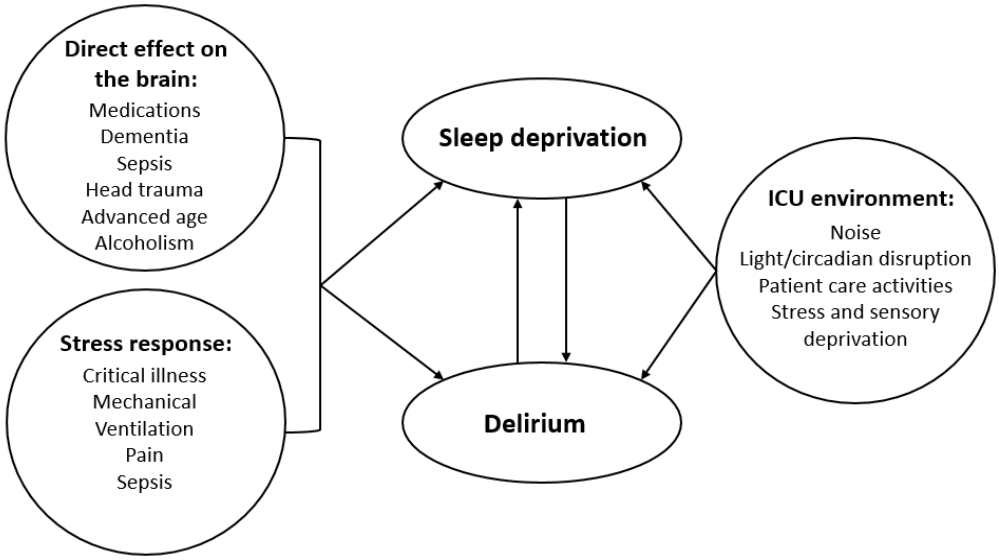


Figure 7.1. A possible relationship between delirium and sleep deprivation and their overlapping risk factors. Figure adapted from Weinhouse GL, Schwab RJ, Watson PL, Patil N, Vaccaro B, Pandharipande P, Ely EW. *Crit Care*. 2009;13(6):234

Benzodiazepines

Benzodiazepines were frequently used in our study, with 56% of patients receiving a non-continuous (i.e. not used for sedative purposes) benzodiazepine one or more nights during their ICU stay. Initiating benzodiazepines as hypnotics did not result in improved sleep quality or duration. These findings lack clear support or contradiction within the existing scientific literature. While benzodiazepines are frequently prescribed as hypnotic, there is limited evidence regarding their efficacy to improve sleep in the ICU. However, several studies have investigated the effects of benzodiazepines on sleep in other populations. Benzodiazepines tend to shorten sleep latency and increase the total sleep time, but negatively affect sleep quality [44, 45]. In the ICU, benzodiazepines can increase the duration of the ICU stay and are associated with an increased risk of developing delirium [46, 47]. Considering these factors, benzodiazepines should be prescribed with caution, and more effective alternatives with less detrimental side effects need to be found.

Chapter 5 provides further grounds for more extensive research into the efficacy of benzodiazepines in the ICU. Furthermore, studies should be conducted to explore and compare non-pharmacological and non-benzodiazepine interventions. Considering that commonly used medications in the ICU, such as benzodiazepines and propofol, may result in suboptimal sleep architecture, it would be interesting to investigate whether sodium oxybate, a medication known to enhance sleep architecture in narcolepsy patients, could be beneficial in the ICU setting. Dexmedetomidine is another medication that shows promise and requires further investigation in the form of a large randomized controlled trial [48, 49].

Sleep and mortality

One year post ICU-discharge mortality was 30%. Higher age, greater disease severity, and female sex were associated were increased post-discharge mortality. The association between decreased sleep quality and increased mortality did not reach statistical significance (p=.070), but should be considered as a potential association that cannot be ruled out at this point. Contrary to findings from previous studies, the association between delirium and increased mortality was not significant in our study [50]. Potential explanations are that some cases of hypoactive delirium may not have been recognized and that cases of delirium were grouped together regardless of severity or duration.

Summary

- In the ICU, sleep duration is short and has poor quality.
- A change in day-night sleep patterns precedes a delirium. More research is required to gain a comprehensive understanding of potential causal pathways.
- No significant direct association between sleep and mortality was found, although further research is needed before definitive conclusions can be drawn.
- Benzodiazepines are commonly prescribed as hypnotics despite the lack of evidence. Further research is needed. As there are known side effect they should be used with caution.
- Use of non-pharmacological interventions and non-benzodiazepine medications in the ICU should be further explored.
- Prospective research involving a combination of methods including PSG, RCSQ and delirium screening could clarify the link between sleep disruption and delirium.

Dream experiences during ICU stay: prevalence, content, vividness, and associated factors

Reports from nurses providing ICU aftercare prompted an investigation of vivid dream experiences in the ICU. The findings are reported in **Chapter 6**. Our findings showed a high prevalence of these experiences, with 79% of participants reporting them. The dreams were not only intense but also frequently had negative emotional associations for participants. Participants reported being impacted both during their stay and even after discharge. The content was often related to the ICU stay itself, and frequently consisted of themes of helplessness and death. We found that it was challenging for both participants and investigators to distinguish dream experiences from hallucinations.

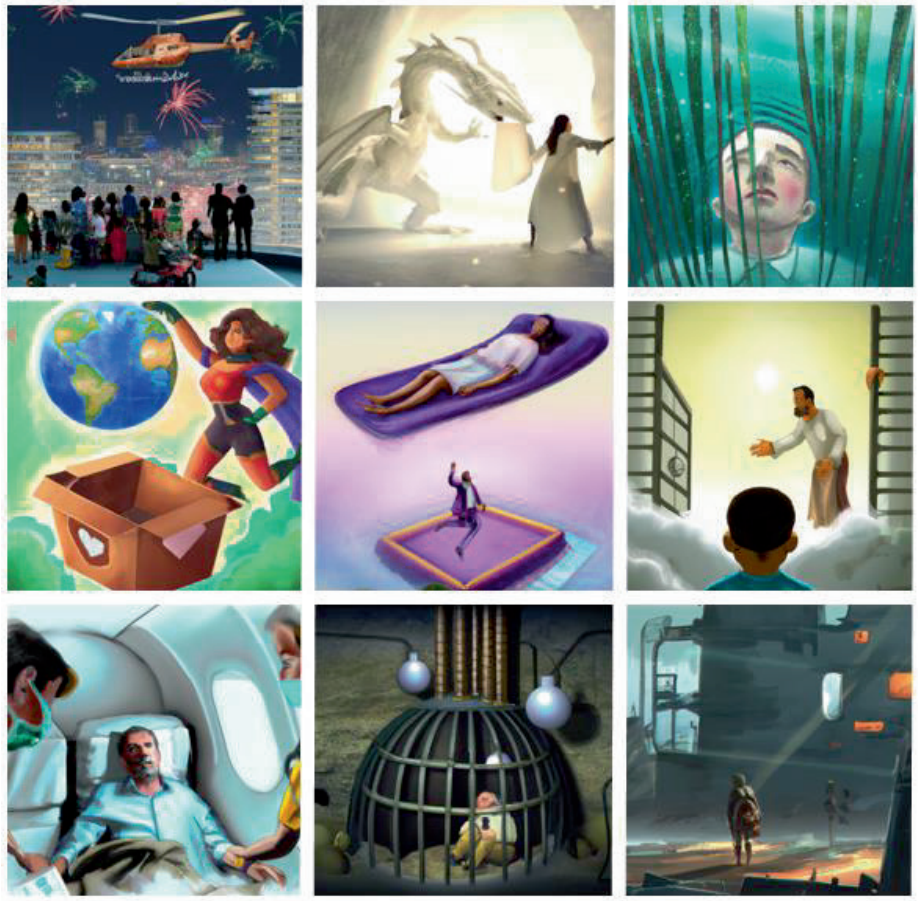


Figure 7.2. Examples of dream experiences mentioned by former ICU patients

Mitigating biases and assessing impact

In future research, it is essential to overcome the limitations of the study design of the study presented in **Chapter 6**. The potential selection and recall bias could partly be avoided by prospectively including ICU patients during their admission and conducting interviews at various intervals throughout their ICU stay and afterwards. This approach also allows for further exploration of the emotional and psychological impact of dream experiences on patients during and after admission. Moreover, a prospective study design could help facilitate a clearer differentiation between perceptual disturbances, hallucinations and dream experiences. It would be easier to relate the occurrence of dream experiences to specific factors when the time of its occurrence is known. Some of the suggestions provided by participants are interesting to further explore, such as whether informing patients admitted for more than a couple of days of the possibility of experiencing vivid dreams has any beneficial impact on their well-being

during and after their ICU stay. Lastly, investigating the occurrence of vivid dream experiences in settings beyond the ICU, such as the general ward, could also help to discern whether this is a specific ICU phenomenon. Such research could contribute to a better understanding of the underlying reasons for vivid dream experiences in the ICU.

Summary

- In our study, the prevalence of (vivid) dream experiences in the ICU is high.
- Dreams are often intense and negative, and can impact patients during and after their ICU stay.
- Dream content is frequently related to the ICU admission and to themes of death and helplessness.
- Differentiating dream experiences from perceptual disturbances is challenging for participants and researchers.
- A future prospective study could overcome some of the limitations from our study, as well as explore the psychological impact of the dream experiences.

Overall conclusion

Sleep constitutes a fundamental aspect of human life, with its disruption significantly impacting both periods of rest and wakefulness, a fact well known by those who have dealt with sleep issues themselves. In conditions like CDH and within the ICU environment, the lines between sleep and wakefulness are blurred. This phenomenon is exemplified by cognitive complaints in CDH and disruptions to circadian rhythms as well as vivid dreams in the ICU. Insights gained into sleep in one context can help shed light on the other. The outcomes of studies explored in this thesis underscore the opportunities both in the field of critical care medicine and central disorders of hypersomnolence. The necessity for refining diagnostic criteria and assessment methods is highlighted as well. By pinpointing areas for future investigation, this thesis sets the stage for more precise diagnostic tools and a deeper understanding of the intricate relationship between sleep and overall well-being, with the ultimate goal of better prevention or treatment and thus better (patient) care.

References

1. Mignot, E., et al., *The Role of Cerebrospinal Fluid Hypocretin Measurement in the Diagnosis of Narcolepsy and Other Hypersomnias.* Archives of Neurology, 2002. **59**(10): p. 1553-1562.
2. Postiglione, E., et al., *Narcolepsy with intermediate cerebrospinal level of hypocretin-1.* Sleep, 2022. **45**(2).
3. Lopez, R., et al., *Rapid eye movement sleep duration during the multiple sleep latency test to diagnose hypocretin-deficient narcolepsy.* Sleep, 2023. **46**(1).
4. Miano, S., et al., *A series of 7 cases of patients with narcolepsy with hypocretin deficiency without HLA DQB1*06:02 allele.* J Clin Sleep Med, 2023.
5. Torstensen, E.W., et al., *Repeated measures of hypocretin-1 in Danish and Italian patients with narcolepsy and in controls.* Sleep Medicine, 2023. **101**: p. 213-220.
6. Andlauer, O., et al., *Predictors of hypocretin (orexin) deficiency in narcolepsy without cataplexy.* Sleep, 2012. **35**(9): p. 1247-55F.
7. Trotti, L.M., B.A. Staab, and D.B. Rye, *Test-retest reliability of the multiple sleep latency test in narcolepsy without cataplexy and idiopathic hypersomnia.* J Clin Sleep Med, 2013. **9**(8): p. 789-95.
8. Ruoff, C., et al., *The MSLT is Repeatable in Narcolepsy Type 1 But Not Narcolepsy Type 2: A Retrospective Patient Study.* J Clin Sleep Med, 2018. **14**(1): p. 65-74.
9. Lammers, G.J. and J.G. van Dijk, *The multiple sleep latency test: a paradoxical test?* Clinical Neurology and Neurosurgery, 1992. **94**: p. 108-110.
10. Gool, J.K., et al., *Data-Driven Phenotyping of Central Disorders of Hypersomnolence With Unsupervised Clustering.* Neurology, 2022. **98**(23): p. e2387-e2400.
11. Pizza, F., et al., *Different sleep onset criteria at the multiple sleep latency test (MSLT): an additional marker to differentiate central nervous system (CNS) hypersomnias.* J Sleep Res, 2011. **20**(1 Pt 2): p. 250-6.
12. Andlauer, O., et al., *Nocturnal rapid eye movement sleep latency for identifying patients with narcolepsy/hypocretin deficiency.* JAMA Neurol, 2013. **70**(7): p. 891-902.
13. Bassetti, C., M. Billiard, and E. Mignot, *Narcolepsy and hypersomnia*. 2006: Informa Health Care.
14. Lindström, M., et al., *Orexin-A measurement in narcolepsy: A stability study and a comparison of LC-MS/MS and immunoassays.* Clin Biochem, 2021. **90**: p. 34-39.
15. Ono, T., et al., *Measurement of cerebrospinal fluid orexin-A (hypocretin-1) by enzyme-linked immunosorbent assay: A comparison with radioimmunoassay.* Psychiatry Clin Neurosci, 2018. **72**(11): p. 849-850.
16. Sarkanen, T., et al., *Misdiagnosis of narcolepsy caused by a false-positive orexin-A/hypocretin-1 enzyme immune assay.* J Clin Sleep Med, 2022. **18**(8): p. 2075-2078.
17. Bårdsen, K., et al., *Considerably Lower Levels of Hypocretin-1 in Cerebrospinal Fluid Is Revealed by a Novel Mass Spectrometry Method Compared with Standard Radioimmunoassay.* Anal Chem, 2019. **91**(14): p. 9323-9329.
18. Hirtz, C., et al., *From radioimmunoassay to mass spectrometry: a new method to quantify orexin-A (hypocretin-1) in cerebrospinal fluid.* Sci Rep, 2016. **6**: p. 25162.
19. Narita, N., et al., *Diurnal Fluctuations of Orexin-A and -B in Cynomolgus Monkey Cerebrospinal Fluid Determined by a Novel Analytical Method Using Antiadsorptive Additive Treatment Followed by Nanoflow Liquid Chromatography-High-Resolution Mass Spectrometry.* ACS Chem Neurosci, 2023. **14**(4): p. 609-618.
20. Grady, S.P., et al., *Diurnal Variation in CSF Orexin-A in Healthy Male Subjects.* Sleep, 2006. **29**(3): p. 295-297.
21. Salomon, R.M., et al., *Diurnal variation of cerebrospinal fluid hypocretin-1 (Orexin-A) levels in control and depressed subjects.* Biological Psychiatry, 2003. **54**(2): p. 96-104.
22. Yoshida, Y., et al., *Fluctuation of extracellular hypocretin-1 (orexin A) levels in the rat in relation to the light–dark cycle and sleep–wake activities.* European Journal of Neuroscience, 2001. **14**(7): p. 1075-1081.
23. Zeitzer, J.M., et al., *Circadian and homeostatic regulation of hypocretin in a primate model: implications for the consolidation of wakefulness.* J Neurosci, 2003. **23**(8): p. 3555-60.
24. Sakai, N., et al., *HPLC analysis of CSF hypocretin-1 in type 1 and 2 narcolepsy.* Sci Rep, 2019. **9**(1): p. 477.
25. Fronczek, R., et al., *Focusing on vigilance instead of sleepiness in the assessment of narcolepsy: high sensitivity of the Sustained Attention to Response Task (SART).* Sleep, 2006. **29**(2): p. 187-91.
26. Thomann, J., et al., *Psychomotor vigilance task demonstrates impaired vigilance in disorders with excessive daytime sleepiness.* J Clin Sleep Med, 2014. **10**(9): p. 1019-24.

27. Lee, I.-S., et al., *The Relationship between psychomotor vigilance performance and quality of life in obstructive sleep apnea.* Journal of clinical sleep medicine : JCSM : official publication of the American Academy of Sleep Medicine, 2011. **7**(3): p. 254-260.
28. Trotti, L.M., et al., *The Psychomotor Vigilance Test as a measure of alertness and sleep inertia in people with central disorders of hypersomnolence.* J Clin Sleep Med, 2022. **18**(5): p. 1395-1403.
29. Mathis, J., et al., *The diagnostic value of sleep and vigilance tests in central disorders of hypersomnolence.* Sleep, 2022. **45**(3).
30. Filardi, M., et al., *Cognitive dysfunction in central disorders of hypersomnolence: A systematic review.* Sleep Med Rev, 2021. **59**: p. 101510.
31. Stuck, A., M.J. Clark, and C.D. Connelly, *Preventing intensive care unit delirium: a patient-centered approach to reducing sleep disruption.* Dimens Crit Care Nurs, 2011. **30**(6): p. 315-20.
32. Pisani, M.A., et al., *Sleep in the intensive care unit.* Am J Respir Crit Care Med, 2015. **191**(7): p. 731-8.
33. Pulak, L.M. and L. Jensen, *Sleep in the Intensive Care Unit:A Review.* Journal of Intensive Care Medicine, 2016. **31**(1): p. 14-23.
34. Friese, R.S., B. Bruns, and C.M. Sinton, *Sleep deprivation after septic insult increases mortality independent of age.* J Trauma, 2009. **66**(1): p. 50-4.
35. Maragno-Correa, J.M., et al., *Sleep deprivation increases mortality in female mice bearing Ehrlich ascitic tumor.* Neuroimmunomodulation, 2013. **20**(3): p. 134-40.
36. Devlin, J.W., et al., *Clinical Practice Guidelines for the Prevention and Management of Pain, Agitation/Sedation, Delirium, Immobility, and Sleep Disruption in Adult Patients in the ICU.* Crit Care Med, 2018. **46**(9): p. e825-e873.
37. Leung, J.M., et al., *Preoperative Sleep Disruption and Postoperative Delirium.* J Clin Sleep Med, 2015. **11**(8): p. 907-13.
38. Oldham, M.A., et al., *Baseline sleep as a predictor of delirium after surgical aortic valve replacement: A feasibility study.* Gen Hosp Psychiatry, 2021. **71**: p. 43-46.
39. Jaiswal, S.J., et al., *Objectively measured sleep fragmentation is associated with incident delirium in older hospitalized patients: Analysis of data collected from an randomized controlled trial.* J Sleep Res, 2021. **30**(3): p. e13205.
40. Weinhouse, G.L., et al., *Bench-to-bedside review: Delirium in ICU patients - importance of sleep deprivation.* Critical Care, 2009. **13**(6): p. 234.
41. Ángeles-Castellanos, M., et al., *Loss of melatonin daily rhythmicity is asociated with delirium development in hospitalized older adults.* Sleep Sci, 2016. **9**(4): p. 285-288.
42. Kamdar, B.B., et al., *The effect of a quality improvement intervention on perceived sleep quality and cognition in a medical ICU.* Crit Care Med, 2013. **41**(3): p. 800-9.
43. Tonna, J.E., et al., *The Effect of a Quality Improvement Intervention on Sleep and Delirium in Critically Ill Patients in a Surgical ICU.* Chest, 2021. **160**(3): p. 899-908.
44. Achermann, P. and A.A. Borbély, *Dynamics of EEG slow wave activity during physiological sleep and after administration of benzodiazepine hypnotics.* Hum Neurobiol, 1987. **6**(3): p. 203-10.
45. Monti, J.M., *Sleep laboratory and clinical studies of the effects of triazolam, flunitrazepam and flurazepam in insomniac patients.* Methods Find Exp Clin Pharmacol, 1981. **3**(5): p. 303-26.
46. Wang, H., et al., *Sedative drugs used for mechanically ventilated patients in intensive care units: a systematic review and network meta-analysis.* Curr Med Res Opin, 2019. **35**(3): p. 435-446.
47. Fraser, G.L., et al., *Benzodiazepine versus nonbenzodiazepine-based sedation for mechanically ventilated, critically ill adults: a systematic review and meta-analysis of randomized trials.* Crit Care Med, 2013. **41**(9 Suppl 1): p. S30-8.
48. Sun, Y.M., et al., *Effect of low-dose dexmedetomidine on sleep quality in postoperative patients with mechanical ventilation in the intensive care unit: A pilot randomized trial.* Front Med (Lausanne), 2022. **9**: p. 931084.
49. Alexopoulou, C., et al., *Effects of dexmedetomidine on sleep quality in critically ill patients: a pilot study.* Anesthesiology, 2014. **121**(4): p. 801-7.
50. Aung Thein, M.Z., et al., *A call to action for delirium research: Meta-analysis and regression of delirium associated mortality.* BMC Geriatrics, 2020. **20**(1): p. 325.

Appendices

A Samenvatting, discussie en toekomstige perspectieven

In deze thesis worden aspecten van slaapaandoeningen bij centrale stoornissen leidend tot hypersomnolentie ("centrale stoornissen van hypersomnolentie" genoemd) en in de intensieve zorgomgeving onderzocht. Diagnostische methoden voor hypersomnolentiestoornissen worden in deel I besproken (**Hoofdstukken 2, 3 en 4**). Besproken worden: de meest optimale afkapwaarde voor hypocretine-1 in hersenvocht (liquor), de waarde van typische kataplexie in het diagnostische proces van narcolepsie type 1 (NT1), het evalueren van de betrouwbaarheid van de hypocretine-1 bepaling in liquor met behulp van radioimmunoassay (RIA), en de waarde van de 'sustained attention to response task' (SART) bij het evalueren en beoordelen van centrale stoornissen van hypersomnolentie. Deel II (**Hoofdstukken 5 en 6**) behandelt slaap in de kritieke zorgomgeving. Dit deel omvat een bespreking van de slaapproblemen die zich voordoen op de Intensive en Medium Care afdelingen, de factoren die slaap beïnvloeden, de relatie tussen slaap en mortaliteit, en de factoren die van invloed kunnen zijn bij levendige en negatieve droomervaringen op de Intensive Care. Hieronder worden de bevindingen van de hoofdstukken samengevat en besproken. Potentiële toekomstige perspectieven worden besproken aan het einde van zowel Deel I als Deel II.

Deel I: Aspecten van Slaap bij Centrale Aandoeningen van Hypersomnolentie

Intermediaire hypocretine-1 niveaus in liquor en typische kataplexie: hun betekenis bij de diagnose van narcolepsie type 1

De meting van hypocretine-1 in hersenvocht, ook wel liquor genoemd, is de gouden standaard bij de diagnose van NT1. De momenteel gebruikte afkapwaarden voor deficiënte en normale niveaus (respectievelijk 110 en 200 pg/mL) werden meer dan twee decennia geleden vastgesteld [1]. Ondanks veranderingen in de diagnostische criteria voor narcolepsie, waarbij individuen nu worden ingedeeld als type 1 of type 2 (NT2) in plaats van met of zonder kataplexie, heeft er geen herbeoordeling van deze afkapwaarden plaatsgevonden. Dit vormt ook een uitdaging voor het diagnosticeren van individuen met intermediaire hypocretine-1 niveaus van 111-200 pg/mL, omdat ze zich in het "grijze gebied" bevinden tussen deficiëntie

en wat als normaal wordt beschouwd. Bovendien, hoewel niet de gouden standaard, is het bepalen van typische kataplexie zeer belangrijk in het diagnostische proces van NT1 vanwege de hoge specificiteit van dit symptoom voor deze stoornis. De expressie van kataplexie kan variëren en atypische variaties komen voor. Echter, de diagnostische waarde van typische vergeleken met atypische kataplexie is tot nu toe niet uitgebreid onderzocht. In **Hoofdstuk 2** vonden we dat degenen met intermediaire hypocretine-1 niveaus in liquor meer kenmerken vertoonden die geassocieerd zijn met NT1 dan degenen met normale niveaus. We rapporteerden dat intermediaire hypocretine-1 niveaus relatief zeldzaam zijn, ongeveer 5% in de onderzochte populatie, in lijn met recent onderzoek [2, 3]. Bovendien, bij het categoriseren van groepen op basis van het type kataplexie, hadden individuen met typische kataplexie meer positieve diagnostische bevindingen voor NT1 dan degenen met atypische of geen kataplexie. Het combineren van zowel typische als atypische kataplexie in één categorie leverde een hogere sensitiviteit op maar lagere specificiteit voor het diagnosticeren van NT1 in vergelijking met het alleen gebruiken van typische kataplexie. Alle HLA-negatieve patiënten zonder kataplexie in de steekproef hadden normale hypocretine-1 niveaus in liquor. Daaruit zou kunnen worden afgeleid dat het bepalen van hypocretine-1 bij deze patiëntengroep grotendeels overbodig is, aangezien de concentratie waarschijnlijk normaal zal zijn. Er werd echter recentelijk in een casusreeks melding gemaakt van zeven gevallen van HLA-negatieve patiënten met hypocretine-1 deficiëntie in liquor. Deze patiënten waren positief voor andere hoogrisico HLA-DQB1-allelen zoals DQB103:01, DQB103:02 en DQB1*02:01 [4]. Het bestuderen van de waarschijnlijkheid van hypocretine-1 deficiëntie bij individuen zonder deze hoogrisico-allelen zou een relevant onderzoeksgebied zijn voor de toekomst.

De Optimale Hypocretine-1 Afkapwaarde

Onze bevindingen geven aan dat een hogere afkapwaarde van 150 pg/mL van hypocretine-1 in liquor de diagnostische nauwkeurigheid voor NT1 zou optimaliseren. Onlangs heeft een andere studie ook de noodzaak benadrukt om de huidige hypocretine-1 afkapwaarde te heroverwegen [5, 6]. Torstensen et al. (2023) hebben onze resultaten gerepliceerd in een Deense en een Italiaanse populatie. Respectievelijk waren de ideale afkapwaarden 179 en 129 pg/mL [5]. Vanwege de toename van de diagnostische nauwkeurigheid bij deze geïdentificeerde hogere afkapwaarden, zou een verhoging van het huidige afkappunt van 110 pg/ml overwogen moeten worden.

De Klinische Relevantie van MSLT- en PSG-bevindingen

We vonden een hypocretine-1 afkapwaarde van 55 pg/mL bij het alleen gebruiken van positieve bevindingen van de multiple sleep latency test (MSLT) en polysomnografie (PSG) om hypocretine-1 deficiëntie te voorspellen, met relatief meer fout-positieve en fout-negatieve resultaten vergeleken met opties die kataplexie omvatten in aanvulling op en/of in plaats van PSG/MSLT bevindingen. Dit is een waarde ver onder de momenteel gebruikte drempel van 110 pg/mL, wat aangeeft dat MSLT/PSG meer afwijkende bevindingen laten zien bij individuen met een meer uitgesproken hypocretine-1 deficiëntie. Het suggereert ook dat de momenteel gebruikte MSLT- en PSG-markers ontoereikend zijn om NT1 voldoende adequaat te diagnosticeren. Deze opvatting wordt ondersteund in recente literatuur [3]. De MSLT heeft een relatief lage sensitiviteit, specificiteit en repliceerbaarheid wanneer de test herhaald wordt, hoewel dit probleem minder uitgesproken is bij de diagnose van NT1 in vergelijking met NT2 en idiopathische hypersomnie (IH) [7, 8]. Bovendien ontbreekt de validatie ervan specifiek voor patiënten met centrale stoornissen van hypersomnolentie (afgezien van NT1-patiënten) [9].

Verschillende recente studies hebben alternatieve MSLT- en PSG-criteria voorgesteld dan die momenteel in gebruik zijn [3, 10, 11]. Recent is de ICSD-3 herzien om enkele van deze zorgen aan te pakken. Nu is, naast de eerdere criteria, de combinatie van kataplexie en een PSG SOREM periode gedurende de nacht voldoende voor een NT1-diagnose. Helaas zal deze toevoeging waarschijnlijk niet alle problemen met de huidige criteria oplossen. Hoewel de aanwezigheid van een PSG SOREM periode zeer specifiek is, ontbreekt het aan sensitiviteit als diagnostische marker voor NT1 [6, 12]. Dus niet alleen het belang van de aanwezigheid van typische kataplexie als diagnostisch criterium en de hypocretine-1 afkapwaarde in liquor moeten worden heroverwogen, maar ook de momenteel gebruikte MSLT- en PSG-criteria.

Metingen van hypocretin-1 in liquor met behulp van radioimmunoassay: betrouwbaarheid binnen en tussen assays en kwantificeringsgrens

Gezien het belang van hypocretine-1 metingen in liquor voor de NT1-diagnose, hebben we als doel gesteld om de betrouwbaarheid van de momenteel gebruikte meettechnieken te beoordelen. **Hoofdstuk 3** evalueert de technische aspecten van hypocretine-1 metingen in liquor met behulp van het RIA-kit van Phoenix Pharmaceuticals Inc., welke wereldwijd veel wordt gebruikt voor dit doel [13]. In dit hoofdstuk worden verschillende belangrijke

bevindingen doorgenomen. Ten eerste was de intra-assay betrouwbaarheid hoog, zelfs rond de momenteel gebruikte afkapwaarde van 110 pg/mL, welke cruciaal is voor een nauwkeurige diagnose. Echter, de inter-assay variabiliteit was te laag als conversie naar Stanford-waarden met behulp van harmonisatiemonsters niet werd toegepast, maar verbeterde aanzienlijk na conversie. Verder onderzoek is nodig om het effect van het omzetten van aanzienlijk verschillende concentraties van monsters te bepalen. De laagste concentratie die betrouwbaar kon worden gemeten, bleek 28 pg/mL te zijn, deze bepaalbaarheidsgrens dient bij toekomstige studies te worden toegepast. Concentraties onder deze grens moeten worden gerapporteerd als 'niet-detecteerbaar'. Bovendien werd gevonden dat de verschillende bereiken van kalibratiecurven sterk varieerden tussen RIAs. Het effect op intra-assay betrouwbaarheid was significant, maar niet klinisch relevant omdat de betrouwbaarheid in elk geval hoog was. Deze bevindingen hebben belangrijke implicaties voor het klinisch gebruik van RIAs bij het meten van hypocretine-1 voor de diagnose van NT1.

Alternatieve methoden: RIA versus massaspectrometrie

RIAs hebben echter nadelen. Naast de hoge inter-assay variabiliteit die gecorrigeerd moet worden met behulp van harmonisatiemonsters, vereisen RIAs gespecialiseerde apparatuur en getraind personeel. Het uitvoeren van een RIA is technisch complex en vereist training en naleving van veiligheidsprotocollen om met de radioactieve isotopen die in het proces worden gebruikt om te gaan. De techniek is duur en de toegankelijkheid beperkt. Bijvoorbeeld, het laboratorium van het Leids Universitair Medisch Centrum (LUMC) fungeert als de locatie waar hypocretine-1 RIAs voor zowel Nederland als België worden uitgevoerd. Vanwege deze nadelen worden alternatieve niet-radioactieve technieken overwogen. Een veelgebruikte techniek, enzymgebonden immunosorbent assay (ELISA), is echter onbetrouwbaar gebleken om hypocretine-1 concentraties in liquor te bepalen [14-16]. Vloeistofchromatografie gecombineerd met massaspectrometrie (LC-MS), lijkt hoopgevender [17, 18]. Bij een studie van Bårdsen et al. (2019) werden dezelfde monsters met LC-MS en RIA gemeten en werd een hoge correlatie tussen de twee gevonden. Interessant genoeg waren de hypocretine-1 concentraties verkregen door LC-MS metingen veel lager en hadden ze een bredere verdeling dan de concentraties verkregen met behulp van RIA [17]. Dit is waarschijnlijk te wijten aan de zeer specifieke aard van LC-MS. Met LC-MS wordt alleen hypocretine-1 gekwantificeerd, terwijl antilichamen die in RIA worden gebruikt ook kunnen reageren met hypocretine-1 fragmenten. De variabiliteit binnen en tussen assays was respectievelijk 8% en 15%. Aangezien hoge inter-assay variabiliteit in RIA vaak wordt genoemd als reden om te zoeken naar

alternatieve methoden [17, 18], is het vermeldenswaardig dat de onderzochte RIAs in onze studie beter presteerden (7,5% variabiliteit tussen RIAs na conversie naar Stanford-waarden). Onlangs zijn scherpe diurnale hypocretine-1 schommelingen in liquor van Java-apen gerapporteerd in een studie waarbij LC-MS werd toegepast, de minimale concentratie was 83% lager dan de maximale concentratie [19]. Dit is veel meer variabiliteit dan de lichte diurnale variatie die eerder werd gevonden met RIA bij gezonde proefpersonen, met slechts een 4% verschil tussen 5 uur 's ochtends en 5 uur' s middags [20] en een variatie van 10% gedurende een periode van 24 uur gevonden in een andere studie [21]. Twee andere dierstudies die wel RIA gebruikten, vertoonden schommelingen van respectievelijk 31% en 27% bij eekhoornapen en ratten [22, 23]. Voorzichtigheid is geboden bij het vergelijken van de resultaten, aangezien de onderzochte populaties en methoden verschillen. Echter, gezien het hoge niveau van specificiteit van de LC-MS-techniek, zou het niet verrassend zijn als deze methode geschikter is voor de detectie van diurnale variatie, terwijl de RIA waarschijnlijk relatief stabieler is door ook inactieve afbraakproducten te meten [24]. Het is belangrijk om te erkennen dat de aanwezigheid van aanzienlijke diurnale fluctuaties van hypocretine-1 niveaus het diagnostische proces van NT1 zou kunnen compliceren, wat een van de redenen is waarom dit fenomeen verder onderzoek vereist voordat implementatie in de klinische praktijk raadzaam is. Een vergelijkende studie waarin zowel RIA als LC-MS bij mensen worden toegepast, zou een duidelijker beeld geven van hun respectieve capaciteiten om diurnale fluctuaties van hypocretine-1 in liquor te meten. Bovendien zou zo'n studie de mogelijkheid bieden om de omvang en timing van deze schommelingen (gemeten met verschillende technieken) in een grotere steekproef te bepalen, wat zou kunnen leiden tot verdere inzichten in de rol en functie van hypocretine-1. Voordat de toepassing van LC-MS in de klinische praktijk kan worden overwogen, zouden alle specifieke factoren die de variabiliteit ervan beïnvloeden volledig moeten worden begrepen. Dit inzicht zal een correcte interpretatie van resultaten mogelijk maken en zou ook kunnen leiden tot gestandaardiseerde meetmethodes. Bovendien moet er een geschikte afkapwaarde voor LC-MS hypocretine-1 in liquor worden gedefinieerd voordat klinische toepassing mogelijk is. Daarom is het noodzakelijk dat, ondanks de voordelen van LC-MS, meer onderzoek gedaan wordt voordat deze methode klinisch kan worden toegepast. Verder is de LC-MS methode relatief duur in vergelijking met RIA, wat een grootschalige klinische implementatie zou kunnen belemmeren. Ondanks de beperkingen blijft de RIA de meest betrouwbare methode om hypocretine-1-deficiëntie in liquor te bepalen in het kader van de diagnose van NT1. Daarom stellen we ook voor dat de inter-assay variabiliteit van

hypocretine-1 metingen in liquor bij lage en hoge concentratieniveaus met behulp van RIA onderzocht moet worden, wat tot nu toe niet is gebeurd.

Toepasbaarheid van de 'Sustained Attention to Response Taks' (SART) bij hypersomnolentie: ervaring en resultaten van een teriair verwijscentrum

We zochten naar aanvullende hulpmiddelen om de impact van centrale stoornissen van hypersomnolentie beter te begrijpen en te monitoren en richtten ons op een onderbelicht aspect van deze stoornissen: vigilantie. In **Hoofdstuk 4** hebben we de toepasbaarheid van de 'Sustained Attention to Response Task' (SART) in de klinische praktijk van een tertiair verwijscentrum voor hypersomnolentie stoornissen onderzocht. Zowel de uitkomsten van personen die reguliere diagnostische procedures ondergingen als degenen die voor een rijbewijskeuring kwamen, werden onderzocht. Associaties tussen SART-resultaten en PSG, MSLT en ESS-resultaten werden beoordeeld. Er is geen universeel geaccepteerde afkapwaarde voor de SART, toch tonen we aan dat de SART verminderde vigilantie kan detecteren op basis van een eerder vastgestelde afkapwaarde van Fronczek et al. (2006) [25]. Echter, de SART kan geen onderscheid maken tussen verschillende hypersomnolentie stoornissen na correctie voor confounders. De SART meet een ander aspect van EDS in vergelijking met de MSLT en PSG, wordt beïnvloed door het tijdstip van de dag, en lijkt niet beïnvloed te worden door symptomen van depressie. Individuen met verhoogde angst vertoonden een neiging om meer omissiefouten te maken. Bovendien werd gevonden dat individuen die een rijbewijskeuring ondergingen minder fouten maakten, mogelijk doordat ze medicamenteus behandeld werden en/of doordat ze meer gemotiveerd waren. Uiteraard is de reden voor de betere prestaties van de rijbewijskeuringsgroep moeilijk te interpreteren omdat er meerdere verschillen zijn tussen de twee groepen en er mogelijk sprake was van een selectiebias (bijv. ernstigere aangedane personen waren mogelijk minder geneigd om een dergelijke keuring te ondergaan). Op basis van de resultaten moeten leeftijd en tijdstip van de dag worden overwogen bij het interpreteren van SART-resultaten.

Barrières voor de beoordeling van alertheid in de klinische praktijk

De resultaten bieden waardevolle inzichten in de toepassing van de SART in de klinische praktijk van een tertiair slaap-waakcentrum. Echter, verder onderzoek is nodig om definitieve conclusies te trekken over de toepasbaarheid van de SART als monitoringsinstrument. Dit

onderzoek zou een meer diverse onderzoekspopulatie moeten omvatten, bestaande uit zowel gezonde personen als personen gediagnosticeerd met hypersomnolentie aandoeningen. Bovendien is het belangrijk om een definitieve afkapwaarde voor de SART te bepalen in een grote, diverse populatie die zowel gezonde personen als personen met verschillende hypersomnolentie aandoeningen omvat. Het effect van leeftijd en tijdstip van de dag moet verder worden onderzocht, ook als het gaat om afkapwaarden, gezien onze resultaten die aangeven dat leeftijd invloed heeft op SART uitkomsten.

Het is duidelijk dat verminderde vigilantie in de context van centrale stoornissen van hypersomnolentie een onderbelicht probleem is. Het is echter nog steeds onduidelijk wat de beste methode is om deze vigilantietekorten te meten. Veel studies gebruiken bijvoorbeeld de 'Psychomotor Vigilance Task' (PVT) in plaats van de SART [26-29]. Het vergelijken van studies wordt uitdagender wanneer de gebruikte methoden erg van elkaar verschillen. Vergelijkende studies waarbij alternatieve meetmethoden worden vergeleken met de SART dienen te worden uitgevoerd. Hierdoor kunnen de respectievelijke doeltreffendheid en betrouwbaarheid van deze testen worden bepaald. In reviews zijn verschillende meetmethodes wel met elkaar vergeleken op basis van artikelen waarin een enkele meetmethode werd onderzocht, maar het ontbreekt aan directe vergelijkingen in klinische onderzoeken [30]. Verder onderzoek op dit gebied is dus nodig.

Deel II: Slaap in de kritieke zorgomgeving

Het verkennen van slaapkwaliteit, slaapduur en gerelateerde factoren op de intensive en medium care afdelingen

Hoofdstuk 5 belicht de prevalentie en negatieve impact van slaapstoornissen bij personen die zijn opgenomen op de Intensive Care (IC) en Medium Care (MC). De schadelijke effecten van slechte slaap op het herstel van patiënten en de associatie met het ontstaan van delier zijn eerder al vastgesteld [31-33]. Delier is een aandoening die vaak leidt tot langere ziekenhuisopnames en verhoogde morbiditeit en mortaliteit [31]. Bovendien hebben dierstudies zelfs een direct causaal verband aangetoond tussen slaapdeprivatie en mortaliteit [34, 35]. Dit roept de vraag op of slaapverstoring de mortaliteit op de IC zou kunnen beïnvloeden, zelfs wanneer gecorrigeerd voor de aanwezigheid van delier. Gezien de bekende negatieve fysieke en mentale effecten van verstoorde slaap, kan het identificeren van factoren geassocieerd met slaapverstoring helpen om negatieve effecten te voorkomen door die specifieke factoren aan te pakken.
We vonden een korte slaapduur gedurende de nacht (van 22.00 uur tot 06.00 uur) van slechts 4,6 uur. De slaapkwaliteit was over het algemeen slecht, met een gemiddelde score van 4,9 uit 10 volgens de door verpleegkundigen gerapporteerde Richards-Campbell Sleep Questionnaire (RCSQ). Een slechte slaapkwaliteit of -duur werd geassocieerd met het gebruik van benzodiazepines, delier en nachtelijke overplaatsingen.

Slaap en delier

De gemiddelde dag-nacht slaapverhouding tijdens de drie dagen voor het begin van een delier was opmerkelijk verschillend (met meer slaap overdag) in vergelijking met het gemiddelde van het totale verblijf op de IC van personen die niet delirant waren geweest. De RCSQ-score was ook lager, maar dit verschil was niet significant. Slaapstoornissen worden momenteel niet vermeld als een risicofactor voor het ontwikkelen van een delier in de klinische richtlijnen voor de behandeling van pijn, agitatie en delier uit 2018 [36]. Desondanks zijn er aanwijzingen die suggereren dat bestaande slaapstoornissen waarschijnlijk geassocieerd zijn met het ontstaan van een delier [37-39]. Bovendien delen delier en slaapdeprivatie vergelijkbare symptomen, risicofactoren en worden ze geacht geassocieerd te zijn met een disfunctie van dezelfde centrale zenuwstelselgebieden (prefrontale cortex, thalamus en achterste pariëtale cortex) [40], wat

verder wijst op een verband. Tot nu toe zijn deze verbanden voornamelijk onderzocht in postoperatieve afdelingen en zorginstellingen niet gericht op intensieve zorg. Bovendien richten studies zich meestal op slaap na het begin van een delier, niet op slaap voorafgaand aan een delier. Slechts enkele studies zijn ontworpen met als doel de potentie van slaapstoornissen als risicofactor voor de ontwikkeling van een delier op de IC te bestuderen. Helton et al. (1980) ontdekten dat personen die slaaptekort hadden meer kans hadden op veranderingen in mentale status en Ángeles-Castellanos et al. (2016) ontdekten dat het gemiddelde melatonineniveau afnam en het melatonineritme verloren ging drie dagen voordat delier werd vastgesteld [41]. Tevens toonden twee studies die IC-brede interventies implementeerden om de slaap te verbeteren een significante toename van delier-vrije dagen [42, 43]. Deze studies ondersteunen onze bevindingen, hoewel er aanvullend onderzoek nodig is om het bi-directionele verband tussen verstoorde slaap en delier beter te begrijpen.

Een prospectief onderzoek waarbij een combinatie van 24-uurs PSG, de RCSQ en delier-screening wordt toegepast, zou ideaal zijn om de aard van de link tussen slaapverstoring en delier te bepalen. Het gebruik van PSG in een grote steekproef is over het algemeen niet haalbaar vanwege de aanzienlijke belasting die dit legt op patiënten en het aanzienlijke werk dat ermee gepaard gaat, maar zou kunnen worden uitgevoerd in een kleinere subset.

Benzodiazepines

Benzodiazepines werden vaak gebruikt in onze studie, waarbij 56% van de patiënten gedurende één of meer nachten tijdens hun IC-verblijf een niet-continue (d.w.z. niet gebruikt voor sedatieve doeleinden) benzodiazepine kreeg. Het starten van benzodiazepines als hypnotica resulteerde niet in verbeterde slaapkwaliteit of -duur. Voor deze bevindingen ontbreekt duidelijke ondersteuning of tegenspraak binnen de bestaande wetenschappelijke literatuur. Hoewel benzodiazepines vaak worden voorgeschreven als hypnotica, is er beperkt bewijs over hun werkzaamheid om de slaap op de IC te verbeteren. Verschillende studies hebben echter de effecten van benzodiazepines op slaap bij andere populaties onderzocht. Benzodiazepinen hebben de neiging om de inslaaptijd te verkorten en de totale slaaptijd te verhogen, maar hebben een negatief effect op de slaapkwaliteit [44, 45]. Op de IC zijn benzodiazepines geassocieerd met een verhoogd risico op het ontwikkelen van delier en langere IC-opnameduur [46, 47]. Gezien deze factoren moeten benzodiazepines met voorzichtigheid worden voorgeschreven en moeten er effectievere alternatieven met minder schadelijke bijwerkingen worden gevonden.

Hoofdstuk 5 biedt verdere gronden voor uitgebreider onderzoek naar de werkzaamheid van benzodiazepines op de IC. Bovendien moeten er studies worden uitgevoerd om niet-farmacologische en niet-benzodiazepine-interventies te verkennen en te vergelijken. Aangezien frequent gebruikte medicijnen op de IC, zoals benzodiazepines en propofol, kunnen leiden tot een suboptimale slaaparchitectuur, zou het interessant zijn om te onderzoeken of natriumoxybaat, een medicijn dat bekend staat om het verbeteren van de slaaparchitectuur bij narcolepsiepatiënten, ook een gunstig effect heeft in de IC-omgeving. Dexmedetomidine is een ander medicijn dat veelbelovend is en verder onderzoek vereist in de vorm van een grote gerandomiseerde gecontroleerde studie [48, 49].

Slaap en mortaliteit

De mortaliteit één jaar na ontslag van de IC was 30%. Hogere leeftijd, grotere ziekte-ernst en het vrouwelijk geslacht werden geassocieerd met verhoogde mortaliteit na ontslag. De associatie tussen verminderde slaapkwaliteit en verhoogde mortaliteit bereikte geen statistische significantie ($p = 0,070$), maar moet worden beschouwd als een mogelijke associatie die niet volledig kan worden uitgesloten. In tegenstelling tot bevindingen uit eerdere studies, was de associatie tussen het hebben van een delier en verhoogde mortaliteit niet significant in onze studie [50]. Mogelijke verklaringen zijn dat sommige gevallen van hypoactief delier wellicht niet zijn herkend en dat gevallen van delier werden gegroepeerd ongeacht de ernst of duur.

Droomervaringen tijdens een verblijf op de Intensive Care: prevalentie, inhoud, levendigheid en geassocieerde factoren

Verslagen van verpleegkundigen die nazorg verlenen op de IC hebben geleid tot een onderzoek naar levendige droomervaringen op de IC aangemoedigd. De uitkomsten worden gerapporteerd in **Hoofdstuk 6**. Onze bevindingen toonden een hoge prevalentie van deze ervaringen aan, 79% van de deelnemers rapporteerde ze namelijk. De dromen waren niet alleen intens, maar hadden ook vaak negatieve emotionele associaties voor de deelnemers. Deelnemers meldden dat ze zowel tijdens hun verblijf als zelfs na ontslag werden beïnvloed door deze droomervaringen. De inhoud was vaak gerelateerd aan het verblijf op de IC en bestond frequent uit thema's van hulpeloosheid en de dood. We ontdekten dat het zowel voor deelnemers als onderzoekers uitdagend was om droomervaringen te onderscheiden van hallucinaties.

Limitaties verminderen en impact beoordelen

Bij toekomstig onderzoek is het essentieel om de beperkingen van de onderzoeksopzet van het onderzoek in **Hoofdstuk 6** te overwinnen. De potentiële selectie- en herinneringsbias zou deels kunnen worden vermeden door prospectief IC-patiënten tijdens hun opname op te nemen en interviews op verschillende tijdstippen tijdens hun verblijf op de IC en daarna uit te voeren. Deze aanpak maakt ook verder onderzoek mogelijk naar de emotionele en psychologische impact van droomervaringen op patiënten tijdens en na opname. Bovendien kan een prospectieve onderzoeksopzet helpen bij een duidelijkere differentiatie tussen waarnemingsstoornissen, hallucinaties en droomervaringen. Het zou gemakkelijker zijn om het optreden van droomervaringen te relateren aan specifieke factoren wanneer het tijdstip van optreden bekend is. Sommige van de suggesties die door deelnemers worden gegeven, zijn interessant om verder te verkennen, zoals of het informeren van patiënten die meer dan een paar dagen zijn opgenomen over de mogelijkheid om levendige dromen te ervaren enig gunstig effect heeft op hun welzijn tijdens en na het verblijf op de IC. Ten slotte zou het onderzoeken van het optreden van levendige droomervaringen in locaties buiten de IC, zoals de algemene afdeling, ook kunnen helpen om te bepalen of dit een specifiek IC-fenomeen is. Dergelijk onderzoek zou kunnen bijdragen aan een beter begrip van de onderliggende redenen voor levendige droomervaringen op de IC.

Algemene conclusie

Slaap vormt een fundamenteel aspect van het menselijk leven, waarvan verstoring een negatieve invloed heeft op zowel rustperiodes als waakzaamheid, een feit dat goed bekend is bij degenen die zelf met slaapverstoring te maken hebben gehad. Bij centrale stoornissen van hypersomnolentie en binnen de IC-omgeving vervagen de grenzen tussen slaap en waak. Dit fenomeen wordt geïllustreerd door cognitieve klachten bij centrale stoornissen van hypersomnolentie en verstoringen van het circadiane ritme, evenals levendige dromen op de IC. Inzichten verkregen in slaap in de ene context kunnen helpen om ook de andere context te verduidelijken. De resultaten van studies die in dit proefschrift worden verkend, benadrukken de kansen zowel in het veld van de intensieve zorg als in centrale stoornissen van hypersomnolentie. De noodzaak om diagnostische criteria en beoordelingsmethoden te verfijnen wordt ook benadrukt. Door gebieden voor toekomstig onderzoek aan te wijzen, legt dit proefschrift de basis voor meer nauwkeurige diagnostische instrumenten, een dieper begrip van de complexe relatie tussen slaap-waak en algeheel welzijn, met als uiteindelijk doel betere preventie of behandeling en daarmee betere (patiënten)zorg.

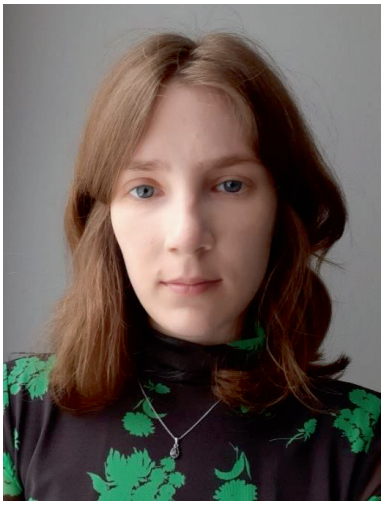
Referenties

1. Mignot, E., et al., *The Role of Cerebrospinal Fluid Hypocretin Measurement in the Diagnosis of Narcolepsy and Other Hypersomnias.* Archives of Neurology, 2002. **59**(10): p. 1553-1562.
2. Postiglione, E., et al., *Narcolepsy with intermediate cerebrospinal level of hypocretin-1.* Sleep, 2022. **45**(2).
3. Lopez, R., et al., *Rapid eye movement sleep duration during the multiple sleep latency test to diagnose hypocretin-deficient narcolepsy.* Sleep, 2023. **46**(1).
4. Miano, S., et al., *A series of 7 cases of patients with narcolepsy with hypocretin deficiency without HLA DQB1*06:02 allele.* J Clin Sleep Med, 2023.
5. Torstensen, E.W., et al., *Repeated measures of hypocretin-1 in Danish and Italian patients with narcolepsy and in controls.* Sleep Medicine, 2023. **101**: p. 213-220.
6. Andlauer, O., et al., *Predictors of hypocretin (orexin) deficiency in narcolepsy without cataplexy.* Sleep, 2012. **35**(9): p. 1247-55F.
7. Trotti, L.M., B.A. Staab, and D.B. Rye, *Test-retest reliability of the multiple sleep latency test in narcolepsy without cataplexy and idiopathic hypersomnia.* J Clin Sleep Med, 2013. **9**(8): p. 789-95.
8. Ruoff, C., et al., *The MSLT is Repeatable in Narcolepsy Type 1 But Not Narcolepsy Type 2: A Retrospective Patient Study.* J Clin Sleep Med, 2018. **14**(1): p. 65-74.
9. Lammers, G.J. and J.G. van Dijk, *The multiple sleep latency test: a paradoxical test?* Clinical Neurology and Neurosurgery, 1992. **94**: p. 108-110.
10. Gool, J.K., et al., *Data-Driven Phenotyping of Central Disorders of Hypersomnolence With Unsupervised Clustering.* Neurology, 2022. **98**(23): p. e2387-e2400.
11. Pizza, F., et al., *Different sleep onset criteria at the multiple sleep latency test (MSLT): an additional marker to differentiate central nervous system (CNS) hypersomnias.* J Sleep Res, 2011. **20**(1 Pt 2): p. 250-6.
12. Andlauer, O., et al., *Nocturnal rapid eye movement sleep latency for identifying patients with narcolepsy/hypocretin deficiency.* JAMA Neurol, 2013. **70**(7): p. 891-902.
13. Bassetti, C., M. Billiard, and E. Mignot, *Narcolepsy and hypersomnia*. 2006: Informa Health Care.
14. Lindström, M., et al., *Orexin-A measurement in narcolepsy: A stability study and a comparison of LC-MS/MS and immunoassays.* Clin Biochem, 2021. **90**: p. 34-39.
15. Ono, T., et al., *Measurement of cerebrospinal fluid orexin-A (hypocretin-1) by enzyme-linked immunosorbent assay: A comparison with radioimmunoassay.* Psychiatry Clin Neurosci, 2018. **72**(11): p. 849-850.
16. Sarkanen, T., et al., *Misdiagnosis of narcolepsy caused by a false-positive orexin-A/hypocretin-1 enzyme immune assay.* J Clin Sleep Med, 2022. **18**(8): p. 2075-2078.
17. Bårdsen, K., et al., *Considerably Lower Levels of Hypocretin-1 in Cerebrospinal Fluid Is Revealed by a Novel Mass Spectrometry Method Compared with Standard Radioimmunoassay.* Anal Chem, 2019. **91**(14): p. 9323-9329.
18. Hirtz, C., et al., *From radioimmunoassay to mass spectrometry: a new method to quantify orexin-A (hypocretin-1) in cerebrospinal fluid.* Sci Rep, 2016. **6**: p. 25162.
19. Narita, N., et al., *Diurnal Fluctuations of Orexin-A and -B in Cynomolgus Monkey Cerebrospinal Fluid Determined by a Novel Analytical Method Using Antiadsorptive Additive Treatment Followed by Nanoflow Liquid Chromatography-High-Resolution Mass Spectrometry.* ACS Chem Neurosci, 2023. **14**(4): p. 609-618.
20. Grady, S.P., et al., *Diurnal Variation in CSF Orexin-A in Healthy Male Subjects.* Sleep, 2006. **29**(3): p. 295-297.
21. Salomon, R.M., et al., *Diurnal variation of cerebrospinal fluid hypocretin-1 (Orexin-A) levels in control and depressed subjects.* Biological Psychiatry, 2003. **54**(2): p. 96-104.
22. Yoshida, Y., et al., *Fluctuation of extracellular hypocretin-1 (orexin A) levels in the rat in relation to the light–dark cycle and sleep–wake activities.* European Journal of Neuroscience, 2001. **14**(7): p. 1075-1081.
23. Zeitzer, J.M., et al., *Circadian and homeostatic regulation of hypocretin in a primate model: implications for the consolidation of wakefulness.* J Neurosci, 2003. **23**(8): p. 3555-60.
24. Sakai, N., et al., *HPLC analysis of CSF hypocretin-1 in type 1 and 2 narcolepsy.* Sci Rep, 2019. **9**(1): p. 477.
25. Fronczek, R., et al., *Focusing on vigilance instead of sleepiness in the assessment of narcolepsy: high sensitivity of the Sustained Attention to Response Task (SART).* Sleep, 2006. **29**(2): p. 187-91.
26. Thomann, J., et al., *Psychomotor vigilance task demonstrates impaired vigilance in disorders with excessive daytime sleepiness.* J Clin Sleep Med, 2014. **10**(9): p. 1019-24.

27. Lee, I.-S., et al., *The Relationship between psychomotor vigilance performance and quality of life in obstructive sleep apnea.* Journal of clinical sleep medicine : JCSM : official publication of the American Academy of Sleep Medicine, 2011. **7**(3): p. 254-260.
28. Trotti, L.M., et al., *The Psychomotor Vigilance Test as a measure of alertness and sleep inertia in people with central disorders of hypersomnolence.* J Clin Sleep Med, 2022. **18**(5): p. 1395-1403.
29. Mathis, J., et al., *The diagnostic value of sleep and vigilance tests in central disorders of hypersomnolence.* Sleep, 2022. **45**(3).
30. Filardi, M., et al., *Cognitive dysfunction in central disorders of hypersomnolence: A systematic review.* Sleep Med Rev, 2021. **59**: p. 101510.
31. Stuck, A., M.J. Clark, and C.D. Connelly, *Preventing intensive care unit delirium: a patient-centered approach to reducing sleep disruption.* Dimens Crit Care Nurs, 2011. **30**(6): p. 315-20.
32. Pisani, M.A., et al., *Sleep in the intensive care unit.* Am J Respir Crit Care Med, 2015. **191**(7): p. 731-8.
33. Pulak, L.M. and L. Jensen, *Sleep in the Intensive Care Unit:A Review.* Journal of Intensive Care Medicine, 2016. **31**(1): p. 14-23.
34. Friese, R.S., B. Bruns, and C.M. Sinton, *Sleep deprivation after septic insult increases mortality independent of age.* J Trauma, 2009. **66**(1): p. 50-4.
35. Maragno-Correa, J.M., et al., *Sleep deprivation increases mortality in female mice bearing Ehrlich ascitic tumor.* Neuroimmunomodulation, 2013. **20**(3): p. 134-40.
36. Devlin, J.W., et al., *Clinical Practice Guidelines for the Prevention and Management of Pain, Agitation/Sedation, Delirium, Immobility, and Sleep Disruption in Adult Patients in the ICU.* Crit Care Med, 2018. **46**(9): p. e825-e873.
37. Leung, J.M., et al., *Preoperative Sleep Disruption and Postoperative Delirium.* J Clin Sleep Med, 2015. **11**(8): p. 907-13.
38. Oldham, M.A., et al., *Baseline sleep as a predictor of delirium after surgical aortic valve replacement: A feasibility study.* Gen Hosp Psychiatry, 2021. **71**: p. 43-46.
39. Jaiswal, S.J., et al., *Objectively measured sleep fragmentation is associated with incident delirium in older hospitalized patients: Analysis of data collected from an randomized controlled trial.* J Sleep Res, 2021. **30**(3): p. e13205.
40. Weinhouse, G.L., et al., *Bench-to-bedside review: Delirium in ICU patients - importance of sleep deprivation.* Critical Care, 2009. **13**(6): p. 234.
41. Ángeles-Castellanos, M., et al., *Loss of melatonin daily rhythmicity is asociated with delirium development in hospitalized older adults.* Sleep Sci, 2016. **9**(4): p. 285-288.
42. Kamdar, B.B., et al., *The effect of a quality improvement intervention on perceived sleep quality and cognition in a medical ICU.* Crit Care Med, 2013. **41**(3): p. 800-9.
43. Tonna, J.E., et al., *The Effect of a Quality Improvement Intervention on Sleep and Delirium in Critically Ill Patients in a Surgical ICU.* Chest, 2021. **160**(3): p. 899-908.
44. Achermann, P. and A.A. Borbély, *Dynamics of EEG slow wave activity during physiological sleep and after administration of benzodiazepine hypnotics.* Hum Neurobiol, 1987. **6**(3): p. 203-10.
45. Monti, J.M., *Sleep laboratory and clinical studies of the effects of triazolam, flunitrazepam and flurazepam in insomniac patients.* Methods Find Exp Clin Pharmacol, 1981. **3**(5): p. 303-26.
46. Wang, H., et al., *Sedative drugs used for mechanically ventilated patients in intensive care units: a systematic review and network meta-analysis.* Curr Med Res Opin, 2019. **35**(3): p. 435-446.
47. Fraser, G.L., et al., *Benzodiazepine versus nonbenzodiazepine-based sedation for mechanically ventilated, critically ill adults: a systematic review and meta-analysis of randomized trials.* Crit Care Med, 2013. **41**(9 Suppl 1): p. S30-8.
48. Sun, Y.M., et al., *Effect of low-dose dexmedetomidine on sleep quality in postoperative patients with mechanical ventilation in the intensive care unit: A pilot randomized trial.* Front Med (Lausanne), 2022. **9**: p. 931084.
49. Alexopoulou, C., et al., *Effects of dexmedetomidine on sleep quality in critically ill patients: a pilot study.* Anesthesiology, 2014. **121**(4): p. 801-7.
50. Aung Thein, M.Z., et al., *A call to action for delirium research: Meta-analysis and regression of delirium associated mortality.* BMC Geriatrics, 2020. **20**(1): p. 325.

B Curriculum Vitae

Adrienne van der Hoeven was born on the 23rd of April 1992 in Haarlem together with her twin brother Ernst. Her parents are Els Nolting and Pieter van der Hoeven. She also has an older brother, Siebrand.
After graduating cum laude from the Eerste Christelijk Lyceum in Haarlem in 2010 she studied Medicine at the University of Amsterdam. In 2013, during the waiting period between her bachelor's and master's degrees in Medicine, she also obtained a propaedeutic degree in Economics and Business Administration. As a student she participated in the student team of UNICEF Amsterdam.
During her medical internships she developed a passion for neurology and, post-graduation in 2018, she started to work as a resident not in training at the Neurology department of the Noordwest hospital in Alkmaar.
She had long desired to delve into scientific research and her interest in neurology and sleep culminated in her admission to a PhD program on narcolepsy and sleep in 2019. Studies were performed at the Sleep-Wake Center of Stichting Epilepsie Instellingen Nederland (SEIN), the Department of Neurology and in the critical care environment of the Leiden University Medical Center (LUMC). During this period, she won a Young Scientist Award at the European Narcolepsy Days in 2023, secured 2 travel grants, and was awarded a Young Talent Fund from the Dutch Society for Sleep-Wake Research (NSWO) in 2022.
Currently, she works as resident not in training at the Neurology department of the Alrijne hospital in Leiderdorp.



C List of publications and articles

van der Hoeven AE, Bijlenga D, van der Hoeven E, Schinkelshoek MS, Hiemstra FW, Kervezee L, van Westerloo DJ, Fronczek R, Lammers GJ. Sleep in the intensive and intermediate care units: Exploring related factors of delirium, benzodiazepine use and mortality. Intensive Crit Care Nurs. 2024 Apr;81:103603.

van der Hoeven AE, Bijlenga D, Bouhuijs P, van Schie MKM, Lammers GJ, Fronczek R. Applicability of the Sustained Attention to Response Task (SART) in hypersomnolence: Experience and results from a tertiary referral center. Sleep Med. 2023 Aug;108:105-113.

Boertien TM, Van Someren EJW, Coumou AD, van den Broek AK, Klunder JH, Wong WY, van der Hoeven AE, Drent ML, Romijn JA, Fliers E, Bisschop PH. Compression of the optic chiasm is associated with reduced photoentrainment of the central biological clock. Eur J Endocrinol. 2022 Nov 4;187(6):809-821.

van der Hoeven AE, van Waaij K, Bijlenga D, Roelandse FWC, Overeem S, Bakker JA, Fronczek R, Lammers GJ. Hypocretin-1 measurements in cerebrospinal fluid using radioimmunoassay: within and between assay reliability and limit of quantification. Sleep. 2022 Jul 11;45(7):zsac103.

van der Hoeven AE, van Waaij K, Bijlenga D, Roelandse FWC, Overeem S, Bakker JA, Fronczek R, Lammers GJ. Hypocretin-1 measurements in cerebrospinal fluid using radioimmunoassay: within and between assay reliability and limit of quantification. Sleep. 2022 Jul 11;45(7):zsac103.

van der Hoeven AE. *Tijd voor slaap: Slaap bevordert herstel patiënten*. Cahier Biowetenschappen en Maatschappij. 2021 March: 48-49

van der Hoeven AE. *Hartlopend: Slapen op de IC*. Cordiaal. 2020, October: 135

D Dankwoord

Ten eerste wil ik natuurlijk de onderzoeksdeelnemers bedanken. Mede dankzij hen heb ik dit proefschrift kunnen schrijven.
Gert Jan en Rolf. Wat een fijne begeleiding heb ik van jullie gehad. COVID kwam net om de hoek kijken toen ik begon, waardoor we elkaar in het begin maar heel weinig in het echt zagen. Ik kon het erg waarderen dat jullie toen je best deden om me ook buiten het werk te leren kennen, bijvoorbeeld door me toen uit te nodigen voor een etentje. Inmiddels zijn we natuurlijk al vele congressen, diners en escaperoom-avonturen verder. En Denise, wij leerden elkaar pas wat later kennen. Dank je voor je goede begeleiding, al je mooie tekstuele aanpassingen en (last but not least) het meermaals organiseren van de eerder genoemde uitjes.
En dan de andere slaappromovendi. Jari, Jingru, Josephine, bedankt voor alle leuke (brainstorm)momenten. Marieke en Veronique, ook jullie bedankt voor de mooie tijd samen, we hebben altijd veel lol gehad en daarnaast kon ik bij jullie ook af en toe mijn hart luchten. Ik had me geen betere paranimfen kunnen wensen.
Mink, dank voor inwerken en ook het samenwerken later. Ik heb veel van je kunnen leren. Mojca, top dat ik zoveel gebruik mocht maken van je SART expertise.
Verder wil ik graag het hele Slaap-Waak team bedanken, wat een positieve en fijne werkomgeving! Het is inspirerend om te zien hoe iedereen in het team zijn eigen unieke bijdrage levert en zo de gezamenlijke doelen bereikt. Pauline, dank je voor je goede adviezen.
Naast Slaap-Waak is er natuurlijk ook de Epilepsiegroep. Hanneke, Silvano, Anouk, Trusjen, we onderzochten misschien iets anders, maar konden altijd samenkomen tijdens een koffiepauze of borrel om ervaringen te delen, zowel de mooie als de minder mooie momenten.
Ook wil ik graag mijn waardering uitspreken voor de collega's van het LUMC. Jaap, zonder jou was het RIA artikel niet tot stand gekomen. David, Laura en Floor, bedankt voor jullie kennis en ondersteuning op IC gebied! Paul en de andere KNF laboranten, hartelijk dank voor de prettige samenwerking. En dan mijn oud-kamergenoten – Yvonne, Nienke, Karin, Annabel, Michel, Nadine - en ook de overburen, Maryam, Esther en Linda: fijn dat we naast een ruimte ook handige PhD tips konden delen.
Het was een voorrecht om met zoveel getalenteerde studenten te mogen samenwerken. Jullie inzet en betrokkenheid hebben echt een verschil gemaakt bij het werk wat we deden en hebben daarnaast ook nog eens hele mooie bachelor- en mastertheses opgeleverd.
Ten slotte wil ik mijn ouders en broers, Siebrand en Ernst, bedanken. Jullie hebben me altijd ondersteund en dat waardeer ik enorm. Anton, jij ook bedankt voor al je steun en specifiek ook voor je hulp bij de lay-out van dit boekje. Ik kan niet genoeg benadrukken hoeveel het voor me betekent om jullie allemaal in mijn hoek te hebben.