Dim-Light Melatonin Onset as a Circadian Marker in Children and Adolescents: Quantification of Melatonin Dynamics

Joanne May Ramil

A Thesis
In the Department of
School of Graduate Studies

Presented in Partial Fulfillment of the
Requirements for the Degree of Master of Science
(Individualised Program, Pure Sciences)
at Concordia University
Montreal, Quebec, Canada

August 2022

CONCORDIA UNIVERSITY
School of Graduate Studies

This is to certify that the thesis prepared

By: Joanne May Ramil
Entitled: Dim-Light Melatonin Onset as a Circadian Marker in Children and Adolescents: Quantification of Melatonin Dynamics

and submitted in partial fulfillment of the requirements for the degree of

Master of Science (Individualised Program – Pure Sciences)

complies with the regulations of the University and meets the accepted standards with respect to originality and quality.

Signed by the final Examining Committee:

Dr. Nizar Bouguila — Examiner

Dr. Lisa Kakinami — Examiner

Dr. Rex Kline — Examiner

Dr. Jennifer McGrath — Thesis Supervisor

Approved by: Dr. Felice Yuen, Graduate Program Director

11 / Aug /2022
dd / mmm / yyyy

Dr. Effrosyni Diamantoudi, Interim Dean of Graduate Studies

Abstract

Dim-Light Melatonin Onset as a Circadian Marker in Children and Adolescents: Quantification of Melatonin Dynamics

Joanne May Ramil

Quantifying melatonin, specifically dim-light melatonin onset (DLMO), in pediatric populations allows us to better understand the implications of circadian rhythm on health. However, current melatonin metrics were developed for adult populations and melatonin collection is typically time and resource intensive. The first goal of this thesis was to identify melatonin metrics traditionally used in pediatric literature and to examine alternative metrics. Data extracted from twenty-two articles identified via scoping review revealed that the most common DLMO metrics focused on a single melatonin dynamic: timing (i.e., the time melatonin concentration surpasses a predetermined threshold). Thus, alternative metrics to quantify timing, concentration, and rate dynamics, such as total melatonin output and rate derived by linear regression, were proposed, and applied to a pediatric secondary dataset. Traditional and alternative timing (r_{avg}=.86) and concentration (r=.93) metrics where highly correlated, while rate metrics were moderately correlated (r=.38). The second goal was to explore potential proxy variables to estimate DLMO. This was achieved by augmenting the secondary dataset using multiple imputation, then by conducting a Classification and Regression Tree analysis to estimate DLMO dynamics. Sleep midpoint and bedtime measured by actigraphy were some of the best proxies for the timing dynamic (R^2=.91) but not for concentration (R^2=.50) and rate (R^2=.41). Thus, other proxies to estimate concentration and rate dynamics, such as temperature, diet, or race, should be explored. Results of this thesis highlight the lack of melatonin metrics for youth populations in the literature and the need to better quantify melatonin and circadian rhythms.

Acknowledgements

This thesis would not have been possible without the help of many people. I would first like to thank Dr. Jennifer McGrath for supervising my thesis; this project would not have been possible without her support, guidance, and encouragements. I could not have asked for a better supervisor. I would also like to thank the members of my supervisory committee, Drs. Nizar Bouguila, Lisa Kakinami, and Rex Kline, for their knowledge and expertise. Additionally, I would like to thank the Fonds de Recherche du Québec Santé (FRQS) for funding my master's degree. I would also like to thank the Pediatric Public Health and Psychology Lab members for collecting and preparing the data used in this thesis, and for providing insightful feedback and advice throughout my thesis. Next, I am grateful for my friends and family who provided emotional support throughout my degree. Finally, a special thank you to my dog, Milo, who sat next to me and kept me company while I wrote for hours and whose barks and nudges always reminded me when it was time to take a break and step away from the computer.

Table of Contents

List of Figures

List of Tables

Thesis Abbreviations & Terminology

Abbreviations

CART: Classification and Regression Tree
DLMO: Dim-light melatonin onset
MESC: Morningness-Eveningness Scale for Children
RMSE: Root mean squared error
SCN: Suprachiasmic nuclei

Common Terminology

Dynamic(s): dimensions of melatonin quantification (i.e., timing, concentration, rate)
Metric(s): specific way a dynamic is quantified, derived from a biomathematical model
Sleep assessment methodology: all instruments (actigraphy, sleep diary, questionnaires) used to measure specific bedtime routine variables (i.e., bedtime and waketime) and circadian markers (sleep midpoint)

Introduction

Fluctuating levels of melatonin are considered the most accurate marker of circadian rhythms in humans. Specifically, the time at which melatonin begins to rise in the evening prior to bedtime in response to dim light (i.e., dim light melatonin onset DLMO) is considered the gold standard circadian marker. While melatonin has been studied over several decades, few have questioned its quantification or refined how it is used to infer a person's circadian timing. The overarching aim of this thesis was to evaluate different ways melatonin can be quantified from a biomathematical perspective and to explore the utility of proxy measures to more easily capture circadian timing. The following introduction begins with a general overview of the dynamics of the neuroendocrine system and their quantification. Key concepts related to circadian rhythms and melatonin are then presented. Next, the methodology to measure melatonin, criticisms of its quantification, and issues salient to its measurement in children and adolescents are reviewed in greater detail. Finally, alternative markers of the endogenous circadian rhythm are discussed as important referents for comparison with melatonin. The introduction concludes with a brief outline of the series of four focused thesis parts and their unique objectives that were conducted sequentially to address the overarching aim of the thesis.

Dynamics of the Neuroendocrine System and their Quantification

The neuroendocrine system is one of several physiological systems that regulate living organisms, are affected by multiple factors, and fluctuate throughout the lifespan. The neuroendocrine system is a network of complex processes, controlled by the hypothalamus, that regulates homeostasis and the production, distribution, absorption, and excretion of hormones, including melatonin and cortisol (Pandi-Perumal et al., 2007). Hormones are produced by neuroendocrine glands and have dynamic patterns, meaning they fluctuate over time and in response to other events (Zavala et al., 2019). Hormone production occurs when the neuroendocrine gland receives a signal triggered by an external or internal cue. For example, cortisol is produced in response to stressor exposure (i.e., external cue) but also in response to daily circadian timing (i.e., internal cue) to yield a diurnal pattern (Wittenbrink et al., 2018). Similarly, melatonin is produced in response to darkness or dimmed light (i.e., external cue) and in response to daily circadian timing (Claustrat et al., 2005). Once hormones are produced, they are released into the blood stream and eventually bind to receptors (i.e., absorption). After hormones perform their function, they are then excreted. By examining the dynamics of these hormones through their production, distribution, binding/absorption, and excretion, we can better understand how they function, and ultimately understand how they impact an individual's health. Precise and accurate quantification of hormone levels is challenging because these processes are constant, ever changing, and simultaneous.

Advances in mathematics and computer science have led to discoveries in how to better study physiological systems with greater detail and precision to improve their quantification. Biomathematical models informed by top-down and bottom-up approaches offer insights into improved ways of quantifying the dynamics of the neuroendocrine system and hormonal levels. There remains much debate about which mathematical approach is preferable (Robeva et al., 2008). Table 1 depicts general measures to quantify each dynamic process. **Timing** (e.g., 24:00 clock time) quantifies the onset of the increase (or offset of decrease) in concentration levels. Timing can also be anchored to exceeding specific pre-determined concentration thresholds. Pre-determined threshold values are defined as the concentration at which a meaningful change in the biological system occurs. Timing is the simplest way to quantify production and excretion of a

hormone. **Concentration** can be quantified as the total amount of hormone produced within a given time interval. The total concentration is calculated as the area under the curve (AUC). Relatedly, accurately sampling the correct medium is essential to assess distribution (e.g., saliva, urine, blood). Maximum concentration can also be derived as an indicator of the highest concentration observed over the time interval. Concentration reflects multiple, simultaneous processes and can be used as a general marker of production, distribution, absorption, and excretion. **Rate** of change is quantified as how quickly the concentration increases (or decreases) over time. Rate is related to not only production, but also absorption and excretion. Steeper slopes, or faster rates of change, indicate that the hormone is being produced faster than it is being absorbed (or vice-versa). To note, the different ways a dynamic is quantified (e.g., total concentration and maximum concentration as ways to quantify the concentration dynamic) are called **metrics** throughout this thesis. In sum, all three dynamics (timing, concentration, rate) provide unique information about melatonin and can be assessed in different ways to yield various metrics. The fluctuations of these dynamics will be the primary focus of this thesis.

Table 1. Systems Quantification Measures

Process	Quantification Dynamic: Timing	Concentration	Rate
Production	■	■	■
Distribution		■	
Absorption		■	■
Excretion	■	■	■

Circadian Rhythms

Circadian rhythms are characterized by patterns that repeat roughly every 24 hours, albeit with slight variations across persons. Circadian rhythms are maintained by the suprachiasmic nuclei (SCN) in the hypothalamus (Robeva et al., 2008). This master internal clock is a pacemaker that coordinates daily processes such as sleep, body temperature, and hormone secretion (Robeva et al., 2008). In healthy individuals, environmental cues known as zeitgeibers synchronize the SCN rhythm to a 24-hour pattern (Pandi-Perumal et al., 2007). The light/dark cycle is the primary zeitgeiber that keeps the SCN synchronized; other zeitgeibers can include daily events like alarm clocks or behaviors (i.e., activity, meals) that happen at the same time every day (Klerman, 2005). In the absence of external cues, the SCN still generates a circadian pattern, but with a period slightly longer than 24 hours (Pandi-Perumal et al., 2007). It is important to make the conceptual distinction between circadian rhythm, which is governed by the endogenous pacemaker, and diurnal rhythm, which is primarily descriptive and refers to the observed daily pattern or one's behavioral routine (e.g., bedtime, waketime; Klerman, 2005). One of the challenges of understanding the sleep-wake cycle is related to accurately disentangling circadian and diurnal rhythms.

Recently, there is renewed interest in measuring circadian rhythms as they supply valuable information regarding a person's physiological processes (Abbott et al., 2020; Reid, 2019; Wittenbrink et al., 2018). Emerging evidence suggests circadian rhythms can help to optimize treatments for certain illnesses when interventions are correctly timed. For example, studies have demonstrated that certain chemotherapy drugs, at the same dosage, are more efficacious depending on the time of day they are administered, suggesting that the internal clock system impacts their pharmacokinetics and pharmacodynamics (Wittenbrink et al., 2018).

Disruption of circadian rhythms has also been linked to poor health outcomes. This disruption occurs when a person's internal endogenous rhythm is not synchronized to their

external clock and is known as circadian misalignment (Baron & Reid, 2014). Circadian misalignment can occur when a person's sleep schedule is not timed appropriately to match their circadian rhythm. Circadian misalignment has been linked to myriad health problems including cardiovascular disease, diabetes, obesity, cancer, as well as several psychiatric disorders, such as depression and schizophrenia (Baron et al., 2017; Baron & Reid, 2014). Additionally, circadian misalignment is particularly prominent during adolescence, when adolescents' sleep patterns typically shift towards later bedtimes, while maintaining early wake times for school (Dolsen & Harvey, 2018). Ultimately, it is important to accurately measure circadian rhythms to better understand their impact on health.

Melatonin

Melatonin is the primary circadian marker and helps to regulate the timing of multiple physiological functions, such as sleep and the immune response (Praninskienė et al., 2012). The production and release of melatonin is regulated by the SCN (Benloucif et al., 2008). Melatonin is synthesized in the pineal gland from serotonin (Claustrat et al., 2005). Once produced, melatonin is released into the bloodstream and acts as a signal for the sleep-wake cycle, where higher levels promote sleep onset and lower levels promote wake. The secretion of melatonin follows a distinct pattern that reflects one's circadian rhythm (Praninskienė et al., 2012); however, it is also sensitive to the light-dark cycle (i.e., day length). Specifically, melatonin concentrations in the body follow a diurnal pattern and are affected by light, effectively illustrating how the circadian rhythm is anchored to the light-dark cycle (Pandi-Perumal et al., 2007). For this reason, melatonin is considered the gold standard to quantify circadian rhythms, specifically circadian phase. Melatonin follows a 24-hr rhythm, typically rising 2-3 hours prior to sleep onset, peaking during sleep, and returning to baseline levels prior to wake (Molina & Burgess, 2011).

While melatonin relies primarily on a person's internal endogenous rhythm, there are external stimuli that also affect its levels in the body. The production of melatonin is sensitive to light; light suppresses melatonin production (Benloucif et al., 2008). Other external cues also influence melatonin production, albeit to a smaller extent. For instance, foods that contain high levels of tryptophan, such as bananas and chocolate, can increase melatonin production as tryptophan is a precursor of serotonin (Claustrat et al., 2005). Thus, to accurately capture the "true" variation of melatonin, external factors should be controlled for (Benloucif et al., 2008).

Melatonin can be measured via blood, urine, and saliva. Because melatonin is released directly into the blood, plasma melatonin provides the most accurate measurement of circulating melatonin (Claustrat et al., 2005). However, repeated samples are required to create a melatonin profile; thus, collecting blood samples is not always feasible as the procedure is invasive. While melatonin concentrations in urine and saliva are lower than in blood (30 to 40% of plasma concentration), they follow the same pattern (Claustrat et al., 2005). Urine sampling is more practical than blood sampling as it can be more easily collected. While urine collection is less invasive, the timing of the samples becomes an issue as it is more difficult to collect urine on demand at regular intervals (Benloucif et al., 2008). Melatonin collection via saliva provides an ideal balance between feasibility and accuracy: samples can be collected via swabs that users place in their mouth, and this can be easily repeated, so long as the participant remains awake.

As previously stated, neuroendocrine systems and hormones can be quantified within three **dynamics**: timing, concentration, and rate. Within each dynamic, there can be multiple **metrics**, which are the different ways each dynamic is quantified. Melatonin was first quantified in the 1970s (Vaughan et al., 1976). Early measures of melatonin focused on the concentration

dynamic, specifically the metric of maximum melatonin concentration, which was also referred to as the amplitude of melatonin. Maximum concentration can be affected easily by several confounders, such as posture and use of non-steroidal anti-inflammatory drugs (NSAIDS; Lewy et al., 2006). In the 1980s, Lewy et al. (Lewy et al., 1980, 1999) proposed a new methodology and related measures to quantify melatonin. This methodology consisted of collecting melatonin samples every few minutes (e.g., 30-60 minutes) in a dimly lit room for up to 5 hours before and 1 hour after bedtime. Light exposure was controlled by the researchers and several measures of melatonin were taken in response to the light cues: Dim Light Melatonin Onset (DLMO), Synthesis Offset (SynOff), and Dim Light Melatonin Offset (DLMOff). Timing was used to capture each of these measures (i.e., the time of melatonin onset in response to dim light exposure). It is important to emphasize that these three measures only capture the timing dynamic of the melatonin curve; concentration and rate are not reflected in these measures. While all of these measures could be used to infer circadian timing, DLMO was thought to be least affected by confounding agents (Lewy et al., 1999). Thus, melatonin is almost ubiquitously quantified using DLMO protocols with pre-defined thresholds of 3 or 4pg/mL (Claustrat et al., 2005; Voultsios et al., 1997). Note that it is essential to accurately assess melatonin because it is the basis for circadian timing. Specifically, circadian phase (or alignment) can be determined by calculating the time difference between DLMO and other indicators of the endogenous clock (e.g., sleep midpoint; Lewy et al., 1999).

Although DLMO remains the predominant method to quantify melatonin in humans, concerns have been raised about tits quantification. Some researchers have criticized the threshold approach as it provides limited information on melatonin and does not account for inter-individual or age differences (Danilenko et al., 2013). Again, the threshold approach is used to identify the time of the sample when melatonin exceeds 3 or 4pg/mL. This criticism is especially relevant for pediatric populations whose baseline melatonin levels are often higher than adult populations (Benloucif et al., 2008; Burgess & Fogg, 2008). In fact, many children exceed the DLMO thresholds at the initial sample taken during melatonin measurement (i.e., first starting sample is already higher than 3 or 4pg/mL; Crowley et al., 2006).

In response to these concerns, others have suggested modifying the onset thresholds to reflect individual differences (Burgess & Fogg, 2008). Instead of constant thresholds across persons (i.e., 3 or 4 pg/mL), the modified onset thresholds are unique to each individual. Namely, the within-person threshold becomes an individual-centric metric that is calculated as two standard deviations above one's mean of their first three melatonin measurement points. Timing is still the basis of measurement: the time at which the melatonin sample taken exceeds this individualised threshold. This revised method provides a more responsive way to quantify melatonin and it is thought to address the concern about individuals who either have extremely low levels of melatonin (i.e., never exceed 3 or 4pg/mL) or individuals whose baseline levels exceed these predetermined thresholds. Whether the original or revised metrics are calculated, both focus exclusively on the timing dynamic.

As an alternative way to measure melatonin, Danilenko et al. (2013) proposed the hockey-stick method to capture rate of change or the rate dynamic. The slope or steepness of the melatonin curve prior to bedtime was calculated. One advantage using a rate or slope approach is that it accounts for inter-individual differences and does not rely on any pre-defined threshold. Additionally, examining the rate at which melatonin increases was thought to yield information distinct from the time of crossing a threshold. In other words, the hockey-stick method provided information that may be missed with the threshold method (i.e., timing dynamic only). In fact, capturing the rate of melatonin surge provides more complete information about the rate dynamic

of melatonin. A steeper rise could reflect larger production and slower absorption or excretion. Few studies have calculated metrics to quantify melatonin rate or concentration dynamics; and to my knowledge, no studies have been conducted to validate these dynamics in children or adolescents.

Current Melatonin Practices in Pediatric Populations

Most studies have measured melatonin in children and adolescents using the same DLMO protocol that is commonly used with adults (Akacem et al., 2015; Crowley et al., 2006, 2014, 2016; Dolsen & Harvey, 2018; LeBourgeois et al., 2013; Mandrell et al., 2018). The DLMO protocol consists of collecting repeated saliva samples in the evening, starting two to three hours before typical bedtime and ending an hour after bedtime. Sampling rates vary from one saliva sample every hour to one sample every 30 minutes. Children are expected to be still (i.e., restricted movement) in a dimly lit room (10-30 lux). A single lux is equivalent to one lumen per square metre and is comparable to a clear night under a full moon. In contrast, typical household lighting at night is around 50 lux (Roemer et al., 2003).

Due to the challenges of measuring DLMO, there are a limited number of studies conducted among children and adolescents. To date (and to the best of my knowledge), there are no reviews of the extant literature that summarize the current melatonin measurement practices in pediatric populations. To build knowledge in the field of child sleep research, it would be beneficial to consolidate: what has been previously done, what DLMO variables are commonly derived, what circadian covariates are frequently observed, and whether this research has been conducted with diverse or representative samples. Further, while melatonin variables derived from the DLMO protocol are considered the gold-standard for accurate measurement of circadian timing, identification of robust proxy variables that can be more easily collected would be a significant contribution to advance both research and clinical practice.

Other Markers of the Endogenous Circadian Rhythm

While melatonin remains the gold standard for identifying circadian timing, other markers also reflect the endogenous circadian rhythm. Standardized questionnaires, such as the *Morningness-Eveningness Questionnaire* (Horne & Östberg, 1976), include validated questions about one's chronotype that focus on the timing of daily behaviors (e.g. preferred time to wake up or go to bed) and circadian preferences (e.g., preferred time to study or take an exam). Chronotype refers to one's preference for morningness (e.g., larks) or eveningness (e.g., owls). Subjective preference of morning or evening chronotype has been previously demonstrated to provide useful information to discern one's circadian timing (Carskadon et al., 1993).

From a practical perspective, standardized questionnaires, self-report measures (e.g., sleep diaries), or activity monitors (e.g., accelerometers) are progressively more time- and resource-intensive, but all are easier to administer than DLMO protocols; and, there is evidence they may provide information that can be used to derive markers of the endogenous circadian rhythm, specifically sleep midpoint (Reid, 2019). To note, self-report measures (questionnaires and sleep diaries) are commonly used to log information regarding children's sleeping habits; sleep diaries are usually completed daily for one to two weeks (Meltzer et al., 2012). However, these self-reports tend to underestimate bed and wake times, and are not the most reliable way to measure sleep (Lamprecht et al., 2015; Meltzer et al., 2012). In contrast, actigraphy is deemed an 'objective' measure of sleep and yields more accurate results (Stone & Ancoli-Israel, 2017). (Note, the basis of my honour's thesis showcased that while actigraphy is referred to as an

'objective' measure of sleep, there is still a modicum of subjective measurement as the researcher must review and score data manually to eliminate artifacts; Ramil, 2018).

Study Objectives

This thesis consists of four sequential parts, each with distinct aims that build upon the prior part. Part 1 was a scoping review to summarize existing findings using DLMO in children and adolescents. The key questions of interest include: (a) what metrics/variables are used to quantify melatonin? (b) what are the means and ranges of values of these metrics observed across samples? and (c) what are the demographics of the study samples (i.e., how diverse or representative are the samples)?

Part 2 compared traditional and alternative metrics for quantifying DLMO using the Healthy Heart Project dataset. Eleven metrics were calculated that were hypothesized to represent the three melatonin dynamics: timing, concentration, and rate. An inter-correlation matrix was constructed to examine the inter-relations among the proposed formulae for alternative DLMO metrics. The Healthy Heart DLMO metrics were compared to those from the existing literature based on the findings of Part 1. The key questions were: (a) are the proposed DLMO metrics comparable to traditional metrics? (b) is the Healthy Heart dataset representative of what is typically found in the literature? I hypothesized that the metrics within each dynamic would be highly correlated with one another and that the Healthy Heart dataset would be similar to the existing literature.

Part 3 strived to capitalize on emerging data science advances to create simulation datasets to more robustly consider the relations among the DLMO dynamics and circadian markers. Findings from the scoping review (Part 1) were used to inform the scope of the parameters for the sample demographics and augment the variable ranges of the Healthy Heart dataset (Part 2) using multiple imputation techniques. The simulation data were then used to answer the question: what are the most representative variables (i.e., candidate marker) for each DLMO dynamic? I hypothesized that the traditional DLMO metrics would be the most representative of each DLMO dynamic.

Part 4 explored the utility of DLMO proxies using variables that can be more easily collected in research and clinical settings. This was achieved by running a classification and regression tree analysis (CART) to estimate each melatonin dynamic (i.e., timing, concentration, rate) from the candidate markers of circadian timing and bedtime routine variables identified, such as bedtime, wake time, sleep midpoint, and chronotype preference, while controlling for relevant covariates (age, sex, BMI, pubertal status, time of year, use of melatonin supplements). The key questions for Part 4 were: (a) are there proxy variables that can be used to estimate DLMO? (b) does the sleep measurement method used affect the accuracy of these estimates? I hypothesized that sleep midpoint and chronotype would be the best proxies to estimate DLMO and that actigraphy would provide the most accurate estimates.

In the following sections, the method and results for each part are presented sequentially. Interpretation of the findings, comparison with the literature, and critical evaluation in relation to the thesis questions are presented later in a singular discussion. Note that occassionally, there are brief interpretation comments in study results that are provided to clarify information that is relevant for decision making for the subsequent study.

Part 1: Scoping Review of DLMO Melatonin Variables in Pediatrics Literature

Method

Part 1 was a scoping review conducted to summarize the way melatonin is quantified and sample characteristics within the pediatrics sleep and melatonin literature. Scoping reviews differ from systematic reviews in that they are typically exploratory and aim to compile information or evidence about a broad topic (Pham et al., 2014). As such, scoping reviews provide an overview of the current state of the literature and potentially reveal any gaps. Several parameters of interest were extracted from identified articles (e.g., melatonin measurement, DLMO metrics, bedtime routine measures, circadian markers, covariates). Extracted data were used to compare observed variable ranges in Part 2 and to inform the parameters of the simulated datasets in Part 3.

Literature Search Strategy

A systematic, electronic literature search was conducted using the APA PsycNet and Ebscohost databases on December 8 2020 and repeated on April 11 2022. Search terms included DLMO, dim-light, or dimlight in the abstract, melatonin in any field, and any variation of child ("child*" or "boy*" or "girl*" or "youth*" or "adolescent*" or "pediatric*" or paediatric*") in any field. A total of 91 potentially relevant studies were identified for full review.

Inclusion Criteria and Study Selection

Any study selected for inclusion had to conduct a DLMO protocol and report at least one melatonin metric in healthy children and/or adolescents (i.e., mean sample age was 8 to <18 years old). Additionally, study results had to be published in English and not previously published in another study. Among the 91 potentially relevant studies, only 22 studies met inclusion criteria and were included in the scoping review.

Scoping Review Data Extraction

Data were extracted from the 21 included studies. Variables were operationally defined and harmonized to ensure consistent data extraction across studies. Operational definitions for all coded variables are presented in Table 2. Coded variables included sample demographics (age, sex, pubertal status, BMI), sleep routine behaviors (bedtime, waketime), circadian markers (chronotype, sleep midpoint), and DLMO metrics (timing: $Time_{LI3}$, $Time_{LI4}$, $Time_{within}$, $Time_{hockey}$; rate: $Rate_{hockey}$, $Rate_{reg}$; concentration: $Conc_{AUC}$, $Conc_{max}$). All time variables were first converted to a 30hr format (e.g., 1:00am was transformed into 25:00) and then to a decimal format (e.g., 21:30 was transformed into 21.50) to ensure accurate mathematical calculations (e.g., averages) and to facilitate comparisons across studies. Time transformations are commonly applied in the sleep literature to accurately calculate several sleep variables. For example, when calculating a jet lag value from a school night bedtime of 22:00 and a weekend night bedtime of 1:00, using 1:00 for bedtime instead of 25:00 would inaccurately lead to a jet lag of -21:00, rather than 3:00. For simplicity, time variables are presented in decimal format throughout this thesis and few time variables were converted back into a time format (i.e., hh:mm) to facilitate the interpretation of those variables in the text.

Table 2. Operational Definitions of Variables Extracted from the Scoping Review.

Variable	k	Operational Definition & Possible Range (if applicable)	Mean (sd) [n, %]	Observed Range
Sample Demographics				
Sample Size	22	Study sample size	59.50 (69.11)	6 to 240
Age	22	Age of participant	14.62 (1.34)	9 to 17 years
Sex	22	Sex of participant (female, male)	[594, 48% female]	0 to 63% female
Puberty	1	Tanner pubertal stage; adrenarche, gonadarche (1=pre-pubertal to 5=post-pubertal)	3.74 (1.32)	n/a
BMI percentile	2	Body mass index percentile (1st to 99th percentile)	n/a	6 to 97th percentile
Time of year	1	Time of year of melatonin sample collection (school year, summer break)	[149, 71.63% school year]	n/a
Melatonin supplement use	0	Use of melatonin supplement during melatonin collection	n/a	n/a
Bedtime Routine				
Bed_T	17	Time attempts to fall asleep; close eyes (hh.mm)	22.84 (.98)	20.13 to 25.53
Bed_S	3	Time attempts to fall asleep on school nights; close eyes (hh.mm)	22.24 (.74)	20.13 to 25.53
Bed_W	3	Time attempts to fall asleep on weekends; close eyes (hh.mm)	23.10 (0.75)	20.13 to 25.53
$Wake_T$	10	Time wakes up in morning; opens eyes (hh.mm)	7.72 (.77)	4.47 to 11.38
$Wake_S$	3	Time wakes up in morning on school days; opens eyes (hh.mm)	6.47 (.73)	4.47 to 11.38
$Wake_W$	3	Time wakes up in morning on weekends; opens eyes (hh.mm)	8.07 (.86)	4.47 to 11.38
Jet lag	0	Difference in bedtimes between weekdays and weekends (hh.mm)	n/a	n/a
Weekend oversleep	0	Difference in waketimes between weekdays and weekends (hh.mm)	n/a	n/a
Circadian Markers				
Sleep Midpoint (hh.mm)	2	Halfway point between bedtime and wake time (hh.mm)	27.64 (.82)	24.00 to 30.23
Chronotype	2	Circadian preference; time of day feel best Morningness-Eveningness Scale for Children[a] 10 (evening type) to 42 (morning type)	37.28 (5.59)	10 to 41

Table 2 (cont.)

Variable	k	Operational Definition & Possible Range (if applicable)	Mean (sd) [n, %]	Observed Range
DLMO Timing Metrics				
$Time_{LI3}$	7	Time of melatonin sample >3pg/mL, calculated by linear interpolation (hh.mm)	21.37 (1.11)	18.11 to 24.09
$Time_{LI4}$	16	Time of melatonin sample >4pg/mL, calculated by linear interpolation (hh.mm)	21.93 (.99)	18.75 to 24.18
$Time_{within}$	1	Time of melatonin sample >2SD + baseline samples (hh.mm)	20.83 (1.15)	19.52 to 23.93
$Time_{reg3}$	0	Time of melatonin sample >3pg/mL, calculated by linear regression (hh.mm)	n/a	n/a
$Time_{reg4}$	0	Time of melatonin sample >4pg/mL, calculated by linear regression (hh.mm)	n/a	n/a

Table 2 (cont.)

Variable	k	Operational Definition & Possible Range (if applicable)	Mean (sd) [n, %]	Observed Range
$Time_{max}$	1	Time of melatonin sample maximum (hh.mm)	25.30 (1.65)	22.03 to 27.02
$Time_{hockey}$	0	Time of maximal slope difference using hockey-stick method (hh.mm)	n/a	n/a
DLMO Rate Metrics				
$Rate_{hockey}$	0	Rate of melatonin surge calculated using the Hockey-Stick formula method (pg/mL/hr)	n/a	n/a
$Rate_{reg}$	0	Rate of melatonin surge calculated using linear regression formula (pg/mL/hr)	n/a	n/a
DLMO Concentration Metrics				
$Conc_{max}$	1	Maximum melatonin concentration (pg/mL)	128.72 (8.16)	20.34 to 310.30
$Conc_{AUC}$	0	Total melatonin output calculated using area under the curve formula (pg/mL)	n/a	n/a

Note: k=number of studies. [a]Carskadon et al, 1993. n/a indicates metric/variable was not reported in any study in the existing literature

Part 1 Results

Findings of the scoping review are summarised above in Table 2. The mean values and range of values observed in the 21 studies are presented in the right columns. Appendix A includes a comprehensive table (see Table 13) with the specific findings extracted from each study.

Participant Demographics

Samples sizes ranged from 6 to 240 participants, with a mean of 60 participants. The age of participants included in the studies ranged from 9 to 17 years old (M=14.62, SD=1.34), with approximately half of samples consisting of female participants. Two of the studies did not include females. Only two studies reported controlling for BMI percentile; however, they did not report the means and standard deviations, only the range of values. A single study controlled for

pubertal status, which was assessed by a physician. Also, only one study controlled for the time of year of melatonin collection (i.e., school year vs summer break). No studies reported controlling for the use of melatonin supplements during the study.

Bedtime Routine Variables

Of the seventeen studies with reported bed- and wake-times, three of them had separate means for school nights and weekends. (As a friendly reminder, the subscript $_S$ refers to school nights, $_W$ refers to weekends, $_T$ refers to total.) On average, bedtime values ranged from 20.13 to 25.53. Bed_T reported was 22.84 (*SD*=.98), while Bed_S and Bed_W were on average 22.24 (*SD*=.74) and 23.10 (*SD*=.75), respectively. Waketime values ranged from 5.78 to 11.50. $Wake_T$ was 7.72 (*SD*=.77), whereas $Wake_S$ and $Wake_W$ were on average 6.47 (*SD*=.73) and 8.07 (*SD*=.86), respectively. Surprisingly, jet lag and weekend oversleep were not reported in any studies.

Circadian Markers

Only two studies reported midpoint values (*M*=27.64, *SD*=.82), with values ranging from 24.00 to 30.23. Two studies also reported using the Morningness-Eveningness Scale for Children to assess chronotype, with scores ranging from 10 to 41.

DLMO Metrics

The most common DLMO measure used was $Time_{LI4}$ (i.e., the time melatonin concentration surpassed a threshold of 4pg/mL using the linear interpolation method $_{LI}$), with 73% of studies using this measure (*M*=21.93, *SD*=.99). Seven studies (32%) used $Time_{LI3}$ (i.e., the time melatonin concentration surpassed a threshold of 3pg/mL; *M*=31.37, *SD*=1.11). Only one study reported using $Time_{within}$ (i.e., the time melatonin concentration surpassed a within-person threshold, calculated as 2 standard deviations above the first 3 melatonin samples). Another study reported $Time_{max}$ (time of maximum melatonin sample). Only one study reported a concentration metric: $Conc_{max}$ (maximum melatonin concentration). No studies reported using any metrics that reflected the rate dynamic of DLMO.

Part 2: Calculating Traditional and Alternative DLMO Melatonin

The objective of Part 2 was to calculate traditional and alternative metrics for melatonin and to compare their inter-relations within and across the three DLMO dynamics: timing, rate, and concentration.

Method

Part 2 was conducted with data from the Healthy Heart Project that was previously collected. Details of the secondary dataset are described below.

Healthy Heart Participants

The original data sample included 223 English-speaking participants, aged 9–17 years (*M*=12.54, *SD*=2.06, 43.90% female) who took part in the larger Healthy Heart Project between 2014 and 2018. Participants were recruited through flyers posted in the community and bookmarks distributed in schools. Participants were first screened via telephone to determine eligibility. Participants were excluded if they had severe medical or psychological disorders, or used medication known to affect cardiovascular health. Participants were scheduled for two laboratory visits to answer questionnaires and take anthropometric measurements. Only relevant elements of the Healthy Heart Project are described.

Healthy Heart Procedure

During the first visit, youth and their parents answered demographic questions (age, sex, pubertal status) and completed sleep questionnaires. Research assistants collected anthropometric data (height, weight, for body mass index). Youth were then fitted with psychophysiology equipment (wrist actigraphy) to be worn for two consecutive weeks. Participants were instructed on the home saliva collection process. They were provided with a DLMO kit, which included a light meter, saliva cotton swabs, blue-light blocking glasses, a timer, a freezer bag, and two nightlights. When the participants returned to the lab for their second visit, they returned all the equipment, questionnaires, and saliva samples. Monetary compensation was provided to families who completed the study.

Healthy Heart DLMO Protocol

Youth were instructed to collect melatonin samples at home on a school night (Wednesday or Thursday, recommended). The DLMO protocol began two hours before habitual bedtime and ended one hour after. Thirty minutes prior to the start of sample collection, participants were instructed to dim the lights in their room until the light meter read a value of under 30 lux. Nightlights included in the DLMO kits were used to ensure that the child was not sitting in complete darkness. Youth were asked to remain in their bedroom throughout the duration of the protocol. If they had to leave the room, they were required to wear the blue-light blocker glasses. Youth were required to rinse their mouths ten minutes prior to each sample collection if they ate or drank.

Saliva samples were collected every 30 minutes for a total of 7 samples. The timer provided in the DLMO kit was used to remind the child when it was time for the next sample. At sample collection, an oral cotton swab was placed under the child's tongue for one to two minutes. The collection time and lux value were recorded on the DLMO collection log. The saliva sample was then placed in the insulated bag that was adjacent to the child. Youth were instructed to store samples in a freezer bag and place it in the freezer until their second visit. When the samples were returned to the lab, they were shipped on dry ice to Salimetrics Laboratory to be immunoassayed. The sensitivity of the assay was 1.37pg/mL and the limit of detection was 0.58pg/mL. The intra- and inter-coefficients of variability were 5.42% and 8.9% respectively. These values are consistent with those previously reported in the literature (Agostini et al., 2017; Akacem et al., 2015; Crowley et al., 2016).

Secondary Data Melatonin Quantification

Three categories for the DLMO dynamics were measured: (i) timing; (ii) concentration; and (iii) rate of surge. These dynamics were derived using four biomathematical models: linear interpolation, linear regression, area under the curve (AUC), and hockey-stick. The models are described below and Table 3 contains a summary of the melatonin metrics and the corresponding names that will be used throughout the thesis.

Timing metrics. Timing of DLMO was determined by calculating the time melatonin concentration surpassed designated thresholds. Linear interpolation and linear regression were used to calculate timing. Linear interpolation consists of predicting values that are between two datapoints, using the following formula:

$$y = y_1 + (x - x_1)\frac{(y_2 - y_1)}{(x_1 - x_2)}$$

where (x, y) represent the time and concentration threshold, respectively; (x_1, y_1) represent the melatonin sample time and concentration preceding the threshold; and (x_2, y_2) represent the sample time and concentration after the threshold was exceeded.

Linear Regression is done using the following formula:

$$y = ax + b$$

where a is the slope, or rate of the melatonin surge, x is the time, b is a constant, and y is the predicted melatonin concentration.

Using the linear interpolation and regression methods, I calculated the time where melatonin concentrations surpassed the pre-determined thresholds (i.e., 3pg/mL, 4pg/mL, 2 standard deviations above the mean of the first three samples). Calculated DLMO timing metrics included Time_{LI3}, Time_{LI4}, $\text{Time}_{\text{reg3}}$, $\text{Time}_{\text{reg4}}$, $\text{Time}_{\text{within}}$, Time_{max}, $\text{Time}_{\text{hockey}}$ (see Table 3).

).

Concentration metrics. Concentration was assessed using the area under the curve method, which allows us to calculate the total melatonin output. This method also allows us to record the highest and lowest melatonin concentrations. Area Under the Curve was calculated using the trapezoidal method (Burgess & Fogg, 2008). This method uses the following formula:

$$AUC = \sum_{k=1}^{k-1} \frac{(s_k + s_{k-1})(t_{k+1} - t_k)}{2}$$

where *k* represents the total number of melatonin samples and *s* represents the melatonin concentration of a particular sample.

The maximum melatonin concentration was also included as a measure of concentration. Calculated DLMO concentration metrics included Conc_{max} and Conc_{AUC} (see Table 3).

Rate metrics. The rate of change was calculated in two ways: (1) using the slope of the regression model previously mentioned, and (2) using the Hockey-Stick Method, proposed by Danilenko et al. (2013):

$$F(t) = f_1(t) \times \theta(t_0 - t) + f_2(t) \times \theta(t - t_0)$$
$$f_1(t) = a_1 t + b_2$$
$$f_1(t) = a_2 t^2 + b_2 t + c_2$$

$$\theta(x) = \begin{cases} 0, & x < 0 \\ \frac{1}{2}, & x = 0 \\ 1, & x > 0 \end{cases}$$

This model consists of fitting the data points to a piecewise function consisting of two parts: a straight line $f_1(t)$ representing the baseline melatonin value and a parabola $f_2(t)$, which plots the rise in melatonin. The time of inflection is represented by t_0 and $\theta(x)$ is denoted by the step function above. Calculated DLMO rate metrics included Rate_{reg} and $\text{Rate}_{\text{hockey}}$ (see Table 3).

Table 3. Melatonin Metrics Calculated

DLMO Dynamic	Metric Abbreviation	Description of Calculation for Metric	Unit
Time	$Time_{LI3}$	Time melatonin >3pg/mL, using linear interpolation	hh.mm
	$Time_{LI4}$	Time melatonin >4pg/mL, using linear interpolation	hh.mm
	$Time_{reg3}$	Time melatonin >3pg/mL, using linear regression	hh.mm
	$Time_{reg4}$	Time melatonin >4pg/mL, using linear regression	hh.mm
	$Time_{within}$	Time melatonin >2 SD of first 3 melatonin samples	hh.mm
	$Time_{max}$	Time of maximum melatonin sample	hh.mm
	$Time_{hockey}$	Time of inflection, using Hockey-Stick Method	hh.mm
Concentration	$Conc_{max}$	Maximum melatonin concentration among all samples	pg/mL
	$Conc_{AUC}$	Total melatonin output, using area under the curve	pg/mL
Rate	$Rate_{reg}$	Slope of melatonin surge, using linear regression	pg/mL/hr.mm
	$Rate_{hockey}$	Difference score of melatonin surge slopes, using Hockey-Stick Method	(pg/mL/hr.mm) difference score

Note: *SD*=standard deviation

Circadian Timing Markers

Two markers of circadian timing were measured: sleep midpoint and chronotype.

Sleep midpoint. Sleep midpoint was calculated from questionnaires, the sleep diary, and actigraphy data. On the questionnaires, youth answered questions about their habitual sleep habits (i.e., "What time do you usually go to bed on week nights/school nights (Sunday to Thursday nights)?", "What time do you usually go to bed on weekend nights (Friday to Saturday nights)?", "What time do you usually wake up on Weekdays/School mornings (Monday to Friday mornings)?", "What time do you usually wake up on weekend mornings (Saturday and Sunday)?"). Sleep diaries were completed by youth daily, for two weeks (Carney et al., 2012). The child reported their bedtime (i.e., "What time did you close your eyes and try to go to sleep last night?") and waketime (i.e., "what time did you get out of bed to start your day today?"); together, these measures were used to calcuate sleep midpoint. Youth simultaneously wore a uniaxial, piezoelectric accelerometer (Actiwatch 2, Philips Respironics, USA) on their non-dominant wrist for up to 14 consecutive days. Youth were instructed to only remove the device when bathing or swimming. Data were sampled in 30-sec epochs. Accelerometry data were scored using a protocol validated by the Pediatric Public Health Psychology Lab (McGrath et al., 2018) and processed using Actiware software (Philips Respironics, USA). Analyzed actigraphy data yield the start- and end-time of the rest interval (i.e., lights out/bedtime, lights on/waketime). Sleep midpoint is defined as the half-way point between bed and wake time. It was calculated using the following formula:

$$Sleep\ midpoint = \frac{bedtime + wake\ time}{2}$$

Similar to the timing data extracted from the scoping review, bedtimes and waketimes were first converted into a 30-hr format and then into decimal format to facilitate calculations. For instance, if a participant went to bed at 23:00 and woke up at 7:00am, no transformations would result erronously in a midpoint of 15:00. Instead, converting the times into a 30-hr decimal format would result in a midpoint of 27.00 hrs, which is equivalent to 3:00am and is the correct midpoint value. Average sleep midpoints were also calculated separately for school nights and weekend nights. For parsimony, total sleep is presented throughout the thesis (weighted average of school nights and weekend nights; readers will note the subscript $_T$). Interested readers can find the complete data stratified by school ($_S$) and weekend ($_W$) nights in the Appendices.

Chronotype. Chronotype was measured with the Morningness-Eveningness Scale for Children (MESC; Carskadon et al., 1993). This is a self-report questionnaire that consists of 10 items assessing timing preferences for certain activities (e.g., taking tests, going to bed, waking up). Items are rated on a Likert scale of 1 to 5 and summed to obtain a total score; higher scores indicate morning preference, and lower scores indicate evening preference. The MESC has been previously validated in pediatric samples, with Cronbach's alpha coefficients ranging from 0.73 to 0.83 (Koscec et al., 2014). The internal consistency of the MESC for the Healthy Heart dataset was 0.74. MESC scores were categorized into three chronotypes: morning preference (morning lark, score > 27), evening preference (night owl, score < 21), or no preference (score 21-27).

Bedtime Routine Variables

Bedtime routine variables refer to an individual's sleeping habits, specifically bedtime and waketime, as well as jet lag and weekend oversleep.

Similar to sleep midpoint, bedtime and wake time were derived from the questionnaires, the sleep diary, and actigraphy data. Bedtime is defined as the time when one attempts to fall asleep (e.g., 'close your eyes and try to go to sleep'). Wake time is defined as the time the participant opens their eyes in the morning with the intent of rising for the day (i.e., not snoozing in bed longer). Average bed and wake times were calculated separately for school nights and weekend nights, excluding the DLMO night. Average bed and wake times were calculated from the sleep diaries and actigraphy device data.

Jet lag and weekend oversleep are measures of sleep variability within a week. Jet lag in sleep research refers to the difference in bedtimes between weekdays and the weekend, whereas weekend oversleep represents the difference in waketimes between weekdays and the weekend (Parsons et al., 2015). Greater jet lag and weekend oversleep indicate larger variability in a person's sleep schedule, which has been linked to health problems, such as obesity (Parsons et al., 2015).

For clarity, bedtime, waketime, and sleep midpoint will follow specific nomenclature described in Table 4. The variable name will be followed by two subscripts, the first subscript will indicate the data used for the mean (i.e., total, school nights, weekend) and the second subscript will indicate the method of data collection (i.e., questionnaire, sleep diary diary, actigraphy). Hence, the format is $Variable_{MM}$, where $_{MM}$ refers to Mean, Method. For example, a total bedtime mean measured by actigraphy would be referred to as $Bed_{TA,}$ whereas the school night mean for midpoint calculated by sleep diary diary would be called $Midpoint_{SD}$.

Table 4. Bedtime, Waketime, and Midpoint Abbreviation Nomenclature

Variable Name	Description
	Variable
Bed	Time participant attempts to fall asleep (i.e., close eyes)
Wake	Time participant opens their eyes to begin day
Midpoint	Halfway point between bedtime and waketime
	Mean (as subscript)
T	Mean all nights (Total)
S	Mean school nights only (Sunday to Thursday)
W	Mean weekend nights only (Friday and Saturday)
	Method (as subscript)
A	Actigraphy measurement (up to 14 days)
D	Diary sleep diary measurement (up to 14 days)
Q	Questionnaire self-report measurement (habitual / routine)

Covariates

Several covariates were assessed, including age, sex, pubertal status, body mass index (BMI), time of year, and use of melatonin supplements (assessed through the sleep diary as "Did you take melatonin supplements today?"). Adrenarche (pubic hair growth) and gonadarche (genital/breast development) are pubertal status development indicators assessed using illustrations corresponding to the five Tanner stages, with 1 representing the beginning of puberty and 5 representing full physical development and are assessed via a self-report questionnaire (Growing and Changing Questionnaire; Golding et al., 2001). Previous studies have demonstrated that the illustrations have high reliability and validity (r= .77-.91) (Dorn et al., 1990; Morris & Udry, 1980; Netherton et al., 2004). Pubertal status was categorised into three groups: pre-pubescent (Tanner stage 1), pubescent (Tanner stages 2-4), and post-pubescent (Tanner stage 5). BMI percentiles were categorized according to obesity status conventions for pediatric samples (i.e., CDC growth curves): underweight (0<5%), normal weight (5 to <85%), overweight (85 to <95%; Kuczmarski et al., 2002), and obese (95 to 100%). Time of year was defined as school year (September 1st to May 30th) or summer (June 1st to August 31st). Analyses in Part 2 used continuous data, unless otherwise noted; categories described were used for decision tree analyses in Part 4.

Data Quality

A data quality and protocol fidelity check was conducted based on the following: (1) sampling compliance (i.e., initials on data collection log); (2) dim light compliance (i.e., lux values <30); (3) weeknight sampling (i.e., verified date of DLMO collection was on a weeknight); (4) saliva sample integrity (e.g., samples remained frozen and were not compromised); and, (5) waking status (i.e., participants stayed awake during collection). A fully compliant subsample (n=161) was created by retaining only those participants who completed <u>all</u> of the following: (1) initialed the data collection log; (2) remained under dim-light conditions and logged the lux values in the data collection log; (3) collected their melatonin samples on a weeknight; (4) appropriately collected and stored the samples; and, (5) stayed awake during the DLMO protocol collection period.

Analyses

T-test were used to determine whether non-compliance compromised the data. Specifically, t-tests were conducted to compare the complete participant sample with the fully compliant subsample. It was decided *a priori* that the results of the t-tests would determine whether the complete secondary dataset should be retained for subsequent analyses of the thesis or whether the fully compliant subsample should be used for Part 3 and 4.

Pearson-product moment correlations were calculated to assess the strength and direction of the associations between variables. Specifically, the associations were examined to assess whether the three DLMO dynamics (i.e., timing, concentration, rate) clustered and were distinguishable from one another, and whether a metric within each dynamic stood out. Pearson correlations were also conducted to assess the association among the circadian markers and the bedtime routine variables.

Part 2 Results

Secondary Data Descriptive Statistics

Descriptive statistics for the complete secondary dataset are presented in Table 5. (For parsimony, only select variables are featured in the thesis tables; however, the interested reader can find analyses for the complete list of variables in Appendix B).

Sample Demographics. The mean age of the Healthy Heart sample was 12.54 (*SD*=2.06), with approximately 40% females. Two thirds of the sample collected melatonin samples during the school year (n=150) and 6% (n=14) reported using melatonin supplements during sample collection.

Bedtime Routine. Reported bedtimes ranged from 20:00 to 03:00 across different sleep assessment methods (i.e., actigraphy, sleep diary, questionnaires) and waketimes ranged from 05:00 to 14:00. On average, there was a 28.20-minute difference (*SD*=0.64) between weekday bedtimes and weekend bedtimes (i.e., jet lag) and a 46.20-minute difference (*SD*=1.02) in waketimes between weekdays and weekends (i.e., weekend oversleep).

Circadian Markers. Chronotype values ranged from 13 to 39, with 53% participants having a morning preference, 40% participants with an evening preference, and 8% participants with no preference. Sleep midpoint values ranged from 24.75 to 31.79 across all different sleep assessment method.

DLMO Metrics. DLMO timing metrics ($Time_{LI3}$, $Time_{LI4}$, $Time_{reg3}$, $Time_{reg4}$, $Time_{within}$, $Time_{max}$, $Time_{hockey}$) averages ranged from 19:33 to 22:58, with $Time_{reg3}$ being the earliest average time and $Time_{max}$ being the latest. DLMO concentration metrics were largely similar. $Conc_{max}$ ranged from 1.20 to 82.54 pg/mL (*M*=28.45, *SD*=12.06) and $Conc_{AUC}$ ranged from 0.30 to 186.10pg/mL (*M*=51.03, *SD*=31.32). There was large variation among the two DLMO rate metrics; however, this is attributable to one rate being the unstandardized beta coefficient of the regression line ($Rate_{reg}$ *M*=6.26, *SD*=5.41) and the other being the difference score between two slopes of the Hockey-Stick method ($Rate_{hockey}$ *M*=0.01, *SD*=0.01).

Finally, there was a curious observation that approximately 20% of the sample from the secondary dataset did not follow the 'typical' melatonin curve of a steady rise in melatonin during the evening prior to sleep onset. Instead, these participants had melatonin values that peaked multiple times throughout the 3-hour collection timeframe (i.e., not a stable, steady rise). This issue will be elaborated on in the Discussion, including consideration of how others have addressed this peculiarity previously in the literature.

Table 5. Secondary and Compliant Dataset Variable Descriptives

	Secondary data (n=223)			Fully compliant data (n=161)			t-test	
Variable	Mean (n)	SD (%)	Min - Max	Mean (n)	SD (%)	Min - Max	t-score	p-value
Age	12.54	2.03	9–17	12.54	2.03	9–17	−0.05	0.96
Sex (female)	(98)	(43.95)	n/a	(74)	(45.96)	n/a	−0.38	0.71
Puberty (gonadarche)	3.15	1.16	1–5	3.15	1.16	1–5	0.17	0.87
Puberty (adrenarche)	3.14	1.34	1–5	3.14	1.34	1–5	0.03	0.98
BMI percentile	54.83	27.61	0.47–99.44	54.83	27.61	3.78–99.44	−0.19	0.85
BMI Z-score	0.20	0.93	−2.6–2.54	0.2	0.93	−1.78–2.54	−0.26	0.79
Time of year (school year)	(73)	(32.74)	n/a	(66)	(34.16)	n/a	−0.28	0.78
Melatonin supplement use	(14)	(6.28)	n/a	(11)	(6.83)	n/a	−0.21	0.83
$Midpoint_{SA}$	27.40	1.11	24.94–31.39	27.36	1.12	25.54–31.39	0.33	0.74
Chronotype	27.35	4.47	13–39	27.31	4.65	13–39	0.09	0.93
$Bedtime_{SA}$	23.15	1.25	20.64–27.55	23.1	1.27	20.86–27.55	0.38	0.70
$Waketime_{SA}$	7.65	1.10	5.24–11.35	7.62	1.11	5.72–11.35	0.23	0.82
Weekend oversleep	0.77	1.02	−1.83–4.02	0.78	1.01	−1.83–4.02	−0.09	0.93
Jet lag	0.48	0.64	−1.45–1.99	0.50	0.65	−1.45–1.93	−0.24	0.81
$Time_{LI3}$	20.94	1.27	19.04–23.53	20.91	1.26	19.04–23.28	0.11	0.91
$Time_{LI4}$	21.13	1.42	19.04–26.27	21.01	1.37	19.04–26.27	0.48	0.63
$Time_{reg3}$	19.55	1.63	11.48–24.12	19.47	1.51	13.43–22.9	0.40	0.69
$Time_{reg4}$	19.73	1.61	11.86–24.31	19.66	1.50	13.77–23.16	0.35	0.72
$Time_{within}$	21.98	1.29	19.57–26.48	21.83	1.17	19.57–24.35	0.90	0.37
$Time_{max}$	22.87	1.37	18.00–26.92	22.81	1.39	18.00–26.58	0.40	0.69
$Time_{hockey}$	21.90	1.47	18.25–26.75	21.89	1.47	18.25–26.75	0.11	0.91
$Conc_{max}$	28.45	15.06	1.20–82.54	28.41	15.15	1.20–82.54	0.03	0.98
$Conc_{AUC}$	51.03	31.32	0.30–186.10	51.75	32.27	0.30–186.09	−0.20	0.84
$Rate_{reg}$	6.26	5.40	−15.40–22.56	6.03	4.92	−9.54–20.42	0.43	0.67
$Rate_{hockey}$	0.01	0.01	0.00–0.07	0.01	0.01	0.00–0.04	0.16	0.87

Data Quality

Data quality and protocol fidelity are summarized in Table 6. Overall, participants demonstrated high compliance; 89-94% of samples were initialed on the DLMO collection sheet. Dim lighting conditions were respected for >91% of all participants' samples. Almost all participants (97%) completed the DLMO protocol during a weeknight. Less than 4% of samples were compromised (i.e., touched, dropped, not frozen after collection). Finally, very few participants (<3%) fell asleep before taking their 7^{th} sample (1 hour past typical bedtime).

Overall, we determined that data quality and fidelity to the protocol were high. Further, sensitivity analyses were planned to conservatively exclude any potentially compromised saliva samples (e.g., no initials, exceeding lux, not frozen, fell asleep), decreasing the sample size from 223 to 161 participants (72% of original sample retained).

The means and standard deviations for the variables calculated using the entire secondary dataset versus the fully compliant subsample were compared. As can be seen in Table 5, there

were no significant differences for any of the sample demographics, sleep routine variables, endogenous circadian markers, DLMO dynamics, nor covariates. Because there were no significant differences between the complete secondary dataset and the fully compliant subsample, the decision was made to use the fully compliant subsample in the subsequent parts of this thesis. While using only the fully compliant subsample reduces the sample size by 30%, eliminating non-compliant samples allowed us to decrease the potential impact of measurement error on the calculated metrics.

Table 6. Data Quality & Non-Compliance to Protocol

	Non-compliant participants per sample (n)						
Saliva Sample	1	2	3	4	5	6	7
No Initials	14	14	14	15	16	18	22
Lux >30	13	12	12	14	16	15	18
Missing lux (not written or does not make sense)	13	12	12	14	16	15	18
Weekend collection	6	6	6	6	6	6	6
Compromised sample (touched/dropped/not frozen)	7	6	4	5	6	3	6
Fell asleep	0	0	1	2	4	4	7
Compliant participants (n)	191	191	190	190	188	183	178

DLMO Dynamics and Metrics

Pearson correlations were examined to evaluate the strength of the association between the metrics of the three DLMO dynamics. Table 7 contains the correlation coefficients of the DLMO metrics, grouped within their respective DLMO dynamics (timing, concentration, rate). It was anticipated that the timing metrics would be more strongly associated with one another, than with the rate or concentration metrics; similar associations were expected within the rate metrics and concentration metrics.

Timing metrics were highly correlated with one another, ranging from r = .75 to .99. $Time_{within}$ was the least correlated with other timing metrics, whereas $Time_{LI3}$ had the highest correlation coefficients. Concentration metrics were highly correlated with one another (r=.91). Rate metrics were only moderately correlated (r=.43).

Timing metrics were moderately negatively correlated with the concentration and rate metrics, suggesting that later DLMO times are associated with smaller concentrations and flatter slopes. Interestingly, $Rate_{reg}$ was highly correlated with the concentration metrics (r=.86-.97) whereas $Rate_{hockey}$ was only moderately correlated with the concentration metrics (r=.47-.58).

In sum, with the exception of the rate metrics, all metrics of the DLMO dynamics where highly correlated with one another.

Table 7. Intercorrelations of DLMO Dynamics and Metrics

		Timing							Concentration		Rate	
	Metrics	LI3	LI4	Reg3	Reg4	Within	Max	Hockey	Max	AUC	Reg	Hockey
Timing	LI3											
	LI4	0.99										
	Reg3	0.97	0.98									
	Reg4	0.98	0.98	.99								
	Within	0.80	0.77	0.80	0.80							
	Max	0.92	0.92	0.94	0.94	0.90						
	Hockey	0.79	0.77	0.81	0.82	0.75	0.78					
Conc	Max	−0.55	−0.55	−0.46	−0.50	−0.39	−0.41	−0.46				
	AUC	−0.63	−0.64	−0.56	−0.59	−0.29	−0.43	−0.45	0.91			
Rate	Reg	−0.46	−0.46	−0.36	−0.40	−0.37	−0.35	−0.43	0.97	0.86		
	Hockey	−0.49	−0.51	−0.46	−0.48	−0.39	−0.47	−0.19	0.58	0.47	0.43	

Concordance Among DLMO Metrics, Circadian Markers, & Bedtime Routine Variables

Figure 1 illustrates the strength of the correlations between all variables, including: DLMO metrics, bedtime routine variables, and circadian markers. Instead of numerical values for the correlation coefficients, the figure contains colored squares according to the strength and direction of the association. Darker blues indicate a correlation approaching 1 and darker reds represent a strong correlation approaching −1. In contrast, the lighter a colour is, the closer the correlation is to 0. This infographic is provided to facilitate interpretation only. (Appendix C contains the numerical correlation matrix with all reported values.)

A quick scan of the figure indicates that the timing metrics were mostly strongly and positively correlated with all sleep midpoints *(r*=.55–.91), and most bedtime routine variables (*r*=.42–.93). Chronotype demonstrated a moderately negative association with the timing metrics (*r*=−.36– −.13), which is expected as lower chronotype values indicate a nighttime preference and are associated with later sleep times. Thus, individuals with lower chronotype values are expected to have later DLMO times. A similar pattern was observed, although with a weaker association, between timing metrics and jet lag (*r*=−.16–−.01). This negative association suggests that individuals with later DLMO timings have smaller jet lag values, indicating that these individuals have more consistent bedtimes between school nights and weekends.

The concentration and rate metrics also followed a distinct pattern from the timing metrics. That is, concentration and rate metrics were more negatively associated with the circadian markers (*r*=−.47–−.07) and bedtime routine variables (*r*=−.53–−.10). Again, an exception to this trend was chronotype, which was weakly correlated with the concentration and rate metrics. This near-zero association is interesting as it suggests that chronotype may not be associated with the amount of melatonin output (concentration dynamic) or the rate of melatonin surge (rate dynamic). Another exception to the negative association trend was jet lag, which was moderately positively correlated with $Conc_{max}$, $Conc_{AUC}$, and $Rate_{reg}$, and weakly correlated with $Rate_{hockey}$. Thus, it seems that individuals who have higher melatonin concentrations and have steeper rates of melatonin rise also have a more variable sleep routine, with bedtimes typically much later on

weekends compared to school nights. It is important to note that all data were collected cross-sectionally; there is no information to imply any temporal order or causality.

Circadian markers, specifically the different sleep assessment methods for midpoint (i.e., actigraphy, sleep diary, questionnaire) were strongly correlated with one another (r=.69-.98). Somewhat surprisingly, there was a moderately negative association between midpoint and chronotype (r=−.38–−.28). I expected to observe a stronger negative association between midpoint and chronotype as I anticipated individuals with a nighttime preference would have markedly later bedtimes and waketimes, thus later midpoints.

Finally, bedtimes and waketimes were generally moderately to strongly associated with one another (r= .48–.99), with the exception of $Wake_{SQ}$ (i.e., habitual waketimes on school nights assessed by a self-report questionnaire), where correlations ranged from .01 to .44. The actigraphy bedtime routine variables (Bed_{TA}, Bed_{SA}, Bed_{WA}, $Wake_{TA}$, $Wake_{SA}$, $Wake_{WA}$) were the most strongly correlated (r=.52–.99). By contrast, the questionnaire bedtime routine variables (Bed_{SQ}, Bed_{WQ}, $Wake_{SQ}$, $Wake_{WQ}$), had the lowest correlations overall (r=.20–.93).

Part 2 Summary

To summarise, the goal of Part 2 was to explore alternative DLMO metrics and assess their inter-relations with traditional metrics across the three DLMO dynamics (timing, concentration, rate). Overall, the metrics were correlated within each dynamic. Additionally, most of the circadian markers and bedtime routine variables were correlated with one another. Chronotype and jet lag were flagged as some correlations with the other variables, particularly with the melatonin concentration and rate metrics, were unexpected. Bedtime and waketime variables measured by actigraphy demonstrated the highest correlations among the sleep variables, while the bedtime and waketime variables measured by self-report questionnaires had the weakest correlations. In Part 3, correlations were done again to see whether the variables maintained similar associations with one another. Finally, the fully compliant subsample was retained in subsequent thesis parts to minimize potential for measurement error and to ensure that the data quality was optimized for the analyses planned.

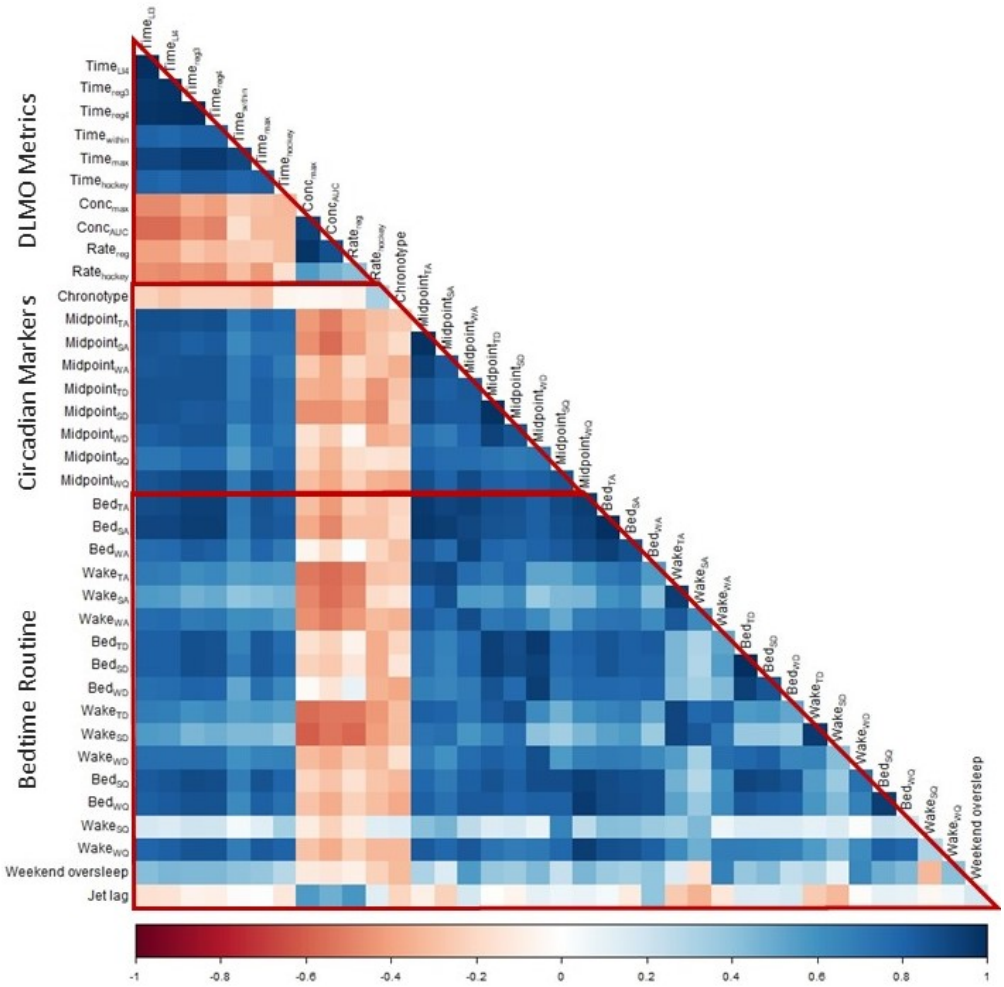

Figure 1. Correlations between DLMO metrics, circadian markers, and sleep routine variables in the secondary subsample

Part 3: Creating Simulated Datasets Informed by Scoping Review to Augment Healthy Heart Secondary Data

The objective of Part 3 was to create simulated datasets that augmented the Healthy Heart secondary data from Part 2 by using parameters obtained from the scoping review in Part 1. The simulated data were then used to identify candidate metrics that were most representative of each DLMO dynamic and would be carried forward to Part 4.

Method

Creating Simulated Datasets

The secondary data source retained from Part 2 (i.e., fully compliant dataset) contained 161 participants. This is considered an extremely large dataset for melatonin data collection and is an outlier among any study in the literature (Appendix B). Nevertheless, I recognize that smaller sample sizes can limit the power of test statistics to yield conclusions that can be drawn for measurement. Part 3 aimed to address this potential limitation by conducting simulation analyses

to more robustly examine the correspondence of the DLMO dynamic metrics with the endogenous circadian timing markers, the sleep routine variables, and the covariates.

First, multiple imputation with the Amelia II R package (version 1.7.6, Honiker et al., 2019) was used to create the simulated datasets. The Amelia II package employs bootstrapping algorithms to produce multiple output datasets from the original secondary dataset. A total of 100 imputed datasets were generated. To inform the creation of the imputed datasets, bounds (i.e., limitation on range of possible values) and priors (i.e., specified means and standard deviations), were imposed on the variables to reflect the distribution of variables reported in the literature from the scoping review. For example, a mean age of 14.62 years was set as a prior as this was the mean age reported in the literature. Furthermore, the age was restricted between 8 and 17 years old as this was the age range from our secondary dataset. When no information was provided for a particular variable (e.g., individual melatonin sample concentrations and times), the bounds and priors were informed from the secondary dataset. Table 8 illustrates the conditions imposed on the variables included in the imputation analyses. Additional variables not presented below can be found in Appendix D.

Table 8. Variable constraints for multiple imputation

Variable	Mean	SD	Min	Max
Age	14.62	1.35	8	17
Sex	1.5	0.5	1	2
$Puberty_{gonadarche}$	3	1.2	1	5
$Puberty_{adrenarche}$	3	1.4	1	5
BMI percentile	50	25	0	100
BMI Z-score	0	1	-3	3
Time of year	-	-	1	2
Melatonin supplement use	-	-	0	1
Chronotype	37.19	5.53	10	42
Bed_{SA}	23.1	0.75	20	26
$Wake_{SA}$	8.07	0.86	4	12

Analyses

Similar to Part 2, Pearson correlations were calculated again to determine whether strength and direction of the associations between the DLMO metrics within each dynamic (timing, concentration, rate) were similar to what was found previously in Part 2. I also examined the association among the DLMO metrics, circadian markers, and bedtime routine variables again to assess whether results from Part 2 could be replicated.

Part 3 Results

Simulated Data Descriptive Statistics

Descriptive statistics for the simulated dataset are presented in Table 9. Means and standard deviations of additional variables not presented below can be found in Appendix B. Overall, the simulated data had largely similar descriptives to the secondary data. There were negligible differences for the maximum value of range for a few variables: $Time_{LI3}$, $Time_{LI4}$, $Time_{within}$, $Time_{max}$, $Conc_{max}$, Jet lag; these differences are not unexpected given the goal to augment the bounds of the secondary dataset as informed by the scoping review.

Sample Demographics. The mean age of the simulated data was 12.54 (*SD*=2.02), with approximately 46% females. The proportion of the simulated data coded as school year and the

proportion of the simulated data coded as using melatonin supplements were 34% and 7%, which were similar to the proportions in the fully compliant subsample.

Bedtime Routine. Reported bedtimes also remained similar, with values ranging from (20.94 to 23.59) across different sleep assessment methods (i.e., actigraphy, sleep diary, questionnaire) and waketimes ranging from 05:00 to 14:00. On average, there was a 28.20-minute difference (*SD*=0.64) between weekday bedtimes and weekend bedtimes (i.e., jet lag) and a 46.20-minute difference (*SD*=1.02) in waketimes between weekdays and weekends (i.e., weekend oversleep). Midpoint values ranged from 24:00 to 09:00 across all measures.

Circadian Markers. Chronotype values ranged from 13 to 39, with 53% participants having a morning preference, 9% participants with an evening preference, and 38% participants with no preference. Sleep midpoint values ranged from 24.94 to 31.39.

Table 9. Fully Compliant Subsample and Simulated Data Variable Descriptives

	Fully compliant subsample (n=161)				**Simulated data (n=16,362)**			
Variable	Mean/ n	SD/ %	Min	Max	Mean/ n	SD/ %	Min	Max
Age	12.54	2.03	9	17	12.55	2.05	9	17
Sex (female)	74	45.96	-	-	7474	45.68	-	-
$Puberty_{gonadarche}$	3.15	1.16	1	5	3.13	1.13	1	5
$Puberty_{adrenarche}$	3.14	1.34	1	5	3.13	1.34	1	5
BMI percentile	54.83	27.61	3.78	99.44	54.2	28.75	0.47	99.44
BMI Z-score	0.20	0.93	-1.78	2.54	0.16	0.98	-2.6	2.54
Time of year (school year)	66	34.16	-	-	5555	33.95	-	-
Melatonin supplement use	11	6.83	-	-	1111	6.79	-	-
$Midpoint_{SA}$	27.36	1.12	25.54	31.39	27.38	1.09	24.94	31.39
Chronotype	27.31	4.65	13	39	27.32	4.46	13	39
$Bedtime_{SA}$	23.1	1.27	20.86	27.55	23.13	1.23	20.64	27.55
$Waketime_{SA}$	7.62	1.11	5.72	11.35	7.63	1.08	5.24	11.35
Weekend oversleep	0.78	1.01	−1.83	4.02	0.79	1	−2.31	4.03
Jet lag	0.50	0.65	−1.45	1.93	0.50	0.65	−2.25	4.22
$Time_{LI3}$	20.91	1.26	19.04	23.28	20.98	1.36	18.73	31.35
$Time_{LI4}$	21.01	1.37	19.04	26.27	21.17	1.44	18.78	31.48
$Time_{reg3}$	19.47	1.51	13.43	22.9	19.16	1.99	12.03	24.97
$Time_{reg4}$	19.66	1.50	13.77	23.16	19.39	2.04	11.77	24.89
$Time_{within}$	21.83	1.17	19.57	24.35	22.01	1.27	19.57	31.24
$Time_{max}$	22.81	1.39	18.00	26.58	22.87	1.40	18	31.50
$Time_{hockey}$	21.89	1.47	18.25	26.75	21.99	1.50	18.25	26.92
$Conc_{max}$	28.41	15.15	1.20	82.54	30.17	15.99	1.2	130
$Conc_{AUC}$	51.75	32.27	0.30	186.09	55.03	33.45	0.3	186.10
$Rate_{reg}$	6.03	4.92	-9.54	20.42	5.89	4.92	-12.57	22.56
$Rate_{hockey}$	0.01	0.01	0.00	0.04	0.01	0.02	0.00	0.08

DLMO Metrics. DLMO timing metrics ($Time_{LI3}$, $Time_{LI4}$, $Time_{reg3}$, $Time_{reg4}$, $Time_{within}$, $Time_{max}$, $Time_{hockey}$) averages ranged from 19.47 to 22:81, with $Time_{reg3}$ being the earliest average time and $Time_{max}$ being the latest. DLMO concentration metrics were largely similar. $Conc_{max}$ ranged from 1.20 to 130 pg/mL (M=30.17, SD=15.99) and $Conc_{AUC}$ ranged from 0.30 to 186.09pg/mL (M=51.75, SD=32.27). Finally, similar to the fully compliant subsample, the DLMO rate metrics were different from one another, with $Rate_{reg}$ (M=5.89, SD=4.92) values ranging between -12.57 to 22.56 and $Rate_{hockey}$ (M=0.01, SD=0.02) values between 0 and .08.

DLMO Dynamics and Metrics

Similar to Part 2, Pearson correlations were examined to evaluate the strength of the association between the metrics of the three DLMO dynamics (timing, concentration, rate). As expected, the Pearson correlations were almost identical to those reported in Part 2. Correlation coefficients are illustrated in Table 10.

Table 10. Intercorrelations of DLMO Metrics in the Simulated Data

		Timing							Concentration		Rate	
	Metrics	LI3	LI4	Reg3	Reg4	Within	Max	Hockey	Max	AUC	Reg	Hockey
Timing	LI3											
	LI4	0.98										
	Reg3	0.93	0.93									
	Reg4	0.94	0.95	0.98								
	Within	0.80	0.80	0.77	0.79							
	Max	0.87	0.86	0.93	0.93	0.84						
	Hockey	0.77	0.74	0.83	0.82	0.72	0.78					
Conc	Max	−0.49	−0.52	−0.37	−0.43	−0.4	−0.30	−0.27				
	AUC	−0.58	−0.61	−0.46	−0.51	−0.32	−0.32	−0.27	0.93			
Rate	Reg	−0.47	−0.49	−0.32	−0.38	−0.43	−0.30	−0.27	0.97	0.89		
	Hockey	−0.29	−0.33	−0.33	−0.35	−0.23	−0.33	−0.16	0.53	0.43	0.38	

Timing metrics were highly correlated with one another, ranging from r = .72 to .99. In the secondary compliant subsample, $Time_{within}$ was the least correlated with other timing metrics, but in the simulated data, $Time_{hockey}$ had the weakest correlations among all timing metrics. Just like in the secondary compliant subsample, $Time_{LI3}$ had the highest correlation coefficients. Concentration metrics were still highly correlated with one another (r=.93), whereas rate metrics were still only moderately correlated (r=.38).

Similar to the results in Part 2, timing metrics were moderately negatively correlated with the concentration and rate metrics. The strong correlation between $Rate_{reg}$ and the concentration metrics (r=.89-.97) were still present in the simulated data, whereas $Rate_{hockey}$ remained only moderately correlated with the concentration metrics (r=.43–.53).

In sum, the simulated datasets created from the fully compliant subsample and informed by the scoping review did not yield any changes in the associations between the different DLMO metrics.

Concordance Among DLMO Metrics, Circadian Markers, & Bedtime Routine Variables

Similar to Part 2, correlations between all DLMO metrics, circadian markers, and bedtime routine variables are illustrated in Figure 2. As a reminder, darker blues indicate a correlation approaching +1 and darker reds represent correlations approaching −1.

Overall, the simulated data had similar correlations than those from the secondary compliant subsample. The DLMO timing metrics were still strongly associated with the midpoint, bedtime, and waketime variables while the correlations between midpoint, bedtime, waketime, and the concentration and rate metrics remained negative. Thus, only differences between both datasets are highlighted below; consistent or similar associations are not described.

The biggest differences between the secondary compliant subsample and the simulated data were the correlations among the bedtime routine variables (Bed, Wake, weekend oversleep, jet lag) and the midpoint variables. In fact, some of the associations changed direction from the secondary dataset to the simulated dataset. This change in association was especially prominent between weekend oversleep and some waketime variables ($Wake_{TA}$, $Wake_{TQ}$, $Wake_{SQ}$), where correlation coefficients went from r= −.16 to .28 in the secondary dataset to r=−.57 to −.35 in the simulated dataset. The association between weekend oversleep and different midpoint variables ($Midpoint_{TA}$, $Midpoint_{SA}$, $Midpoint_{TD}$, $Midpoint_{SD}$, $Midpoint_{SQ}$) also varied between datasets, with coefficients ranging from r=.14–.38 in the secondary dataset to r=−.34–−.14 in the simulated dataset. This difference in correlations seemed to apply only to the total and school night means; weekend wake and midpoint variables were weakly, positively correlated with weekend oversleep (r=.11–.31). These results indicate that participants who woke up earlier during the week, tended to have more variability in their bedtime routine and slept in longer on weekends. Similar negative correlations were identified between jet lag and the same variables (i.e., total and school night waketime and midpoint). This result is interesting, particularly for the waketime variables, because jet lag is related to bedtime, not waketime.

Candidate Variables

DLMO Metrics. Candidate DLMO metrics were selected based on the results of the scoping review in Part 1 and supported by the correlations done in Parts 2 and 3. The goal was to select a candidate metric that would represent each dynamic.

For the timing dynamic, $Time_{LI3}$ was selected as the candidate metric to retain in Part 4. Although, $Time_{LI4}$ was more used in the literature, $Time_{LI3}$ was also a common metric and was more strongly correlated with the other timing metrics. Given the lack of information on the use of AUC in the literature to quantify melatonin, $Conc_{max}$ was retained as the candidate metric. When comparing both rate metrics, $Rate_{reg}$ was more correlated with the other DLMO metrics. As such, I decided to retain $Rate_{reg}$ as the candidate metric moving forward.

Sleep Assessment Methodology Variables. Actigraphy measures were selected as candidate variables to include in Part 4 as they were the most highly correlated with one another and actigraphy is more precise than sleep diaries and questionnaires (Short et al., 2012). Additionally, melatonin production is linked to the sleep-wake cycle and melatonin measures are more accurate when an individual's sleep pattern is more stable (i.e., they wake up and go to bed at the same time every day) (Benloucif et al., 2005, 2008). Therefore, I decided to retain the school night means as candidate variables because bedtimes and waketimes are more stable than weekends or total nights, and are expected to provide more accurate measurements (Parsons et al., 2015).

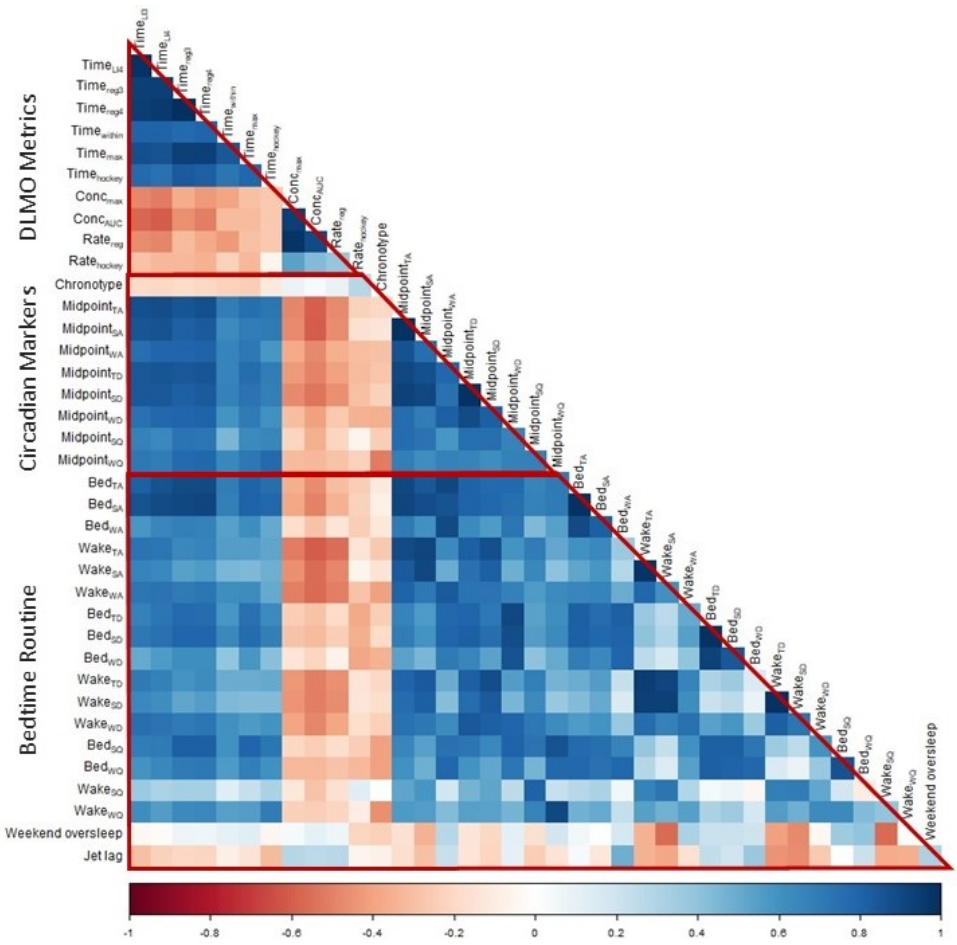

Figure 2. Correlations between DLMO metrics, circadian markers, and sleep routine variables in simulated data

Part 3 Summary

To summarise, the goal of Part 3 was to create simulated datasets to more robustly assess the relations among the DLMO dynamics, circadian markers, and bedtime routine variables. Overall, most variables had comparable correlations to the secondary compliant subsample in Part 2. The main difference pertained to weekend oversleep and jet lag, which had negative correlations with both midpoint ($Midpoint_{TA}$, $Midpoints_{A}$, $Midpoint_{TD}$, $Midpoints_{D}$, $Midpoints_{Q}$) and waketime ($Wake_{TA}$, $Wake_{TQ}$, $Wake_{SQ}$). Finally, candidate metrics were identified within each dynamic to use in Part 4: $Time_{LI3}$, $Conc_{max}$, and $Rate_{reg}$ and school night actigraphy means were selected for sleep midpoint ($Midpoints_{A}$), bedtime ($Beds_{A}$), and waketime ($Wake_{SA}$).

Part 4: Exploring the Utility of Alternate Variables as DLMO Proxies

The objective of Part 4 was to explore alternative variables (i.e., circadian markers, bedtime routine variables; more easily measured) that could potentially be used as proxies to estimate DLMO. Part 4 was intended to address the challenges of conducting DLMO protocols, which are time- and resource-intensive. Identifying proxies would be of great utility to both researchers and clinicians. The conceptual plan for Part 4 was twofold: first, to consider the candidate metrics and variables from Part 3 to estimate DLMO; and second, to consider whether using less

methodologically demanding variables (from actigraphy to sleep diary to questionnaire) could still produce comparable results. Decision tree analyses were used to analyze whether the candidate variables could accurately estimate DLMO dynamics (i.e., timing, concentration, rate).

Method

Classification and Regression Tree analyses (CART) were conducted. This statistical technique was selected because it permits a model to be created with scale (e.g., age), ordinal (e.g., puberty status), and/or categorical (e.g., sex) variables to estimate a continuous outcome (i.e., DLMO metric) and does not rely on parametric assumptions (i.e., normal distributions), yielding a resulting model robust against skewed or non-normal distribution. DLMO metrics vary and their distributions are more likely to be skewed or censored rather than normally distributed (e.g., concentration metrics cannot be lower than zero; timing metrics largely occur within a limited timeframe). Thus, given its forgiving statistical assumptions, CART analyses were deemed appropriate to explore possible DLMO proxies.

A regression tree is built via recursive partitioning by generating progressive binary splits. In other words, a tree starts as the root node, which is the mean value of the outcome variable before any splits are made. For example, if $Time_{LI3}$ is used as the outcome variable, the root node of the regression tree would be the mean of $Time_{LI3}$, 20.98 (see Table 2). Next, the tree will be split from this root node into two child nodes, which in turn become the parent nodes for the next split, and so on, until a specified splitting criterion is reached. The final nodes of the tree are referred to as leaves. When the outcome variable is continuous, the lowest residual sum of squares dictates the way the tree is split. To note, while multiple variables may be included as predictors, not all variables will necessarily be included in the final tree. To assess the relevance of a given variable in a particular model, the relative variable importance is used. Relative variable importance is reported as a percentage and simply indicates the amount a given predictor variable contributes to the overall model. For instance, if age has a variable importance of 30%, this means that age contributed 30% to the model. Thus, variable importance helps us determine which variables are better DLMO proxies. To note, while multiple variables can be added as predictors in a regression tree, they may not all be selected in the final model. For instance, sex might be included as a predictor but may have a low variable importance and be dropped from the final model. Pruning is then used to remove nodes to ensure that the model does not overfit the data. Model goodness of fit was assessed by examining the R^2 of the model (i.e., how much variance in the DLMO metric could be explained by the variables in the model) and comparing the RMSE between the estimated DLMO metric value and its actual value. In short, higher R^2 indicates that greater outcome variance is accounted for by the predictors, while smaller RMSE indicates that the model's estimations, on average, are closer to the actual values. RMSE can also be compared to the standard deviation of a given metric; if the RMSE is smaller than the standard deviation, the model is considered a good estimator of the outcome variable. For example, if the RMSE of a model is .5 and the standard deviation of the outcome variable is 1, the model would be considered as accurate as it can estimate values within one standard deviation of the actual values. However, if the RMSE is 1 and the standard deviation of the outcome variable is .5, the model would be considered to have a poor estimation of the outcome variable.

CART models

Three separate CART models were computed, one for each DLMO dynamic (timing, concentration, rate). For each model, the DLMO candidate metric identified in Part 3 was used as the outcome ($Time_{LI3}$ for timing, $Conc_{max}$ for concentration, and $Rate_{reg}$ for rate); the designated

predictors included four demographic variables (age, sex, $puberty_{adrenarche}$, BMI), four sleep routine variables (bedtime, waketime, jet lag, weekend oversleep) and two endogenous circadian timing markers (sleep midpoint, chronotype). Covariates included time of year (school year or summer) and melatonin supplement use. Nodes in the tree were constrained to have a minimum size of 20 records in parent nodes and 20 records in final child nodes. Ten-fold cross-validation was used to assess the predictive ability of each tree model. To note, two types of CART models will be described in the results section: candidate CART models, which consist of the candidate metric and predictors listed above, and final CART models, which only contain the metric outcome and predictors variables retained in the model. In other words, the candidate CART models included the variables I selected as DLMO proxies, based on information from Parts 1 to 3, whereas the final CART model only contained the variables retained from the analysis.

Exploratory Model Comparisons Across Alternative DLMO Proxies and Metrics

After the final CART models were created for the three DLMO dynamics from the candidate metrics and aforementioned variables, I explored the robustness of these models for alternative DLMO metrics and proxies. Specifically, I was interested in exploring two ideas. First, I was interested in assessing whether the final CART models generated could also estimate other metrics in a given dynamic. For example, in a CART model where the timing outcome metric was $Time_{LI3}$, I substituted $Time_{LI4}$ to determine how accurate the CART model estimations would be. Second, I was interested in exploring whether substituting the sleep routine variables and circadian timing markers measured by actigraphy with different sleep assessment methodology variables would affect the CART estimations. Consider, while actigraphy provides a reliable and valid estimate for sleep routine variables, these devices are costly, scoring their data is time-consuming, and they may not be readily available to clinicians and researchers (Meltzer et al., 2012). Instead, others may ask children to fill out weekly sleep diaries or use questionnaires asking about habitual bedtime and waketime, and consequently sleep midpoint. As such, the accuracy of the CART models was also tested when substituting actigraphy for other sleep assessment methodology (i.e., questionnaire or sleep diaries). CART models were re-run with these alternative proxy variables. Re-running the models in these exploratory test conditions permitted me to determine whether the models were robust and could be applied in different scenarios when DLMO metrics or other predictor variables are not available.

Part 4 Results

Interpreting the CART models

Each tree is read from top to bottom. Each rectangle with rounded corners represents a node, which contains the mean DLMO metric for that split, followed by the condition for the following split. For example, in Figure 3, the root node (at the top of the tree) includes a time of 20:55, indicating that the mean of the timing metric ($Time_{LI3}$) was 20:55, and a splitting condition of $Midpoint_{SA} < 27:13$, indicating that the first split in the tree was based on a $Midpoint_{SA}$ value of 27:13. Following the splits, if $Midpoint_{SA}$ was earlier than 27:13, the resulting value of the timing metric was estimated to be 20:05, if no further splits were made. Alternatively, if $Midpoint_{SA}$ was greater than 27:13, the estimated timing metric value was 22:07, again assuming no further splits were made. Following each of these splits and nodes leads us to the leaves of the tree, which are represented by the rectangles with a bolded outline at the bottom of the tree. As such, an example of an interpretation of the tree, using the timing model depicted in Figure 3, would be as follows: an individual who has a sleep midpoint earlier than 27:13 and a bedtime earlier than 22:10, and who is over 9.5 years old, has an estimated DLMO timing metric of 19:28.

If instead, the individual has a sleep midpoint earlier than 27:13 and a bedtime earlier than 22:10 but is less than 9.5 years old, their estimated DLMO timing metric would be 20:38. The regression trees for the concentration and rate models are illustrated in Appendix .

CART Models

The R^2, RMSE, and relative variable importance for each DLMO dynamic model are summarized in Table 11. To note, while the same predictor variables were included in each model, the relative variable importance varied and different variables were retained in the final models, suggesting that each DLMO dynamic potentially have different proxy variables. As expected, sleep midpoint was considered a good proxy for all DLMO dynamics. However, chronotype was surprisingly not retained in the final models described below.

Timing Model. The timing model had the highest R^2 (.91) and the lowest RMSE (.20) among all three final models generated. In other words, the predictor variables accounted for 91% of the variance in the timing metric and estimated the DLMO timing metric ($Time_{LI3}$) within .20 hours (i.e., 12 minutes) on average from the actual value. The proxy variables included in the candidate model with the highest relative variable importance were, in descending order, $Midpoint_{SA}$, Bed_{SA}, and $Wake_{SA}$, age, jet lag, and time of year. Although $Wake_{SA}$ was deemed an important variable, it was not included in the final CART model, suggesting that $Wake_{SA}$ was masked by another variable (most likely midpoint and/or bedtime). Thus, the final timing model included $Time_{LI3}$ as the timing outcome variable and $Midpoint_{SA}$, Bed_{SA}, age, jet lag, and time of year as predictor variables. In other words, the final timing CART model estimated $Time_{LI3}$, on average, approximately .20 hours (i.e., 12 minutes) from the actual value using $Midpoint_{SA}$, Bed_{SA}, age, jet lag, and time of year; these variables accounted for 91% of the variance in $Time_{LI3}$.

Concentration Model. The RMSE for the concentration model was 10.90 and the R^2 was .50. The most important variables that were retained in the final model were, in decreasing order of relative importance, $Midpoint_{SA}$, weekend oversleep, $Wake_{SA}$, Bed_{SA}, sex, and $puberty_{adrenarche}$. Thus, the final concentration model included $Conc_{max}$ as the concentration outcome variable and $Midpoint_{SA}$, Bed_{SA}, age, jet lag, weekend oversleep, and time of year as predictor variables. In other words, the final concentration CART model estimated $Conc_{max}$, on average, approximately 10.90pg/mL from the actual value from $Midpoint_{SA}$, weekend oversleep, $Wake_{SA}$, Bed_{SA}, sex, and $puberty_{adrenarche}$; these variables accounted for 50% of the variance in $Conc_{max}$, which is a lot lower than the timing model.

Rate Model. The RMSE for the concentration model was 3.79 and the R^2 was .41. The most important variables retained in the final model were, in decreasing order of relative importance, Bed_{SA}, $Midpoint_{SA}$, $Wake_{SA}$, age, weekend oversleep, and jet lag. Thus, the final rate model included $Rate_{reg}$ as the rate outcome variable and Bed_{SA}, $Midpoint_{SA}$, $Wake_{SA}$, age, weekend oversleep, and jet lag as predictor variables. In other words, the final rate CART model estimated $Rate_{reg}$, on average, approximately 3.79pg/mL/hr from the actual value from Bed_{SA}, $Midpoint_{SA}$, $Wake_{SA}$, age, weekend oversleep, and jet lag; these variables accounted for 41% of the variance in $Rate_{reg}$, which is considerably lower than both the final timing and concentration models.

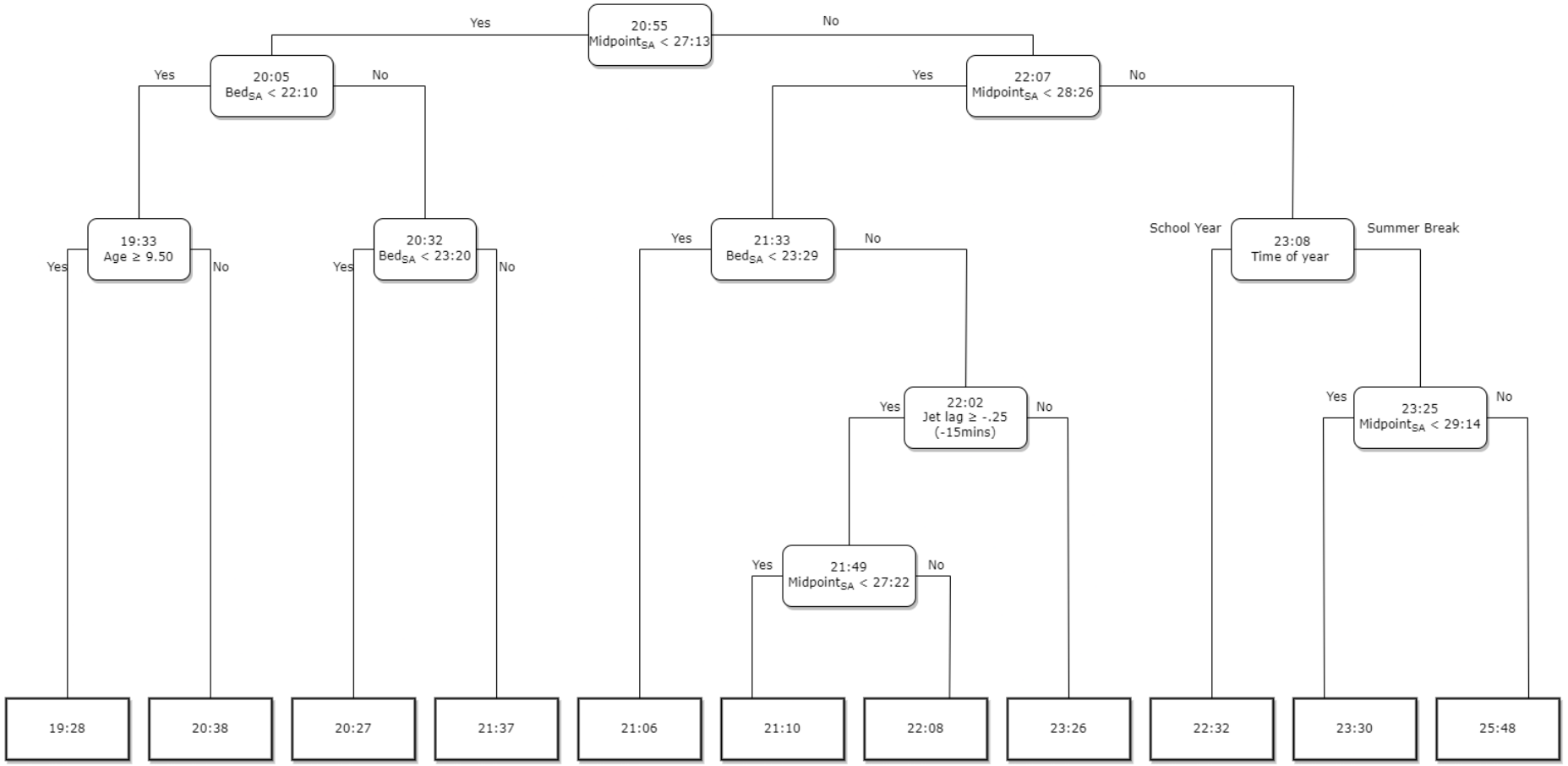

Figure 3. CART Timing Model.

Table 11. DLMO Dynamics CART Model Summaries

	Dynamic	Timing	Concentration	Rate
Model	R^2	.91	.50	.41
Performance	RMSE	.20	10.90	3.79
Variable	Age	12†	3	8†
Relative	Sex	-	6†	-
Importance	$Puberty_{adrenarche}$	-	2†	6†
(%)	BMI	-	-	-
	Time of year	5†	3	1
	Use of melatonin supplement	-	1	-
	Chronotype	4	2	-
	$Midpoint_{SA}$	27†	23†	21†
	Bed_{SA}	24†	16†	35†
	$Wake_{SA}$	16	19†	15†
	Weekend oversleep	1	20†	8†
	Jet lag	9†	5	7†

†These variables were retained in the final CART models generated.

Evaluating alternative DLMO proxies and metrics

As a reminder, the goal of this exploratory analysis was to assess whether *(1)* the final CART models could accurately estimate other metrics within the same dynamic, and *(2)* whether using different sleep assessment methodologies impact the estimations. The first goal was achieved by comparing the estimated values of a given model with the actual values of another metric within the same dynamic. Using the final timing model as an example, I compared the estimated DLMO timing values derived from the model with the actual values of all other timing metrics. We would expect to observe larger discrepancies (i.e., larger RMSEs) by doing this as each model was built to estimate a specific metric ($Time_{LI3}$ in the case of the timing model). However, the goal of this exploratory analysis was to simply observe how close the estimations were to other metrics, in the event that a clinician or researcher would want to estimate a different metric than the one used in the model. The second goal was achieved by using variables (bedtime, waketime, sleep midpoint) derived from alternative sleep measurement methodologies (i.e., sleep diary and questionnaire). In other words, instead of using the variables measured by actigraphy (Bed_{SA}, $Wake_{SA}$, $Midpoint_{SA}$), the final models were rerun using variables measured by a different sleep assessment methodology (e.g., Bed_{SD}, $Wake_{SD}$, $Midpoint_{SD}$). To illustrate how this was done conceptually, we can use Figure 3 as an example and replace the sleep assessment methodology used in the model from actigraphy to sleep diary. The root node of the model makes a split based on a $Midpoint_{SA}$ value of 27:13. Instead of using an individual's $Midpoint_{SA}$, we would use their $Midpoint_{SD}$ to decide which split to follow. If the individual's $Midpoint_{SD}$ is greater than 27:13, the following split would be based on their bedtime. Again, instead of using Bed_{SA}, we would use Bed_{SD} to decide which split to follow. This process was repeated for all models and for each sleep assessment methodology.

Only the resulting RMSEs are summarised in Table 12. Different combinations of alternative DLMO metrics and sleep assessment methodology were explored to determine whether a pattern was discernible. To note, the last column in Table 12 contains the standard deviations of each DLMO metric from the simulated data. As previously mentioned, standard

deviations can be used to assess whether a model is a good estimate of a given outcome variable; if the RMSE is less than the standard deviation, the model is typically considered accurate.

Alternative DLMO Metrics. For simplicity, the alternative DLMO metrics in the Actigraphy column of Table 12 are be described below. Similar patterns are found within the different sleep assessment methodology (sleep diary and questionnaire).

In the timing model, substituting the outcome variable ($Time_{LI3}$) for the other timing metrics ($Time_{LI4}$, $Time_{within}$, $Time_{reg3}$, $Time_{re\ Vg4}$, $Time_{max}$, $Time_{hockey}$), resulted in increased RMSEs, ranging from .53 (32 minutes) to 2.30 (138 minutes), indicating that the estimated values were less accurate. In fact, when comparing the RMSEs to their corresponding standard deviations, we can see that the RMSE was higher than the standard deviation for all timing metrics, except for $Time_{LI3}$ (which was used in the final model) and $Time_{LI4}$. Thus, the timing model produced good estimations for $Time_{LI4}$ only.

The concentration model also resulted in a higher RMSE when $Conc_{max}$ was replaced with $Conc_{AUC}$; which was also higher than the standard deviation. The model estimated, on average a $Conc_{AUC}$ value that was approximately 36.37pg/mL from the actual value, which is higher than the standard deviation for this metric (33.61). Thus, the concentration model did not produce a good estimation for $Conc_{AUC}$.

Finally, the rate model estimated on average $Rate_{hockey}$ values approximately within 6pg/mL/hr of the actual values. Considering $Rate_{hockey}$ values fluctuated between 0 and 0.8 and had a standard deviation of .02, this estimation was less precise.

Table 12. Accuracy of DLMO Dynamic Models with Proxy Variables and Metrics

		Root Mean Square Error (RMSE)			Melatonin Metric Standard Deviation
Dynamic	Metric	Sleep assessment Methodology			
		Actigraphy	Sleep Diary	Questionnaire	(for comparison)
Timing	**$Time_{LI3}$**	0.20	0.55	0.73	1.35
	$Time_{LI4}$	0.53	0.71	0.87	1.39
	$Time_{reg3}$	2.74	2.84	2.95	1.90
	$Time_{reg4}$	2.67	2.80	2.94	1.99
	$Time_{within}$	1.36	1.62	1.81	1.20
	$Time_{max}$	2.30	2.65	3.04	1.41
	$Time_{hockey}$	1.71	1.96	2.32	1.49
Concentration	**$Conc_{max}$**	10.90	16.84	17.12	15.81
	$Conc_{AUC}$	36.37	37.07	38.44	33.61
Rate	**$Rate_{reg}$**	3.79	5.25	5.51	4.77
	$Rate_{hockey}$	6.63	7.65	7.84	0.02

Note: bolded variables were the metrics and proxies included in the final models ($Time_{LI3}$, $Conc_{max}$, $Rate_{reg}$, Bed_{SA}, $Wake_{SA}$, $Midpoint_{SA}$).

Alternative Sleep Measurement Methodology. When comparing RMSEs across rows in Table 12, we can observe an increase of the RMSE, regardless of the metric. Additionally, all metrics, except $Time_{LI3}$ and $Time_{LI4}$, had RMSEs that were higher than their corresponding standard deviation for both sleep diary and questionnaire, indicating that the estimations are not nearly as

good as estimations derived with actigraphy. In all cases, actigraphy provided the most accurate estimations, followed by sleep diary, and finally questionnaire.

If we take the metric $Time_{LI3}$ as an example, using actigraphy as the sleep measurement methodology resulted in an RMSE of .20 hours (i.e., 12 minutes). When actigraphy was replaced with sleep diary, the RMSE increased to .55, indicating that on average, the model estimated $Time_{LI3}$ approximately .55 hours (33 minutes) from the actual value. Then, using questionnaire as the sleep measurement methodology instead resulted in an RMSE of .73 (i.e., on average, estimated values were approximately 44 minutes from the actual values). Thus, for $Time_{LI3}$, estimations increased on average by approximately 12 minutes when replacing a sleep measurement methodology with the subsequent one (i.e., from actigraphy to sleep diary, and from sleep diary to questionnaire) In fact, across all timing metrics, the estimation increased on average 12 minutes when going from actigraphy, to sleep diary, and finally to questionnaire. For example, if an individual used actigraphy as the sleep measurement methodology and a $Time_{LI3}$ was estimated to be 20:30, using a sleep diary instead would be expected to result in an estimated $Time_{LI3}$ of approximately 20:42, and 20:54 if a questionnaire was used.

In the concentration model, RMSEs also increased across the sleep measurement methodologies. However, unlike the timing model, there was no discernable pattern. For example, when going from actigraphy to sleep diary, the estimated $Conc_{max}$ increased by almost six points, but the difference between sleep diary and questionnaire was less than one.

For the rate model, a similar pattern to the concentration model was observed, where there was a bigger difference in RMSE when going from actigraphy to sleep diary (1.46 difference in $Rate_{reg}$ estimation), than from sleep diary to questionnaire (.26 difference). Based on these results, it seems that actigraphy is the best sleep assessment methodology to use for concentration and rate metrics. However, if actigraphy is not available, both sleep diary and questionnaire will produce similar estimations to one another. For example, if an individual used actigraphy as their sleep assessment methodology and had an estimated $Conc_{max}$ of 50pg/mL and a $Rate_{reg}$ of 2pg/mL/hr, using either a sleep diary or questionnaire instead would result in a $Conc_{max}$of approximately 56pg/mL and a $rate_{reg}$ of approximately 3.4pg/mL/hr.

Discussion

My thesis was organized into four sequential parts to address its overarching aim, which was to examine the quantification of melatonin from a biomathematical perspective and to consider whether alternative measures that are more easily assessed could be used as proxies of circadian timing. The ability to use proxy variables to accurately measure circadian timing would make an important scientific contribution because this would reduce the costs (time and financial) and burden of DLMO protocol procedures on participants and researchers or clinicians.

Scoping Review and Current State of the Literature

As evidenced by the results of the scoping review, limited information regarding melatonin in pediatric populations is available.

Demographics. Minimal information regarding sample demographics were reported in the reviewed studies. While all studies reported sex and age, few mentioned other variables that may affect melatonin production, such as puberty. In fact, only one study included information regarding the participants' pubertal status (Crowley & Carskadon, 2010), which used a physician to assess pubertal status. Only one study reported participants' race (Crowley & Carskadon, 2010) and this sample was comprised of three groups: Caucasians (43%), African-Americans (36%), and Other (21%). Lastly, sample sizes were typically small, with most comprised of

typically 50 to 60 participants, but sex distributions were usually equal across males and females; no studies described gender. Thus, based on the limited demographics reported in the literature, it may be pertinent to further explore the impact of these demographics on DLMO.

Protocol. For DLMO protocols, studies varied slightly in the duration of melatonin collection, but most started collection 1-2 hours before bedtime and up to 1 hour afterward. Most studies collected samples every 30 minutes and kept them in freezers until assaying. Nearly all studies in the scoping review (n=21, 91%) conducted melatonin saliva collection in a laboratory setting to control for lighting and ensure sample compliance. The amount of effort and resources required highlights the difficulty in measuring melatonin and the need to have simpler ways to accurately assess DLMO.

Circadian Markers. Just like the demographic variables, circadian markers (chronotype and midpoint) were scarcely reported in the literature, with only two studies including them.

Bedtime Routine Variables. Bedtime routine variables were commonly reported and some studies stratified data by school and weekend nights. However, no studies reported jet lag nor weekend oversleep, which are variables that provide additional information regarding the variability of an individual's sleep routine.

DLMO Timing. In the literature, melatonin is typically quantified with timing variables, with $Time_{LI4}$ being the most common. Researchers typically report an approximate fifteen minute difference between the timing metrics (i.e., $Time_{LI4}$, $Time_{LI3}$, $Time_{LIwithin}$; Crowley et al., 2016). Additionally, while the hockey-stick method has not yet been validated in children, Danilenko et al. (2014) found that there was approximately a 30-minute difference between DLMO metrics derived from the hockey-stick method ($Time_{Hockey}$) and those calculated using the 3pg/mL threshold ($Time_{LI3}$).

DLMO Concentration and Rate. Only one study identified in the scoping review reported melatonin concentration and used $Conc_{max}$ as the measure. Unfortunately, no studies have examined total melatonin output ($Conc_{AUC}$) as a measure or used a linear regression to estimate melatonin measures ($Time_{reg3}$, $Time_{reg4}$, $Rate_{reg}$). The results of the scoping review highlight the limited information about potential salient DLMO dynamics in the literature.

Comparing the Secondary Dataset to the Literature

Demographics. Demographics of the secondary dataset were largely similar to that reported in the literature in terms of age distribution and sex composition. Also, while not previously addressed in this thesis, the secondary dataset was comprised mostly of Caucasian participants. Not enough information is available in the literature to determine whether the racial composition of the secondary data is similar to other studies of similar caliber.

Protocol. The Healthy Heart DLMO protocol was nearly identical to studies included in the scoping review, with one important exception: it was conducted at <u>home</u>. Including light meters, blue-blocker eyewear, and compliance logs enabled the DLMO protocol to be completed at home for the Healthy Heart Project. Despite this difference in location of data collection (lab vs home), the traditional DLMO metrics calculated for this thesis were similar to those reported in the literature; this held true for both the complete sample and the fully compliant subsample.

Circadian Markers. The circadian marker values derived from the secondary dataset were similar to those reported in the literature.

Bedtime Routine Variables. Bedtimes and waketimes were similar between the secondary dataset and the literature. In fact, most studies used actigraphy to calculate bedtimes and waketimes. Additionally, as there was no information regarding weekend oversleep and jet lag, it is difficult to determine whether the Healthy Heart data was comparable to the literature.

DLMO Timing Metric. In the secondary dataset, I found approximately fifteen-minute differences between timing metrics (i.e., $Time_{LI4}$, $Time_{LI3}$, $Time_{LIwithin}$), which is similar to what has been reported in the literature. The difference between $Time_{Hockey}$ and $Time_{LI3}$ in the secondary data was approximately 1 hour, which was 30 minutes more than what was found in the literature.

DLMO Concentration and Rate Metrics. Only one study identified in the scoping review used $Conc_{max}$ as a measure (M=128.72, SD=8.16) and the values were a lot higher than those in the secondary data (M=30.17, SD=15.99). Unfortunately, the lack of DLMO concentration and rate metrics limit the interpretations we can make regarding the new metrics proposed ($Conc_{AUC}$, $Time_{reg3}$, $Time_{reg4}$, $Rate_{reg}$).

Variability in Melatonin Profile Observed in Dataset

The traditional DLMO metrics calculated with the secondary data were largely similar to those reported in the literature. However, as previously mentioned in Part 2, a portion of participants in the secondary dataset did not follow a typical rise in melatonin concentration throughout sample collection. Instead, these individuals had multiple melatonin peaks during the collection period. Variability in the melatonin profiles was not described in other studies (although this does not mean it did not exist). It is plausible that the variation observed in the secondary data was due to sampling compliance issues. For example, if the child did not consistently remain in the dimly-lit room or did not immediately freeze the saliva sample after collection, incorrect melatonin values may have occurred. However, even among the fully-compliant subsample, approximately 15% displayed fluctuating melatonin values during the collection period. Alternative explanations merit consideration. First, the 'typical' melatonin curve may not be 'typical' after all, especially in pediatric populations with limited evidence of DLMO protocols in the literature. Prior work has suggested that children have higher melatonin concentration values than adults, so it stands that the melatonin profile itself may differ than the 'typical' expected. Thus, using a concentration or rate metric may allow us to have a better understanding of the melatonin profile in pediatric populations as these metrics may offer additional information not captured by DLMO timing metrics. Second, certain individuals may be more at risk for having an atypical melatonin curve, especially those with sleep problems, such as parasomnias or insomnia (Goril et al., 2016). Participants in the Healthy Heart secondary dataset were not screened for sleep disorders, and some sleep problems, such as nightmares and night terrors, are more prevalent in children than in adults. Third, puberty may also alter melatonin concentrations. As youth go through puberty, melatonin secretions tend to decrease (Crowley & Carskadon, 2010). Future research should consider further investigating subsamples with atypical melatonin profiles.

Calculation of Traditional and Alternative Metrics for DLMO Dynamics

In this thesis, I proposed DLMO mathematical models that have not been previously used to quantify melatonin, specifically linear regression ($Time_{reg3}$, $Time_{reg4}$, $Rate_{reg}$) and total area under a curve ($Conc_{AUC}$). I extended the use of the Hockey-Stick method ($Time_{hockey}$, $Rate_{hockey}$) to children, which was originally proposed by Danilenko et al. (2013) in adults and has not yet been validated in children. Additionally, I considered metrics relating to the maximum melatonin concentration ($Time_{max}$, $Conc_{max}$).

While $Time_{reg3}$ and $Time_{reg4}$ were highly correlated with the other traditional timing metrics ($Time_{LI3}$, $Time_{LI4}$), there was over an hour difference between these two proposed timing metrics and the traditional metrics, despite using the same thresholds (i.e., 3 and 4pg/mL). In the sleep

literature, an hour is an extremely large span of time. (Consider, previous studies comparing bedtimes on weight status found approximately a six-minute difference in bedtimes for one pound of weight; Jarrin et al., 2013). Linear regression may therefore not be an ideal calculation to capture DLMO timing. $Time_{hockey}$ also was not as strongly associated with the timing metrics, suggesting that the Hockey-stick method may need to be reconsidered for pediatric populations. Alternatively, it is possible that the portion of our secondary subsample with fluctuating melatonin values may have affected the rate metrics calculations. It may be pertinent to assess whether removing these participants "cleans" the rate metrics.

$Conc_{AUC}$ was highly correlated with $Conc_{max}$, and surprisingly, $Rate_{reg}$ as well. The high correlations between the concentration and rate metrics suggests that rate and concentration may be linked. A possible explanation for this association may be mathematical given that the maximum melatonin sample ($Conc_{max}$) could be driving rate that is derived by linear regression ($Rate_{reg}$). Another explanation may be that concentration and rate metrics are not separate and may measure the same construct. $Rate_{hockey}$ was not strongly associated with any other DLMO metric. It is worth noting that while the hockey-stick method calculates the rate of melatonin surge, the researchers who developed the method (Danilenko et al., 2013) only used the time of inflection ($Time_{hockey}$) as a melatonin measure and did not specify the relevance of the rate calculated. Thus, further studies should be done to better understand DLMO concentration and rate metrics. For instance, a factor analysis or latent variable analyses could be used to further evaluate whether there are three unique DLMO dynamics or whether there are only two.

Estimating DLMO Dynamics Using Proxy Variables

Three CART models were generated, one for each DLMO dynamic (timing, concentration, rate). While the same predictor variables were included in the candidate models (age, sex, $puberty_{adrenarche}$, BMI, time of year, use of melatonin supplements, chronotype, $Midpoint_{SA}$, Bed_{SA}, $Wake_{SA}$, weekend oversleep, jet lag), different combinations of variables were retained in each model. As anticipated, $Midpoint_{SA}$ was retained in all three models and can be considered a good proxy for all DLMO dynamics. However, much to our surprise, chronotype was not a significant predictor for the different melatonin dynamics, and was not retained in the final model. I anticipated chronotype to be a more robust proxy for DLMO but the observed results suggest that an individual's preference for time of day was not associated with their melatonin profile. It is possible that even though a child has a certain preferred chronotype, the rigidity of their routine and school schedule does not afford them the flexibility to go to bed or wake up as they would like. In adults, chronotype has been found to be reliably associated with melatonin (Sun et al., 2020). Finally, other demographic variables and covariates, such as BMI and use of melatonin supplements, were not retained in the final models.

Timing Model. The timing model had the best fit out of the three proposed. Curiously, while $Wake_{SA}$ had a high relative importance value, it was not included in the final timing model. This occurs in CART analysis when a variable is a surrogate of another variable, meaning that it is highly correlated with another predictor variable and does not add additional information to the model that is not already attributed to the other variable it is correlated with. In this case, $Wake_{SA}$ was highly correlated with Bed_{SA} and $Midpoint_{SA}$. Thus, based on the results of the regression tree, one only needs Bed_{SA}, $Midpoint_{SA}$, jet lag, time of year, and age to accurately estimate a timing dynamic. In fact, the model estimated on average a timing value within 12 minutes of the actual value.

Concentration Model. The proxy variables included in the model only accounted for 50% of the variance in the concentration metric, which signifies that perhaps other proxy variables may

estimate DLMO concentration metrics better. To note, sex and puberty were identified as important variables in the model, which may suggest an interaction between sex and puberty. While puberty is known to after melatonin levels (Crowley & Eastman, 2018), no studies identified addressed sex-differences in melatonin concentrations. It may therefore be pertinent to explore the association between sex and melatonin concentration. Additionally, weekend oversleep was considered an important variable in the model, suggesting that variable wake times may have an impact on melatonin concentrations.

Rate Model. The predictor variables only accounted for 41% of the variance in the rate metric, which means that there may be other variables that affect rate. The variables identified as important in the rate model were similar to those included in the concentration model. The models, along with the correlations described earlier suggest that rate and concentration may be tapping into similar constructs.

Exploratory Alternative Metrics and Proxy Measures for DLMO

Alternative Sleep Assessment Methodology. Overall, models with sleep variables measured by actigraphy produced the most accurate estimates, followed by sleep diary and questionnaire sleep variables respectively. This result supports current findings in the literature about the accuracy of actigraphy, compared to self-report measures (Short et al., 2012). Interestingly, for the timing metrics, there was on average a 12-minute difference in the estimated values between each subsequent sleep assessment methodology (i.e., from actigraphy, to sleep diary, to questionnaire), potentially indicating that the loss of accuracy is consistent across all three methodologies. In contrast, for both concentration and rate metrics, specifically $Conc_{max}$ and $Rate_{reg}$, the difference in model accuracy (RMSE) was greater when moving from actigraphy to sleep diary, than from sleep diary to questionnaire. These results may indicate that if actigraphy is not available, using either sleep diary or questionnaire will result in similar estimations. Thus, actigraphy should be used for the most accurate estimates but depending on the precision required by the researcher or clinician (e.g., if they don't mind a 12-minute difference in time or a 6pg/mL difference in concentration), sleep diary or questionnaire may still be adequate to use.

Other Timing Metrics. The timing model generated from the CART analysis most accurately estimated $Time_{LI4}$, where estimated values were approximately 32 minutes from the actual value on average. This result is expected, given the high correlation between $Time_{LI3}$ and $Time_{LI4}$. Additionally, both metrics were derived using the same mathematical model. The timing model was less accurate in measuring all the other timing variables, where the estimations ranged, on average, between one and three hours from the actual value.

Other Concentration Metrics. The concentration model was able to estimate $Conc_{AUC}$ within 36.37pg/mL, on average, of the actual value. While this estimation was greater than the standard deviation of $Conc_{AUC}$ (33.61pg/mL), and would typically be considered a poor estimation, we should consider whether a difference of approximately 3pg/mL is clinically relevant. Based on these results, the concentration model was able to estimate $Conc_{AUC}$ within about one standard deviation. As such, depending on the individual's needs, this estimation may still be pertinent and accurate.

Other Rate Metrics. The rate model did not accurately estimate $Rate_{hockey}$; the RMSE was 6.63, which is large considering $Rate_{hockey}$ values ranged from 0 and .8 and the standard deviation was .01. These results further support a previous statement that the Hockey-Stick method may not be appropriate for pediatric populations.

Finally, it should be noted that DLMO metrics are used as continuous variables throughout the literature. However, it may be worth exploring whether clinical cut-offs exist within each

dynamic. This thesis treated all DLMO metrics as continuous variables but it might be more relevant to clinicians and researchers if results of regression trees were categories (e.g., DLMO timing metrics split into early, normal, or late timing; DLMO concentration metrics split into low, medium or high concentrations; DLMO rate metrics split into flat or steep rates). Doing so may have a more practical use for clinicians or researchers who may want to easily categorise patients or participants.

Study Limitations and Strengths

Five limitations and a corresponding strength are described below.

The first and main limitation of this study was the lack of endogenous circadian markers that follow a diurnal pattern similar to that of melatonin, such as temperature (Reid, 2019). Core body temperatures are higher during the day and decrease during sleep (Van Someren & Nagtegaal, 2007). Core body temperature is typically measured using an anal probe or internal ear thermistor. However, advances in measurement permit the assessment of peripheral temperature (e.g., thermistor on inner wrist surface) that can be recorded easily without disrupting a person's daily routine. Peripheral temperature is more easily affected by ambient temperature and environmental factors; thus, there is greater error or noise when using peripheral body temperature (i.e., difficult to measure accurately in a naturalistic setting). It is worth mentioning again that while sleep-wake patterns (e.g., bedtime, waketime) are linked to circadian rhythms, they are not equivalent. Circadian rhythms are controlled by an internal rhythm governed by the SCN pacemaker, whereas sleep-wake cycles are also greatly influenced by external demands. Thus, to validate measures of melatonin, it is necessary to include other physiological markers of circadian timing (e.g., temperature), in addition to daily sleep measures (e.g., bedtime, wake time), and personal chronotype preference to triangulate circadian rhythm. To date (and to the best of my knowledge), no studies have examined the alignment of multiple circadian markers. Greater alignment across markers would provide necessary evidence to support plausible alternative methods to quantify melatonin. However, a strength of this study was the use of sleep midpoint as a circadian marker, which was a fairly robust proxy measure and is easy to collect, even more so than temperature.

A second study limitation was how sample compliance was managed. All participants with non-compliant samples were removed from the sample. However, it may have been interesting to examine whether the different reasons for non-compliance were linked to differences in the resulting DLMO metrics. For instance, perhaps the participants who fell asleep before completing the protocol had steeper rate metrics, which prompted them to sleep sooner. However, compliance can also be viewed as a study strength at 70% of our sample was compliant, which is impressive given that participants were asked to complete the DLMO protocol at home, without the assistance or supervision of a trained research assistant. In fact, of all the studies from the scoping review, only one had DLMO protocols completed at home (Higuchi et al., 2016) this study did not report any measures of compliance. Additionally, this study had a much smaller sample size (41) relative to our secondary dataset, even when all non-compliant samples were removed. Thus, while laboratory settings may increase compliance, home studies tend to be more ecologically valid (Burgess et al., 2015).

The third limitation of this thesis was the use of Pearson correlations. Instead, partial correlations, controlling for key covariates (e.g., age, sex, puberty), may have been more useful for the analyses. It is possible that some of the associations expected to be observed would have been observed if partial correlations had been used. For example, based on past literature,

chronotype was expected to be highly correlated with the bedtime routine variables (Stone et al., 2021). However, chronotype was not as strongly correlated with any variable.

The fourth limitation pertains to the CART analysis. While CART models are robust against non-normal data, they are known to overfit data (Hayes et al., 2015). The models may therefore not perform as well if we test them on a different dataset. Random forest is another method that also creates decision trees and is typically more accurate than CART. However, Random Forests are less intuitive, and the results are less interpretable (Hayes et al., 2015). As such, for the purpose of this thesis, a CART analysis was deemed appropriate.

Fifth, while the secondary data was comparable to other studies, our sample consisted of predominantly Caucasian participants. In fact, most melatonin studies have been done on Caucasian participants. Few studies have been done to examine racial differences in melatonin production in adult populations and some researchers suggest that there are racial differences. For instance, Bhatti et al. (2013) found that Asian nightshift workers had more stable melatonin profiles compared to Caucasian workers. However, to our knowledge, no such studies have been done in pediatric populations. Additionally, while it was not assessed in this study, diet has been shown to affect melatonin production, particularly if certain foods contain elements that are precursors of melatonin (Peuhkuri et al., 2012). As such, if certain foods are easily accessible in a certain region and part of a typical diet, melatonin levels may be higher in those individuals. Thus, examining racial differences may help develop some of the melatonin measures explored in this thesis, particularly the concentration and rate measures. However, a strength of this study was the inclusion of variables, particularly puberty, time of year, jet lag, and weekend overlseep (i.e., school year vs summer break) that are known to affect melatonin production (Crowley et al., 2006, 2014; Crowley & Carskadon, 2010) but are rarely controlled for in pediatric studies. In fact, jet lag especially was not reported in any of the studies included in the scoping review, yet it was an important variable in the regression tree models.

Finally, a strength of this thesis was the exploration of new melatonin metrics. While more studies still need to be done to better understand and refine the metrics proposed, I examined novel ways of quantifying melatonin to capture information that may otherwise missed when we focus solely on one or two metrics. The work done in this thesis can therefore potentially be a starting point for future studies aiming to understand different dimensions of melatonin.

Implications and Future Research

Findings from this thesis highlighted the lack of research on melatonin quantification in pediatric populations. While there is currently a broad interest in sleep and circadian timing within health research, the existing methodology primarily relies on metrics designed for adult populations. To our knowledge, this is the first study that attempted to identify different melatonin metrics used in pediatric populations and estimate those metrics with proxy variables, such as bedtime, sleep midpoint, and other covariates.

Most studies focus on timing melatonin measures. More studies should explore the other melatonin dynamics suggested (concentration, rate). The models generated to estimate the concentration and rate variables were not as accurate as the timing model, suggesting that other variables not included in this thesis may be better proxy measures for the concentration and rate metrics. Examining other metrics to quantify melatonin that do not centre solely on timing will improve our understanding of the link between child and adolescent health and circadian misalignment, which has implications on sleep research and can update our understanding of sleep in youth, and effectively informing current policies on sleep (e.g., recommended sleep duration).

Additionally, as mentioned in the limitations section, more studies should be done to explore the possible effect of race or ethnicity on melatonin production in children. Doing so can allow us to have more precise measures of melatonin and circadian timing. In fact, developing a more accurate and precise metric for circadian phase has scalable implications for health technology, particularly the signal processing used with commercially available wearable technology (e.g., smart watches), which have exponentially increased in popularity and use among the public.

Finally, one of the goals of this thesis was to propose proxy variables to estimate DLMO without having to collect melatonin, which may be a solution for clinicians and researchers who may not have the time or resources to complete the DLMO protocol. The timing model was fairly accurate and only required five proxy variables. Thus, if a researcher or clinician wants to estimate an individual's DLMO timing dynamic, they would simply need to know the person's bedtime, sleep midpoint (which can be calculated from bedtime and waketime), their jet lag (again, can be calculated from bedtime) whether they are in school or not, and their age. While actigraphy provides the most accurate bed- and wake-times to estimate DLMO timings, sleep diaries and questionnaires, which are much less intensive, provide adequate proxies as well, depending on the individual's needs. While the concentration and rate models were not as accurate as the timing model, it is worth noting that both models also did not require many variables to estimate their respective DLMO dynamic. For concentration, bedtime, waketime, midpoint, weekend oversleep, sex and pubertal status can be used to estimate the concentration dynamic, these same variables, along with age and jet lag can be used to estimate the rate dynamic. However, with R^2 being lower for both models, it may be pertinent to explore whether other variables can also serve as proxies for concentration (e.g., diet) and rate (e.g., temperature fluctuations) dynamics.

Conclusion

This thesis had two overarching goals: first, to identify the ways melatonin is measured in pediatric populations and apply them in a secondary dataset; and second, to estimate melatonin measures with proxy variables, such as self-reported bed- and wake-times. Three types of melatonin dynamics were proposed: timing, concentration, and rate. Traditional melatonin metrics were identified from a scoping review ($Time_{LI3}$, $Time_{LI4}$, $Time_{within}$, $Time_{max}$, $Time_{hockey}$, $Conc_{max}$, $Rate_{hockey}$) and alternative metrics derived from two mathematical models, linear regression and area under the curve, were proposed ($Time_{reg3}$, $Time_{reg4}$, $Conc_{AUC}$, $Rate_{reg}$). Proxy variables (e.g., bedtime, age, sex, chronotype) were used to estimate DLMO dynamics using classification and regression tree analysis. It was found that same variables could be used as proxies for the different DLMO dynamics (e.g., sleep midpoint) but there are likely other proxy variables not included in the present thesis that could be used to estimate the dynamics, particularly concentration and rate dynamics. Future research should focus on further examining and validating alternative melatonin dynamics in pediatric settings, such as concentration and rate, and additional proxy variables not included in this thesis, such as temperature and race. Findings from this thesis can help improve our understanding of melatonin in children and adolescents, and how it is linked to circadian rhythms and health.

Bibliography

Abbott, S. M., Malkani, R. G., & Zee, P. C. (2020). Circadian disruption and human health: A bidirectional relationship. *European Journal of Neuroscience*, *51*(1), 567–583. https://doi.org/10.1111/ejn.14298

Agostini, A., Carskadon, M. A., Dorrian, J., Coussens, S., & Short, M. A. (2017). An experimental study of adolescent sleep restriction during a simulated school week: Changes in phase, sleep staging, performance and sleepiness. *Journal of Sleep Research*, *26*(2), 227–235. https://doi.org/10.1111/jsr.12473

Akacem, L. D., Simpkin, C. T., Carskadon, M. A., Wright, K. P., Jr, Jenni, O. G., Achermann, P., & LeBourgeois, M. K. (2015). The Timing of the Circadian Clock and Sleep Differ between Napping and Non-Napping Toddlers. *PloS One*, *10*(4), e0125181. MEDLINE. https://doi.org/10.1371/journal.pone.0125181

Baron, K. G., & Reid, K. J. (2014). Circadian misalignment and health. *International Review of Psychiatry*, *26*(2), 139–154. https://doi.org/10.3109/09540261.2014.911149

Baron, K. G., Reid, K. J., Kim, T., Van Horn, L., Attarian, H., Wolfe, L., Siddique, J., Santostasi, G., & Zee, P. C. (2017). Circadian timing and alignment in healthy adults: Associations with BMI, body fat, caloric intake and physical activity. *International Journal of Obesity (2005)*, *41*(2), 203–209. PubMed. https://doi.org/10.1038/ijo.2016.194

Benloucif, S., Burgess, H. J., Klerman, E. B., Lewy, A. J., Middleton, B., Murphy, P. J., Parry, B. L., & Revell, V. L. (2008). Measuring Melatonin in Humans. *Journal of Clinical Sleep Medicine : JCSM : Official Publication of the American Academy of Sleep Medicine*, *4*(1), 66–69. https://doi.org/10.5664/jcsm.27083

Benloucif, S., Guico, M., Reid, K., Wolfe, L., L'Hermite, M., & Zee, P. (2005). Stability of Melatonin and Temperature as Circadian Phase Markers and Their Relation to Sleep Times in Humans. *Journal of Biological Rhythms*, *20*, 178–188. https://doi.org/10.1177/0748730404273983

Bhatti, P., Mirick, D. K., & Davis, S. (2013). Racial Differences in the Association Between Night Shift Work and Melatonin Levels Among Women. *American Journal of Epidemiology*, *177*(5), 388–393. https://doi.org/10.1093/aje/kws278

Burgess, H. J., & Fogg, L. F. (2008). Individual differences in the amount and timing of salivary melatonin secretion. *PloS One*, *3*(8).

Burgess, H. J., Wyatt, J. K., Park, M., & Fogg, L. F. (2015). Home Circadian Phase Assessments with Measures of Compliance Yield Accurate Dim Light Melatonin Onsets. *Sleep*, *38*(6), 889–897. MEDLINE. https://doi.org/10.5665/sleep.4734

Carney, C. E., Buysse, D. J., Ancoli-Israel, S., Edinger, J. D., Krystal, A. D., Lichstein, K. L., & Morin, C. M. (2012). The consensus sleep diary: Standardizing prospective sleep self-monitoring. *Sleep*, *35*(2), 287–302. https://doi.org/10.5665/sleep.1642

Carskadon, M. A., Vieira, C., & Acebo, C. (1993). Association between puberty and delayed phase preference. *Sleep*, *16*(3), 258–262. https://doi.org/10.1093/sleep/16.3.258

Carskadon, M. A., Wolfson, A. R., Acebo, C., Tzischinsky, O., & Seifer, R. (1998). Adolescent Sleep Patterns, Circadian Timing, and Sleepiness at a Transition to Early School Days. *Sleep*, *21*(8), 871–881. https://doi.org/10.1093/sleep/21.8.871

Claustrat, B., Brun, J., & Chazot, G. (2005). The basic physiology and pathophysiology of melatonin. *Sleep Medicine Reviews*, *9*(1), 11–24. https://doi.org/10.1016/j.smrv.2004.08.001

Crowley, S. J., Acebo, C., Fallone, G., & Carskadon, M. A. (2006). Estimating dim light melatonin onset (DLMO) phase in adolescents using summer or school-year sleep/wake schedules. *Sleep*, *29*(12), 1632–1641. MEDLINE.

Crowley, S. J., & Carskadon, M. A. (2010). Modifications to weekend recovery sleep delay circadian phase in older adolescents. *Chronobiology International*, *27*(7), 1469–1492. https://doi.org/10.3109/07420528.2010.503293

Crowley, S. J., & Eastman, C. I. (2017). Human adolescent phase response curves to bright white light. *Journal of Biological Rhythms*, *32*(4), 334–344. https://doi.org/10.1177/0748730417713423

Crowley, S. J., & Eastman, C. I. (2018). Free-running circadian period in adolescents and adults. *Journal of Sleep Research*, *27*(5), 1–8. https://doi.org/10.1111/jsr.12678

Crowley, S. J., Fournier, C. L., & Eastman, C. I. (2018). Late bedtimes prevent circadian phase advances to morning bright light in adolescents. *Chronobiology International*, *35*(12), 1748–1752. https://doi.org/10.1080/07420528.2018.1504784

Crowley, S. J., Suh, C., Molina, T. A., Fogg, L. F., Sharkey, K. M., & Carskadon, M. A. (2016). Estimating the dim light melatonin onset of adolescents within a 6-h sampling window: The impact of sampling rate and threshold method. *Sleep Medicine*, *20*, 59–66. MEDLINE. https://doi.org/10.1016/j.sleep.2015.11.019

Crowley, S. J., Van Reen, E., LeBourgeois, M. K., Acebo, C., Tarokh, L., Seifer, R., Barker, D. H., & Carskadon, M. A. (2014). A Longitudinal Assessment of Sleep Timing, Circadian Phase, and Phase Angle of Entrainment across Human Adolescence. *PLoS ONE*, *9*(11), 1–11. https://doi.org/10.1371/journal.pone.0112199

Danilenko, K., Verevkin, E., Antyufeev, V., Wirz-Justice, A., & Cajochen, C. (2013). Hockey-stick method to estimate evening dim light melatonin onset (DLMO). *Chronobiology International*, *31*. https://doi.org/10.3109/07420528.2013.855226

Dolsen, M. R., & Harvey, A. G. (2018). Dim Light Melatonin Onset and Affect in Adolescents With an Evening Circadian Preference. *The Journal of Adolescent Health : Official Publication of the Society for Adolescent Medicine*, *62*(1), 94–99. PubMed. https://doi.org/10.1016/j.jadohealth.2017.07.019

Dolsen, M. R., Wyatt, J. K., & Harvey, A. G. (2019). Sleep, circadian rhythms, and risk across health domains in adolescents with an evening circadian preference. *Journal of Clinical Child and Adolescent Psychology*, *48*(3), 480–490. https://doi.org/10.1080/15374416.2017.1416620

Dorn, L. D., Susman, E. J., Nottelmann, E. D., Inoff-Germain, G., & Chrousos, G. P. (1990). Perceptions of puberty: Adolescent, parent, and health care personnel. *Developmental Psychology*, *26*(2), 322–329. https://doi.org/10.1037/0012-1649.26.2.322

Figueiro, M. G., & Rea, M. S. (2010). Evening daylight may cause adolescents to sleep less in spring than in winter. *Chronobiology International*, *27*(6), 1242–1258. https://doi.org/10.3109/07420528.2010.487965

Golding, J., Pembrey, M., Jones, R., & ALSPAC Study Team. (2001). ALSPAC--the Avon Longitudinal Study of Parents and Children. I. Study methodology. *Paediatric and Perinatal Epidemiology*, *15*(1), 74–87. https://doi.org/10.1046/j.1365-3016.2001.00325.x

Goldman, S. E., Adkins, K. W., Calcutt, M. W., Carter, M. D., Goodpaster, R. L., Wang, L., Shi, Y., Burgess, H. J., Hachey, D. L., & Malow, B. A. (2014). Melatonin in children with autism spectrum disorders: Endogenous and pharmacokinetic profiles in relation to sleep. *Journal of Autism and Developmental Disorders*, *44*(10), 2525–2535. https://doi.org/10.1007/s10803-014-2123-9

Goril, S., Zalai, D., Scott, L., & Shapiro, C. M. (2016). Sleep and melatonin secretion abnormalities in children and adolescents with fetal alcohol spectrum disorders. *Sleep Medicine*, *23*, 59–64. https://doi.org/10.1016/j.sleep.2016.06.002

Hayes, T., Usami, S., Jacobucci, R., & McArdle, J. J. (2015). Using Classification and Regression Trees (CART) and Random Forests to Analyze Attrition: Results From Two Simulations. *Psychology and Aging*, *30*(4), 911–929. https://doi.org/10.1037/pag0000046

Higuchi, S., Lee, S., Kozaki, T., Harada, T., & Tanaka, I. (2016). Late circadian phase in adults and children is correlated with use of high color temperature light at home at night. *Chronobiology International*, *33*(4), 448–452. https://doi.org/10.3109/07420528.2016.1152978

Horne, J. A., & Östberg, O. (1976). A self-assessment questionnaire to determine morningness-eveningness in human circadian rhythms. *International Journal of Chronobiology*, *4*, 97–110.

James, H., King, G., & Blackwell, M. (2019). *Amelia: A program for missing data* (1.7.6) [R package]. https://cran.r-project.org/web/packages/Amelia/index.html

Jarrin, D. C., McGrath, J. J., & Drake, C. L. (2013). Beyond sleep duration: Distinct sleep dimensions are associated with obesity in children and adolescents. *International Journal of Obesity (2005)*, *37*(4), 552–558. https://doi.org/10.1038/ijo.2013.4

Klerman, E. B. (2005). Clinical Aspects of Human Circadian Rhythms. *Journal of Biological Rhythms*, *20*(4), 375–386. https://doi.org/10.1177/0748730405278353

Koscec, A., Radosevic-Vidacek, B., & Bakotic, M. (2014). Morningness-eveningness and sleep patterns of adolescents attending school in two rotating shifts. *Chronobiology International*, *31*(1), 52–63. https://doi.org/10.3109/07420528.2013.821128

Kuczmarski, R. J., Ogden, C. L., Guo, S. S., Grummer-Strawn, L. M., Flegal, K. M., Mei, Z., Wei, R., Curtin, L. R., Roche, A. F., & Johnson, C. L. (2002). 2000 CDC Growth Charts for the United States: Methods and development. *Vital and Health Statistics. Series 11, Data from the National Health Survey*, *246*, 1–190.

Laberge, L., Carrier, J., Lespérance, P., Lambert, C., Vitaro, F., Tremblay, R. E., & Montplaisir, J. (2000). Sleep and circadian phase characteristics of adolescent and young adult males in a naturalistic summertime condition. *Chronobiology International*, *17*(4), 489–501. https://doi.org/10.1081/CBI-100101059

Lamprecht, M. L., Bradley, A. P., Tran, T., Boynton, A., & Terrill, P. I. (2015). Multisite accelerometry for sleep and wake classification in children. *Physiological Measurement*, *36*(1), 133–147. https://doi.org/10.1088/0967-3334/36/1/133

LeBourgeois, M. K., Carskadon, M. A., Akacem, L. D., Simpkin, C. T., Wright, K. P., Achermann, P., & Jenni, O. G. (2013). Circadian Phase and Its Relationship to Nighttime Sleep in Toddlers. *Journal of Biological Rhythms*, *28*(5), 322–331. https://doi.org/10.1177/0748730413506543

Lewy, A. J., Cutler, N. L., & Sack, R. L. (1999). The Endogenous Melatonin Profile as a Marker for Circadian Phase Position. *Journal of Biological Rhythms*, *14*, 227–236. https://doi.org/10.1177/074873099129000641

Lewy, A. J., Emens, J., Jackman, A., & Yuhas, K. (2006). Circadian Uses of Melatonin in Humans. *Chronobiology International: The Journal of Biological & Medical Rhythm Research*, *23*(1/2), 403–412. Academic Search Ultimate.

Lewy, A. J., Wehr, T. A., Goodwin, F. K., Newsome, D. A., & Markey, S. P. (1980). Light suppresses melatonin secretion in humans. *Science (New York, N.Y.)*, *210*(4475), 1267–1269. https://doi.org/10.1126/science.7434030

Mandrell, B. N., Avent, Y., Walker, B., Loew, M., Tynes, B. L., & Crabtree, V. M. (2018). In-home salivary melatonin collection: Methodology for children and adolescents. *Developmental Psychobiology*, *60*(1), 118–122.

McGrath, J. J., Noel, N., & Burdayron, R. (2018). Development and validation of an actigraphy scoring protocol for sleep to define lights off/on. *Psychophysiology*, *55*(S1), S60.

Meltzer, L. J., Montgomery-Downs, H. E., Insana, S. P., & Walsh, C. M. (2012). Use of actigraphy for assessment in pediatric sleep research. *Sleep Medicine Reviews*, *16*(5), 463–475. https://doi.org/10.1016/j.smrv.2011.10.002

Molina, T. A., & Burgess, H. J. (2011). Calculating the Dim Light Melatonin Onset: The Impact of Threshold and Sampling Rate. *Chronobiology International*, *28*(8), 714–718. https://doi.org/10.3109/07420528.2011.597531

Morris, N. M., & Udry, J. R. (1980). Validation of a self-administered instrument to assess stage of adolescent development. *Journal of Youth and Adolescence*, *9*(3), 271–280. https://doi.org/10.1007/BF02088471

Netherton, C., Goodyer, I., Tamplin, A., & Herbert, J. (2004). Salivary cortisol and dehydroepiandrosterone in relation to puberty and gender. *Psychoneuroendocrinology*, *29*(2), 125–140. https://doi.org/10.1016/S0306-4530(02)00150-6

Pandi-Perumal, S. R., Smits, M., Spence, W., Srinivasan, V., Cardinali, D. P., Lowe, A. D., & Kayumov, L. (2007). Dim light melatonin onset (DLMO): A tool for the analysis of circadian phase in human sleep and chronobiological disorders. *Progress in Neuro-Psychopharmacology & Biological Psychiatry*, *31*(1), 1–11. Academic Search Ultimate.

Parsons, M. J., Moffitt, T. E., Gregory, A. M., Goldman-Mellor, S., Nolan, P. M., Poulton, R., & Caspi, A. (2015). Social jetlag, obesity and metabolic disorder: Investigation in a cohort study. *International Journal of Obesity (2005)*, *39*(5), 842–848. https://doi.org/10.1038/ijo.2014.201

Peixoto, C. A. T., da Silva, A. G. T., Carskadon, M. A., & Louzada, F. M. (2009). Adolescents living in homes without electric lighting have earlier sleep times. *Behavioral Sleep Medicine*, *7*(2), 73–80. https://doi.org/10.1080/15402000902762311

Peuhkuri, K., Sihvola, N., & Korpela, R. (2012). Dietary factors and fluctuating levels of melatonin. *Food & Nutrition Research*, *56*, 10.3402/fnr.v56i0.17252. https://doi.org/10.3402/fnr.v56i0.17252

Pham, M. T., Rajić, A., Greig, J. D., Sargeant, J. M., Papadopoulos, A., & McEwen, S. A. (2014). A scoping review of scoping reviews: Advancing the approach and enhancing the consistency. *Research Synthesis Methods*, *5*(4), 371–385. https://doi.org/10.1002/jrsm.1123

Praninskienė, R., Dumalakienė, I., Kemežys, R., Mauricas, M., & Jučaitė, A. (2012). Diurnal melatonin patterns in children: Ready to apply to practice? *Pediatric Neurology*, *46*(2), 70–76. https://doi.org/10.1016/j.pediatrneurol.2011.11.018

Ramil, J. (2018). *Placement of actigraph and measurement of sleep: A psychometric comparison in children*. Concordia University.

Reid, K. J. (2019). Assessment of Circadian Rhythms. *Neurologic Clinics*, *37*(3), 505–526. https://doi.org/10.1016/j.ncl.2019.05.001

Robeva, R., Kirkwood, J. R., Davies, R. L., Farhy, L., Kovatchev, B., Straume, M., & Johnson, M. L. (2008). *An invitation to biomathematics*. Elsevier.

Roemer, H. C., Griefahn, B., Kuenemund, C., Blaszkewicz, M., & Gerngroβ, H. (2003). The Reliability of Melatonin Synthesis as an Indicator of the Individual Circadian Phase Position. *Military Medicine*, *168*(8), 674–678. https://doi.org/10.1093/milmed/168.8.674

Short, M. A., Gradisar, M., Lack, L. C., Wright, H., & Carskadon, M. A. (2012). The discrepancy between actigraphic and sleep diary measures of sleep in adolescents. *Sleep Medicine*, *13*(4), 378–384. https://doi.org/10.1016/j.sleep.2011.11.005

Stone, J. E., Phillips, A. J. K., Chachos, E., Hand, A. J., Lu, S., Carskadon, M. A., Klerman, E. B., Lockley, S. W., Wiley, J. F., Bei, B., & Rajaratnam, S. M. W. (2021). In-person vs home schooling during the COVID-19 pandemic: Differences in sleep, circadian timing, and mood in early adolescence. *Journal of Pineal Research*, *71*(2), e12757. https://doi.org/10.1111/jpi.12757

Stone, K. L., & Ancoli-Israel, S. (2017). *Actigraphy* (6th ed.). Elsevier. https://doi.org/10.1016/B978-0-323-24288-2.00174-4

Sun, X., Gustat, J., Bertisch, S., Redline, S., & Bazzano, L. (2020). The association between sleep chronotype and obesity among black and white participants of the Bogalusa Heart Study. *Chronobiology International*, *37*(1), 123–134. https://doi.org/10.1080/07420528.2019.1689398

Taylor, D. J., Jenni, O. G., Acebo, C., & Carskadon, M. A. (2005). Sleep tendency during extended wakefulness: Insights into adolescent sleep regulation and behavior. *Journal of Sleep Research*, *14*(3), 239–244. https://doi.org/10.1111/j.1365-2869.2005.00467.x

van der Heijden, K. B., Blok, M. J., Spee, K., Archer, S. N., Smits, M. G., Curfs, L. M., & Gunning, W. B. (2005). No evidence to support an association of PER3 clock gene polymorphism with ADHD-related idiopathic chronic sleep onset insomnia. *Biological Rhythm Research*, *36*(5), 381–388. https://doi.org/10.1080/09291010500218449

Van Maanen, A., Meijer, A. M., Smits, M. G., & Oort, F. J. (2011). Melatonin and sleep effects on health, behavior problems and parenting stress. *Sleep and Biological Rhythms*, *9*(3), 165–171. https://doi.org/10.1111/j.1479-8425.2011.00502.x

van Maanen, A., Meijer, A. M., Smits, M. G., van der Heijden, K. B., & Oort, F. J. (2017). Effects of melatonin and bright light treatment in childhood chronic sleep onset insomnia with late melatonin onset: A randomized controlled study. *Sleep: Journal of Sleep and Sleep Disorders Research*, *40*(2), 1–11.

Van Someren, E. J. W., & Nagtegaal, E. (2007). Improving melatonin circadian phase estimates. *Sleep Medicine*, *8*(6), 590–601. https://doi.org/10.1016/j.sleep.2007.03.012

Vaughan, G. M., Pelham, R. W., Pang, S. F., Loughlin, L. L., Wilson, K. M., Sandock, K. L., Vaughan, M. K., Koslow, S. H., & Reiter, R. J. (1976). Nocturnal elevation of plasma melatonin and urinary 5-hydroxyindoleacetic acid in young men: Attempts at modification by brief changes in environmental lighting and sleep and by autonomic drugs. *The Journal of Clinical Endocrinology and Metabolism*, *42*(4), 752–764. https://doi.org/10.1210/jcem-42-4-752

Voultsios, A., Kennaway, D. J., & Dawson, D. (1997). Salivary Melatonin as a Circadian Phase Marker: Validation and Comparison to Plasma Melatonin. *Journal of Biological Rhythms*, *12*(5), 457–466. https://doi.org/10.1177/074873049701200507

Wittenbrink, N., Ananthasubramaniam, B., Münch, M., Koller, B., Maier, B., Weschke, C., Bes, F., de Zeeuw, J., Nowozin, C., Wahnschaffe, A., Wisniewski, S., Zaleska, M., Bartok, O., Ashwal-Fluss, R., Lammert, H., Herzel, H., Hummel, M., Kadener, S., Kunz, D., & Kramer, A. (2018). High-accuracy determination of internal circadian time from a single blood sample. *The Journal of Clinical Investigation*, *128*(9), 3826–3839. PubMed. https://doi.org/10.1172/JCI120874

Zavala, E., Wedgwood, K. C. A., Voliotis, M., Tabak, J., Spiga, F., Lightman, S. L., & Tsaneva-Atanasova, K. (2019). Mathematical Modelling of Endocrine Systems. *Trends in Endocrinology & Metabolism*, *30*(4), 244–257. https://doi.org/10.1016/j.tem.2019.01.008

Appendices

Appendix A

Table 13. Variable Means and Standard Deviations Extracted from the Scoping Review

	Demographics				Bedtime Routine		Circadian Markers			DLMO Metrics				
Author (year)	N (female)	Age (years)	Puberty (tanner)	BMI (percentile)	Bedtime (hh.mm)	Wake Time (hh.mm)	Sleep Midpoint (hh.mm)	Chronotype (MESC score)	Phase angle (min)	$Time_{LI3}$ (hh.mm)	$Time_{LI4}$ (hh.mm)	$Time_{within}$ (hh.mm)	$Time_{max}$ (hh.mm)	$Conc_{max}$ (pg/mL)
Agostini (2017)	12 (6)	16.1 (0.9)			T: 22.67(0.85) S: 22.24 (0.74) W: 23.1 (0.75)	T: 7.97 (1.23) S: 6.90 (3.53) W: 9.04 (0.66)					21.91 (0.80)			
Carskadon (1998)	40 (25)	15 (0.5)			T: 23.13 (0.76) S: 22.67 (0.66) W: 23.69 (0.88)	T: 7.26 (0.69) S: 6.22 (0.39) W: 8.51 (1.06)					20.72 (1.05)			
Crowley (2006)	149 (75)				22.26 (0.88)	7.95 (0.74)	2.83 (0.71)		82.75 (51.28)		20.88 (0.89)			
Crowley (2010)	12 (4)	15.9 (0.5)			22.95 (0.38)	6.27 (0.38)				21.17 (1.04)				
Crowley (2014)	94 (43)		3.74 (1.32)		T: 22.78 (0.72) S: 22.39 (0.63) W: 23.17 (0.82)	T: 7.14 (0.66) S: 6.52 (0.52) W: 7.76 (0.80)		37.37 (5.64)	98.00 (51.00)		20.76 (0.87)			
Crowley (2016)	66 (40)			6-97th percentile						21.26 (1.27)	21.48 (1.30)	20.83 (1.15)		
Crowley (2017)	44 (23)	16.2 (1)		6-97th percentile							21.07 (0.15)			
Crowley (2018a)	46 (24)	15.2 (1)			23.67 (1.18)	8.67 (1.18)	4.45 (1.22)	36.8 (5.3)			21.06 (1.15)			
Crowley (2018)	6 (3)	14.75 (2.12)			23.55 (1.58)	9.15 (1.48)					21.82 (1.38)			
Dolsen (2018)	176 (94)	14.7 (1.80)			23.10 (1.03)	7.57 (0.77)				21.32 (1.07)				
Dolsen (2019)	163 (93)	14.7 (1.80)			22.92 (1.05)					21.52 (1.13)				
Figueiro (2010)	12 (5)				22.39 (0.52)						20.28 (0.72)			
Goldman (2014)	9 (0)				21.4 (0.78)					21.42 (0.53)			25.3 (1.65)	128.72 (8.16)

Table 13 (Cont.)

		Demographics			Bedtime Routine		Circadian Markers			DLMO Metrics				
Author (year)	N (female)	Age (years)	Puberty (tanner)	BMI (percentile)	Bedtime (hh.mm)	Wake Time (hh.mm)	Sleep Midpoint (hh.mm)	Chronotype (MESC score)	Phase angle (min)	$Time_{LI3}$ (hh.mm)	$Time_{LI4}$ (hh.mm)	$Time_{within}$ (hh.mm)	$Time_{max}$ (hh.mm)	$Conc_{max}$ (pg/mL)
Higuchi (2016)	20 (13)	9.7 (1.4)			21.63 (0.57)					20.92 (0.73)				
Laberge (2000)	19 (0)	15.7 (1)			23.88 (0.63)	9.85 (0.85)					20.08 (0.63)			
Peixoto (2009)	37 (14)								-103.8 (79.2)		20.40 (1.05)			
Short (2012)	64 (32)										21.14 (0.75)			
Stone (2021)	26 (14)	13.1 (1.7)			22.27 (0.67)	7.20 (0.57)					20.58 (0.92)			
Taylor (2005)	34 (14)	15.91 (0.86)									20.97 (0.83)			
van der Heijden (2005)	59 (33)	12.8 (0.4)			21.65 (0.88)						21.10 (1.19)			
Van Maanen (2011)	41 (17)				21.12 (1.12)					20.97 (.83)				
van Maanen (2017)	84 (33)				22.03 (0.77)						21.91 (0.80)			
Mean	59.50	14.62	3.74	-	T: 22.84 S: 22.24 W: 23.1	T: 7.72 S: 6.47 W: 8.07	27.64	37.09	25.65	21.37	21.93	20.83	25.3	128.72
SD	69.11	1.35	1.32	-	T: 0.98 S: 0.74 W: 0.75	T: 0.77 S: 0.73 W: 0.86	.82	5.47	60.49	1.11	0.99	1.15	1.65	8.16
Min	9	12	1	6	20.13	4.47	24.00	10	−184.2	18.11	18.75	19.52	22.03	20.34
Max	176	17	5	97	25.53	11.38	30.23	40	220	24.09	24.18	23.93	27.02	310.30

Table 14. DLMO Collection Methodology

Author (year)	Collection timeframe	Frequency	Location	Collection conditions	Assay	Sleep measurement
Agostini (2017)	17:00 to 22:30	Every 30mins	Lab	<30lux	intra: <10%; inter: 7.3%/13.4%	Questionnaire
Carskadon (1998)		Every 30mins OR every 60mins	Lab	<50 lux	Detection limit: .75pg/mL; Intra: 12.1%, 5.7%, 9.8%; inter: 13.2%, 8.4%, 9.2%	Actigraphy
Crowley (2006)	5hrs before bedtime to 30mins after	Every 30mins (12 to 14 samples)	Lab	<50 lux, no caffeine, NSAIDs, chocolate	limit of detection: .2pg/mL	Sleep diary
Crowley (2010)		Every 30mins	Lab	<20 lux	Intra: 4.1% and 4.8%; inter 6.6% amd 8.4%	
Crowley (2014)	From 5hrs before bed to 30mins after	Every 30mins (total 12 samples)	Lab	<40 lux	Radioimmunoassay (limit of detection: .9 pg/mL; CV=4.1% for low and 4.8% for high)	
Crowley (2016)	5hrs before bedtime to 1hr after	Every 30mins (13 samples) OR Every 1hr (7 samples)	Lab	<20 lux	Intra: 4.1% (low) and 4.8% (high); inter: 6.6% (low) and 8.4% (high), limit of detection: .9pg/mL	
Crowley (2017)				<5 lux, no food or drink 10mins before each sampling	Intra: 4.1% and 4.8%; inter 6.6% amd 8.4%	
Crowley (2018a)	15:30 to 3:30	Every 30mins	Lab	<5 lux, no food or drink 10mins before each sampling	Intra: 4.1% and 4.8%; inter 6.6% amd 8.4%	Actigraphy
Crowley (2018b)						Actigraphy
Dolsen (2018)	5.5hrs before bedtime to 30mins after	Every 30mins	Lab	<50 lux	Sensitivity: .3pg/mL; limit of detection: .05pg/mL; intra: 7.9%; inter: 9.4%	Sleep diary
Dolsen (2019)	5.5hrs before bedtime to 30mins after	Every 30mins	Lab	<50 lux	Sensitivity 0.3pg/mL; minimum detectable limit 0/05 pg/mL; intra 7.9%; inter 9.4%	Sleep diary
Figueiro (2010)	18:00 to 22:00	Every 30mins	Lab	<10 lux, no food/drink	Intra: 1.1%; inter: 13.2%; sensitivity 0.7pg/mL	Actigraphy, sleep diary
Goldman (2014)			Lab	<30 lux		Actigraphy
Herz (2018)		Every 30mins	Lab		Sensitivity: 0.9pg/mL; intra: 7.9%; inter: 11.7%	Self-reported
Higuchi (2016)	19:00 to 1hr after habitual bedtime	Every 30mins	Home	<30 lux		Sleep diary, actigrpahy
Laberge (2000)	From wake to sleep	Every 120mins	Lab	no chocolate, caffeine, alcohol, <50 lux	limit of detection: 1pg/mL; intra: 10.8%, 9.5%, 12.9%; inter: 15.5%, 16.1%, 14.9%	Actigraphy, sleep diary
Peixoto (2009)	19:15 to 22:15	Every 30mins	Lab	<10lux		
Short (2018)	17:00 to 23:00	Every 30mins	Lab	<30 lux	intra: <10%; inter: 7.3%, 13.4%	
Taylor (2005)			Lab			
Van de Heijden (2007)	18:00 to 23:00	Every 60mins	Lab	Curtains closed		
van Maanen (2011)	19:00 to 23:00	Every 60mins	Lab			
van Maanen (2017)	19:00 to 23:00	Every 60mins	Lab	Closed curtains, only 1 dim light allowed		

Appendix B

Table 15. Means and Standard Deviations of the Secondary, Fully Compliant, and Simulated Datasets

	Secondary Dataset (n=223)		**Fully Compliant Dataset (n=161)**		**Simulated Dataset (n=16362)**	
Variables	Mean/ n	SD/%	Mean/ n	SD/%	Mean/ n	SD/%
Age	12.53	2.06	12.54	2.03	12.54	2.02
Sex (female)	98	43.95	74	45.96	7474	45.68
$Puberty_{gonadarche}$	3.17	1.16	3.15	1.16	3.11	1.13
$Puberty_{adrenarche}$	3.15	1.37	3.14	1.34	3.14	1.3
BMI percentile	54.29	28.91	54.83	27.61	54.85	27.47
BMI z-score	0.17	0.99	0.20	0.93	0.20	0.92
Time of year (school year)	73	32.74	66	34.16	5555	33.95
Use of melatonin supplement	14	6.28	11	6.83	1111	6.79
Bed_{TA} (hh.mm)	23.28	1.22	23.24	1.23	23.22	1.20
$Wake_{TA}$ (hh.mm)	7.86	1.02	7.83	1.03	7.82	1.01
Bed_{SA} (hh.mm)	23.15	1.25	23.1	1.27	23.09	1.24
$Wake_{SA}$ (hh.mm)	7.65	1.09	7.62	1.11	7.60	1.10
Bed_{WA} (hh.mm)	23.65	1.26	23.62	1.23	23.59	1.22
$Wake_{WA}$ (hh.mm)	8.43	1.18	8.42	1.16	8.40	1.15
Bed_{SQ} (hh.mm)	22.04	1.20	21.97	1.20	21.97	1.19
$Wake_{SQ}$ (hh.mm)	6.99	0.83	6.94	0.81	6.94	0.81
Bed_{WQ} (hh.mm)	23.26	1.64	23.23	1.64	23.23	1.64
$Wake_{WQ}$ (hh.mm)	9.08	1.81	9	1.76	8.99	1.75
Bed_{TD} (hh.mm)	22.49	1.09	22.42	1.06	22.42	1.06
$Wake_{TD}$ (hh.mm)	7.85	1.05	7.80	1.07	7.80	1.06
Bed_{SD} (hh.mm)	22.31	1.10	22.23	1.07	22.23	1.07
$Wake_{SD}$ (hh.mm)	7.56	1.13	7.52	1.16	7.52	1.15
Bed_{WD} (hh.mm)	22.96	1.26	22.94	1.22	22.94	1.22
$Wake_{WD}$ (hh.mm)	8.60	1.23	8.56	1.23	8.56	1.23
Weekend oversleep (hr)	0.77	1.02	0.78	1.01	0.79	1
Jet lag (hr)	0.48	0.64	0.50	0.65	0.50	0.65
$Midpoint_{TA}$ (hh.mm)	27.57	1.07	27.54	1.07	27.52	1.05
$Midpoint_{SA}$ (hh.mm)	27.40	1.10	27.36	1.12	27.34	1.10
$Midpoint_{WA}$ (hh.mm)	28.04	1.14	28.02	1.10	27.99	1.09
$Midpoint_{TD}$ (hh.mm)	27.17	0.96	27.11	0.95	27.11	0.95
$Midpoint_{SD}$ (hh.mm)	26.94	0.99	26.88	0.99	26.88	0.98
$Midpoint_{WD}$ (hh.mm)	27.78	1.12	27.75	1.10	27.75	1.10
$Midpoint_{SQ}$ (hh.mm)	26.51	0.79	26.45	0.77	26.45	0.77
$Midpoint_{SQ}$ (hh.mm)	28.17	1.48	28.12	1.47	28.11	1.47
Chronotype	27.35	4.47	27.31	4.65	27.31	4.63

Table15. (cont.)

	Secondary Dataset		Fully Compliant Dataset		Simulated Dataset	
Variables	Mean	SD	Mean	Mean	SD	Mean
Melatonin sample 1 Time (hh.mm)	20.36	1.27	20.28	1.27	20.28	1.27
Melatonin sample 2 Time (hh.mm)	20.87	1.29	20.79	1.28	20.79	1.28
Melatonin sample 3 Time (hh.mm)	21.38	1.29	21.29	1.28	21.29	1.28
Melatonin sample 4 Time (hh.mm)	21.88	1.30	21.8	1.29	21.80	1.28
Melatonin sample 5 Time (hh.mm)	22.38	1.30	22.3	1.29	22.30	1.29
Melatonin sample 6 Time (hh.mm)	22.96	1.53	22.85	1.44	22.85	1.43
Melatonin sample 7 Time (hh.mm)	23.49	1.55	23.36	1.45	23.36	1.45
Melatonin sample 1 Concentration (pg/mL)	9.38	8.38	9.10	7.28	9.16	7.26
Melatonin sample 2 Concentration (pg/mL)	11.48	11.05	11.08	9.38	11.12	9.35
Melatonin sample 3 Concentration (pg/mL)	14.61	11.80	14.41	11.75	14.34	11.69
Melatonin sample 4 Concentration (pg/mL)	18.30	13.50	18	13.31	18	13.22
Melatonin sample 5 Concentration (pg/mL)	22.04	14.56	21.94	15.05	21.92	15
Melatonin sample 6 Concentration (pg/mL)	24.14	14.97	24.01	14.87	24.10	14.84
Melatonin sample 7 Concentration (pg/mL)	25.71	15.11	25.72	14.85	25.89	14.81
$Time_{LI3}$ (hh.mm)	20.94	1.27	20.91	1.26	20.99	1.35
$Time_{LI4}$ (hh.mm)	21.13	1.42	21.01	1.37	21.08	1.39
$Time_{reg3}$ (hh.mm)	19.55	1.63	19.47	1.51	19.18	1.90
$Time_{reg4}$ (hh.mm)	19.73	1.61	19.66	1.50	19.39	1.99
$Time_{within}$ (hh.mm)	21.98	1.29	21.83	1.17	21.91	1.20
$Time_{max}$ (hh.mm)	22.87	1.37	22.81	1.39	22.83	1.41
$Time_{hockey}$ (hh.mm)	21.90	1.47	21.89	1.47	21.93	1.49
$Conc_{max}$ (pg/mL)	28.45	15.06	28.41	15.15	29.23	15.81
$Conc_{AUC}$ (pg/mL)	51.03	31.32	51.75	32.27	53.49	33.61
$Rate_{reg}$ (pg/mL/hh.mm)	6.26	5.40	6.03	4.92	5.95	4.77
$Rate_{hockey}$ (pg/mL/hh.mm)	0.01	0.01	0.01	0.01	0.01	0.02
$Time_{LI3}$ phase angle (hh.mm)	2.43	0.98	2.48	0.78	2.09	0.70
$Time_{LI4}$ phase angle (hh.mm)	2.52	0.85	2.51	0.79	2.08	0.77
$Time_{reg3}$ phase angle (hh.mm)	3.50	1.73	3.52	1.51	3.85	1.87
$Time_{reg4}$ phase angle (hh.mm)	3.32	1.71	3.33	1.49	3.65	1.97
$Time_{within}$ phase angle (hh.mm)	1.70	1.01	1.68	1.02	1.15	1.09
$Time_{max}$ time phase angle (hh.mm)	0.26	1.19	0.26	1.22	0.25	1.21
$Time_{hockey}$ phase angle (hh.mm)	0.22	1.26	0.27	1.09	1.16	1.18

Appendix C

Table 16. Correlations between DLMO Metrics, Circadian Markers, and Bedtime Routine Variables in the Secondary Subsample

	DLMO Metrics										
Variable	**$Time_{LI3}$**	**$Time_{LI4}$**	**$Time_{reg3}$**	**Timereg4**	**$Time_{within}$**	**Timemax**	**$Time_{hockey}$**	**$Conc_{max}$**	**$Conc_{AUC}$**	**$Rate_{reg}$**	**$Rate_{hockey}$**
Chronotype	−.23	−.28	−.23	−.23	−.23	−.29	−.06	−.04	−.05	−.07	.32
$Midpoint_{TA}$	.87	.87	.86	.87	.69	.80	.77	−.43	−.52	−.38	−.31
$Midpoint_{SA}$	.85	.84	.82	.83	.63	.74	.74	−.46	−.56	−.41	−.28
$Midpoint_{WA}$	.84	.85	.86	.86	.69	.79	.7	−.27	−.39	−.21	−.29
$Midpoint_{TD}$	.86	.86	.86	.86	.69	.82	.75	−.36	−.39	−.27	−.45
$Midpoint_{SD}$	.86	.85	.83	.84	.72	.81	.74	−.46	−.46	−.39	−.47
$Midpoint_{WD}$	.82	.83	.85	.85	.6	.78	.72	−.16	−.26	−.05	−.37
$Midpoint_{SQ}$	.72	.72	.78	.78	.55	.72	.77	−.22	−.34	−.18	−.15
$Midpoint_{WQ}$	.85	.87	.91	.91	.7	.87	.82	−.31	−.39	−.25	−.34
Weekend oversleep	.17	.15	.2	.19	.04	.12	.33	−.11	−.23	−.11	.12
Jet lag	.79	.81	.85	.85	.69	.83	.8	−.29	−.37	−.24	−.33
Bed_{TA}	.42	.45	.44	.44	.4	.38	.27	−.12	−.14	−.09	−.23
Bed_{SA}	−.16	−.15	−.07	−.09	−.02	−.01	−.13	.57	.48	.58	.12
Bed_{WA}	.88	.89	.93	.93	.71	.86	.82	−.31	−.43	−.24	−.29
$Wake_{TA}$	.91	.91	.94	.94	.7	.85	.83	−.38	−.49	−.3	−.3
$Wake_{SA}$	.76	.77	.83	.82	.64	.79	.7	−.06	−.21	.01	−.22
$Wake_{WA}$	.70	.69	.61	.64	.54	.56	.55	−.53	−.57	−.51	−.29
Bed_{TD}	.57	.55	.47	.5	.39	.42	.45	−.49	−.55	−.48	−.19
Bed_{SD}	.76	.77	.7	.72	.61	.62	.55	−.46	−.52	−.43	−.32
Bed_{WD}	.82	.82	.87	.86	.67	.84	.76	−.17	−.23	−.07	−.37
$Wake_{TD}$	.82	.82	.87	.86	.71	.84	.78	−.23	−.27	−.14	−.37
$Wake_{SD}$	.75	.76	.8	.79	.51	.75	.64	−.03	−.15	.09	−.35
$Wake_{WD}$	.69	.68	.6	.63	.54	.57	.53	−.57	−.53	−.53	−.44
Bed_{SQ}	.57	.55	.45	.48	.44	.44	.38	−.61	−.54	−.58	−.42
Bed_{WQ}	.75	.75	.75	.75	.6	.66	.67	−.32	−.37	−.26	−.32
$Wake_{SQ}$	.83	.85	.89	.88	.69	.86	.79	−.22	−.29	−.16	−.28
$Wake_{WQ}$	.81	.83	.86	.86	.63	.82	.74	−.29	−.37	−.23	−.31

Table 17. Correlations between Circadian Markers and Bedtime Routine Variables in the in the Secondary Subsample

Variables	Chronotype	$Midpoint_{TA}$	$Midpoint_{SA}$	$Midpoint_{WA}$	$Midpoint_{TD}$	$Midpoint_{SD}$	$Midpoint_{WD}$	$Midpoint_{SQ}$	$Midpoint_{WQ}$	Weekend oversleep	Jet lag	Bed_{TA}	Bed_{SA}	Bed_{WA}	$Wake_{TA}$	$Wake_{SA}$	$Wake_{WA}$	Bed_{TD}	Bed_{SD}	Bed_{WD}	$Wake_{TD}$	$Wake_{SD}$	$Wake_{WD}$	Bed_{SQ}	Bed_{WQ}	$Wake_{SQ}$	$Wake_{WQ}$
Chronotype	1																										
$Midpoint_{TA}$	−.26	1																									
$Midpoint_{SA}$	−.19	.98	1																								
$Midpoint_{WA}$	−.36	.93	.84	1																							
$Midpoint_{TD}$	−.29	.86	.81	.84	1																						
$Midpoint_{SD}$	−.25	.88	.83	.83	.98	1																					
$Midpoint_{WD}$	−.33	.75	.69	.79	.92	.85	1																				
$Midpoint_{SQ}$	−.18	.80	.76	.76	.73	.7	.72	1																			
$Midpoint_{WQ}$	−.38	.87	.80	.89	.82	.79	.81	.85	1																		
Weekend oversleep	.15	.41	.46	.25	.14	.17	.09	.66	.26	1																	
Jet lag	−.33	.83	.78	.84	.73	.69	.71	.78	.94	.3	1																
Bed_{TA}	−.32	.30	.14	.57	.38	.35	.44	.22	.45	−.33	.41	1															
Bed_{SA}	−.22	−.09	−.23	.17	−.01	−.06	.1	.06	.08	−.04	.05	.17	1														
Bed_{WA}	−.22	.95	.91	.93	.87	.86	.82	.89	.94	.42	.87	.36	.04	1													
$Wake_{TA}$	−.20	.96	.94	.9	.86	.85	.81	.87	.92	.41	.86	.33	−.11	.99	1												
$Wake_{SA}$	−.30	.84	.76	.92	.79	.76	.8	.83	.89	.36	.82	.39	.39	.93	.87	1											
$Wake_{WA}$	−.27	.89	.91	.76	.70	.77	.52	.52	.61	.33	.62	.16	−.26	.71	.75	.56	1										
Bed_{TD}	−.14	.79	.87	.57	.55	.62	.36	.44	.45	.44	.48	−.16	−.34	.59	.64	.42	.94	1									
Bed_{SD}	−.35	.83	.76	.88	.71	.74	.62	.5	.69	.07	.69	.67	−.13	.73	.74	.62	.84	.63	1								
Bed_{WD}	−.22	.74	.67	.77	.93	.88	.95	.77	.82	.14	.69	.38	.16	.85	.82	.83	.45	.3	.52	1							
$Wake_{TD}$	−.14	.74	.66	.78	.93	.89	.92	.76	.8	.15	.68	.43	.15	.85	.81	.82	.45	.29	.55	.99	1						
$Wake_{SD}$	−.39	.68	.65	.7	.86	.77	.94	.72	.78	.14	.67	.24	.17	.78	.77	.79	.43	.33	.44	.93	.85	1					
$Wake_{WD}$	−.31	.82	.8	.71	.83	.88	.62	.47	.61	.1	.57	.28	−.26	.66	.69	.51	.91	.78	.81	.57	.58	.51	1				
Bed_{SQ}	−.32	.73	.75	.58	.67	.76	.4	.34	.45	.14	.43	.1	−.33	.53	.57	.36	.91	.84	.71	.37	.37	.33	.95	1			
Bed_{WQ}	−.17	.69	.60	.76	.82	.78	.87	.57	.68	.01	.62	.64	−.02	.7	.69	.63	.55	.33	.75	.78	.81	.65	.66	.42	1		
$Wake_{SQ}$	−.33	.78	.70	.82	.86	.81	.88	.88	.93	.22	.82	.49	.1	.89	.87	.86	.47	.28	.6	.91	.89	.85	.55	.36	.74	1	
$Wake_{WQ}$	−.38	.81	.74	.83	.83	.79	.82	.82	.95	.2	.78	.44	.09	.89	.88	.86	.53	.37	.62	.84	.82	.8	.58	.43	.67	.94	1

Table 18. Correlations between DLMO Metrics, Circadian Markers, and Bedtime Routine Variables in the Simulated Dataset

	DLMO Metrics										
Variable	**$Time_{LI3}$**	**$Time_{LI4}$**	**$Time_{reg3}$**	**Timereg4**	**$Time_{within}$**	**Timemax**	**$Time_{hockey}$**	**$Conc_{max}$**	**$Conc_{AUC}$**	**$Rate_{reg}$**	**$Rate_{hockey}$**
Chronotype	−.2	−.21	−.19	−.21	−.23	−.25	−.12	.08	.03	.09	.27
$Midpoint_{TA}$	.87	.88	.85	.87	.69	.75	.72	−.49	−.6	−.48	−.23
$Midpoints_{A}$	.85	.86	.81	.83	.63	.69	.7	−.48	−.61	−.47	−.16
$Midpoint_{WA}$	.75	.79	.79	.8	.66	.71	.58	−.38	−.48	−.36	−.31
$Midpoint_{TD}$	.84	.85	.84	.85	.66	.76	.69	−.43	−.5	−.41	−.32
$Midpoints_{D}$	.84	.84	.81	.83	.67	.73	.68	−.48	−.54	−.46	−.3
$Midpoint_{WD}$	.74	.78	.79	.8	.59	.7	.6	−.31	−.41	−.27	−.35
$Midpoints_{Q}$	.67	.64	.73	.72	.45	.63	.65	−.22	−.36	−.22	−.06
$Midpoint_{WQ}$	.71	.69	.77	.78	.64	.71	.77	−.32	−.33	−.29	−.23
Weekend oversleep	−.04	−.02	.07	.06	.1	.12	−.05	.03	.1	.06	−.22
Jet lag	−.3	−.23	−.21	−.22	−.09	−.15	−.32	.28	.26	.27	−.06
Bed_{TA}	.82	.86	.89	.89	.69	.8	.75	−.36	−.47	−.31	−.25
$Beds_{A}$	.86	.89	.9	.91	.68	.8	.78	−.39	−.5	−.34	−.23
Bed_{WA}	.58	.65	.68	.67	.54	.61	.5	−.18	−.29	−.14	−.25
$Wake_{TA}$	.73	.72	.64	.66	.55	.55	.52	−.52	−.61	−.55	−.15
$Wakes_{A}$	.66	.64	.54	.57	.45	.44	.48	−.46	−.57	−.49	−.06
$Wake_{WA}$	.72	.73	.7	.72	.61	.62	.5	−.51	−.56	−.51	−.29
Bed_{TD}	.67	.72	.77	.76	.58	.74	.63	−.25	−.29	−.19	−.38
$Beds_{D}$	.71	.74	.79	.79	.64	.76	.68	−.29	−.32	−.25	−.36
Bed_{WD}	.5	.57	.61	.61	.39	.58	.41	−.14	−.22	−.07	−.39
$Wake_{TD}$	.7	.67	.59	.62	.49	.49	.5	−.44	−.51	−.46	−.15
$Wakes_{D}$	.64	.61	.52	.55	.44	.43	.43	−.45	−.51	−.47	−.14
$Wake_{WD}$	.75	.73	.7	.72	.6	.57	.59	−.4	−.49	−.39	−.18
$Beds_{Q}$	.69	.7	.82	.82	.6	.79	.72	−.21	−.25	−.18	−.25
Bed_{WQ}	.59	.64	.7	.71	.56	.68	.6	−.33	−.33	−.29	−.33
$Wakes_{Q}$	.36	.31	.33	.33	.13	.22	.31	−.13	−.29	−.16	.13
$Wake_{WQ}$	.61	.55	.63	.64	.54	.55	.69	−.24	−.25	−.22	−.1

Table 19. Correlations between Circadian Markers and Bedtime Routine Variables in the in the Simulated Dataset

Variables	Chronotype	$Midpoint_{TA}$	$Midpoint_{SA}$	$Midpoint_{WA}$	$Midpoint_{TD}$	$Midpoint_{SD}$	$Midpoint_{WD}$	$Midpoint_{SQ}$	$Midpoint_{WQ}$	Weekend oversleep	Jet lag	Bed_{TA}	Bed_{SA}	Bed_{WA}	$Wake_{TA}$	$Wake_{SA}$	$Wake_{WA}$	Bed_{TD}	Bed_{SD}	Bed_{WD}	$Wake_{TD}$	$Wake_{SD}$	$Wake_{WD}$	Bed_{SQ}	Bed_{WQ}	$Wake_{SQ}$	$Wake_{WQ}$
Chronotype	1																										
$Midpoint_{TA}$	−.21	1																									
$Midpoint_{SA}$	−.14	.98	1																								
$Midpoint_{WA}$	−.30	.86	.75	1																							
$Midpoint_{TD}$	−.28	.9	.87	.78	1																						
$Midpoint_{SD}$	−.23	.92	.9	.74	.98	1																					
$Midpoint_{WD}$	−.36	.76	.69	.81	.89	.81	1																				
$Midpoint_{SQ}$	−.25	.76	.73	.58	.75	.75	.66	1																			
$Midpoint_{WQ}$	−.52	.68	.61	.67	.66	.61	.67	.67	1																		
Weekend oversleep	−.24	−.16	−.34	.27	−.14	−.21	.11	−.17	.18	1																	
Jet lag	−.08	−.18	−.3	.22	−.14	−.22	.1	−.25	−.15	.31	1																
Bed_{TA}	−.12	.90	.84	.90	.80	.77	.77	.66	.7	.05	.03	1															
Bed_{SA}	−.09	.91	.88	.86	.80	.78	.75	.67	.7	−.01	−.13	.98	1														
Bed_{WA}	−.13	.70	.60	.89	.63	.57	.72	.44	.54	.17	.48	.89	.81	1													
$Wake_{TA}$	−.27	.89	.91	.63	.82	.87	.59	.69	.51	−.35	−.35	.6	.64	.36	1												
$Wake_{SA}$	−.15	.83	.90	.49	.74	.81	.48	.64	.4	−.57	−.39	.52	.58	.28	.97	1											
$Wake_{WA}$	−.41	.80	.71	.83	.73	.73	.67	.57	.63	.31	−.16	.65	.66	.49	.78	.6	1										
Bed_{TD}	−.25	.66	.56	.78	.79	.71	.90	.53	.61	.21	.24	.81	.76	.82	.37	.26	.5	1									
Bed_{SD}	−.19	.69	.59	.79	.81	.75	.88	.55	.63	.21	.19	.82	.78	.8	.4	.29	.54	.98	1								
Bed_{WD}	−.34	.53	.45	.70	.65	.53	.86	.40	.50	.20	.37	.69	.64	.79	.26	.17	.39	.92	.84	1							
$Wake_{TD}$	−.21	.79	.83	.49	.82	.88	.55	.68	.46	−.41	−.45	.49	.55	.22	.94	.91	.66	.3	.35	.16	1						
$Wake_{SD}$	−.18	.76	.81	.42	.76	.83	.45	.63	.37	−.48	−.47	.43	.5	.16	.93	.93	.61	.21	.26	.08	.98	1					
$Wake_{WD}$	−.24	.74	.71	.62	.84	.82	.79	.71	.62	−.04	−.26	.57	.59	.37	.76	.67	.74	.53	.58	.35	.82	.72	1				
Bed_{SQ}	−.41	.65	.54	.74	.72	.64	.8	.73	.86	.35	.01	.77	.75	.66	.38	.23	.6	.82	.82	.72	.35	.25	.59	1			
Bed_{WQ}	−.41	.6	.48	.78	.59	.51	.72	.46	.78	.39	.21	.78	.74	.78	.28	.14	.54	.82	.8	.78	.16	.08	.37	.86	1		
$Wake_{SQ}$	.01	.52	.58	.18	.46	.52	.25	.8	.22	−.56	−.37	.27	.31	.05	.66	.72	.29	.04	.07	−.05	.68	.69	.51	.18	−.1	1	
$Wake_{WQ}$	−.46	.57	.55	.43	.54	.53	.46	.66	.9	−.01	−.36	.47	.50	.23	.54	.47	.54	.31	.35	.17	.55	.48	.63	.63	.43	.39	1

Appendix D

Table 20. Complete List of Variables Constraints for Multiple Imputations

Variable	Mean	SD	Min	Max
Age	14.62	1.35	9	17
Sex	-	-	1	2
$Puberty_{gonadarche}$	3	1.20	1	5
$Puberty_{adrenarche}$	3	1.40	1	5
BMI percentile	50	25	0.47	99.44
BMI z-score	0	1	-2.60	2.54
Time of year	-	-	1	2
Use of melatonin supplement	-	-	0	1
Bed_{SQ} (hh.mm)	22.24	0.74	20	26
$Wake_{SQ}$ (hh.mm)	6.47	0.73	5.17	11
Bed_{WQ} (hh.mm)	23.10	0.75	20.5	28
$Wake_{WQ}$ (hh.mm)	8.07	0.86	1	12.50
Bed_{TA} (hh.mm)	22.84	0.98	20.64	27.40
$Wake_{TA}$ (hh.mm)	7.72	0.77	5.21	11.59
Bed_{SA} (hh.mm)	22.24	0.74	20.64	27.55
$Wake_{SA}$ (hh.mm)	6.47	0.73	5.24	11.35
Bed_{WA} (hh.mm)	23.10	0.75	20.61	27.80
$Wake_{WA}$ (hh.mm)	8.07	0.86	5.12	14.78
Bed_{TD} (hh.mm)	22.49	1.09	20	26
$Wake_{TD}$ (hh.mm)	7.85	1.05	4	12
Bed_{SD} (hh.mm)	22.31	1.10	20	26
$Wake_{SD}$ (hh.mm)	7.57	1.13	4	12
Bed_{WD} (hh.mm)	22.95	1.26	20	26
$Wake_{WD}$ (hh.mm)	8.60	1.23	4	12
Chronotype	37.19	5.53	10	42
Melatonin sample 1 Time (hh.mm)	20.36	1.27	17	26
Melatonin sample 2 Time (hh.mm)	20.88	1.29	17	26
Melatonin sample 3 Time (hh.mm)	21.39	1.29	17	26
Melatonin sample 4 Time (hh.mm)	21.89	1.3	17	26
Melatonin sample 5 Time (hh.mm)	22.38	1.3	17	26
Melatonin sample 6 Time (hh.mm)	22.96	1.53	17	26
Melatonin sample 7 Time (hh.mm)	23.49	1.55	17	26
Melatonin sample 1 Concentration (pg/mL)	9.41	8.38	0	130
Melatonin sample 2 Concentration (pg/mL)	11.52	11.06	0	130
Melatonin sample 3 Concentration (pg/mL)	14.61	11.8	0	130
Melatonin sample 4 Concentration (pg/mL)	18.37	13.49	0	130
Melatonin sample 5 Concentration (pg/mL)	22.09	14.58	0	130
Melatonin sample 6 Concentration (pg/mL)	24.18	15	0	130
Melatonin sample 7 Concentration (pg/mL)	25.74	15.15	0	130

Appendix E

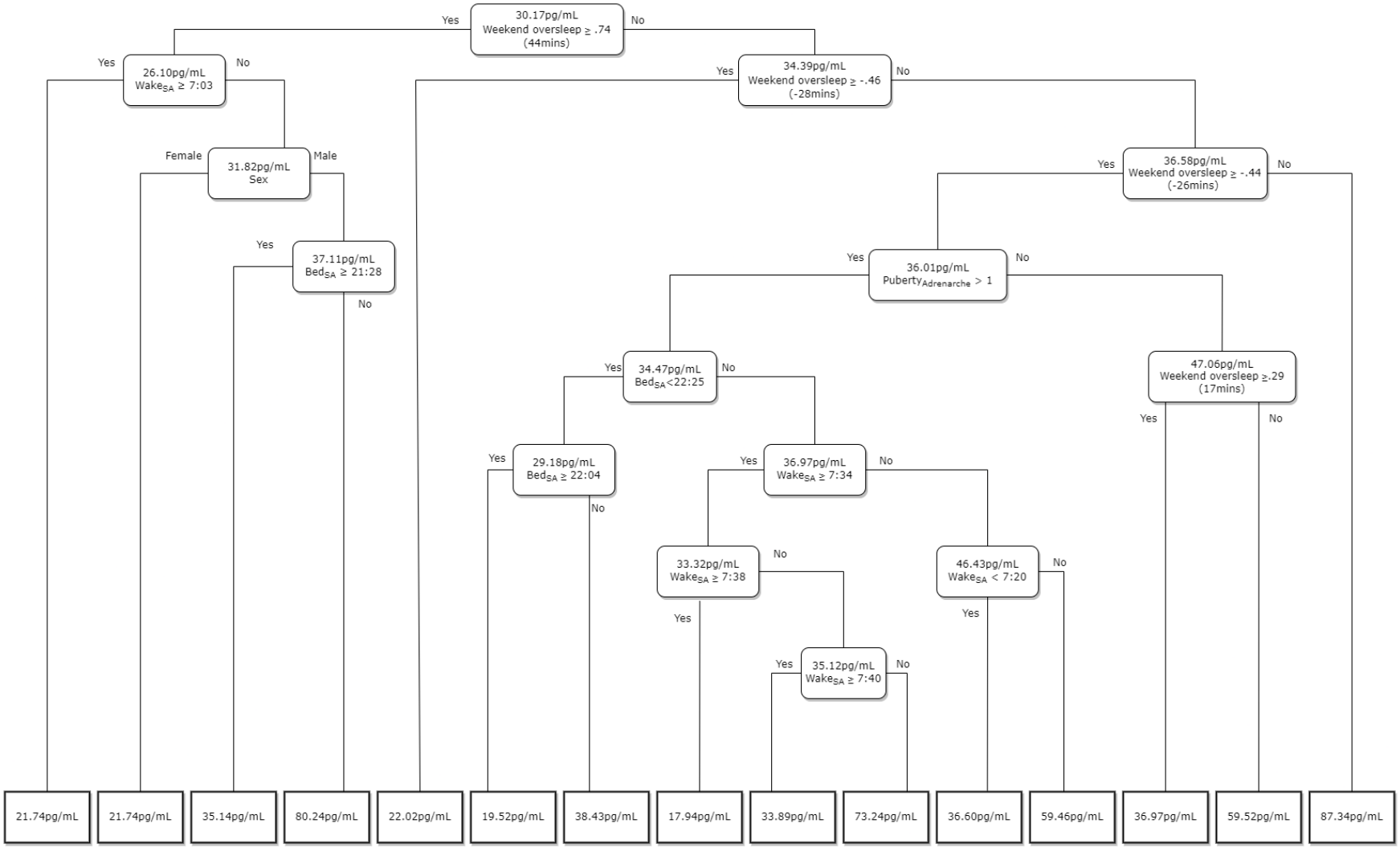

Figure 4. CART Concentration Model

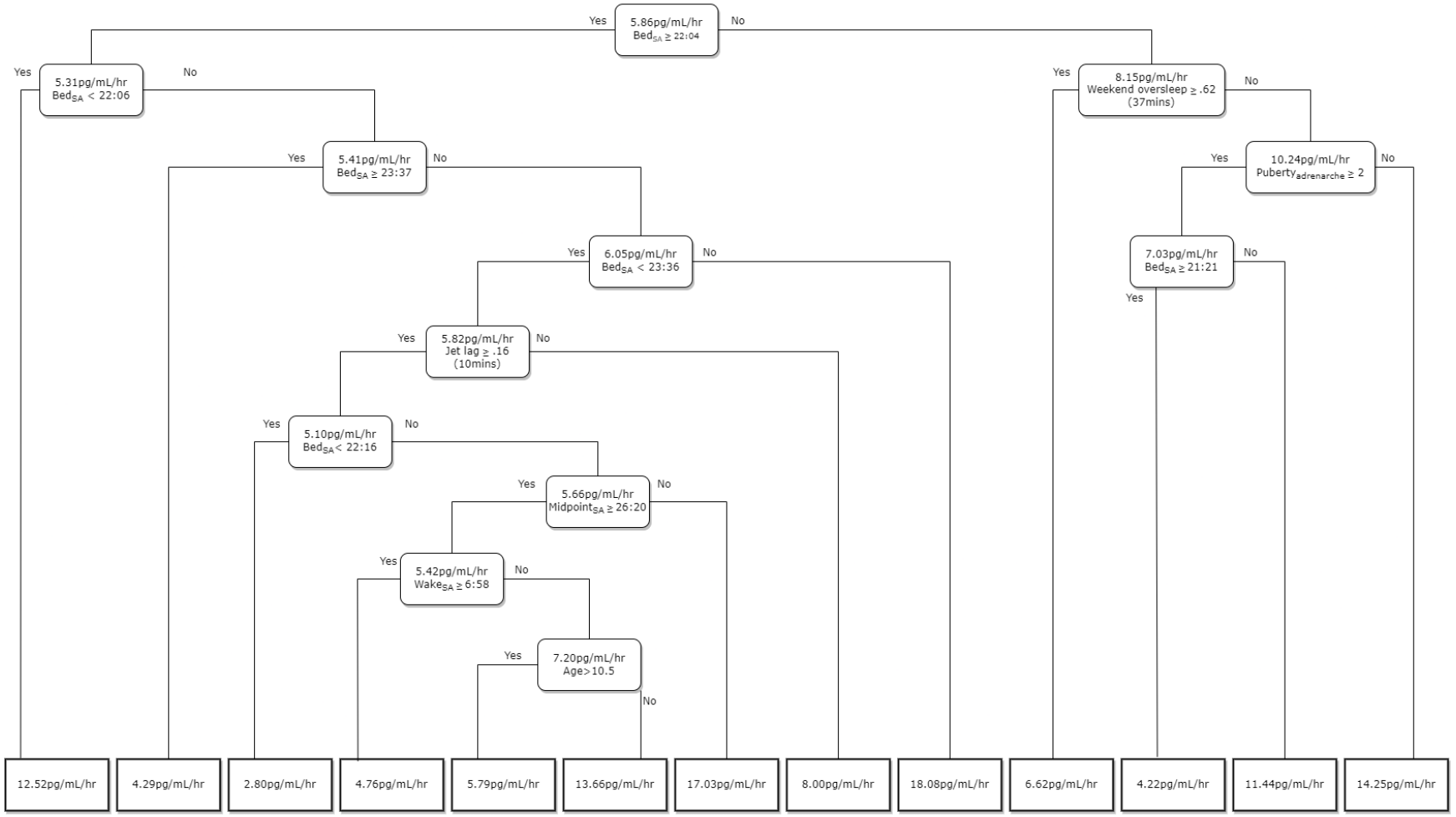

Figure 5. CART Rate Model