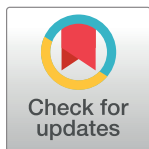REVIEW

Using routinely collected clinical data for circadian medicine: A review of opportunities and challenges

Laura Kervezee [1][‡]*, **Hassan S. Dashti**[2], **Luísa K. Pilz**[3,4], **Carsten Skarke**[5,6,7], **Marc D. Ruben**[8][‡]*

1 Group of Circadian Medicine, Department of Cell and Chemical Biology, Leiden University Medical Center, Leiden, the Netherlands, **2** Department of Anesthesia, Critical Care and Pain Medicine, Massachusetts General Hospital and Harvard Medical School, Boston, Massachusetts, United States of America, **3** Department of Anesthesiology and Intensive Care Medicine CCM / CVK, Charité–Universitätsmedizin Berlin, corporate member of Freie Universität Berlin and Humboldt Universität zu Berlin, Berlin, Germany, **4** ECRC Experimental and Clinical Research Center, Charité–Universitätsmedizin Berlin, corporate member of Freie Universität Berlin and Humboldt Universität zu Berlin, Berlin, Germany, **5** Institute for Translational Medicine and Therapeutics (ITMAT), University of Pennsylvania Perelman School of Medicine, Philadelphia, Pennsylvania, United States of America, **6** Chronobiology and Sleep Institute (CSI), University of Pennsylvania Perelman School of Medicine, Philadelphia, Pennsylvania, United States of America, **7** Department of Medicine, University of Pennsylvania Perelman School of Medicine, Philadelphia, Pennsylvania, United States of America, **8** Divisions of Pulmonary and Sleep Medicine and Biomedical Informatics, Cincinnati Children's Hospital Medical Center, Cincinnati, Ohio, United States of America

☯ These authors contributed equally to this work.
‡ These authors share senior authorship on this work.
* l.kervezee@lumc.nl (LK); Marc.Ruben@cchmc.org (MDR)

Abstract

A wealth of data is available from electronic health records (EHR) that are collected as part of routine clinical care in hospitals worldwide. These rich, longitudinal data offer an attractive object of study for the field of circadian medicine, which aims to translate knowledge of circadian rhythms to improve patient health. This narrative review aims to discuss opportunities for EHR in studies of circadian medicine, highlight the methodological challenges, and provide recommendations for using these data to advance the field. In the existing literature, we find that data collected in real-world clinical settings have the potential to shed light on key questions in circadian medicine, including how 24-hour rhythms in clinical features are associated with—or even predictive of—health outcomes, whether the effect of medication or other clinical activities depend on time of day, and how circadian rhythms in physiology may influence clinical reference ranges or sampling protocols. However, optimal use of EHR to advance circadian medicine requires careful consideration of the limitations and sources of bias that are inherent to these data sources. In particular, time of day influences almost every interaction between a patient and the healthcare system, creating operational 24-hour patterns in the data that have little or nothing to do with biology. Addressing these challenges could help to expand the evidence base for the use of EHR in the field of circadian medicine.

Introduction

Thousands of data points may be generated and digitally stored each day a patient is in the hospital [1]. These electronic health records (EHR) often include vital signs, imaging and laboratory results, diagnoses, medications, medical history, demographic information, and clinical notes. Although EHR, by definition, are built for patient care [2], the data are increasingly used to address a wide range of biomedical research questions.

One area of research that may benefit from routinely collected clinical data is the burgeoning field of circadian medicine, which aims to translate findings from circadian biology to clinical practice. Owing to their high resolution and longitudinal nature, EHR data have been referred to as a "treasure trove of information" that can support all facets of research in circadian medicine [3]. However, repurposing real-world clinical data for biomedical research requires a great deal of consideration. In addition to the general challenges facing all EHR-based research [4–6], efforts to understand the role of time of day present unique challenges, which have received only limited attention.

In this narrative review, our goal is to (i) provide an overview of the potential for EHR to advance circadian medicine; (ii) to highlight the challenges pertaining to responsible use of these data; and (iii) offer recommendations for researchers who are planning to use EHR data in the field of circadian medicine.

Literature search methodology

This narrative review was prepared by identifying peer-reviewed articles in electronic databases (PubMed and Google Scholar) using search terms including "circadian" and "electronic health records" or "routinely-collected clinical data" and variants of those terms. Inclusion criteria were research articles that describe 24-hour patterns in physiological readouts collected using routinely collected clinical data, focusing on EHR data from inpatient hospital stays since these patients are monitored around the clock. However, many of the points we discuss are applicable to EHR data from outpatient clinics, disease registries, and other sources. In addition, to provide an overview of the limitations and challenges related to the use of electronic health records, we searched PubMed and Google Scholar for peer-reviewed articles on this topic. In principle, all literature available on these topics were considered without an explicit date range, although priority was given to recent literature due to constraints in word count and number of references set by the journal.

Daily rhythms in physiology

Human physiology and behavior change profoundly between day and night. The most obvious example is the sleep–wake cycle, but also a broad range of other behaviors (e.g., food intake, mood) and physiology including cardiac (e.g., blood pressure, heart rate, thrombus formation), metabolic (e.g., glucose and lipid metabolism, insulin secretion, metabolic rate), immune and inflammatory (e.g., cytokine secretion, circulating leukocytes), gastrointestinal (e.g., digestion, absorption, and electrolyte balance), thermoregulatory, and hormonal functions vary with predictable patterns over 24 hours [7–12]. Consequently, many clinical tests or measures may depend on the time of day of data collection.

These daily fluctuations in physiology are in part driven by the circadian clock [13], an internal timing system built from 24-hour molecular oscillators that exist in virtually every cell in the body [14]. These oscillators drive circa 24-hour rhythms in gene transcription and translation that synchronize with salient 24-hour rhythms in the environment, such as light–dark or feeding-fasting cycles. In this way, the circadian system anticipates predictable changes in the environment and helps to enact the right physiology at the right time, while also keeping

an internal temporal order [15]. For example, blood pressure follows a 24-hour rhythm characterized by a morning rise, afternoon peak, and overnight dip, with a typical daily variation from 10% to 20% over the course of a day [16]. This variation is thought to reflect the differing demands for blood flow during the day (physical and mental exertion) versus the night (recovery, repair, efficiency) [17].

Circadian rhythms can be uncovered in highly controlled laboratory conditions using constant routine or forced desynchrony protocols [18,19]. For example, core body temperature exhibits endogenous circadian rhythmicity that continues to oscillate with a period of approximately 24 hours, even in the absence of external timing cues [20,21]. Under entrained conditions, i.e., in the presence of cues from a 24-hour light–dark cycle, circadian-regulated physiology will cycle with a period of exactly 24 hours. However, the endogenous circadian cycle is not the only contributor to day–night variation in physiology. Changes in activity, posture, meals, sleep, room temperature, light, and potentially many other factors can contribute to daily patterns in physiology [19]. For example, any periodic fluctuation in core body temperature observed in daily life may be caused by diet-induced thermogenesis or physical activity, leading to an increase in core body temperature during the active phase when food intake occurs and a decrease during the resting phase. In the context of distal skin temperature rhythmicity, the peak-to-trough difference of the endogenous circadian rhythm, measured in highly controlled conditions, is about 3.5˚C, while this is 6˚C in freely living conditions, likely conditioned by behavioral and environmental components [22]. In clinical settings, medical activities such as drug administration, mechanical ventilation, nutritional support, or physical therapy as well as excess sound and artificial light may impose daily rhythms on physiology or completely mask or abolish them. Although the relevance of the interaction between circadian rhythms and behavioral and environmental factors remains to be determined, conceptually, EHR data offer an opportunity to study this at scale and close to the clinical reality. Patient-derived data are inherently noisier than data collected from well-controlled experiments but may contain time-specific signals that are modulated by the circadian clock-behavioral-environmental interaction. The question is: Can novel variables be extracted from this parameter space and would these be indicative of the patient's clinical status? Here, the massive amounts of data stored in EHR offer opportunities to evaluate this in the context of clinical care.

Electronic health records for research

EHR databases store all of the electronic data collected in clinical care. Although EHR are primarily used for patient care (e.g., to inform clinical decision-making) and administrative purposes (e.g., billing and insurance), these data are increasingly being applied to biomedical research questions. The advantages of EHR seem clear: (1) data are readily available, reducing costs and burden; (2) data are systematically captured from all patients within the healthcare system, reducing inclusion bias and potentially improving generalizability; (3) databases are large and rapidly grow with time, enabling the ascertainment of case–control patient cohorts, assessment of interactions, and subgroup analyses; (4) data are collected over prolonged periods of time, enabling longitudinal analyses; (5) data are collected in real-world conditions, improving ecological validity [23]; and (6) data are multimodal, enabling analyses along many health measures—medical and nonmedical. However, such advantages should be viewed with caution as EHR are not primarily intended for research.

Many factors can compromise the validity and generalizability of EHR-based research, as has been reviewed elsewhere [5,24]. As patients interact with a healthcare system, EHR documents what happens and when it happens. However, EHR does not explicitly document the why. Hospital admissions, tests, treatment, and discharge are the result of complex decisions

made by the patient, clinical staff, and healthcare system. This underlies a key challenge in EHR analysis: The decision-making (data-generating) process is often unknown. This may result in sample selection bias, whereby the presence of data is related to its outcome, as a measurement is only made with clinical justification. Similarly, sampling frequency is inevitably correlated with disease severity; monitoring is more frequent for sicker patients. Furthermore, information within the EHR may be imprecisely defined or related to billing rather than clinical purpose resulting in patient misclassification. For example, a diagnostic code for sleep apnea is necessary for the reimbursement of a sleep apnea exam but is not always indicative of sleep apnea onset [25]. These forms of bias and confounding should be considered when undertaking EHR-based research; a multidisciplinary team with clinical, epidemiological, and statistical expertise is indispensable.

Despite these challenges, EHR-based research is increasing dramatically, in part owing to the release of public databases with deidentified data [26] and technological advances that allow researchers to use EHR from their own or collaborating institutions. Notable examples of open-access databases in the intensive care domain are the MIMIC-IV database [27], the eICU [28], the HiRID database [29], and the AmsterdamUMCdb [30]. Before embarking on a research project that involves setting up a new database derived from EHR, it is important to be aware that this process is time-consuming, labor-intensive, and will undoubtedly raise nontrivial and unexpected issues, as recently described in an honest account of the challenges encountered while preparing EHR data for secondary use [31]. Extraction and curation of EHR data requires a specific set of skills, including expertise with (i) usability; (ii) data quality and validation; (iii) standards for data structuring; (iv) governance (including ethical, security, and privacy aspects related to the responsible use of clinical data); and (v) data integration [32,33]. Therefore, the ease with which one can tap into EHR data within an institute depends on the infrastructure and (programming) support that is available as well as the local legal and privacy regulations [34].

Opportunities for the use of EHR in circadian medicine

EHR can be used to address key questions in circadian medicine (Fig 1). Variables that are routinely monitored in patients show robust 24-hour rhythms in healthy individuals, such as vital signs (e.g., heart rate, blood pressure, and core body temperature) as well as biochemical,
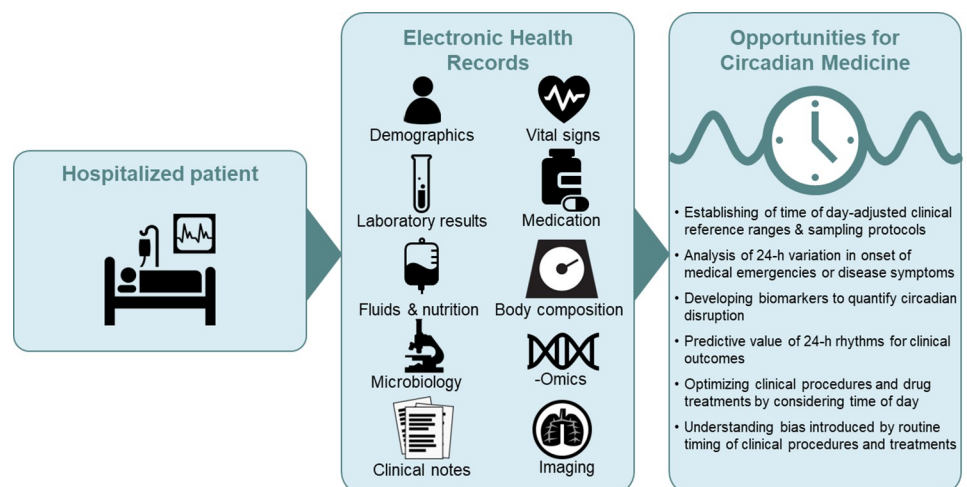

Fig 1. Overview of opportunities for the use of EHR for circadian medicine.

https://doi.org/10.1371/journal.pdig.0000511.g001

metabolic, and hematological factors [35–40]. Using EHR, researchers can evaluate these features in any disease context, providing real-world evidence for the hypothesis that circadian rhythms are a marker of healthy physiological functioning and that altered rhythms indicate poor health [41]. For example, an abnormal blood pressure rhythm, in particular the loss of overnight dipping, has been associated with an increased risk of death and cardiovascular events [42]. An EHR-based study of approximately 60,000 primary care patients found that nighttime systolic blood pressure is more informative about the risk of death than clinic or daytime ambulatory blood pressure, underscoring the value of large 24-hour datasets [43].

The effects of disease onset and hospitalization on 24-hour physiology have historically been evaluated in small, well-defined groups of patients [44–47]. Recent studies are tackling this question using large-scale, open-access EHR databases. For example, a study encompassing nearly 200,000 ICU patients spanning multiple hospitals in the United States of America and the United Kingdom detected population-level 24-hour rhythms in systolic blood pressure, heart rate, respiratory rate, and body temperature in the 24 hours prior to patient discharge. These rhythms were comparable to rhythms in healthy individuals, albeit with reduced peak-to-trough excursions [48]. This raises many questions, potentially also evaluable with EHR. For example, do these peak-to-trough excursions correlate with patient health status, and if so, can they be used as part of the decision-making process in patient discharge? EHR studies suggest that reduced day–night differences in vital signs are associated with increased mortality in hospitalized patients [49–53], but not all studies agree [54]. Future research is needed to understand if and how 24-hour physiology tracks health and predicts outcomes.

EHR-based studies of biomarker variation may also help to refine reference ranges [55]. Analysis of approximately 500,000 blood draws found that the reference interval for thyroid hormone (the range covering the central 95% of results) differs based on age, sex, ethnicity, and time of day [56]. Others have suggested that the time of day of sampling may be an important factor when assessing glucose control in ICU patients [57,58] or to help minimize under- or overdiagnosis in the context of adrenal gland disorders [59]. To speculate, the information gained from any laboratory test about the clinical status of a patient may improve with an understanding of the expected 24-hour dynamics, analogous to the case for blood pressure discussed above. EHR offers huge amounts of data to evaluate this, although analysis of time of day for blood draws is not without its challenges, as discussed in a subsequent section.

A key goal of circadian medicine is to optimize the timing of medication and clinical procedures [3]. Understanding how time of day modifies treatment efficacy and safety may lead to cost-effective and noninvasive timing-based strategies [60,61]. Although randomized controlled trials (RCTs) will remain the gold standard in this regard, current RCTs have some limitations. A survey of RCTs comparing the effects of medication taken at different times of day [62] pointed out that many studies (i) did not account for individual variation in circadian phase (e.g., 8 AM is not morning-time for a shift worker); (ii) often compare only 2 dosing times, morning versus evening; and (iii) were small and single-sited, leading to concerns about their replicability. Exploratory studies using EHR offer a practical solution to some of these shortcomings and may be used to guide confirmatory RCTs. For example, using EHR from a pediatric hospital, it was found that the clinical response to hydralazine, an acute antihypertensive, is greatest at night [63].

Further, by analyzing breakthrough Coronavirus disease 2019 (COVID-19) infection in 1.5 million patients as a function of time of vaccination, late morning to early afternoon was found to offer better protection than evening vaccination [64]. The authors of this study were able to construct a continuous efficacy curve over a 12-hour period using EHR data. It is difficult to imagine an RCT achieving similar temporal resolution at this scale. These examples

highlight the potential of everyday healthcare data to deliver insights into the impact of time on treatment and many other health measures.

In addition to considering time of day as a source of biological variation, the interaction between time of day and clinical routines, or healthcare processes, is another aspect of circadian medicine. Although clinical care takes place around the clock in hospitals, specific procedures or treatments are provided at specific times of day and not at other times. The timing of medication orders and administration are influenced by hospital rounding and may not be driven by the time of greatest clinical need [65,66]. Similarly, sampling for laboratory tests typically occurs in the early morning hours [67,68] and other care activities are—not surprisingly—unevenly distributed around the clock [69]. Given that all this information is typically documented in EHR, there is an opportunity to study how healthcare processes interact with biological variation.

We expect EHR-based circadian medicine research to grow as technological integration evolves. Wearable devices, for example, extend close monitoring of physiological and behavioral variables beyond the clinic into patient communities under real-world conditions. First evidence for this seamless integration of EHR with wearable devices are emergent [70,71], including wrist accelerometers [72] and sleep-monitoring technologies [73]. Wearable technology for longitudinal recording of physiological or behavioral variables is extensively adopted in the field of chronobiology to characterize 24-hour patterns in variables of interest. Some examples include activity trackers [74,75], continuous glucose monitoring [76], wrist temperature [22], and heart rate [76,77]. The relevance of using wearables for circadian monitoring in patient populations is exemplified by studies showing the predictive value of circadian metrics and clinical outcomes in patient populations, such as circadian features of heart rate as a prognostic marker for postoperative recovery scores [78] and rest–activity cycles as a predictor for patient survival [79] in oncology patients.

Furthermore, integration of molecular information with EHR will also pave the way for precision medicine [80]. For example, variation in circadian genes or genetic profiles related to circadian rhythms could inform treatment and risk prediction. In addition, further validation of blood-based circadian phenotyping using a single blood sample may allow for accurate determination of a patient's internal circadian time based on transcriptional biomarkers [80,82]. How best to incorporate those data with EHR and how to make this clinically actionable remains to be determined.

Challenges for the use of EHR for circadian medicine

EHR studies use data that were not collected to address predefined research questions and thus require a great deal of care in analysis and interpretation. Besides pitfalls that pertain to research with EHR in general, the use of these databases for circadian medicine purposes introduces multiple challenges (**Table 1**). The COVID-19 vaccine timing study [64] offers a good example. In that study, an individual's occupation may have independently influenced both the time of vaccination and the risk of COVID-19 infection (**Fig 2A**). To account for this, the authors carefully estimated a baseline infection rate for each comparison group by calculating COVID-19 positivity in the 14 days after the first vaccine dose, before any protection from the vaccine is expected. Although confounding is impossible to rule out, considering all of the potential causal relationships in the data and measuring them whenever possible increases the power and interpretability of EHR analyses.

It is easy to underestimate the degree to which time of day influences the interaction between a patient and the healthcare system. Two recent studies highlight how treatment [65] and diagnosis [68] follow well-established 24-hour patterns of care in the hospital (**Fig 2B**).

Table 1. Challenges and recommendations on applying EHR to studies of circadian medicine.

Challenges	Recommendations	Guiding questions
Sample selection bias EHR are often collected from specific healthcare systems or patient populations. As such, it may not represent the entire population.	Always remember that data collection was not guided by research questions and is greatly influenced by the functioning of the healthcare institution.	• Have we properly discussed external validity and generalizability?
Heterogeneity in data collection Procedures and type of data collected (e.g., which patients, sampling frequency, different technologies) heavily depend on clinical routines and administrative purposes.	Consider potential biases by drawing out causal diagrams, and engaging domain and statistical expertise before, during, and after analysis.	• Could daily patterns of hospital operation have imprinted daily patterns of clinical values? • Are there hidden variables (e.g., disease severity) that may have influenced both the time of day of measurement and the outcome of interest (e.g., white blood cell count)? • Which of all the potentially causal variables can we account for (and not account for) in our analysis?
Variability in data storage/management Data management practices may vary significantly across institutions.	Confirm that event date-time stamps are reliable.	• Were date-times shifted to protect privacy? • Could there be a significant date-time lag between the event and its electronic entry?
	Consider whether the clinical features important to your study may be imprecisely defined.	• Could the diagnosis code that we used to define a cohort have been assigned primarily for billing purposes? How might this affect my results?
Data structure/types Multilevel data: Data may involve a single measure from each of N patients or multiple measures from each of N patients.	Consider the independence of observations in your dataset.	• Have we accounted for hierarchical structure in our dataset? For example, repeated hospital admissions for the same patient cannot be treated as independent observations. Similarly, repeated measures on a patient during an admission cannot be treated as independent observations.
Time of day is a circular variable: It represents a cyclical phenomenon.	Avoid discretizing circular variables.	• Can we treat time of day as a circular variable in order to avoid binning observations into wide intervals (e.g., day vs. night)?
Reproducibility-related challenges EHR data has sensitive patient information, and strict privacy regulations have to be followed when assessing, using and sharing these data for research purposes. Data are often stored in different formats and systems, which make integration difficult. Preprocessing of data can be undertaken in different ways.	Follow best-practice statements developed specifically for EHR-based studies.	• Was CODE-EHR or a similar framework used when planning and conducting the study?
	Determine how to share data, code, and software needed for someone else to recreate findings and figures.	• How do we legally and responsibly share the data and code? • How do we future-proof this access, for example, how do we prevent broken web links? • How do we meet the requirements for journals' data/code availability statements?
	When drafting a manuscript, follow appropriate reporting guidelines.	• Have we checked RECORD (observational) or CONSORT-ROUTINE (intervention studies)? • Did we openly discuss the biases we are aware of (there are always plenty of them) in our manuscript?

https://doi.org/10.1371/journal.pdig.0000511.t001

For example, blood and urine specimens are routinely collected in the early morning hours so that the test results are available for clinical team rounding that typically occurs 3 to 4 hours later. Of course, not all specimens are collected in this routine window. In particular, a patient with acutely concerning signs or symptoms (and, therefore, a higher likelihood of abnormal test values compared to more stable patients) may receive off-peak testing. To this point, the time of day that a lab test is ordered predicts mortality, in some cases even more powerfully than the test result itself [67]. In other words, the decision to acquire data is influenced by clinicians' perception of clinical need in a background of operational routine (**Fig 2C and 2D**). As a consequence, it may be easy to mistake patterns in hospital operations for rhythms in biology (**Fig 2D**). For example, 24-hour variation in glucose levels in EHR data from critically ill patients was initially attributed to circadian variation in glucose control [57] but was later shown to be—at least partially—caused by more intensive sampling in sicker patients [58]. When using time-weighted averages, the time of day effect was greatly attenuated [58].
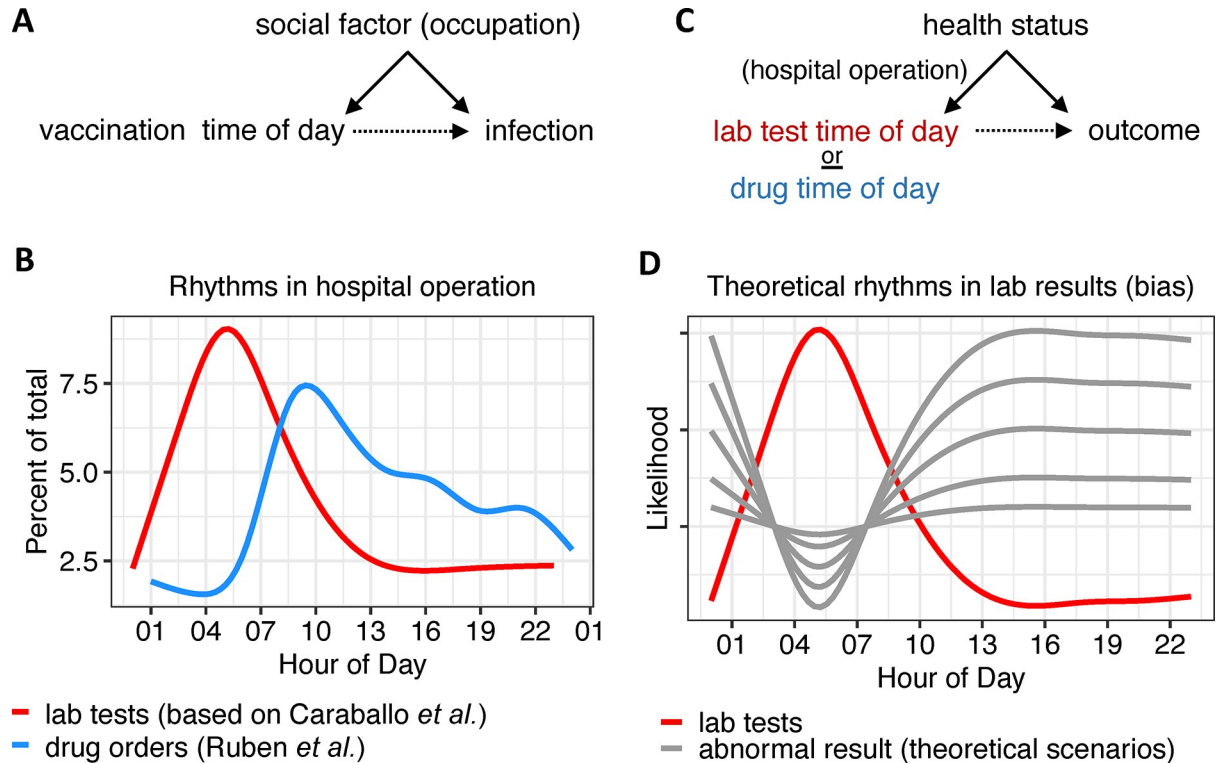
Fig 2. Considering causes in EHR studies of biological rhythms. (A) An individual's occupation may independently "cause" both the time of vaccination (scheduling/availability) and the risk of COVID-19 infection (exposure). A way to think about cause in these diagrams is that the variable being pointed to "listens to" the variable pointing to it. **(B)** Twenty-four-hour rhythms in hospital operation based on data from Ruben and colleagues [65] and Caraballo and colleagues [68]. **(C)** Proposed causal model for the effect of time of day of lab test (or drug administration) on outcome. This model assumes that health status is a confounder because it independently influences both the time of lab test (or drug administration) and the outcome of interest. **(D)** Given the causal model for lab results in panel C, we propose that the probability of an abnormal result at a particular time is inversely related to the probability of testing at that time. Theoretical curves (gray color) reflect different scaling, i.e., proposed relationships between the probability of testing and probability of obtaining an abnormal result. Ground truth is unknown.

https://doi.org/10.1371/journal.pdig.0000511.g002

Considering the potential influences on an outcome, drawing causal diagrams [83], and engaging domain expertise can help to clarify the clinical process and determine the best statistical approach [84]. Propensity score matching may be used to ensure patient subgroups are comparable for analyses [85]. Still, propensity matching requires (1) awareness of potential causal factors and (2) a way to measure them. This can be hard. For example, we may be aware that health status causes the collection of lab values and, therefore, try to "match" health status between comparison groups. But how do we measure health status? Can we disentangle these factors, given that time of day of testing predicts health status, in some cases even better than the test results themselves [67]? This is an interesting form of confounding that makes it difficult to ever fully control for health status when trying to study biological rhythms in lab values. A similar issue may present in studies on medication timing (**Fig 2B**) where treatment received "off peak" may describe different types of patient.

Inpatients may receive automated 24-hour continuous monitoring of clinical parameters, including blood pressure, heart rate, body temperature, respiratory rate, and oxygen saturation. Automation mitigates some but not all concerns of the bias discussed above. Consider that not all inpatients receive 24-hour monitoring. It is a decision, influenced by operational routine and perceived clinical need. The patient population that is being closely monitored may not be representative of all individuals with the condition of interest. In addition, the use

of routine dosing times of drugs may impose 24-hour rhythms in monitored parameters; for example, the administration of vasodilators or vasopressors during a specific time window may produce a 24-hour pattern in blood pressure. It is possible to exclude data points collected within the time frame that drug administration occurs to limit this effect [48]; however this may exclude a substantial amount of data, especially in settings like the ICU where medication use is high.

In the vaccination study above, and other EHR-based assessments of timing effects [48,50], data from many individuals are compressed onto a single hypothetical 24-hour cycle and analyzed. This may involve a single measure from each of N patients or multiple measures from each of N patients. Either strategy helps to "cover the clock" when individual sampling is sparse but can create imbalances in statistical power across the 24-hour cycle. Mixed-effects models or other statistical methods that take into account repeated sampling may be useful to account for interindividual variability in baseline or responses in case patients contribute multiple measures [86]. In addition, it should be noted that time of day is circular and, although it may be tempting to discretize observations into "morning" or "evening" (or something similar), time is best handled as a circular predictor to avoid information loss. The development of statistical methods to describe periodic processes is an active area of research [75,87,88].

As a final point, precise definition of the variables under study is crucial. In circadian medicine, it is important to verify that time of day is correctly specified. This may sound trivial but open-access EHR databases may shift dates and times to protect patient anonymity. For example, while time of day information is preserved in the MIMIC-IV, dates have been shifted by a random number of days, precluding seasonal or day of week analyses [27]. In some databases, time may be parameterized relative to when the patient was first admitted, making it impossible to study time of day effects [30].

Conclusions

The promise of circadian medicine is clear: There are numerous opportunities to improve diagnostics, treatments, and patient care by incorporating circadian principles. A goal of future research is to expand the evidence base for these approaches in real-world clinical settings. With this goal in mind, we offer a few recommendations for the use of EHR in circadian medicine (see **Table 1**).

First, we recommend evaluating the suitability of variables for circadian research based on data collection methods rather than data type. EHR-based circadian research has mostly focused on vital signs like heart rate, blood pressure, and body temperature. However, as discussed throughout this review, the EHR holds data on numerous other variables that could offer valuable insights into circadian effects on health and disease. These variables include information about patients' diet, exercise, posture, sleep, light exposure, urinary output, nurse call button usage, pain levels, and much more. The collection process is not always transparent but almost always impacts data quality, resolution, and potential bias, as exemplified by the distinction between automated and nonautomated data collection. The process of clinical data collection is evolving, with technological advances allowing continuous recording of physiology that is currently measured only at limited time points [76,79]. These advances offer promising opportunities for circadian medicine research.

Moreover, it is important to know the intricacies of a database before analyzing its content. Random shuffling of times and dates for deidentification can go unnoticed, jeopardizing studies. Second, responsible sharing of deidentified raw patient-level data according to FAIR (Findable, Accessible, Interoperable, Reusable) principles [90] enhances transparency, accelerates the development of data-driven methods, and allows benchmarking across multiple

cohorts, as is done in the field of sleep research [91]. Institutional or national regulations may preclude data sharing, especially in the case of sensitive medical data, although informative examples of data sharing strategies that comply with national privacy regulations exist [30].

To ensure reproducibility, it is considered good practice to share and maintain any code or applications used to process and analyze the data [92]. When embarking on a study involving EHR, we encourage authors to follow best-practice statements that have been specifically developed for this purpose, such as the CODE-EHR framework [93]. Likewise, when drafting a manuscript, reporting guidelines intended for studies using routinely collected clinical data should be followed, such as the RECORD guidelines for observational studies [2] or the CON-SORT-ROUTINE guidelines for intervention studies [94]. This will enhance transparency and reproducibility of the study and facilitate systematic reviews and meta-analyses.

In general, we expect that the opportunities, challenges, and recommendations discussed in this review will help to future-proof the research in the field of circadian medicine. EHR may indeed be a treasure trove for the field, but—like real treasure troves—only if they are handled with care.

Acknowledgments

Author Contributions

Conceptualization: Laura Kervezee.

Visualization: Laura Kervezee, Marc D. Ruben.

Writing – original draft: Laura Kervezee, Marc D. Ruben.

Writing – review & editing: Laura Kervezee, Hassan S. Dashti, Luísa K. Pilz, Carsten Skarke, Marc D. Ruben.

References

1. Sanchez-Pinto LN, Luo Y, Churpek MM. Big Data and Data Science in Critical Care. Chest. 2018; 154 (5):1239–1248. https://doi.org/10.1016/j.chest.2018.04.037 PMID: 29752973

2. Benchimol EI, Smeeth L, Guttmann A, Harron K, Moher D, Petersen I, et al. The REporting of studies Conducted using Observational Routinely-collected health Data (RECORD) statement. PLoS Med. 2015; 12(10):e1001885. https://doi.org/10.1371/journal.pmed.1001885 PMID: 26440803

3. Kramer A, Lange T, Spies C, Finger AM, Berg D, Oster H. Foundations of circadian medicine. PLoS Biol. 2022; 20(3):e3001567. https://doi.org/10.1371/journal.pbio.3001567 PMID: 35324893

4. Abul-Husn NS, Kenny EE. Personalized Medicine and the Power of Electronic Health Records. Cell. 2019; 177(1):58–69. https://doi.org/10.1016/j.cell.2019.02.039 PMID: 30901549

5. Sauer CM, Chen LC, Hyland SL, Girbes A, Elbers P, Celi LA. Leveraging electronic health records for data science: common pitfalls and how to avoid them. Lancet Digit Health. 2022; 4(12):e893–e898. https://doi.org/10.1016/S2589-7500(22)00154-6 PMID: 36154811

6. Zhang J, Symons J, Agapow P, Teo JT, Paxton CA, Abdi J, et al. Best practices in the real-world data life cycle. PLoS Digit Health. 2022; 1(1):e0000003. https://doi.org/10.1371/journal.pdig.0000003 PMID: 36812509

7. Wang C, Lutes LK, Barnoud C, Scheiermann C. The circadian immune system. Sci Immunol. 2022; 7 (72):eabm2465. https://doi.org/10.1126/sciimmunol.abm2465 PMID: 35658012

8. Rijo-Ferreira F, Takahashi JS. Genomics of circadian rhythms in health and disease. Genome Med. 2019; 11(1):82. https://doi.org/10.1186/s13073-019-0704-0 PMID: 31847894

9. Voigt RM, Forsyth CB, Keshavarzian A. Circadian rhythms: a regulator of gastrointestinal health and dysfunction. Expert Rev Gastroenterol Hepatol. 2019; 13(5):411–424. https://doi.org/10.1080/17474124.2019.1595588 PMID: 30874451

10. Crnko S, Du Pre BC, Sluijter JPG, Van Laake LW. Circadian rhythms and the molecular clock in cardio-vascular biology and disease. Nat Rev Cardiol. 2019; 16(7):437–447. https://doi.org/10.1038/s41569-019-0167-4 PMID: 30796369

11. Reinke H, Asher G. Crosstalk between metabolism and circadian clocks. Nat Rev Mol Cell Biol. 2019; 20(4):227–241. https://doi.org/10.1038/s41580-018-0096-9 PMID: 30635659

12. Roenneberg T, Merrow M. The Circadian Clock and Human Health. Curr Biol. 2016; 26(10):R432–R443. https://doi.org/10.1016/j.cub.2016.04.011 PMID: 27218855

13. Allada R, Bass J. Circadian Mechanisms in Medicine. N Engl J Med. 2021; 384(6):550–561. https://doi.org/10.1056/NEJMra1802337 PMID: 33567194

14. Takahashi JS. Transcriptional architecture of the mammalian circadian clock. Nat Rev Genet. 2017; 18 (3):164–179. https://doi.org/10.1038/nrg.2016.150 PMID: 27990019

15. Finger AM, Kramer A. Mammalian circadian systems: Organization and modern life challenges. Acta Physiol (Oxford). 2021; 231(3):e13548. https://doi.org/10.1111/apha.13548 PMID: 32846050

16. Guo YF, Stein PK. Circadian rhythm in the cardiovascular system: chronocardiology. Am Heart J. 2003; 145(5):779–786. https://doi.org/10.1016/S0002-8703(02)94797-6 PMID: 12766733

17. Thosar SS, Shea SA. Circadian control of human cardiovascular function. Curr Opin Pharmacol. 2021; 57:89–97. https://doi.org/10.1016/j.coph.2021.01.002 PMID: 33610933

18. Duffy JF, Dijk DJ. Getting through to circadian oscillators: why use constant routines? J Biol Rhythm. 2002; 17(1):4–13. https://doi.org/10.1177/074873002129002294 PMID: 11837947

19. Wang W, Yuan RK, Mitchell JF, Zitting KM, St Hilaire MA, Wyatt JK, et al. Desynchronizing the sleep—wake cycle from circadian timing to assess their separate contributions to physiology and behaviour and to estimate intrinsic circadian period. Nat Protoc. 2023; 18(2):579–603. https://doi.org/10.1038/s41596-022-00746-y PMID: 36376588

20. Dijk DJ, Duffy JF. Novel Approaches for Assessing Circadian Rhythmicity in Humans: A Review. J Biol Rhythm. 2020; 35(5):421–438. https://doi.org/10.1177/0748730420940483 PMID: 32700634

21. Shea SA, Hilton MF, Hu K, Scheer FA. Existence of an endogenous circadian blood pressure rhythm in humans that peaks in the evening. Circ Res. 2011; 108(8):980–984. https://doi.org/10.1161/CIRCRESAHA.110.233668 PMID: 21474818

22. Brooks TG, Lahens NF, Grant GR, Sheline YI, FitzGerald GA, Skarke C. Diurnal rhythms of wrist tem-perature are associated with future disease risk in the UK Biobank. Nat Commun. 2023; 14(1):5172. https://doi.org/10.1038/s41467-023-40977-5 PMID: 37620332

23. Rudrapatna VA, Butte AJ. Opportunities and challenges in using real-world data for health care. J Clin Invest. 2020; 130(2):565–574. https://doi.org/10.1172/JCI129197 PMID: 32011317

24. Sherman RE, Anderson SA, Dal Pan GJ, Gray GW, Gross T, Hunter NL, et al. Real-World Evidence—What Is It and What Can It Tell Us? N Engl J Med. 2016; 375(23):2293–2297. https://doi.org/10.1056/NEJMsb1609216 PMID: 27959688

25. Cade BE, Hassan SM, Dashti HS, Kiernan M, Pavlova MK, Redline S, et al. Sleep apnea phenotyping and relationship to disease in a large clinical biobank. Jamia Open. 2022; 5(1):ooab117. https://doi.org/10.1093/jamiaopen/ooab117 PMID: 35156000

26. Sauer CM, Dam TA, Celi LA, Faltys M, de la Hoz MAA, Adhikari L, et al. Systematic Review and Com-parison of Publicly Available ICU Data Sets-A Decision Guide for Clinicians and Data Scientists. Crit Care Med. 2022; 50(6):e581–e588. https://doi.org/10.1097/CCM.0000000000005517 PMID: 35234175

27. Johnson AEW, Bulgarelli L, Shen L, Gayles A, Shammout A, Horng S, et al. MIMIC-IV, a freely accessi-ble electronic health record dataset. Sci Data. 2023; 10 (1):1. https://doi.org/10.1038/s41597-022-01899-x PMID: 36596836

28. Pollard TJ, Johnson AEW, Raffa JD, Celi LA, Mark RG, Badawi O. The eICU Collaborative Research Database, a freely available multi-center database for critical care research. Sci Data. 2018; 5:180178. https://doi.org/10.1038/sdata.2018.178 PMID: 30204154

29. Hyland SL, Faltys M, Huser M, Lyu X, Gumbsch T, Esteban C, et al. Early prediction of circulatory failure in the intensive care unit using machine learning. Nat Med. 2020; 26(3):364–373. https://doi.org/10.1038/s41591-020-0789-4 PMID: 32152583

30. Thoral PJ, Peppink JM, Driessen RH, Sijbrands EJG, Kompanje EJO, Kaplan L, et al. Sharing ICU Patient Data Responsibly Under the Society of Critical Care Medicine/European Society of Intensive Care Medicine Joint Data Science Collaboration: The Amsterdam University Medical Centers Database (AmsterdamUMCdb) Example. Crit Care Med. 2021; 49(6):e563–e577. https://doi.org/10.1097/CCM.0000000000004916 PMID: 33625129

31. Maletzky A, Bock C, Tschoellitsch T, Roland T, Ludwig H, Thumfart S, et al. Lifting Hospital Electronic Health Record Data Treasures: Challenges and Opportunities. JMIR Med Inform. 2022; 10(10):e38557. https://doi.org/10.2196/38557 PMID: 36269654

32. Holmes JH, Beinlich J, Boland MR, Bowles KH, Chen Y, Cook TS, et al. Why is the electronic health record so challenging for research and clinical care? Methods Inf Med. 2021; 60(1–02):32–48. https://doi.org/10.1055/s-0041-1731784 PMID: 34282602

33. Cowie MR, Blomster JI, Curtis LH, Duclaux S, Ford I, Fritz F, et al. Electronic health records to facilitate clinical research. Clin Res Cardiol. 2017; 106(1):1–9. https://doi.org/10.1007/s00392-016-1025-6 PMID: 27557678

34. Richesson RL, Marsolo KS, Douthit BJ, Staman K, Ho PM, Dailey D, et al. Enhancing the use of EHR systems for pragmatic embedded research: lessons from the NIH Health Care Systems Research Collaboratory. J Am Med Inform Assoc. 2021; 28(12):2626–2640. https://doi.org/10.1093/jamia/ocab202 PMID: 34597383

35. Kanabrocki EL, Sothern RB, Scheving LE, Vesely DL, Tsai TH, Shelstad J, et al. Reference values for circadian rhythms of 98 variables in clinically healthy men in the fifth decade of life. Chronobiol Int. 1990; 7(5–6):445–461. https://doi.org/10.3109/07420529009059156 PMID: 2097078

36. Scheer FA, Hu K, Evoniuk H, Kelly EE, Malhotra A, Hilton MF, et al. Impact of the human circadian system, exercise, and their interaction on cardiovascular function. Proc Natl Acad Sci U S A. 2010; 107 (47):20541–20546. https://doi.org/10.1073/pnas.1006749107 PMID: 21059915

37. Sennels HP, Jorgensen HL, Hansen AL, Goetze JP, Fahrenkrug J. Diurnal variation of hematology parameters in healthy young males: the Bispebjerg study of diurnal variations. Scand J Clin Lab Invest. 2011; 71(7):532–541. https://doi.org/10.3109/00365513.2011.602422 PMID: 21988588

38. Sennels HP, Jorgensen HL, Goetze JP, Fahrenkrug J. Rhythmic 24-hour variations of frequently used clinical biochemical parameters in healthy young males—the Bispebjerg study of diurnal variations. Scand J Clin Lab Invest. 2012; 72(4):287–295. https://doi.org/10.3109/00365513.2012.662281 PMID: 22397709

39. Chellappa SL, Vujovic N, Williams JS, Scheer F. Impact of Circadian Disruption on Cardiovascular Function and Disease. Trends Endocrinol Metab. 2019; 30(10):767–779. https://doi.org/10.1016/j.tem.2019.07.008 PMID: 31427142

40. Grant LK, Czeisler CA, Lockley SW, Rahman SA. Time-of-day and Meal Size Effects on Clinical Lipid Markers. J Clin Endocrinol Metab. 2021; 106(3):e1373–e1379. https://doi.org/10.1210/clinem/dgaa739 PMID: 33051649

41. Oldham MA, Lee HB, Desan PH. Circadian Rhythm Disruption in the Critically Ill: An Opportunity for Improving Outcomes. Crit Care Med. 2016; 44(1):207–217. https://doi.org/10.1097/CCM.0000000000001282 PMID: 26308428

42. Fagard RH, Thijs L, Staessen JA, Clement DL, De Buyzere ML, De Bacquer DA. Night-day blood pressure ratio and dipping pattern as predictors of death and cardiovascular events in hypertension. J Hum Hypertens. 2009; 23(10):645–653. https://doi.org/10.1038/jhh.2009.9 PMID: 19225527

43. Staplin N, de la Sierra A, Ruilope LM, Emberson JR, Vinyoles E, Gorostidi M, et al. Relationship between clinic and ambulatory blood pressure and mortality: an observational cohort study in 59 124 patients. Lancet. 2023; 401(10393):2041–2050. https://doi.org/10.1016/S0140-6736(23)00733-X PMID: 37156250

44. Tweedie IE, Bell CF, Clegg A, Campbell IT, Minors DS, Waterhouse JM. Retrospective study of temperature rhythms of intensive care patients. Crit Care Med. 1989; 17(11):1159–1165. https://doi.org/10.1097/00003246-198911000-00012 PMID: 2791594

45. Paul T, Lemmer B. Disturbance of circadian rhythms in analgosedated intensive care unit patients with and without craniocerebral injury. Chronobiol Int. 2007; 24(1):45–61. https://doi.org/10.1080/07420520601142569 PMID: 17364579

46. Gazendam JAC, Van Dongen HPA, Grant DA, Freedman NS, Zwaveling JH, Schwab RJ. Altered circadian rhythmicity in patients in the ICU. Chest. 2013; 144(2):483–489. https://doi.org/10.1378/chest.12-2405 PMID: 23471224

47. Knauert MP, Haspel JA, Pisani MA. Sleep Loss and Circadian Rhythm Disruption in the Intensive Care Unit. Clin Chest Med. 2015; 36(3):419–429. https://doi.org/10.1016/j.ccm.2015.05.008 PMID: 26304279

48. Davidson S, Villarroel M, Harford M, Finnegan E, Jorge J, Young D, et al. Vital-sign circadian rhythms in patients prior to discharge from an ICU: a retrospective observational analysis of routinely recorded physiological data. Crit Care. 2020; 24(1):181. https://doi.org/10.1186/s13054-020-02861-2 PMID: 32345354

49. Gao Y, Wang Q, Li J, Zhang J, Li R, Sun L, et al. Impact of Mean Arterial Pressure Fluctuation on Mortality in Critically Ill Patients. Crit Care Med. 2018; 46(12):e1167–e1174. https://doi.org/10.1097/CCM.0000000000003435 PMID: 30247271

50. Beyer SE, Salgado C, Garcao I, Celi LA, Vieira S. Circadian rhythm in critically ill patients: Insights from the eICU Database. Cardiovasc Digit Health J. 2021; 2(2):118–125. https://doi.org/10.1016/j.cvdhj.2021.01.004 PMID: 35265899

51. Davidson S, Villarroel M, Harford M, Finnegan E, Jorge J, Young D, et al. Day-to-day progression of vital-sign circadian rhythms in the intensive care unit. Crit Care. 2021; 25(1):156. https://doi.org/10.1186/s13054-021-03574-w PMID: 33888129

52. van Goor HMR, van Loon K, Breteler MJM, Kalkman CJ, Kaasjager KAH. Circadian patterns of heart rate, respiratory rate and skin temperature in hospitalized COVID-19 patients. PLoS ONE. 2022; 17(7): e0268065. https://doi.org/10.1371/journal.pone.0268065 PMID: 35797369

53. Yang Z, Xie X, Zhang X, Li L, Bai R, Long H, et al. Circadian rhythms of vital signs are associated with in-hospital mortality in critically ill patients: A retrospective observational study. Chronobiol Int. 2023; 40 (3):262–271. https://doi.org/10.1080/07420528.2022.2163656 PMID: 36597185

54. Boots R, Xue G, Tromp D, Rawashdeh O, Bellapart J, Townsend S, et al. Circadian Rhythmicity of Vital Signs at Intensive Care Unit Discharge and Outcome of Traumatic Brain Injury. Am J Crit Care. 2022; 31(6):472–482. https://doi.org/10.4037/ajcc2022821 PMID: 36316179

55. Feriel J, Tchipeva D, Depasse F. Effects of circadian variation, lifestyle and environment on hematological parameters: A narrative review. Int J Lab Hematol. 2021; 43(5):917–926. https://doi.org/10.1111/ijlh.13590 PMID: 34019728

56. Ehrenkranz J, Bach PR, Snow GL, Schneider A, Lee JL, Ilstrup S, et al. Circadian and Circannual Rhythms in Thyroid Hormones: Determining the TSH and Free T4 Reference Intervals Based Upon Time of Day, Age, and Sex. Thyroid. 2015; 25(8):954–961. https://doi.org/10.1089/thy.2014.0589 PMID: 26061389

57. Egi M, Bellomo R, Stachowski E, French CJ, Hart G, Stow P. Circadian rhythm of blood glucose values in critically ill patients. Crit Care Med. 2007; 35(2):416–421. https://doi.org/10.1097/01.CCM.0000253814.78836.43 PMID: 17205020

58. Vogelzang M, van der Horst IC, Zijlstra F, Nijsten MW. Circadian variation of glucose levels: biology or timing of measurements? Crit Care Med. 2007;35(7):1800–1801author reply 1–2. https://doi.org/10.1097/01.CCM.0000269406.40845.B5 PMID: 17581376

59. Thosar SS, Rueda JF, Berman AM, Lasarev MR, Herzig MX, Clemons NA, et al. Separate and interacting effects of the endogenous circadian system and behaviors on plasma aldosterone in humans. Am J Phys Regul Integr Comp Phys. 2019; 316(2):R157–R164. https://doi.org/10.1152/ajpregu.00314.2018 PMID: 30521366

60. Cederroth CR, Albrecht U, Bass J, Brown SA, Dyhrfjeld-Johnsen J, Gachon F, et al. Medicine in the Fourth Dimension. Cell Metab. 2019; 30(2):238–250. https://doi.org/10.1016/j.cmet.2019.06.019 PMID: 31390550

61. Walton JC, Walker WH 2nd, Bumgarner JR, Melendez-Fernandez OH, Liu JA, Hughes HL, et al. Circadian Variation in Efficacy of Medications. Clin Pharmacol Ther. 2021; 109(6):1457–1488. https://doi.org/10.1002/cpt.2073 PMID: 33025623

62. Ruben MD, Francey LJ, Wu G, Smith DF, FitzGerald GA, Hughey JJ, Hogenesch JB. Factors Associated with Medicine Timing Effects: A Meta-analysis. medRxiv. 2021:2021.10.24.21265348. https://doi.org/10.1101/2021.10.24.21265348

63. Ruben MD, Smith DF, FitzGerald GA, Hogenesch JB. Dosing time matters. Science. 2019; 365 (6453):547–549. https://doi.org/10.1126/science.aax7621 PMID: 31395773

64. Hazan G, Duek OA, Alapi H, Mok H, Ganninger A, Ostendorf E, et al. Biological rhythms in COVID-19 vaccine effectiveness in an observational cohort study of 1.5 million patients. J Clin Invest. 2023; 133 (11):e167339. https://doi.org/10.1172/JCI167339 PMID: 37053011

65. Ruben MD, Francey LJ, Guo Y, Wu G, Cooper EB, Shah AS, et al. A large-scale study reveals 24-h operational rhythms in hospital treatment. Proc Natl Acad Sci U S A. 2019; 116(42):20953–20958. https://doi.org/10.1073/pnas.1909557116 PMID: 31575744

66. Loput CM, Saltsman CL, Rahm RC, Roberts WD, Sharma S, Borum C, et al. Evaluation of medication administration timing variance using information from a large health system's clinical data warehouse. Am J Health Syst Pharm. 2022; 79(Suppl 1):S1–S7. https://doi.org/10.1093/ajhp/zxab378 PMID: 34653239

67. Agniel D, Kohane IS, Weber GM. Biases in electronic health record data due to processes within the healthcare system: retrospective observational study. BMJ. 2018; 361:k1479. https://doi.org/10.1136/bmj.k1479 PMID: 29712648

68. Caraballo C, Mahajan S, Murugiah K, Mortazavi BJ, Lu Y, Khera R, et al. Timing of Blood Draws Among Patients Hospitalized in a Large Academic Medical Center. JAMA. 2023; 329(3):255–257. https://doi.org/10.1001/jama.2022.21509 PMID: 36648476

69. Schlosser KR, Fiore GA, Smallwood CD, Griffin JF, Geva A, Santillana M, et al. Noninvasive Ventilation Is Interrupted Frequently and Mostly Used at Night in the Pediatric Intensive Care Unit. Respir Care. 2020; 65(3):341–346. https://doi.org/10.4187/respcare.06883 PMID: 31551282

70. Genes N, Violante S, Cetrangol C, Rogers L, Schadt EE, Chan YY. From smartphone to EHR: a case report on integrating patient-generated health data. NPJ Digit Med. 2018; 1:23. https://doi.org/10.1038/s41746-018-0030-8 PMID: 31304305

71. Biersteker T, Hilt A, van der Velde E, Schalij MJ, Treskes RW. Real-World Experience of mHealth Implementation in Clinical Practice (the Box): Design and Usability Study. JMIR Cardio. 2021; 5(2): e26072. https://doi.org/10.2196/26072 PMID: 34642159

72. Dinh-Le C, Chuang R, Chokshi S, Mann D. Wearable Health Technology and Electronic Health Record Integration: Scoping Review and Future Directions. JMIR Mhealth Uhealth. 2019; 7(9):e12861. https://doi.org/10.2196/12861 PMID: 31512582

73. Perez-Pozuelo I, Zhai B, Palotti J, Mall R, Aupetit M, Garcia-Gomez JM, et al. The future of sleep health: a data-driven revolution in sleep science and medicine. NPJ Digit Med. 2020; 3:42. https://doi.org/10.1038/s41746-020-0244-4 PMID: 32219183

74. Ancoli-Israel S, Cole R, Alessi C, Chambers M, Moorcroft W, Pollak CP. The role of actigraphy in the study of sleep and circadian rhythms. Sleep. 2003; 26(3):342–392. https://doi.org/10.1093/sleep/26.3.342 PMID: 12749557

75. Bowman C, Huang Y, Walch OJ, Fang Y, Frank E, Tyler J, et al. A method for characterizing daily physiology from widely used wearables. Cell Rep Methods. 2021; 1(4):100058. https://doi.org/10.1016/j.crmeth.2021.100058 PMID: 34568865

76. Phillips NE, Collet TH, Naef F. Uncovering personalized glucose responses and circadian rhythms from multiple wearable biosensors with Bayesian dynamical modeling. Cell Rep Methods. 2023; 3 (8):100545. https://doi.org/10.1016/j.crmeth.2023.100545 PMID: 37671030

77. Wiersema JM, Kamphuis AEP, Rohling JHT, Kervezee L, Akintola AA, Jansen SW, et al. The association between continuous ambulatory heart rate, heart rate variability, and 24-h rhythms of heart rate with familial longevity and aging. Aging (Albany NY). 2022; 14(18):7223–7239. https://doi.org/10.18632/aging.204219 PMID: 35980264

78. van den Eijnden MAC, van der Stam JA, Bouwman RA, Mestrom EHJ, Verhaegh WFJ, van Riel NAW, et al. Machine Learning for Postoperative Continuous Recovery Scores of Oncology Patients in Perioperative Care with Data from Wearables. Sensors (Basel). 2023; 23(9):4455. https://doi.org/10.3390/s23094455 PMID: 37177659

79. Levi F, Dugue PA, Innominato P, Karaboue A, Dispersyn G, Parganiha A, et al. Wrist actimetry circadian rhythm as a robust predictor of colorectal cancer patients survival. Chronobiol Int. 2014; 31 (8):891–900. https://doi.org/10.3109/07420528.2014.924523 PMID: 24927369

80. Li R, Chen Y, Ritchie MD, Moore JH. Electronic health records and polygenic risk scores for predicting disease risk. Nat Rev Genet. 2020; 21(8):493–502. https://doi.org/10.1038/s41576-020-0224-1 PMID: 32235907

81. Wittenbrink N, Ananthasubramaniam B, Munch M, Koller B, Maier B, Weschke C, et al. High-accuracy determination of internal circadian time from a single blood sample. J Clin Invest. 2018; 128(9):3826–3839. https://doi.org/10.1172/JCI120874 PMID: 29953415

82. Braun R, Kath WL, Iwanaszko M, Kula-Eversole E, Abbott SM, Reid KJ, et al. Universal method for robust detection of circadian state from gene expression. Proc Natl Acad Sci U S A. 2018; 115(39): E9247–E9256. https://doi.org/10.1073/pnas.1800314115 PMID: 30201705

83. Tennant PWG, Murray EJ, Arnold KF, Berrie L, Fox MP, Gadd SC, et al. Use of directed acyclic graphs (DAGs) to identify confounders in applied health research: review and recommendations. Int J Epidemiol. 2021; 50(2):620–632. https://doi.org/10.1093/ije/dyaa213 PMID: 33330936

84. Pearl J, Mackenzie D. The book of why: the new science of cause and effect. Basic Books; 2018.

85. Yadav P, Steinbach M, Kumar V, Simon G. Mining Electronic Health Records (EHRs): A Survey. ACM Comput Surv. 2018; 50(6):1–40. https://doi.org/10.1145/3127881

86. Hickey GL, Mokhles MM, Chambers DJ, Kolamunnage-Dona R. Statistical primer: performing repeated-measures analysis. Interact Cardiovasc Thorac Surg. 2018; 26(4):539–544. https://doi.org/10.1093/icvts/ivy009 PMID: 29596693

87. Doyle MM, Murphy TE, Miner B, Pisani MA, Lusczek ER, Knauert MP. Enhancing cosinor analysis of circadian phase markers using the gamma distribution. Sleep Med. 2022; 92:1–3. https://doi.org/10.1016/j.sleep.2022.01.015 PMID: 35306404

88. Baum L, Johns M, Poikela M, Moller R, Ananthasubramaniam B, Prasser F. Data integration and analysis for circadian medicine. Acta Physiol (Oxford). 2023; 237(4):e13951. https://doi.org/10.1111/apha.13951 PMID: 36790321

89. Upton TJ, Zavala E, Methlie P, Kampe O, Tsagarakis S, Oksnes M, et al. High-resolution daily profiles of tissue adrenal steroids by portable automated collection. Sci Transl Med. 2023; 15(701):eadg8464. https://doi.org/10.1126/scitranslmed.adg8464 PMID: 37343084

90. Wilkinson MD, Dumontier M, Aalbersberg IJ, Appleton G, Axton M, Baak A, et al. The FAIR Guiding Principles for scientific data management and stewardship. Sci Data. 2016; 3(1):160018. https://doi.org/10.1038/sdata.2016.18 PMID: 26978244

91. Palotti J, Mall R, Aupetit M, Rueschman M, Singh M, Sathyanarayana A, et al. Benchmark on a large cohort for sleep-wake classification with machine learning techniques. NPJ Digit Med. 2019; 2:50. https://doi.org/10.1038/s41746-019-0126-9 PMID: 31304396

92. Ostropolets A, Albogami Y, Conover M, Banda JM, Baumgartner WA, Blacketer C, et al. Reproducible variability: assessing investigator discordance across 9 research teams attempting to reproduce the same observational study. J Am Med Inform Assoc. 2023; 30(5):859–868. https://doi.org/10.1093/jamia/ocad009 PMID: 36826399

93. Kotecha D, Asselbergs FW, Achenbach S, Anker SD, Atar D, Baigent C, et al. CODE-EHR best-practice framework for the use of structured electronic health-care records in clinical research. Lancet Digit Health. 2022; 4(10):e757–e764. https://doi.org/10.1016/S2589-7500(22)00151-0 PMID: 36050271

94. Kwakkenbos L, Imran M, McCall SJ, McCord KA, Frobert O, Hemkens LG, et al. CONSORT extension for the reporting of randomised controlled trials conducted using cohorts and routinely collected data (CONSORT-ROUTINE): checklist with explanation and elaboration. BMJ. 2021; 373:n857. https://doi.org/10.1136/bmj.n857 PMID: 33926904