Review

AI-Based Monitoring Devices for the Identification of Circadian Parameters as Digital Predictive Biomarkers of Neurodegenerative Diseases in Older Adults: Systematic Review

J Manuel Herrera-Zamora[1*], PhD; Jesus Monroy-Rodriguez[2*], PhD; Jesus Garcia-Mancilla[3*], MSc; Pedro C Santana-Mancilla[4*], Prof Dr

[1]University Center for Biomedical Research, Facultad de Medicina, Universidad de Colima, Colima, Colima, Mexico

[2]Department of Basic Health Sciences, Centro Universitario del Sur, Universidad de Guadalajara, Ciudad Guzmán, Jalisco, Mexico

[3]Engineering, Sports Illustrated Tickets, Houston, TX, United States

[4]School of Telematics, Universidad de Colima, Colima, Colima, Mexico

[*]all authors contributed equally

Corresponding Author:
Pedro C Santana-Mancilla, Prof Dr
School of Telematics
Universidad de Colima
Av. Universidad 333
Colima, Colima, 28040
Mexico
Phone: 52 3123161075
Email: psantana@ucol.mx

Abstract

Background: Neurodegenerative diseases such as Alzheimer disease, Parkinson disease, and related dementias are a growing health burden among older adults. Circadian disruption is increasingly recognized as both a consequence and a potential early marker of neurodegenerative progression. Wearable sensing combined with AI enables continuous, noninvasive, and ecologically valid monitoring of circadian parameters as potential digital biomarkers.

Objective: This systematic review synthesizes evidence on AI-enhanced wearable and objective sensing devices used to monitor circadian parameters in older and cognitively vulnerable adults and evaluates their potential as digital predictive biomarkers for neurodegenerative diseases.

Methods: Following PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) 2020 guidelines, an original search of PubMed, Scopus, and JMIR journals (June 2025) yielded 12 studies, 7 of which met updated eligibility criteria on rescreening. A complementary search of PubMed, IEEE Xplore, and the ACM Digital Library (March 24, 2026) added 3 eligible studies and 3 sensitivity references (excluded records retained to delimit the corpus boundary). Two reviewers (JMH-Z and JM-R) independently assessed risk of bias with the Mixed Methods Appraisal Tool. Heterogeneous end points precluded meta-analysis; the narrative synthesis covered 4 prespecified domains: device class, circadian construct, AI and machine learning (ML) method family, and population and neurocognitive context.

Results: The final corpus comprised 13 records (10 eligible studies and 3 sensitivity references) published from 2015 to 2025. Wrist actigraphy predominated, and samples ranged from community-dwelling older adults to nursing home residents with advanced dementia (n=16-7297). Circadian parameters included interdaily stability, intradaily variability, relative amplitude, midline estimating statistic of rhythm, acrophase, and circadian stimulus. AI/ML methods included supervised learning (random forests, gradient boosting, and regression models), unsupervised clustering, and algorithm development or validation; 7 studies applied deterministic algorithms without fitting a predictive model. Circadian fragmentation was associated with functional decline, dementia risk, and cognitive performance. Most studies were descriptive or observational with small, nondiverse samples; no AI model was externally validated, and no study followed an AI-specific reporting guideline.

Conclusions: AI-enhanced wearable devices show meaningful potential for noninvasive measurement of circadian disruption in older and cognitively vulnerable populations. Evidence certainty is moderate for actigraphy-derived fragmentation-function associations in institutional settings and low for predictive biomarker claims in community populations. From 6 recurring challenges

XSL•FO
RenderX

along the sensing-to-biomarker pipeline, we propose a 4-layer reference architecture (sensing and acquisition, circadian parameter extraction, AI/ML inference, and a clinical and validation interface) with a layered validation pathway. Longitudinal, multisite validation studies using standardized circadian metrics, diverse populations, and explainable AI remain necessary before clinical deployment as neurodegenerative biomarkers.

(JMIR Aging 2026;9:e83031) doi: 10.2196/83031

KEYWORDS

circadian rhythms; wearable devices; AI; neurodegenerative diseases; older adults

Introduction

Background

Neurodegenerative diseases are a heterogeneous group of disorders characterized by the progressive loss of neuronal structure and function, significantly affecting memory, cognition, behavior, and sensorimotor function. These conditions, which include Alzheimer disease (AD), Parkinson disease (PD), and Lewy body dementia, represent a growing global public health challenge. The global number of people living with dementia is projected to reach 153 million by 2050 [1], underscoring the urgency of developing tools for earlier detection and monitoring.

Disruptions in circadian rhythms are increasingly recognized as a hallmark feature of neurodegeneration, appearing not only as a functional consequence of advancing disease but also as a potential early biological signal that may precede clinical onset [2,3]. The circadian system functions as the body's intrinsic biological clock, coordinating cellular processes, sleep-wake cycles, hormonal secretion, and behavioral rhythms with the 24-hour light-dark environment [4]. In older adults and those with neurodegenerative conditions, this system progressively deteriorates, resulting in measurable fragmentation of rest-activity rhythms, altered sleep architecture, and disrupted melatonin secretion [5]. Recent conceptual frameworks, developed in populations other than older adults, treat sleep and circadian disruption as the joint product of circadian, neuroendocrine, and behavioral stressors rather than as a single timing signal [6]; this multidimensional view is relevant to aging, where circadian decline coexists with endocrine and behavioral change.

Recent advances in wearable technologies and AI offer new approaches to continuously and ecologically monitoring circadian parameters outside clinical or laboratory settings. Wrist-worn actigraphs, electroencephalography (EEG) headbands, multisensor wearable patches, wearable light sensors, and ambient smart-home monitoring systems can collect physiological and behavioral data from individuals in their natural environments. When combined with AI and machine learning (ML) algorithms, these data streams can be processed to extract circadian parameters, such as interdaily stability (IS), intradaily variability (IV), relative amplitude (RA), midline estimating statistic of rhythm (MESOR), and acrophase, enabling continuous characterization of individual circadian profiles [7].

Despite the promise of this technology, the evidence base remains fragmented. Existing reviews have addressed wearable sensing and AI in aging broadly, without systematically evaluating the extent to which included studies operationalize genuine circadian constructs, apply AI methods rigorously, and target populations with neurodegenerative relevance. This review addresses that gap by synthesizing evidence from studies that explicitly combine objective sensing, AI, and circadian monitoring in older or cognitively vulnerable populations, with structured attention to the specificity of the circadian construct, the AI method family used, and the clinical relevance of reported outcomes. Drawing on this synthesis, we identify 6 structural challenges at the sensing-to-biomarker interface and propose a 4-layer reference architecture with a layered validation pathway intended to guide future system design and empirical evaluation in this field.

Contributions of This Review

This systematic review makes the following scholarly contributions:

- We synthesize evidence combining AI-enhanced wearable and objective sensing devices with circadian monitoring in older and cognitively vulnerable adults, organized across 4 prespecified analytical domains: device class, circadian construct, AI/ML method family, and population/neurocognitive context.
- We identify and characterize 6 structural challenges that currently limit the translation of circadian monitoring into validated predictive biomarkers for neurodegenerative disease, situating each challenge within the pipeline stage it most directly constrains.
- Drawing on the recurrence of these 4 analytical domains across the reviewed corpus, we propose a 4-layer reference pipeline architecture together with a layered validation pathway that maps each challenge to architectural responses and implementation propositions. This proposal represents a conceptual framework derived from the reviewed evidence; empirical validation through prospective field studies or multisite testbeds is outside the scope of this review but constitutes an explicit priority for future work.
- We articulate 4 implementation propositions grounded in the evidence base, including harmonized circadian end-point reporting, explainable AI (XAI) integration at model design time, population-specific device validation, and demographic representativeness as an operational validation criterion.

XSL•FO
RenderX

Methods

Overview

This systematic review was conducted in accordance with the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) 2020 guidelines [8] to ensure methodological transparency and rigor. The completed PRISMA 2020 checklist is available in Multimedia Appendix 1. The review protocol was prospectively deposited in Zenodo prior to data collection, in accordance with Findable, Accessible, Interoperable, and Reusable data principles and open science standards recommended by PRISMA 2020 [1]. The Zenodo record is available at [9].

During peer review, reviewers noted that the original search (PubMed, Scopus, and JMIR journals) underrepresented the engineering, biomedical informatics, and human-computer interaction literature, where AI-based sensing work is often published. To address this, we conducted a complementary systematic search across IEEE Xplore, the ACM Digital Library, and an updated PubMed query, applying the same prespecified eligibility criteria without modification. Because this search postdated the protocol deposited in Zenodo, we treat it as a documented amendment to the protocol rather than a change to the review question or eligibility rules; the deposited record reflects only the initial search.

The aim of this systematic review was to evaluate the current use of wearable devices that monitor circadian parameters using AI and to explore their potential as digital biomarkers for predicting the onset of neurodegenerative diseases in older adults, thereby contributing to improved health outcomes in aging populations. As a second, synthesis-driven aim, we sought to translate the analytical domains and structural challenges recurring across the reviewed evidence into a 4-layer reference pipeline architecture and an accompanying layered validation pathway, intended to guide the design and empirical evaluation of clinically deployable AI-circadian biomarker systems. This second aim is interpretive: the architecture and validation pathway were derived by the authors from the synthesized evidence and are not findings of the systematic review itself.

Data Sources and Search Strategy

Original Search

A primary literature search was conducted in June 2025 across PubMed, Scopus, and JMIR journals, targeting studies published between January 2015 and June 2025. Search terms were organized around three conceptual blocks: (1) wearable and objective sensing technologies, (2) circadian parameters and 24-hour behavioral constructs, and (3) aging or neurocognitive populations. This search identified 8441 records and, after duplicate removal and screening, yielded 12 included studies.

Complementary Search

Following peer-review recommendations, a complementary systematic search was conducted on March 24, 2026, across 3 additional bibliographic databases: PubMed (to capture publications beyond the original cutoff), IEEE Xplore, and the ACM Digital Library. These databases were selected to capture engineering, biomedical informatics, and human-computer interaction literature not adequately represented in the original search. Two parallel queries were run per database: a main search and a neurocognitive sensitivity search. Full database-specific Boolean search strings, including all field tags, filters (peer-reviewed journals only, 2015 onward), and date limits, are reported in Multimedia Appendix 2 in accordance with PRISMA-S (Preferred Reporting Items for Systematic Reviews and Meta-Analyses Literature Search Extension) guidelines [10]. This search identified 162 records; 28 duplicates were removed, yielding 134 records for screening. Full-text assessment was conducted on 27 records. The publication window (2015-2025) was the same as in the original search; the later search date extended database coverage, not the period of eligible publications.

Eligibility Criteria

Studies were included if they met all the following criteria:

- Original peer-reviewed journal article.
- Study population comprising adults aged ≥45 years, or individuals of any age with a documented neurodegenerative condition relevant to this review (mild cognitive impairment [MCI], AD, PD, Lewy body dementia, or related dementia syndromes).
- Use of at least 1 wearable or objective sensing device for physiological or behavioral data collection.
- Application of at least 1 AI or ML method in data analysis, pattern detection, feature extraction, or predictive modeling.
- Report of at least 1 explicit circadian parameter (IS, IV, RA, MESOR, acrophase, cosinor-derived parameter, circadian stimulus, or circadian rhythm strength) or a 24-hour behavioral construct interpreted within a circadian or chronobiological framework.

Studies were excluded if they were reviews, meta-analyses, editorials, letters, conference abstracts, or protocols; used exclusively nonhuman participants; lacked full-text access; or did not incorporate an AI or ML component in the analysis.

Sensitivity references were defined as records excluded at full-text assessment that failed exactly 1 eligibility criterion, met all the others, and addressed a methodological element directly relevant to the review question: validation of a wearable sleep-wake algorithm against a reference standard, AI-based prediction from continuous sensor data, or extraction of circadian parameters in an institutional population. Sensitivity references are not counted as eligible studies. They are identified by a footnote in all tables, reported separately in the risk-of-bias assessment, and used to delimit the methodological boundary of the corpus; no conclusion of this review rests on them alone.

Study Selection Process

Records were imported into Zotero for deduplication and screening. Title and abstract screening were conducted using Rayyan Qatar Computing Research Institute (original search) and a standardized manual protocol (complementary search). Full-text eligibility assessment was performed by 2 independent reviewers (JMH-Z and JM-R for the original search; JMH-Z and PCS-M for the complementary search). Disagreements were resolved through consensus or by a third reviewer (JG-M).

XSL•FO
RenderX

The 12 studies from the original search were rescreened against the updated eligibility criteria during revision. In total, 5 were excluded: 3 because no AI or ML method was used in the analysis [11-13], and 2 because the enrolled population did not meet the population criterion or no explicit circadian parameter was analyzed [14,15]. A total of 7 studies were retained from the original corpus.

Data Extraction and Synthesis

Data were extracted from each included study using a standardized form covering authors, year and country, study design, population characteristics (sample size, age, condition), sensing device and modality, circadian construct and parameter names, AI/ML method and family, analytic approach, and main findings. Two reviewers (JMH-Z and JM-R) extracted data independently; discrepancies were resolved by discussion. For each study, we also extracted how computational methods were reported: validation scheme, unit of data partitioning, hyperparameter reporting, availability of code and data, explainability or variable-importance reporting, and adherence to AI-specific reporting guidelines (eg, TRIPOD+AI [Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis+Artificial Intelligence] [16], CLAIM [Checklist for Artificial Intelligence in Medical Imaging] [17], DECIDE-AI [Developmental and Exploratory Clinical Investigations of Decision Support Systems Driven by Artificial Intelligence] [18], and CONSORT-AI [Consolidated Standards of Reporting Trials-Artificial Intelligence] [19]). A codebook with controlled vocabularies separated absent reporting (not reported) from inapplicability (not applicable and when no predictive model was fitted). A large language model–based agentic workflow (powered by Anthropic's Claude) produced a first-pass extraction from the full texts, recording a verbatim evidence locator for every value other than not reported; 2 reviewers (JMH-Z and PCS-M) then independently verified each value against the full text and supplementary materials of each source article, resolving disagreements by discussion or by a third reviewer (JG-M). Because AI assistance in scholarly evaluation can speed up the work but raises concerns about the depth and authenticity of the resulting judgments [20], the assistant's output was treated as a draft only, and no extracted value was retained without human verification. Consistent with the eligibility criterion, which counts feature extraction as a qualifying use of AI or ML, studies applying deterministic algorithms to sensor data were eligible even when they fitted no predictive model; this distinction is reported for every study in the Results section.

Narrative synthesis was organized around four prespecified analytical domains: (1) device class and sensing modality, (2) circadian construct and parametric specificity, (3) AI/ML method family, and (4) population and neurocognitive context. Studies were first cataloged descriptively within each domain, then analyzed comparatively to identify convergent evidence, methodological gaps, and key contributions. Three sensitivity-reference studies [21-23] were used to interpret the methodological boundary of the corpus and are indicated by a footnote in all tables.

Risk of Bias Assessment

Risk of bias was assessed using the Mixed Methods Appraisal Tool (MMAT) [24], which provides design-specific quality criteria for quantitative descriptive (type 4), quantitative nonrandomized (type 3), and quantitative randomized controlled trial (RCT; type 2) designs. Two reviewers (JMH-Z and JM-R) independently assessed each included study, applying the MMAT screening questions (S1 and S2) and 5 design-specific quality criteria (Q1-Q5). Disagreements were resolved by consensus. Domain-level judgments (Yes/No/Cannot tell) and overall appraisal (low, moderate, or high risk) are reported in the next section.

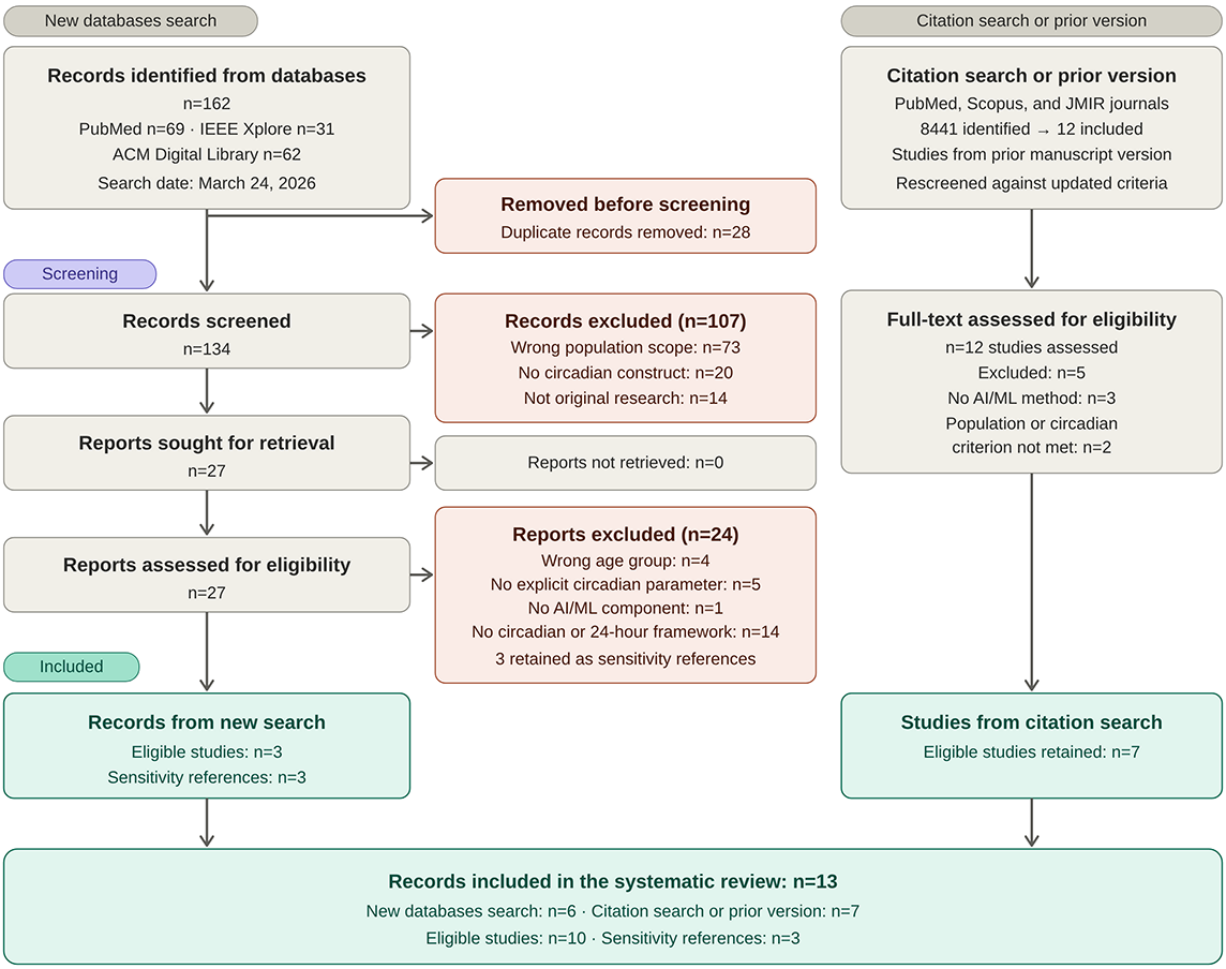
Results

Study Selection

The 2-stage search process yielded a final corpus of 13 records: 10 eligible studies and 3 sensitivity references (Figure 1). From the original search corpus, 7 of 12 previously included studies were retained after rescreening; 5 were excluded. From the complementary search, 162 records were identified across 3 databases (PubMed: n=69; IEEE Xplore: n=31; ACM Digital Library: n=62). After removing 28 duplicates, 134 records were screened at the title and abstract; 107 were excluded. In total, 27 records were assessed at full text; 24 were excluded (wrong age group: n=4; no explicit circadian parameter measured: n=5; no AI/ML component: n=1; no circadian or 24-hour framework: n=14). A total of 3 records met all eligibility criteria and were included; 3 of the 24 excluded records met the definition of sensitivity references and were retained as such. Figure 1 presents the PRISMA flow diagram with the full search strategy.

XSL•FO
RenderX

Figure 1. PRISMA flowchart for study selection (publications 2015-2025; searches run in June 2025 and March 2026). Of the 8603 records identified, 13 were included in the final corpus: 10 met all eligibility criteria, and 3 were retained as sensitivity references. ML: machine learning; PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses.



Study Characteristics

The 13 records (10 eligible studies and 3 sensitivity references) were published between 2015 and 2025. They were conducted in the United States (n=7), the United Kingdom (n=2), Taiwan (n=2), the Netherlands (n=1), and Finland (n=1). Sample sizes ranged from 16 to 7297 participants. In total, 6 studies enrolled populations with documented or high-risk neurodegenerative conditions; 4 enrolled community-dwelling or population-based samples without documented cognitive impairment; and 3 used middle-aged or mixed-age samples as methodological anchors. Full characteristics are presented in Table 1.

XSL•FO
RenderX

Table 1. Characteristics of included studies on AI-enhanced wearable devices for circadian monitoring in aging populations.

Author (year)	Population and setting	Device	Circadian parameters	AI/ML[a] method	Key findings
Chuang et al (2024) [25]	Taiwan; n=78 adults (mean 41.5 years; 59% women; 29 obese, 16 overweight, 33 controls)	Wrist actigraphy+Rhythm smartphone app	IS[b], social jet lag, sleep indicators	Actigraphy-based algorithm; multivariate regression	IS strongly correlated with adiposity. App-based measurement provided stronger associations than traditional actigraphy; IS from app correlated with obesity index, unlike social jet lag.
Razjouyan et al (2017) [26]	United States; n=21 adults (mean 50.8, SD 12.8 years; 47.8% women; poor sleep)	Chest sensor (BioPatch), wrist Actiwatch, PSG[c] (reference)	Sleep-wake epoch; implicit rest-activity rhythm	Posture+acceleration algorithm; logistic regression	Chest sensor achieved 85.8% accuracy versus 79.8% for Actiwatch. Combined posture and acceleration improved sleep/wake epoch detection.
Gabb et al (2025) [27]	United Kingdom; n=41 older adults (mean 70.9 years); 10 AD[d], 11 LBD[e], 20 controls	Axivity AX3, Dreem 2 EEG[f] headband, MyDignio app	Sleep continuity, circadian rhythm, cognitive task performance	GGIR[g] automated pipeline; Dreem EEG algorithm	Remote longitudinal monitoring feasible; 98% retention, 79%-97% data completeness. Multimodal approach viable in populations with MCI[h] and dementia.
Cavaillès et al (2025) [28]	United States (MrOS[i] cohort); n=2667 older men	SleepWatch-O wrist actigraph	24-hour activity profiles; sleep/circadian phenotypes	Unsupervised ML clustering (MCGHD[j]+PCA[k])	Three phenotypes: active healthy (64%), fragmented poor (14.1%), and long nappers (21.9%). Fragmented poor sleepers showed higher dementia and CVD[l] event risk.
Wang et al (2024) [29]	Taiwan; n=191 community adults ≥50 years	Actiwatch (3-axis accelerometer+light+temperature sensors), sleep diaries	IS, IV[m], RA[n], chronotype	Kuo actigraphy algorithm; ANOVA+regression	Older adults had lower activity, worse sleep efficiency, earlier chronotype. Women showed greater IS and RA. Age-gender interactions for life satisfaction and RA.
Shim et al (2023) [30]	United States (NHANES[o]); n=7297 adults	ActiGraph GT3X+wrist accelerometer	24-hour activity profiles as digital biomarkers (inflammation, biological age, mortality)	Hierarchical clustering (Ward method); recursive feature elimination	Five distinct 24-hour activity profiles identified. Circadian misalignment linked to systemic inflammation, accelerated biological age, and increased mortality risk.
Jones et al (2025) [31]	United States; n=40 older adults (mean 67.2 years; 72.5% women); 19 MCI, 21 controls	Apple Watch+myRhythmWatch platform	IS, IV, RA, cosinor (MESOR[p], acrophase, rhythm strength)	myRW[q] algorithm; cosinor analysis; linear regression	Later activity onset associated with worse memory. Fragmented rhythms linked to slower processing speed. Platform feasible for circadian assessment in MCI.

XSL•FO
RenderX

Author (year)	Population and setting	Device	Circadian parameters	AI/ML[a] method	Key findings
Beiglary et al (2025) [32]	United States; n=29 older adults with dementia; 2 assisted-living facilities	Wearable light sensor calibrated versus spectrophotometer	Circadian stimulus, personal circadian lighting exposure	Random forest (ML); laboratory calibration+field deployment	Random forest accurately predicted circadian stimulus. Substantial interindividual variability in circadian light exposure. In total, 40% device loss; validated against spectrophotometric ground truth.
Merilahti et al (2016) [23][r]	Finland; n=16 nursing home residents (mean 90.7 years; 7 with dementia)	Wrist-worn device with integrated activity sensor	MESOR, amplitude, acrophase, IS, IV, RA, CRS[s]	Single cosinor function (24 hours); nonparametric RAS[t] analysis	Physical functioning associated with rhythm strength and stability. Functional decline tracked with reduced activity, lower MESOR/amplitude, and increased IV.
Froggatt et al (2020) [33]	United Kingdom; n=32 nursing home residents with advanced dementia (mean 81.5 years)	ActiGraph wrist actigraphy	IS, IV (circadian fragmentation and synchronization)	Nonparametric actigraphy analysis (IS/IV computation)	Feasibility cluster RCT[u]. Highly inactive population with fragmented rhythms (IS: mean 0.08; IV: mean 0.7). Actigraphy feasible in advanced dementia.
Sprint et al (2022) [34]	United States; n=14 amnestic MCI (mean 74.4 years) in smart-home environment	Smart-home sensors+EMMA tablet app+text markers	Circadian rhythm strength, 24-hour circadian rhythm, regularity indices	Multimodal fusion; joint prediction (regression/ML)	Moderate-to-large prediction correlations (r=0.601-0.871) for 10 clinical assessments. Joint prediction outperformed independent modality models.
de Zambotti et al (2015) [21][r]	United States; n=28 midlife women (mean 50.1 years); sleep laboratory	Jawbone UP consumer tracker vs PSG (reference)	Sleep-wake rhythm (nightly sleep metrics)	Proprietary consumer algorithm; PSG comparison	High sleep sensitivity, poor wake specificity; overestimated TST[v], underestimated WASO[w]. Potentially useful for daily sleep-wake rhythm tracking over multiple days.

XSL•FO
RenderX

Author (year)	Population and setting	Device	Circadian parameters	AI/ML[a] method	Key findings
Robben et al (2017) [22][r]	Netherlands; n=13 older adults (mean 80, SD 7 years); independent living	Ambient smart-home sensors (PIR[x], pressure, switches); no on-body device	24-hour behavioral structure as temporal proxy	Delta-feature temporal modeling; random forest prediction	Delta features predicted functional health change, outperforming static baselines. Temporal behavioral regularity exploited as implicit circadian proxy.

[a]ML: machine learning.
[b]IS: interdaily stability.
[c]PSG: polysomnography.
[d]AD: Alzheimer disease.
[e]LBD: Lewy body dementia.
[f]EEG: electroencephalography.
[g]GGIR: R-package for accelerometry.
[h]MCI: mild cognitive impairment.
[i]MrOS: Osteoporotic Fractures in Men Study.
[j]MCGHD: multiple coalesced generalized hyperbolic distribution.
[k]PCA: principal component analysis
[l]CVD: cardiovascular disease.
[m]IV: intradaily variability.
[n]RA: relative amplitude.
[o]NHANES: National Health and Nutrition Examination Survey.
[p]MESOR: midline estimating statistic of rhythm.
[q]myRW: myRhythmWatch.
[r]Sensitivity-reference study.
[s]CRS: circadian rhythm strength.
[t]RAS: rest-activity rhythm analysis.
[u]RCT: randomized controlled trial.
[v]TST: total sleep time.
[w]WASO: wake after sleep onset.
[x]PIR: passive infrared.

Risk of Bias and Quality Assessment

Risk of bias was assessed with the MMAT (version 2018) [24], as described in the Methods section. The complete dataset to enhance transparency is available as Multimedia Appendix 3.

MMAT results are presented in Table 2. In total, 8 studies were classified as quantitative descriptive (type 4), 4 as quantitative nonrandomized (type 3), and 1 as a quantitative RCT (type 2). A total of 6 studies received an overall low-risk appraisal. Wang et al [29] was rated moderate risk due to a nonrepresentative sample (Q2: No). Beiglary et al [32] was rated moderate risk because sample representativeness could not be determined (Q2: Cannot tell), and Merilahti et al [23] because low reporting bias could not be confirmed (Q4: Cannot tell). Sprint et al [34] was rated as moderate-to-high risk, given a small sample size (n=14), no control of confounders, and uncertain completeness of outcome data. Froggatt et al [33] received a high-risk appraisal, as expected for a complex-care feasibility RCT with nonblinded outcomes; this does not diminish the descriptive validity of the actigraphy findings. All 3 sensitivity-reference studies were rated moderate risk.

XSL•FO
RenderX

Table 2. Mixed Methods Appraisal Tool (MMAT) results for included studies[a].

Author (year)	MMAT type	S1	S2	Q1	Q2	Q3	Q4	Q5	Appraisal
Chuang et al (2024) [25]	Type 4	Y[b]	Y	Y	Y	Y	Y	Y	Low risk
Razjouyan et al (2017) [26]	Type 4	Y	Y	Y	Y	Y	Y	Y	Low risk
Gabb et al (2025) [27]	Type 4	Y	Y	Y	Y	Y	Y	Y	Low risk
Cavaillès et al (2025) [28]	Type 3	Y	Y	Y	Y	Y	Y	Y	Low risk
Wang et al (2024) [29]	Type 4	Y	Y	Y	N[c]	Y	Y	Y	Moderate
Shim et al (2023) [30]	Type 3	Y	Y	Y	Y	Y	Y	Y	Low risk
Jones et al (2025) [31]	Type 4	Y	Y	Y	Y	Y	Y	Y	Low risk
Beiglary et al (2025) [32]	Type 4	Y	Y	Y	C[d]	Y	Y	Y	Moderate
Merilahti et al (2016) [23][e]	Type 4	Y	Y	Y	Y	Y	C	Y	Moderate
Froggatt et al (2020) [33]	Type 2	Y	Y	Y	N	C	N	C	High risk[f]
Sprint et al (2022) [34]	Type 3	Y	Y	C	Y	C	N	Y	Moderate-high
de Zambotti et al (2015) [21][e]	Type 4	Y	Y	Y	C	Y	Y	Y	Moderate
Robben et al (2017) [22][e]	Type 3	Y	Y	Y	Y	Y	C	Y	Moderate

[a]Type 2 (randomized controlled trial): Q1=randomization appropriately performed; Q2=groups comparable at baseline; Q3=complete outcome data; Q4=outcome assessors blinded; Q5=participants adhered to the assigned intervention. Type 3 (nonrandomized): Q1=participants representative of the target population; Q2=appropriate measurement of outcome and exposure; Q3=complete outcome data; Q4=confounders accounted for in design and analysis; Q5=exposure occurred as intended. Type 4 (descriptive): Q1=sampling strategy relevant to the research question; Q2=sample representative of the target population; Q3=appropriate measurements; Q4=low risk of nonresponse bias; Q5=appropriate statistical analysis.

[b]Y: yes.

[c]N: no.

[d]C: cannot tell.

[e]Sensitivity-reference study.

[f]High risk is expected and acceptable for a feasibility cluster randomized controlled trial.

Synthesis by Domain

Device Class and Sensing Modality

Wrist actigraphy was the most used modality, enabling derivation of nonparametric circadian metrics and cosinor parameters from continuous rest-activity data recorded over days to months. Devices deployed included ActiGraph GT3X+ [30], SleepWatch-O [28], Axivity AX3 [27], Actiwatch [29], Apple Watch with the myRhythmWatch platform [31], an unspecified integrated activity-sensing wrist device [23], and a spectrophotometer-calibrated wearable light sensor [32]. Device validation varied widely: Beiglary et al [32] performed explicit laboratory calibration against a spectrophotometric reference before field deployment; de Zambotti et al [21] compared consumer tracker output with polysomnography; and Razjouyan et al [26] validated a novel algorithm against polysomnography in a sleep laboratory setting. Most other studies used devices without concurrent within-study validation. Multisensor platforms that combine actigraphy with EEG headbands, apps, and pulse oximetry were used by Gabb et al [27]. Smart-home ambient sensor arrays without on-body devices were used in Sprint et al [34] and Robben et al [22]. A smartphone-actigraphy combination formed the sensing platform in Chuang et al [25].

Circadian Construct and Parametric Specificity

Studies differed substantially in how explicitly they operationalized circadian biology. Six studies derived or analyzed named chronobiological metrics as primary or clearly defined outcomes: IS, IV, RA, MESOR, acrophase, and circadian stimulus [23,25,29,31-33]. Four studies embedded temporal or 24-hour rhythm features within predictive or clustering frameworks without foregrounding them as primary end points [27,28,30,34]. Two studies used continuous sleep-wake rhythm monitoring as a functional proxy for circadian structure without computing formal circadian parameters [21,26]. One study treated ambient behavioral regularity as an implicit temporal proxy [22].

This gradient from explicit to implicit circadian operationalization has important implications for interpreting studies as evidence for circadian biomarkers. Only studies reporting formal chronobiological metrics provide direct evidence for the circadian construct. Studies that use sleep-wake rhythm or behavioral regularity as proxies provide contextual support but cannot serve as substitutes for direct circadian measurement.

AI/ML Method Family

Four AI/ML families were identified (Table 3). Supervised predictive modeling was applied in 3 studies: a random forest

XSL•FO
RenderX

predicting circadian light exposure from a calibrated wearable light sensor [32]; decision tree, random forest, and gradient-boosted regressors predicting clinical assessment scores from multimodal behavioral markers in MCI [34]; and regression forests predicting change in functional health from ambient sensor data [22]. Unsupervised phenotyping identified distinct 24-hour activity profiles in 2 large-sample studies, using hierarchical clustering after recursive feature elimination [30] and generalized hyperbolic mixture modeling on principal components of actigraphy variables [28]. Algorithm development or validation was the primary contribution of 2 studies: a chest-worn sleep-wake classifier developed against polysomnography [26] and the polysomnographic validation of a consumer tracker's proprietary scoring algorithm [21]. The remaining 6 studies applied deterministic algorithms (nonparametric rest-activity indices, cosinor fitting, and automated sleep-wake scoring) followed by conventional statistical inference to relate circadian metrics to functional, clinical, or feasibility outcomes [23,25,27,29,31,33].

Reporting of these methods was uneven (Table 3). No model was validated on an independent external cohort, and no study reported following an AI-specific reporting guideline; 3 cited general reporting guidelines (STROBE [Strengthening the Reporting of Observational Studies in Epidemiology] [27,28], TIDieR [Template for Intervention Description and Replication], and CONSORT [Consolidated Standards of Reporting Trials] [33]). Seven studies fitted no predictive model, so validation, partitioning, and hyperparameter criteria did not apply to them [21,23,25,27,29,31,33]. Of the 4 studies with a training-testing distinction, 2 stated that data were partitioned at the participant level [22,26]; 1 reported leave-one-out cross-validation over a participant-level feature matrix without stating the partition unit [34]; and 1 described a train-test split and 5-fold cross-validation for its linear models without stating the unit or proportion of the split, while its deposited random forest code fitted the complete training set without resampling [32]. The 2 unsupervised studies performed no held-out evaluation: one assessed cluster stability by repeated subsampling [28], and the other reported no validation of the clustering solution [30]. In total, 5 studies reported hyperparameter values partially [22,26,28,30,34], and 1 relied on package defaults that are visible only in its deposited code [32]; none reported every setting needed to reproduce its model. A total of 3 studies deposited analysis code, in each case covering only part of the analysis [28,32,34], and 3 made their data openly available [28,30,32].

Explainability was addressed only through model-level variable importance. One study applied a post-hoc attribution method, permutation importance for its random forest [32], and 3 reported model-internal importance measures: Gini importance [22], variables ranked by their contribution to cluster separation [28], and random forest importance within the feature-selection step that preceded clustering, rather than for the clustering itself [30]. No study used local attribution methods such as Shapley Additive Explanations (SHAP) or Local Interpretable Model-Agnostic Explanations (LIME), and none explained individual predictions.

XSL•FO
RenderX

Table 3. Reporting of computational methods in the included studies[a].

Author (year)	AI/ML[b] family	Validation	Partition unit	Hyperparameters	Code	Data	XAI[c]
Chuang et al (2024) [25]	Pattern association	Reference standard[d]	N/A[e]	N/A	NR[f]	Request	N/A
Razjouyan et al (2017) [26]	Algorithm validation	CV	Participant	Partial	NR	NR	NR
Gabb et al (2025) [27]	Pattern association	N/A	N/A	N/A	NR	Request	N/A
Cavaillès et al (2025) [28]	Unsupervised	Resubstitution[g]	N/A	Partial	Open[h]	Open	Importance
Wang et al (2024) [29]	Pattern association	N/A	N/A	N/A	NR	Request	N/A
Shim et al (2023) [30]	Unsupervised	NR	N/A	Partial	NR	Open	Importance[i]
Jones et al (2025) [31]	Pattern association	N/A	N/A	N/A	NR	NR	N/A
Beiglary et al (2025) [32]	Supervised	Unclear[j]	NR	Defaults[k]	Open[h]	Open	Post hoc
Merilahti et al (2016) [23][l]	Pattern association	N/A	N/A	N/A	NR[m]	NR	N/A
Froggatt et al (2020) [33]	Pattern association	N/A	N/A	N/A	NR	Request	N/A
Sprint et al (2022) [34]	Supervised	CV	Unclear[n]	Partial	Open[h]	NR	NR
de Zambotti et al (2015) [21][l]	Algorithm validation[o]	Reference standard	N/A	N/A	NR[m]	NR	N/A
Robben et al (2017) [22][l]	Supervised	CV	Participant	Partial	NR	NR	Importance[p]

[a]AI/ML family: supervised=supervised predictive modeling; unsupervised=unsupervised phenotyping or clustering; algorithm validation=algorithm development or validation; pattern association=pattern-association modeling with conventional statistical inference. Validation: CV=cross-validation; reference standard=comparison against a concurrent reference standard; resubstitution=performance reported on the fitting data. Partition unit: participant=data partitioned at the participant level. Hyperparameters: partial=some values reported; defaults=package defaults used. Code and data: open=publicly available; request=available on request. XAI: post hoc=post hoc attribution method; importance=variable importance without a named attribution framework; unclear=reporting insufficient to determine. No study reported external validation on an independent cohort or adherence to an AI-specific reporting guideline.

[b]ML: machine learning.

[c]XAI: explainable AI.

[d]Reference standard was research-grade actigraphy, not polysomnography.

[e]N/A: not applicable (no fitted predictive model).

[f]NR: not reported.

[g]Internal stability only (100 subsamples excluding 5% of participants; adjusted Rand index reported); no held-out performance evaluation.

[h]Deposited code covers only part of the analysis: cluster identification [28], the circadian light model [32], and digital marker extraction [34].

[i]Variable importance reported for the random forest used in recursive feature elimination, not for the clustering model.

[j]Five-fold cross-validation and a train-test split are described for the linear regression models; the deposited random forest code fits the full training set without resampling.

[k]Package defaults evident in the deposited code; not stated in the text; no random seed.

[l]Sensitivity-reference study.

[m]The device algorithm is stated to be proprietary; no statement on the authors' analysis code.

[n]Leave-one-out cross-validation over a participant-level feature matrix; partition unit not stated explicitly.

[o]Validation of a commercial algorithm against polysomnography; no algorithm was developed by the authors.

[p]The authors also state that the learned decision rules are interpretable.

Population and Neurocognitive Context

Six studies enrolled populations with documented or high-risk neurodegenerative conditions. Three targeted residents of nursing homes or assisted-living facilities with advanced dementia [23,32,33], where circadian fragmentation was most severe (IS mean ~0.08, IV mean ~0.7 in Froggatt et al [33]), and actigraphy feasibility was explicitly evaluated. Two enrolled participants with amnestic MCI [31,34], with circadian metrics linked to cognitive test scores and clinical assessment predictions. One enrolled a mixed MCI/dementia cohort alongside a cognitively normal control [27]. Four studies used

XSL•FO
RenderX

community-dwelling or population-based samples without documented cognitive impairment [22,28-30], examining 24-hour activity profiles in relation to dementia risk, biological age, or functional health as outcome proxies. Three studies provided methodological contributions using middle-aged or mixed-age samples [21,25,26].

Sex composition was notably imbalanced: Cavaillès et al [28] enrolled exclusively older men (Osteoporotic Fractures in Men Study [MrOS] cohort), and de Zambotti et al [21] enrolled exclusively women. Racial and ethnic diversity was rarely reported. No study was conducted in low- or middle-income country settings, despite the global burden of neurodegeneration in those populations.

The 4 analytical domains examined across this synthesis do not operate independently. Their recurrence across the included studies reveals a set of structural dependencies: sensing modality choices constrain the circadian constructs that can be extracted; construct operationalization specificity determines the inferential ceiling available to AI/ML methods; and the characteristics of the target population bound the generalizability of any biomarker claim. Taken together, these dependencies constitute the functional logic of a pipeline, not merely a descriptive taxonomy.

Across domains, the evidence shows a consistent pattern: continuous circadian-related data capture is technically feasible, but biomarker interpretation remains constrained by construct heterogeneity, limited external validation, and population-level generalizability gaps. The most robust evidence comes from actigraphy-derived rest-activity fragmentation in institutional or cognitively vulnerable populations, whereas AI-based prediction in community and preclinical settings remains exploratory. The Evidence-Derived Proposal: Toward Clinically Deployable AI-Circadian Biomarker Systems: Structural Challenges, a Reference Pipeline Architecture, and Validation Priorities section translates that logic into a proposed reference architecture and layered validation pathway grounded in the evidence synthesized earlier.

Evidence-Derived Proposal: Toward Clinically Deployable AI-Circadian Biomarker Systems: Structural Challenges, a Reference Pipeline Architecture, and Validation Priorities

The structural challenges, reference architecture, and validation pathway that follow were derived by the authors from the synthesized evidence; they are not outcomes of the screening and appraisal process described in the Methods section.

Structural Challenges at the Sensing-to-Biomarker Interface

The 13-study corpus reveals a field in which technical capacity has outpaced analytical standardization. Six structural challenges constrain the transition from ecologically valid continuous monitoring to validated predictive biomarkers for neurodegenerative disease, and each aligns with a specific stage of the pipeline that connects sensing hardware to clinical interpretation.

At the sensing layer, device heterogeneity is compounded by inconsistent within-study validation. Only 3 studies [21,26,32] performed explicit concurrent validation against a reference standard (polysomnography or laboratory spectrophotometry), while the remaining studies assumed device validity in populations for which independent performance evidence is often lacking. This limitation intensifies in advanced dementia, where motor disorganization, behavioral symptoms, and adherence constraints may invalidate calibration obtained from younger or healthier samples, as illustrated by the 40% device loss reported by Beiglary et al [32].

At the parameter extraction layer, circadian end-point heterogeneity prevents cross-study integration. Studies using nonparametric metrics (IS, IV, and RA), cosinor-derived parameters (MESOR, acrophase, and rhythm strength), circadian stimulus, and unsupervised activity phenotypes differ not only in end-point choice but also in epoch length, analysis software, and time window, precluding quantitative synthesis despite consistent directional findings. A parallel concern is the gradient from explicit chronobiological operationalization to implicit proxies based on sleep-wake rhythm or behavioral regularity; the distinction matters because only the former supports direct circadian biomarker inference; yet, it is frequently blurred in cross-study interpretation.

At the AI/ML inference layer, 2 interlocking challenges stand out. The first is the absence of external validation, compounded by incomplete reporting of how models were validated. Every model in the corpus was developed within a single dataset, and none was evaluated on an independent cohort; in 2 of the 4 studies with a training-testing distinction, the unit of data partitioning could not be determined from the report; no study reported every hyperparameter needed for reproduction; and none followed an AI-specific reporting guideline such as TRIPOD+AI [16] (Table 3). Even the large-sample clustering studies [28,30] relied on retrospective datasets without prospective biomarker validation designs. The same pattern has been documented for AI-based risk prediction in sports medicine, where only 5 of 39 injury prediction studies reported external validation [35], which points to a methodological gap shared across physiological monitoring domains rather than one specific to circadian research. The second challenge is the shallowness of explainability. Four studies reported variable importance [22,28,30,32], but only at the level of the whole model: no study used local attribution methods such as SHAP [36] or LIME [37], and none explained individual predictions, which is the level at which a clinician would act on a biomarker result. Converging regulatory expectations require transparency and human oversight before AI-based biomarkers are deployed clinically; for circadian models intended to inform neurodegeneration risk, this gap is a direct barrier to regulatory clearance and clinician trust.

At the population and deployment layer, 2 further challenges emerge. All included studies were conducted in high-income countries, with limited reporting of racial and ethnic diversity and pronounced sex imbalances, with exclusively male [28] or exclusively female [21] samples. Given that the global burden of neurodegeneration falls disproportionately on populations underrepresented in this corpus, biomarkers validated in this

XSL•FO
RenderX

sample frame cannot be assumed to transfer to the populations for whom early detection is most urgent. Compounding this, the descriptive-to-predictive gap limits clinical actionability: 7 of the 13 studies fitted no predictive model, applying deterministic algorithms for feature extraction followed by conventional statistical inference, and the 6 that did were small, single-site, or retrospective (Table 3).

These 6 challenges are not independent; they constrain one another. Device validation failures propagate into parameter extraction errors, which in turn exacerbate AI model overfitting and confound the association with clinical outcomes. Addressing them individually is insufficient; they must be addressed in structural relation to the pipeline that generates a circadian biomarker (Table 4).

Table 4. Structural challenges at the sensing-to-biomarker interface and their architectural responses.

Identified challenge/gap	Why it matters	Architectural/design response	Future validation implication
Device validation heterogeneity across populations	Calibration from younger or healthier samples may not hold in advanced dementia, where device loss and motion artifact are substantial	Layer 1: concurrent validation against reference standards in the target population as a precondition, not an add-on	Population-specific device validation required before biomarker claims
Circadian end-point heterogeneity (metrics, epoch lengths, and pipelines)	Prevents quantitative synthesis and direct effect-size comparison across studies	Layer 2: harmonized parameter set, standardized analysis pipelines, explicit distinction between direct circadian measurement and behavioral proxies	Minimum standard reporting of end points, epochs, and pipelines in biomarker studies
Absence of external validation and incomplete AI reporting	All included AI models are developed within a single dataset; validation procedures and hyperparameters are incompletely reported, and no study follows an AI-specific reporting guideline	Layer 3: external replication on independent cohorts as a design requirement	Cross-cohort generalization evidence and guideline-concordant reporting (eg, TRIPOD+AI[a]) before clinical translation claims
Explainable AI (XAI) gap	Regulatory deployment and clinical trust require transparency; importance reporting in the corpus is model-level only, with no explanation of individual predictions	Layer 3: explainability mechanisms integrated at design time, with model outputs mapped to chronobiologically interpretable parameters	Prespecified XAI reporting, including explanations of individual predictions (eg, SHAP[b] and LIME[c]), as a standard element of biomarker studies
Population diversity deficit	High-income, demographically narrow samples do not represent the global burden of neurodegeneration	Layer 4: demographic representativeness as an architectural concern, not a discussion caveat	Quantified transferability across sex, race, ethnicity, and country-income setting as a validation criterion
Descriptive-to-predictive gap	In total, 7 of 13 studies fit no predictive model; the 6 that do are small, single-site, or retrospective	Layer 3: principled distinction between algorithmic processing and learned representation, with implications for evidence strength	Prospective biomarker validation designs rather than retrospective cross-sectional association

[a]TRIPOD+AI: Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis + Artificial Intelligence

[b]SHAP: Shapley Additive Explanations.

[c]LIME: Local Interpretable Model-Agnostic Explanations.

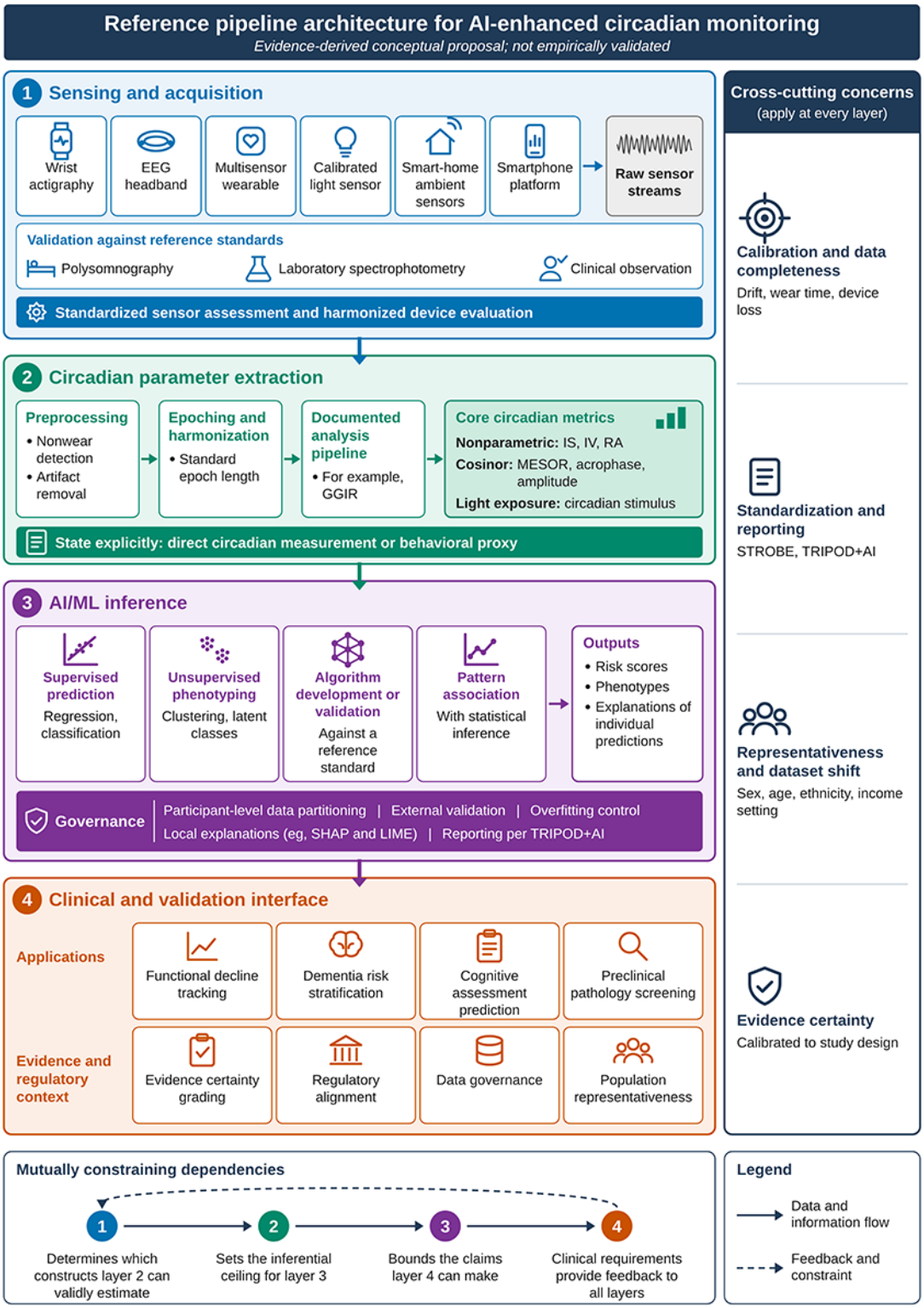
A Reference Pipeline Architecture for AI-Enhanced Circadian Monitoring

Overview

The recurrence of 4 analytical domains across the reviewed corpus supports a 4-layer reference architecture for AI-enhanced circadian biomarker pipelines (as Figure 2 shows). The architecture is conceptual: it does not prescribe a specific implementation but articulates the functional dependencies that any credible biomarker system in this field must respect.

XSL•FO
RenderX

Figure 2. Reference pipeline architecture for AI-enhanced circadian monitoring. EEG: electroencephalography; GGIR: R package for processing raw accelerometer data; IS: interdaily stability; IV: intradaily variability; LIME: Local Interpretable Model-Agnostic Explanations; MESOR: midline estimating statistic of rhythm; ML: machine learning; RA: relative amplitude; SHAP: Shapley Additive Explanations; STROBE: Strengthening the Reporting of Observational Studies in Epidemiology; TRIPOD+AI: Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis+Artificial Intelligence.



Sensing and Acquisition Layer

This layer encompasses the modalities observed in the corpus: wrist actigraphy as the dominant channel, supplemented by EEG headbands, multisensor wearables, calibrated wearable light sensors, ambient smart-home sensor arrays, and smartphone-based platforms. Its function is to produce continuous, ecologically valid physiological and behavioral data streams with explicit, population-specific validation against reference measurements. The cross-cutting concerns at this layer are device calibration, data completeness under real-world conditions, and feasibility in the populations for whom the

XSL•FO
RenderX

biomarker is intended, particularly institutionalized individuals with advanced dementia, where adherence, motion artifact, and device loss are nontrivial [32,33]. Calls to standardize how sensor technologies and the algorithms built on them are assessed have also been made in applied fields outside medicine [38]; in circadian biomarker pipelines, this standardization matters most at layers 1 and 2, because device validation and parameter extraction set the ceiling for every inference downstream.

Circadian Parameter Extraction Layer

This layer converts raw sensor streams into chronobiologically interpretable metrics. The corpus supports a core parameter set comprising nonparametric metrics (IS, IV, and RA), cosinor-derived parameters (MESOR, acrophase, and rhythm strength), and specialized metrics such as circadian stimulus for light exposure studies. Its function is to produce reproducible, comparable circadian end points. The cross-cutting concerns are harmonized epoch length, standardized analysis pipelines (such as GGIR or equivalently documented alternatives), explicit operationalization of whether a study measures circadian biology directly or uses behavioral proxies, and transparent reporting that enables downstream meta-analytic integration.

AI/ML Inference Layer

This layer applies computational reasoning to extracted parameters or, in more advanced designs, directly to raw sensor streams. The corpus demonstrates 4 viable method families: supervised predictive modeling, unsupervised phenotyping, algorithm development and validation, and pattern-association modeling combined with conventional statistical inference. Its function is to extract clinically meaningful representations beyond the reach of single-metric description. The cross-cutting concerns are external validation across independent datasets, explainability mechanisms compatible with regulatory expectations, overfitting control under small-sample conditions, and the principled distinction between genuine learned representations and algorithmic feature engineering, a distinction that directly affects how results should be interpreted as biomarker evidence.

Clinical and Validation Interface Layer

This layer situates pipeline outputs within clinical, epidemiological, and regulatory contexts. Its function is to connect circadian biomarker outputs to actionable constructs: functional decline tracking, dementia risk stratification, cognitive assessment prediction, and preclinical pathology screening. The cross-cutting concerns are population representativeness across sex, race, ethnicity, socioeconomic status, and geography; regulatory alignment (including XAI requirements and medical device pathways); governance of longitudinal biometric data; and the evidence certainty calibration that distinguishes moderate-certainty institutional findings from lower-certainty community preclinical claims.

The layers are not merely sequential; they are mutually constraining. Layer 1 decisions determine the circadian constructs that layer 2 can produce. Layer 2 decisions determine the inferential ceiling available to layer 3. Layer 3 decisions determine the evidence strength with which layer 4 claims can be made. This dependency structure is the architectural translation of the cross-layer failure propagation described in the Structural Challenges at the Sensing-to-Biomarker Interface subsection.

Conceptual Validation Logic and Implementation Propositions

The proposed architecture suggests a layered validation logic that may guide subsequent empirical work. Rather than evaluating AI-circadian biomarker systems solely by end-point predictive accuracy, each layer should be assessed against criteria appropriate to its function, with the overall pipeline's evidence certainty bounded by its weakest layer.

At layer 1, validation should establish device performance in the specific population of interest, with concurrent comparison against an accepted reference (polysomnography, laboratory spectrophotometry, or supervised clinical observation, depending on modality) and transparent reporting of data completeness, device loss, and calibration drift over the monitoring period. At layer 2, validation should demonstrate that extracted circadian parameters are reproducible under standardized pipelines and interpretable within chronobiological theory, with an explicit statement of whether the study measures circadian biology directly or uses a behavioral proxy. At layer 3, validation should require external replication on independent cohorts, prespecified explainability reporting, and honest characterization of overfitting risk, particularly when sample sizes approach those observed in most included clinical studies. At layer 4, validation should evidence diagnostic or prognostic utility in prospectively followed, demographically representative cohorts, with certainty grading that respects the gradient already visible in the present corpus: moderate certainty for actigraphy-derived fragmentation-function associations in institutional settings and low certainty for AI-based predictive claims in community or preclinical populations.

From this logic, 4 implementation propositions follow directly from the evidence base. First, prospective longitudinal multisite cohort studies using harmonized circadian end points (nonparametric metrics and cosinor parameters computed with consistent epoch lengths and transparently reported pipelines) are required across cognitively normal aging, MCI, and dementia populations to move the field beyond retrospective and cross-sectional evidence. Second, AI models intended for clinical deployment should incorporate explainability mechanisms at design time rather than retrofit them, with explicit mapping between model outputs and chronobiologically interpretable parameters. Third, device and algorithm validation should be treated as population-specific rather than generic: performance in younger or healthier samples does not warrant extrapolation to institutionalized older adults or individuals with advanced dementia, where the biomarker is most clinically consequential. Fourth, demographic diversity should be operationalized as a validation criterion (not an ethical aspiration) with biomarker transferability quantified across sex, race, ethnicity, and country-income setting before generalizability is claimed.

XSL•FO
RenderX

Taken together, the challenge synthesis, reference architecture, and layered validation logic articulate what a clinically deployable AI–circadian biomarker system would require, and where the existing evidence in this review stops short of those requirements. Circadian disruption remains a biologically plausible, technically measurable, and increasingly computable candidate biomarker for neurodegeneration. Its maturation into a validated clinical tool depends less on additional sensing innovation than on structural coherence across the pipeline from device to clinical interface, on external and population-specific validation of AI components, and on explainability and governance frameworks commensurate with the clinical stakes.

Discussion

Principal Findings

Building on the cross-domain analysis, 3 convergent findings emerge from the 13 studies examining AI-enhanced wearable and objective sensing systems for circadian monitoring in older or cognitively vulnerable populations. First, circadian fragmentation (measurable as elevated IV, reduced IS, and attenuated RA from wrist actigraphy) is consistently documented in populations with advanced dementia and is associated with concurrent functional decline [23,33]. Second, AI and ML contribute meaningful value beyond descriptive actigraphy by enabling automated phenotyping of 24-hour activity profiles at a population scale [28,30], calibrated prediction of circadian light exposure [32], and multimodal fusion of behavioral markers into clinically interpretable predictions [34]. Third, the field's technical capacity to collect ecologically valid continuous data has advanced faster than its ability to standardize circadian end points, harmonize device-agnostic reporting, and independently validate AI models across diverse populations.

Contextual Interpretation and Recent Evidence

The association between actigraphy-derived circadian fragmentation and dementia risk is supported by prospective cohort evidence extending well beyond this review's corpus. Lim et al [39] and Musiek et al [40] demonstrated that rest-activity fragmentation predicts incident dementia and tracks preclinical Alzheimer pathology, supporting the biological plausibility of IS and IV as early biomarkers. The present review advances this framing by showing that AI approaches (particularly unsupervised clustering [28,30] and supervised learning [32]) can extract richer, phenotypically meaningful representations from these data streams that go beyond single-metric descriptors.

The multimodal sensing approach in Gabb et al [27] and Sprint et al [34] aligns with emerging frameworks for remote monitoring of cognitive decline that combine multiple behavioral data streams to improve predictive validity. However, neither study achieved external validation, and both were small. In Sprint et al [34], leave-one-out cross-validation on 14 participants leaves substantial overfitting risk; Gabb et al [27] fitted no predictive model, so its contribution is feasibility evidence rather than model performance.

Explainability was addressed only partially. One study applied permutation importance to a random forest [32], and 3 reported model-internal variable importance [22,28,30]; no study used local attribution methods such as SHAP [36] or LIME [37], and none explained individual predictions. Regulatory frameworks, including the US Food and Drug Administration's action plan for AI/ML-based software as a medical device [41], emphasize that clinical deployment of AI-based biomarkers requires transparency, interpretability, and human oversight, expectations that already apply to a growing number of approved AI/ML-enabled devices [42]. Knowing which hours of the day or which light-sensor channels drive a model helps during development, but a clinician deciding whether one patient's rest-activity rhythm signals elevated risk needs to know why the model flagged that patient. Circadian biomarker models that cannot give that account are unlikely to earn clinical trust or regulatory clearance in neurodegenerative contexts.

Population diversity remains a critical shortcoming. The Lancet Commission on Dementia has underscored that the global burden of neurodegenerative disease falls disproportionately on populations underrepresented in biomarker research [43]. All included studies were conducted in high-income countries, with limited reporting of racial and ethnic diversity. Wearable-based circadian biomarkers validated exclusively in these populations cannot be assumed to generalize to the populations most affected by neurodegeneration worldwide. In ML terms, this is a problem of dataset shift [44]: a model learns the relationship between sensor signals and outcomes in the population it was trained on, and that relationship may not hold when age, sex, ethnicity, comorbidity, or device-wearing behavior are distributed differently. Performance in the development sample is therefore not evidence of performance in underrepresented groups.

Limitations of the Evidence and of This Review

Sample sizes were predominantly small (n=16 to ~50 in clinical studies), with adequate statistical power only in the 2 large-scale epidemiological studies [28,30] that used retrospective data without prospective biomarker validation designs. Sex imbalances (male-only in MrOS [28]; female-only in de Zambotti et al [21]) limit generalizability. In total, 7 of the 13 studies fitted no predictive model, so conclusions about AI model performance rest on 6 studies [22,26,28,30,32,34], each developed within a single dataset and none externally validated. The assessment of methods reporting relied on the published articles and their supplementary materials; details available only in protocols or repositories not referenced by the articles were not captured. Circadian end-point heterogeneity across studies (different parametric conventions, epoch lengths, device positions, analysis software, and time windows) prevented quantitative synthesis and precluded direct cross-study comparison of effect estimates. Certainty in the evidence is moderate for actigraphy-derived fragmentation-function associations in institutional settings and low to very low for AI-based circadian biomarker claims in community or preclinical populations.

Regarding the review itself, the original search did not cover IEEE Xplore or the ACM Digital Library, which were added in the complementary search following peer-review recommendations. The complementary search is a documented

XSL•FO
RenderX

amendment to the deposited protocol, as described in the Methods. The reference architecture and validation pathway are an interpretive synthesis by the authors; they were not produced by the systematic review process and have not been empirically validated.

Implications and Future Directions

Four research priorities are indicated for the field to progress toward clinically deployable AI-based circadian biomarkers. First, prospective longitudinal cohort studies with standardized circadian metric reporting (including IS, IV, RA, and cosinor parameters calculated with harmonized epoch lengths and analysis pipelines) are needed across cognitively normal aging, MCI, and dementia populations. Second, AI models should incorporate explainability mechanisms to facilitate clinical interpretation and regulatory compliance. Third, device validation must be population-specific, accounting for the performance differences expected in very old, institutionalized, or demented populations. Fourth, population diversity in sex, race, ethnicity, and socioeconomic status must be actively prioritized to support generalizable biomarker development.

The structural interdependence of these 4 priorities (across sensing, parameter extraction, AI/ML inference, and clinical validation) motivates the reference architecture and layered validation pathway developed in the previous section, which maps each challenge to its corresponding pipeline layer and proposes implementation-oriented responses grounded in the reviewed evidence.

Conclusions

The 13-study corpus reviewed here documents a field in productive transition. AI-enhanced wearable and objective sensing devices can measure circadian fragmentation, derive 24-hour activity phenotypes, estimate circadian light exposure, and predict clinical assessments in older and cognitively vulnerable populations. Current evidence is strongest in institutional settings using actigraphy-derived nonparametric circadian metrics, where fragmentation is most pronounced and most consistently associated with functional outcomes. Circadian disruption, quantified through objective wearable sensing and AI-based analysis, constitutes a biologically plausible and technically measurable candidate biomarker for neurodegeneration. Transitioning this candidate into a validated predictive tool requires prospective longitudinal validation, cross-device harmonization, and transparent XAI approaches that the existing literature does not yet provide. The structural challenges, reference architecture, and layered validation pathway developed from this synthesis provide a conceptual map for researchers and system designers working toward that transition.

Acknowledgments

During the preparation of this manuscript, the authors used Claude (Anthropic) and Grammarly to refine language and improve clarity. An agentic workflow using Claude was also used to produce a first-pass draft of the extraction of computational-method reporting summarized in Table 3, as described in the Methods section; every extracted value was verified independently by 2 authors (JMH-Z and PCS-M) against the source articles and their supplementary materials. The authors reviewed and edited all AI-generated content and take full responsibility for the final version of the manuscript.

Funding

The authors declared no financial support was received for this work.

Data Availability

All data generated or analyzed during this study are included in this published article and its supplementary information files (Multimedia Appendices 1-3). The title-abstract and full-text screening records are available from the corresponding author on reasonable request.

Authors' Contributions

All authors contributed equally to the conception and design of this review. The literature search and analysis were performed collaboratively. The manuscript was written by JMH-Z, JM-R, JG-M, and PCS-M, and critically reviewed and edited by all authors. All authors read and approved the final version of the manuscript.

Conflicts of Interest

PCS-M is an associate editor of *JMIR Human Factors* at the time of this publication. All other authors declare no conflicts of interest.

Multimedia Appendix 1

PRISMA checklist.
[PDF File (Adobe PDF File), 254 KB-Multimedia Appendix 1]

XSL•FO
RenderX

Multimedia Appendix 2

Search strategy (PRISMA-S).
[PDF File (Adobe PDF File), 112 KB-Multimedia Appendix 2]

Multimedia Appendix 3

Mixed Methods Appraisal Tool.
[XLSX File (Microsoft Excel File), 15 KB-Multimedia Appendix 3]

References

1. GBD 2019 Dementia Forecasting Collaborators. Estimation of the global prevalence of dementia in 2019 and forecasted prevalence in 2050: an analysis for the Global Burden of Disease Study 2019. Lancet Public Health. 2022;7(2):e105-e125. [FREE Full text] [doi: 10.1016/S2468-2667(21)00249-8] [Medline: 34998485]
2. Hood S, Amir S. Neurodegeneration and the circadian clock. Front Aging Neurosci. 2017;9:170. [FREE Full text] [doi: 10.3389/fnagi.2017.00170] [Medline: 28611660]
3. Aubourg T, Demongeot J, Provost H, Vuillerme N. Exploitation of outgoing and incoming telephone calls in the context of circadian rhythms of social activity among elderly people: observational descriptive study. JMIR Mhealth Uhealth. 2020;8(11):e13535. [FREE Full text] [doi: 10.2196/13535] [Medline: 33242018]
4. Dose B, Yalçin M, Dries SPM, Relógio A. TimeTeller for timing health: the potential of circadian medicine to improve performance, prevent disease and optimize treatment. Front Digit Health. 2023;5:1157654. [FREE Full text] [doi: 10.3389/fdgth.2023.1157654] [Medline: 37153516]
5. Fifel K, Videnovic A. Circadian and sleep dysfunctions in neurodegenerative disorders—an update. Front Neurosci. 2020;14:627330. [FREE Full text] [doi: 10.3389/fnins.2020.627330] [Medline: 33536872]
6. Dhahbi W, Chamari K, Dergaa I, Trabelsi K, BaHammam AS, Ben Saad H. Sleep under fire: towards a tripartite circadian-neuroendocrine framework for conflict-zone athletes. Br J Sports Med. 2026;60(19):1453-1456. [doi: 10.1136/bjsports-2026-111904] [Medline: 42242882]
7. Reddy A, Reddy RP, Roghani AK, Garcia RI, Khemka S, Pattoor V, et al. Artificial intelligence in parkinson's disease: early detection and diagnostic advancements. Ageing Res Rev. 2024;99:102410. [doi: 10.1016/j.arr.2024.102410] [Medline: 38972602]
8. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ. 2021;372:n71. [FREE Full text] [doi: 10.1136/bmj.n71] [Medline: 33782057]
9. Herrera JM, Monroy-Rodriguez J, Santana-Mancilla PC. Artificial intelligence-based monitoring devices for the identification of circadian parameters as digital predictive biomarkers of neurodegenerative diseases in older adults. Zenodo. 2025. URL: https://zenodo.org/records/16883416 [accessed 2025-11-10]
10. Rethlefsen ML, Kirtley S, Waffenschmidt S, Ayala AP, Moher D, Page MJ, et al. PRISMA-S Group. PRISMA-S: an extension to the PRISMA Statement for Reporting Literature Searches in Systematic Reviews. Syst Rev. 2021;10(1):39. [FREE Full text] [doi: 10.1186/s13643-020-01542-z] [Medline: 33499930]
11. Wei J, Boger J. Sleep detection for younger adults, healthy older adults, and older adults living with dementia using wrist temperature and actigraphy: prototype testing and case study analysis. JMIR Mhealth Uhealth. 2021;9(6):e26462. [FREE Full text] [doi: 10.2196/26462] [Medline: 34061038]
12. Kuula L, Gradisar M, Martinmäki K, Richardson C, Bonnar D, Bartel K, et al. Using big data to explore worldwide trends in objective sleep in the transition to adulthood. Sleep Med. 2019;62:69-76. [doi: 10.1016/j.sleep.2019.07.024] [Medline: 31563008]
13. Cai R, Chao J, Gao C, Gao L, Hu K, Li P. Association between sleep duration and cognitive frailty in older Chinese adults: prospective cohort study. JMIR Aging. 2025;8:e65183. [FREE Full text] [doi: 10.2196/65183] [Medline: 40267503]
14. Tsanas A, Woodward E, Ehlers A. Objective characterization of activity, sleep, and circadian rhythm patterns using a wrist-worn actigraphy sensor: insights into posttraumatic stress disorder. JMIR Mhealth Uhealth. 2020;8(4):e14306. [FREE Full text] [doi: 10.2196/14306] [Medline: 32310142]
15. Hu B, Dixon PC, Jacobs JV, Dennerlein JT, Schiffman JM. Machine learning algorithms based on signals from a single wearable inertial sensor can detect surface- and age-related differences in walking. J Biomech. 2018;71:37-42. [doi: 10.1016/j.jbiomech.2018.01.005] [Medline: 29452755]
16. Collins GS, Moons KGM, Dhiman P, Riley RD, Beam AL, Van Calster B, et al. TRIPOD+AI statement: updated guidance for reporting clinical prediction models that use regression or machine learning methods. BMJ. 2024;385:e078378. [FREE Full text] [doi: 10.1136/bmj-2023-078378] [Medline: 38626948]
17. Tejani AS, Klontzas ME, Gatti AA, Mongan JT, Moy L, Park SH, et al. Checklist for Artificial Intelligence in Medical Imaging (CLAIM): 2024 update. Radiol Artif Intell. 2024;6(4):e240300. [FREE Full text] [doi: 10.1148/ryai.240300] [Medline: 38809149]

XSL•FO
RenderX

18. Vasey B, Nagendran M, Campbell B, Clifton DA, Collins GS, Denaxas S, et al. Reporting guideline for the early-stage clinical evaluation of decision support systems driven by artificial intelligence: DECIDE-AI. Nat Med. 2022;28(5):924-933. [doi: 10.1038/s41591-022-01772-9] [Medline: 35585198]
19. Liu X, Cruz Rivera S, Moher D, Calvert MJ, Denniston AK, SPIRIT-AICONSORT-AI Working Group. Reporting guidelines for clinical trial reports for interventions involving artificial intelligence: the CONSORT-AI extension. Nat Med. 2020;26(9):1364-1374. [FREE Full text] [doi: 10.1038/s41591-020-1034-x] [Medline: 32908283]
20. Ben Saad H, Dergaa I, Ghouili H, Ceylan H, Chamari K, Dhahbi W. The assisted technology dilemma: a reflection on AI chatbots use and risks while reshaping the peer review process in scientific research. AI Soc. 2025;40(7):5649-5656. [doi: 10.1007/s00146-025-02299-6]
21. de Zambotti M, Claudatos S, Inkelis S, Colrain IM, Baker FC. Evaluation of a consumer fitness-tracking device to assess sleep in adults. Chronobiol Int. 2015;32(7):1024-1028. [FREE Full text] [doi: 10.3109/07420528.2015.1054395] [Medline: 26158542]
22. Robben S, Englebienne G, Krose B. Delta features from ambient sensor data are good predictors of change in functional health. IEEE J Biomed Health Inform. 2017;21(4):986-993. [doi: 10.1109/JBHI.2016.2593980] [Medline: 27455530]
23. Merilahti J, Viramo P, Korhonen I. Wearable monitoring of physical functioning and disability changes, circadian rhythms and sleep patterns in nursing home residents. IEEE J Biomed Health Inform. 2016;20(3):856-864. [doi: 10.1109/JBHI.2015.2420680] [Medline: 25861091]
24. Hong QN, Fàbregues S, Bartlett G, Boardman F, Cargo M, Dagenais P, et al. The Mixed Methods Appraisal Tool (MMAT) version 2018 for information professionals and researchers. EFI. 2018;34(4):285-291. [doi: 10.3233/EFI-180221]
25. Chuang H, Lin C, Lee L, Chang H, She G, Lin Y. Comparing human-smartphone interactions and actigraphy measurements for circadian rhythm stability and adiposity: algorithm development and validation study. J Med Internet Res. 2024;26:e50149. [FREE Full text] [doi: 10.2196/50149] [Medline: 38838328]
26. Razjouyan J, Lee H, Parthasarathy S, Mohler J, Sharafkhaneh A, Najafi B. Improving sleep quality assessment using wearable sensors by including information from postural/sleep position changes and body acceleration: a comparison of chest-worn sensors, wrist actigraphy, and polysomnography. J Clin Sleep Med. 2017;13(11):1301-1310. [FREE Full text] [doi: 10.5664/jcsm.6802] [Medline: 28992827]
27. Gabb VG, Blackman J, Morrison H, Li H, Kendrick A, Turner N, et al. Longitudinal remote sleep and cognitive research in older adults with mild cognitive impairment and dementia: prospective feasibility cohort study. JMIR Aging. 2025;8:e72824. [FREE Full text] [doi: 10.2196/72824] [Medline: 40435500]
28. Cavaillès C, Wallace M, Leng Y, Stone KL, Ancoli-Israel S, Yaffe K. Multidimensional sleep profiles via machine learning and risk of dementia and cardiovascular disease. Commun Med (Lond). 2025;5(1):306. [FREE Full text] [doi: 10.1038/s43856-025-01019-x] [Medline: 40696147]
29. Wang C, Li C, Liang S. The lifestyle of new middle-aged and older adults in Taiwan described by wearable device: age and gender differences. Eur J Ageing. 2024;21(1):27. [doi: 10.1007/s10433-024-00824-y] [Medline: 39340574]
30. Shim J, Fleisch E, Barata F. Wearable-based accelerometer activity profile as digital biomarker of inflammation, biological age, and mortality using hierarchical clustering analysis in NHANES 2011-2014. Sci Rep. 2023;13(1):9326. [FREE Full text] [doi: 10.1038/s41598-023-36062-y] [Medline: 37291134]
31. Jones CD, Wasilko R, Zhang G, Stone KL, Gujral S, Rodakowski J, et al. Detecting sleep/wake rhythm disruption related to cognition in older adults with and without mild cognitive impairment using the myRhythmWatch platform: feasibility and correlation study. JMIR Aging. 2025;8:e67294. [FREE Full text] [doi: 10.2196/67294] [Medline: 40194310]
32. Beiglary S, Feng Y, Wang N, Ghaeili N, Jao Y, Liao Y, et al. Using wearable sensors to measure and predict personal circadian lighting exposure in nursing home residents: model development and validation. JMIR Aging. 2025;8:e72338. [FREE Full text] [doi: 10.2196/72338] [Medline: 41055091]
33. Froggatt K, Best A, Bunn F, Burnside G, Coast J, Dunleavy L, et al. A group intervention to improve quality of life for people with advanced dementia living in care homes: the namaste feasibility cluster RCT. Health Technol Assess. 2020;24(6):1-140. [FREE Full text] [doi: 10.3310/hta24060] [Medline: 31971506]
34. Sprint G, Cook DJ, Schmitter-Edgecombe M, Holder LB. Multimodal fusion of smart home and text-based behavior markers for clinical assessment prediction. ACM Trans Comput Healthc. 2022;3(4):41. [FREE Full text] [doi: 10.1145/3531231] [Medline: 36381500]
35. Dhahbi W, Jebabli N, Souaifi M, Ceylan H, Ben Saad H, Chamari K, et al. Application of artificial intelligence for predicting sports injuries and customizing personalized prevention strategies: a scoping review. Bioengineering (Basel). 2026;13(6):692. [FREE Full text] [doi: 10.3390/bioengineering13060692] [Medline: 42351936]
36. Lundberg SM, Lee SI. Unified approach to interpreting model predictions. ArXiv. Preprint posted online on November 25, 2017. 2017. [doi: 10.48550/arXiv.1705.07874]
37. Ribeiro MT, Singh S, Guestrin C. "Why Should I Trust You?": explaining the predictions of any classifier. 2016. Presented at: KDD '16: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining; August 13-17, 2016:1135-1144; San Francisco, CA, United States. [doi: 10.1145/2939672.2939778]

XSL•FO
RenderX

38. Dhahbi W, Chamari K. The algorithmic athlete: a call to standardize assessment of sensor technologies and artificial intelligence. Int J Sports Physiol Perform. 2026;21(4):505-506. [FREE Full text] [doi: 10.1123/ijspp.2025-0547] [Medline: 41493778]
39. Lim ASP, Yu L, Costa MD, Leurgans SE, Buchman AS, Bennett DA, et al. Increased fragmentation of rest-activity patterns is associated with a characteristic pattern of cognitive impairment in older individuals. Sleep. 2012;35(5):633-640B. [FREE Full text] [doi: 10.5665/sleep.1820] [Medline: 22547889]
40. Musiek ES, Bhimasani M, Zangrilli MA, Morris JC, Holtzman DM, Ju YS. Circadian rest-activity pattern changes in aging and preclinical alzheimer disease. JAMA Neurol. 2018;75(5):582-590. [FREE Full text] [doi: 10.1001/jamaneurol.2017.4719] [Medline: 29379963]
41. Artificial Intelligence/Machine Learning (AI/ML)-Based Software as a Medical Device (SaMD) Action Plan. Food & Drug Administration. 2021. URL: https://www.fda.gov/media/145022/download [accessed 2026-10-27]
42. Joshi G, Jain A, Araveeti SR, Adhikari S, Garg H, Bhandari M. FDA-approved artificial intelligence and machine learning (AI/ML)-enabled medical devices: an updated landscape. Electronics. Jan 24, 2024;13(3):498. [doi: 10.3390/electronics13030498]
43. Livingston G, Huntley J, Liu KY, Costafreda SG, Selbæk G, Alladi S, et al. Dementia prevention, intervention, and care: 2024 report of the Lancet standing Commission. Lancet. Aug 2024;404(10452):572-628. [doi: 10.1016/s0140-6736(24)01296-0]
44. Finlayson SG, Subbaswamy A, Singh K, Bowers J, Kupke A, Zittrain J, et al. The clinician and dataset shift in artificial intelligence. N Engl J Med. Jul 15, 2021;385(3):283-286. [FREE Full text] [doi: 10.1056/NEJMc2104626] [Medline: 34260843]

Abbreviations

AD: Alzheimer disease
CLAIM: Checklist for Artificial Intelligence in Medical Imaging
CONSORT: Consolidated Standards of Reporting Trials
CONSORT-AI: Consolidated Standards of Reporting Trials-Artificial Intelligence
DECIDE-AI: Developmental and Exploratory Clinical Investigations of Decision Support Systems Driven by Artificial Intelligence
EEG: electroencephalography
IS: interdaily stability
IV: intradaily variability
LIME: Local Interpretable Model-Agnostic Explanations
MESOR: midline estimating statistic of rhythm
ML: machine learning
MMAT: Mixed Methods Appraisal Tool
MrOS: Osteoporotic Fractures in Men Study
PD: Parkinson disease
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PRISMA-S: Preferred Reporting Items for Systematic Reviews and Meta-Analyses Literature Search Extension
RA: relative amplitude
RCT: randomized controlled trial
SHAP: Shapley Additive Explanations
STROBE: Strengthening the Reporting of Observational Studies in Epidemiology
TIDieR: Template for Intervention Description and Replication
TRIPOD+AI: Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis+Artificial Intelligence
XAI: explainable AI

XSL•FO
RenderX

Edited by N Schütz; submitted 26.Aug.2025; peer-reviewed by S Mitra, W Dhahbi; comments to author 20.Nov.2025; revised version received 21.Sep.2026; accepted 19.Oct.2026; published 20.Nov.2026

Please cite as:
Herrera-Zamora JM, Monroy-Rodriguez J, Garcia-Mancilla J, Santana-Mancilla PC
AI-Based Monitoring Devices for the Identification of Circadian Parameters as Digital Predictive Biomarkers of Neurodegenerative Diseases in Older Adults: Systematic Review
JMIR Aging 2026;9:e83031
URL: https://aging.jmir.org/2026/1/e83031
doi: 10.2196/83031
PMID:

©J Manuel Herrera-Zamora, Jesus Monroy-Rodriguez, Jesus Garcia-Mancilla, Pedro C Santana-Mancilla. Originally published in JMIR Aging (https://aging.jmir.org), 20.Nov.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX