Do Objective Circadian-Alignment Scores Track Emergency Physicians' Own Sense of a "Good" Schedule?

A derivation-cohort study of the Circadian Alignment Index (CAI) using blinded, forced-choice schedule comparisons

Michael Zimmer, MD, FACEP

CAIER Insights, AIER Technologies Inc.

Correspondence: hello@aiertechnologies.com · research.caier.ai

Abstract

Background. Emergency-department (ED) schedules are widely believed to differ in how fatiguing they are, but scheduling software rarely tests whether its own "fairness" or "quality" scores agree with what practicing physicians actually feel about a schedule. The Circadian Alignment Index (CAI) is a proprietary composite score intended to quantify this, informed by established circadian-biology and fatigue-risk research; the specific factors it weighs and how they are combined are not disclosed here. We report the first ("v1") cohort of an ongoing public research platform built to test whether CAI agrees with physician preference.

Methods. 50 board-certified/board-eligible emergency physicians and advanced practice clinicians (57 enrolled; 87.7% completion) completed blinded, forced-choice comparisons between pairs of synthetic single-physician monthly schedules (8-hour and 12-hour shift profiles, ~30 days each), selecting the schedule they would personally prefer to work. Sessions embedded obvious attention-check pairs, near-threshold ("low-delta") pairs for just-noticeable-difference estimation, and order-swapped repeat pairs for test–retest reliability. The primary endpoint, agreement between participant choice and the higher-CAI schedule, was analyzed with a pre-specified group-sequential design (O'Brien–Fleming boundaries; Gwet's AC1 with participant-clustered intervals).

Results. Across 626 forced-choice comparisons (576 non-attention-probe trials analyzed here), participants agreed with the higher-CAI schedule on ~69% of trials (raw agreement 68.2%, 393/576; participant-clustered signal $Z \approx 3.4$). This directional signal did not cross the pre-specified early-stopping efficacy boundary at this sample size. Agreement rose with the size of the CAI gap between schedules, from 60.2% at $|\Delta CAI|$ 1–9 to 78.9% at $|\Delta CAI|$ 60–69, and with participants' self-reported confidence (35–78% from confidence 1 to 5). Order-swapped repeat items showed 82.0% test–retest agreement (41/50). Attention-check pairs were answered as expected 86.0% of the time. This cohort's data were also used, by design, to derive a revised scoring engine (CAI v2.1); v1→v2.1 rank correlation was $\rho \approx 0.95$–0.97.

Conclusions. In this exploratory derivation cohort, physician schedule preference tracked an objective circadian-alignment score well above chance, with a plausible dose–response gradient and reasonable short-term reliability — but the pre-registered efficacy threshold was not met, and the same data were used to tune the index, which is a circular basis for confirmation. These findings should be read as hypothesis-generating pilot evidence, not as validation of CAI, pending an independent confirmatory cohort against a frozen scoring engine.

Keywords: *emergency medicine; physician scheduling; fatigue risk management; circadian rhythm; shift work; forced-choice comparison; construct validity*

1. Introduction

Emergency physicians routinely describe some monthly schedules as “brutal” and others as “fine,” even when both satisfy the same coverage requirements and total-hours targets. A substantial body of sleep and circadian-biology literature links shift sequencing, direction of rotation, consecutive-night runs, and quick turnarounds to measurable fatigue and performance decrements in shift workers generally and emergency clinicians specifically (Sack et al., 2007; Morgenthaler et al., 2007; Boivin, Boudreau & Kosmadopoulos, 2021; Wickwire et al., 2017; Barger et al., 2018; Fowler et al., 2022; Kontrick & Agarwal, 2024). Scheduling software increasingly attaches a single composite “quality” or “fairness” score to a generated roster, but it is uncommon for the developers of such a score to test, prospectively and blinded, whether it agrees with the people who actually have to work the schedule.

The Circadian Alignment Index (CAI) is one such composite score, developed by AIER Technologies Inc. for its emergency-medicine scheduling product. CAI draws on established circadian-biology and fatigue-risk research to rank schedules by their expected physiologic burden; the specific factors it weighs and how it combines them are proprietary and not described in this report. CAIER Insights is a standalone, public research instrument (research.caier.ai), separate from the commercial product, built specifically to test whether CAI’s rankings agree with the gut judgment of practicing emergency physicians, and to estimate how large a CAI difference has to be before physicians notice it.

This report describes the platform’s first data collection wave (referred to throughout as the “v1” or “derivation” cohort): 50 completed physician sessions and 626 blinded, forced-choice schedule comparisons, collected between July and August 2026 under the original CAI scoring engine. We report it here explicitly as an exploratory, hypothesis-generating study — for two reasons that we think are important to state plainly rather than bury in a limitations paragraph. First, the platform’s own pre-registered sequential analysis plan evaluated this cohort against a formal efficacy boundary, and the observed signal, while clearly directional, did not cross it. Second, this same cohort’s comparisons were used, by design, to fit a revised version of the scoring engine (CAI v2.1) — which means the cohort cannot also serve as independent confirmation that the revised engine is correct. Readers should treat the estimates below as a derivation / pilot dataset, not as a validated result.

2. Methods

2.1 Design overview

CAIER Insights is an open, uncompensated (except where noted) online research platform. Visitors complete an e-consent flow and then a series of forced-choice comparisons: on each trial, two candidate monthly schedules are shown side by side, with position (left/right, i.e., “A/B”) randomized, and the participant is asked which schedule they would personally prefer to work. CAI scores are never shown to participants during the task. Each session presents a batch of 10 comparison trials. The trial set for each session mixes: (a) primary comparison trials drawn from a bank of scored schedule pairs; (b) attention-check probes, in which the CAI gap between the two schedules is large and unambiguous (CAI ≥ 90 vs. ≤ 20), used to screen for inattentive responding; (c) low-delta probes, in which the CAI gap is small (target range 2–8 points), used to characterize the smallest CAI difference physicians reliably notice; and (d) a small number of repeated trials, in which an earlier pair is re-shown later in the session with the two schedules’ left/right position swapped, used to estimate within-session test–retest reliability independent of position bias.

2.2 Participants

Eligible participants were board-certified or board-eligible emergency physicians, urgent-care physicians, or advanced practice clinicians (APCs) working in emergency/urgent-care settings, recruited primarily through a compensated physician-research panel (35/50 completers) with a smaller number arriving directly or via unlabeled referral (15/50). Fifty-seven sessions were started; 50 (87.7%) were completed and form the analysis sample. All participants self-identified their specialty as emergency medicine (two also indicated urgent care), predominantly with 10+ years in practice, across a mix of community, academic, and trauma-center settings (Table 1). No protected health information was collected; sessions were pseudonymous, tied only to a participant token.

2.3 Stimuli

Comparison stimuli were synthetically generated single-physician monthly schedules (not full departmental rosters), covering two shift-length profiles — 8-hour shifts (~121 hours/month) and 12-hour shifts (~132 hours/month) — each spanning approximately 30 days. Each schedule was scored on generation by the CAI engine, producing a single composite score; the engine's internal sub-metrics are proprietary and not broken out in this report. Schedule pairs shown to participants were selected to span a range of CAI gaps, from near-threshold (a few points) to large and unambiguous (50+ points).

2.4 Outcomes and statistical analysis

The pre-specified primary endpoint was agreement between a participant's stated preference and the schedule with the higher CAI score, analyzed using Gwet's AC1 statistic with 95% confidence intervals clustered by participant, restricted to each participant's first completed session, and evaluated against a pre-specified group-sequential (O'Brien–Fleming) efficacy boundary defined in the platform's internal, pre-registered analysis plan. Attention-check probes were excluded from the primary endpoint (they test compliance, not the hypothesis) but reported separately as a data-quality check. We additionally report, as descriptive/secondary analyses computed directly from the underlying trial-level data: (i) agreement rate by the magnitude of the CAI gap between the two schedules (dose–response), (ii) agreement rate by participants' self-reported confidence, (iii) agreement rate restricted to low-delta (near-threshold) probes, (iv) test–retest agreement on order-swapped repeated trials, and (v) the frequency of free-text reason codes participants attached to their choices. The headline primary-endpoint statistics (Gwet's AC1, its clustered confidence interval, and the sequential-boundary test result) are reported as published by the platform's own locked analysis pipeline rather than re-derived here; all secondary/descriptive statistics were computed directly from the underlying database for this report.

2.5 Data availability

Aggregate results are published on an ongoing basis at research.caier.ai/results. De-identified trial-level data underlying this report can be made available on reasonable request, consistent with the platform's participant consent and privacy terms.

3. Results

3.1 Participants and data quality

Table 1 summarizes the analysis sample (n = 50 completed sessions).

Characteristic	n	%
Years in practice: 10+	33	66%
Years in practice: 3–9	16	32%
Years in practice: not reported	1	2%
Practice setting includes: community	24	48%
Practice setting includes: academic	17	34%
Practice setting includes: rural / critical access	8	16%
Practice setting includes: trauma center (any level)	20	40%
Recruitment: physician research panel	35	70%
Recruitment: direct / unlabeled	15	30%

Note: practice-setting categories are non-exclusive (participants could select multiple).

Completion and engagement. 57 sessions were started and 50 (87.7%) completed the full 10-trial batch. Completers spent a median of 54 seconds per comparison trial (mean 72 seconds).

Attention checks. Participants selected the expected (unambiguously higher-CAI) schedule on 43/50 (86.0%) attention-check trials, supporting adequate task engagement.

3.2 Primary endpoint: agreement with the CAI ranking

Across 626 total forced-choice comparisons (576 non-attention-probe trials), participants chose the higher-CAI schedule in 68.2% of trials (393/576; participant-clustered bootstrap 95% CI 64.1–72.5%). Using the platform's pre-specified primary analysis (Gwet's AC1 with participant-clustered intervals), the directional signal corresponded to $Z \approx 3.4$ — clearly above chance in direction, but the pre-specified group-sequential efficacy boundary for this sample size was not crossed. Under the platform's own governance rules, this cohort is therefore reported as directional / hypothesis-generating rather than as a confirmed positive result.

3.3 Dose–response by CAI gap size, and confidence calibration

Agreement increased with the absolute CAI gap between the two schedules shown, from 60.2% at $|\Delta CAI|$ of 1–9 points to 78.9% at $|\Delta CAI|$ of 60–69 points, with intermediate bins following the same broad trend (Figure 1A; Table 2). Agreement also increased with participants' self-reported confidence in their choice, from 35.3% at confidence level 1 (n = 17) to 77.9% at confidence level 5 (n = 172) (Figure 1B). Both patterns are consistent with participants tracking a real, graded signal rather than responding at random.

Figure 1. Descriptive agreement patterns in the v1 derivation cohort (non-attention-probe trials, n=576)

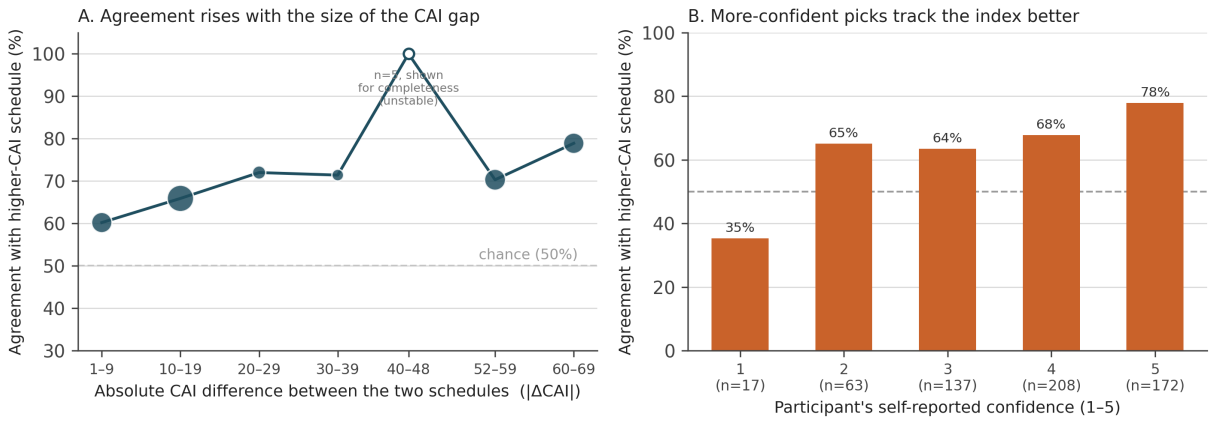

Figure 1. (A) Agreement with the higher-CAI schedule by the size of the CAI gap between the two schedules compared (marker size ∝ trial count; the 40–48 bin, n = 5, is shown hollow as unstable). (B) Agreement by participant's self-reported confidence in their choice. Dashed line = chance (50%).

|ΔCAI| bin	Trials (n)	Agreement (%)	Mean confidence
1–9	93	60.2%	3.48
10–19	264	65.9%	3.77
20–29	25	72.0%	3.76
30–39	14	71.4%	3.79
40–48[1]	5	100.0%	4.40
52–59	101	70.3%	3.72
60–69	95	78.9%	4.01

[1] Sparse bin (n = 5); interpret with caution.

3.4 Near-threshold agreement (just-noticeable-difference probes)

Trials specifically constructed with a small CAI gap (low-delta probes, target range 2–8 points, observed mean gap 4.8 points, n = 78 across all sessions / n = 76 among completers) still showed above-chance agreement, at 62.8–63.2%. This suggests physicians can detect and act on CAI differences considerably smaller than the large gaps used in attention checks, though the precise just-noticeable-difference threshold is not yet pinned down by this sample.

3.5 Test–retest reliability

Fifty trials were re-presented later in the same session with the two schedules' left/right position swapped, to separate genuine preference reliability from simple position bias. Comparing which physical schedule (not which screen position) was chosen, participants gave the same answer on both presentations in 82.0% of repeated trials (41/50).

3.6 Participants' stated reasons

Participants could optionally tag their choice with one or more structured reason codes describing what drove their pick. The most commonly cited themes — including the predictability of the shift rhythm, weekend workload, and how night shifts were spaced and recovered from — were broadly consistent with well-established circadian-fatigue risk factors in the literature cited above. We report this as general, reassuring face validity for the category of score CAI represents, not as a description of CAI's internal scoring logic, which we do not disclose here, and it does not substitute for the quantitative agreement analysis above.

3.7 Downstream use: derivation of CAI v2.1

Consistent with its role as a derivation cohort, the v1 comparison data were used to refine the scoring engine (CAI v2.1). Adjustments were informed by where v1 physicians and CAI disagreed most; the specific scoring architecture and parameters involved are proprietary and not disclosed in this report. Rank correlation between schedules' CAI v1 and CAI v2.1 scores was high ($\rho \approx 0.95$–0.97), indicating the revision is a refinement rather than a wholesale change in what the index rewards. CAI v2.1 has since been locked (frozen) as the scoring engine for subsequent, independent data collection.

4. Discussion

In this first, exploratory cohort of 50 emergency physicians and 626 blinded forced-choice comparisons, stated schedule preference agreed with an objective circadian-alignment score on roughly two-thirds to 69% of trials — well above the 50% expected by chance — with a graded dose–response relationship to the size of the score gap, reasonable short-term test–retest reliability (82%), and stated reasons that were broadly consistent with known circadian-fatigue factors. Taken together, these are encouraging signs of construct validity for CAI as a proxy for physician-perceived schedule burden.

4.1 How good is a 68–69% agreement rate, in plain terms?

What "Z ≈ 3.4" means. A Z-score measures how far a result sits from what pure chance would produce, in standard-deviation units. As an everyday reference point, most single-look studies treat $Z \geq 1.96$ (roughly a 1-in-20 chance the result is a fluke) as "statistically significant." Our $Z \approx 3.4$ clears that everyday bar comfortably — on its own, a result that size would usually be reported as a positive finding. But this study did not use the everyday bar. Because it was designed with several planned looks at the data rather than one (see §4.2), the pre-registered rule for declaring this specific look a success was deliberately stricter than $Z \geq 1.96$, precisely so an early, small look can't oversell itself before more data arrives. Under that stricter, pre-committed bar, this look fell short — even though the same number would clear the bar most single-look studies use.

What "68–69% agreement" means next to other things clinicians agree (or disagree) on. It is tempting to hear "69%" and assume that is a weak result because it is well short of 100%. In practice, agreement rates in medicine are rarely near 100%, even for judgments that feel more clear-cut than "which schedule would I rather work." A few reference points help calibrate this:

• On the highly structured end: two clinicians triaging the same patient with a well-defined, protocol-driven 5-level scale (the Emergency Severity Index) have been shown to agree very highly with one another — weighted kappa around 0.94, "almost perfect" on the standard scale below (Jafari-Rouhi et al., 2013).

• On the more open-ended end: a Mayo Clinic study of 286 patients who sought a second opinion found the second physician fully confirmed the original diagnosis in only about 12% of cases — 66% were "refined" and 21% were substantively different (Van Such et al., 2017). Even physician-to-physician agreement on "what is wrong with this patient" clusters well below 100% once the judgment becomes more open-ended.

• The standard scale researchers use to label agreement/reliability coefficients (Landis & Koch, 1977) treats 0.61–0.80 as “substantial” agreement and 0.41–0.60 as “moderate,” reserving “almost perfect” for 0.81 and above (Table 3).

Coefficient range	Landis & Koch (1977) label
< 0	Poor
0.00 – 0.20	Slight
0.21 – 0.40	Fair
0.41 – 0.60	Moderate
0.61 – 0.80	Substantial
0.81 – 1.00	Almost perfect

Table 3. Standard benchmark scale for interpreting chance-corrected agreement coefficients, shown here for orientation only — our headline figures (raw agreement, Z) are not on this 0–1 coefficient scale, but the qualitative bands (“moderate,” “substantial”) are the common shorthand this field uses to describe results in the same broad range as ours.

Read against that backdrop, a ~69% agreement rate between a purely computed circadian-fatigue score and physicians' personal, gut-level “which one would I rather work” judgment is a genuinely encouraging number for a first pass. It sits well above chance, on a task — subjective preference, not a structured clinical protocol — that tends to produce lower agreement than more procedural judgments like triage. It simply is not, by itself, enough to clear the stricter bar this particular study pre-committed to.

What the v1→v2.1 correlation (ρ ≈ 0.95–0.97) means. This rank correlation says that schedules the original formula scored well, the revised formula also tends to score well, and vice versa — the ordering of “better” and “worse” schedules is largely preserved. A useful analogy is recalibrating a bathroom scale: a reading that used to say 8.2 lb might now say 8.4 lb, but heavy things are still heavy and light things are still light. CAI v2.1 was a tune-up of where the needle sits, not a rebuild of the scale — aimed specifically at closing the physician-agreement gap described above.

4.2 Why v1 is a derivation cohort, not a validation

Before any data were collected, the study’s analysis plan committed to a specific decision rule rather than leaving that decision to be made after seeing the results — a common and well-documented source of bias in research ("peeking" and moving the goalposts). The plan specified a group-sequential design with formal looks at the data at N = 50, 100, and 250 completed sessions, each using a pre-set statistical boundary (O’Brien–Fleming-style) that gets easier to clear at each later look, so that an early, small sample cannot declare victory on noise alone.

Separately, governance rules set before data collection specified that using v1 data to revise CAI — for any reason — would convert the affected sessions from prospective validation data into derivation data, precluding the same data from both motivating a change to the engine and later being used to validate it. This is a general data-hygiene rule rather than a specific pre-commitment that a boundary miss at N = 50 would, by itself, trigger a revision.

In practice, a pre-registered exploratory diagnostic was run on these 626 comparisons to identify which fatigue and circadian axes CAI v1 was under-weighting or missing entirely. That diagnostic informed the derivation of CAI v2.1 (§3.7); because that revision drew on v1 data, the governance rule above applied, and the v1 sessions were retained as derivation data rather than treated as a validation of CAI v1. CAI v2.1 is now locked/frozen, and a new, independent cohort is collecting data against it.

We think this sequence — a pre-registered governance rule, an honest miss against a strict boundary, a diagnostic-driven revision, and a frozen index carried into the next test — is itself part of the result worth reporting, rather than a footnote to it. It is also why we describe v1 throughout this report as a derivation cohort rather than a validation: v1's job was to tell us whether CAI v1 was good enough to stand on its own (by the pre-registered standard, it was not) and, if not, what specifically to fix. It did that job.

4.3 What this cohort does and does not show

To be precise about what remains open: the Z ≈ 3.4 signal described above is real and directional, but it is not yet a statistically confirmed result by this study's own pre-specified standard (§4.2). Separately, and just as importantly, this cohort's comparisons were themselves used to tune CAI v2.1. Using the same data both to derive a model and to claim that the model is correct is circular; it can produce apparently strong agreement even when the underlying construct validity is weaker than it looks. For both reasons, we present these findings as hypothesis-generating pilot data — useful for informing engine design and for estimating effect sizes and sample-size requirements for the next look — rather than as evidence that CAI has been validated.

Other limitations bear on interpretation. The sample was recruited substantially through a single paid physician-research panel, is modest in size (n = 50), and skews toward experienced (10+ year) community and academic emergency physicians; it may not represent newer clinicians, other specialties who staff EDs (e.g., family medicine-trained APCs outside our inclusion criteria), or physicians outside the recruitment channels used. Stimuli were single-physician monthly schedules rather than full departmental rosters, so results may not generalize to how physicians judge schedules embedded in a real, group-level rota with swap dynamics and colleague-specific context. Participation was uncompensated for a subset of participants, which may select for physicians with stronger opinions about scheduling. Finally, because CAI scores were withheld during the task but the schedules themselves were synthetic and machine-generated, we cannot rule out that some other correlated feature of "CAI-favored" schedules (rather than the construct CAI is meant to measure) drove part of the observed agreement.

We think the appropriate reading of this dataset is as Study 1 in a two-part design: a derivation phase (reported here) used to characterize physician response patterns and refine the scoring engine, intended to be followed by an independent confirmatory phase using a frozen version of the engine. We report v1 here on its own terms, and encourage readers evaluating CAI or similar composite schedule-quality scores to look for that independent confirmatory evidence before treating agreement estimates like the ones above as validated.

5. Conclusion

Emergency physicians' stated schedule preferences tracked an objective circadian-alignment index (CAI) well above chance in this exploratory, 50-physician derivation cohort, with a plausible dose–response gradient and reasonable short-term reliability. The pre-specified efficacy threshold for this cohort was not met, and the same data were used to tune the index under evaluation. We present these results as pilot, hypothesis-generating evidence supporting further, independent study of whether composite fatigue/circadian scores can serve as a usable proxy for physician-perceived schedule quality.

Declarations

Conflict of interest. The author is the founder of AIER Technologies Inc., which develops the CAI scoring engine and the commercial scheduling product (AIER Schedule / "ER Schedule") that CAI also informs, and operates the CAIER Insights research platform described here. This is a direct financial and professional conflict of interest in evaluating whether CAI performs well; readers should weigh the results accordingly, and independent replication is encouraged.

Ethics. Participation was voluntary, preceded by informed e-consent, pseudonymous, and did not involve protected health information or patient data. This study has not been submitted for or reviewed by an Institutional Review Board. Given its minimal-risk, anonymous, non-clinical, survey-style design, the author believes it would likely qualify as IRB-exempt; a formal exemption or approval determination has not yet been sought and will be pursued before further dissemination or journal submission.

Funding. This work was self-funded by AIER Technologies Inc. No external funding was received.

Data and materials. Aggregate, continuously updated results are published at research.caier.ai/results. Study methods are described at research.caier.ai/study-methods.

Acknowledgments. The author thanks the emergency physicians and advanced practice clinicians who volunteered their time to this study.

References

American Academy of Sleep Medicine. (2020). Position statement on sleep deprivation and physician burnout.

American Academy of Sleep Medicine. (2024). Position statement on sleep, circadian health, and work scheduling.

American Academy of Sleep Medicine. (2026). Management of shift work disorder: Clinical practice guideline.

American Academy of Sleep Medicine & Sleep Research Society. Guiding principles for work shift duration, performance, safety, and health.

Association of Salaried Medical Specialists (ASMS). (2016). Shift work, scheduling, and risk factors: A literature review.

Barger, L. K., et al. (2018). Effect of fatigue training on safety, fatigue, and sleep in emergency medical services. Prehospital Emergency Care.

Boivin, D. B., Boudreau, P., & Kosmadopoulos, A. (2021). Disturbance of the circadian system in shift work and its health impact. Journal of Biological Rhythms.

CDC/NIOSH. Work hours, shift work, and worker fatigue (resource collection).

Coelho, J., et al. (2023). Sleep timing, workplace well-being, and mental health in healthcare workers. Sleep Medicine.

Fowler, L. A., et al. (2022). Objective assessment of sleep and fatigue risk in emergency medicine physicians. Academic Emergency Medicine.

Fox, A., Dall'Ora, C., & Young, E. (2025). Fatigue risk management in healthcare: A scoping literature review. International Journal of Nursing Studies.

Harrison, Y., et al. (2019). Circadian profiles and scheduling effects on sleep and performance in emergency departments. Journal of Emergency Medicine.

Hsu, C. C., Obermeyer, Z., & Tan, C. (2023). Clinical notes reveal physician fatigue.

Jafari-Rouhi, A. H., et al. (2013). The Emergency Severity Index, version 4, for pediatric triage: A reliability study in Tabriz Children's Hospital, Tabriz, Iran. International Journal of Emergency Medicine, 6, 36.

Klinefelter, Z., et al. (2023). Emergency physician perspectives on fatigue and shift work. Clocks & Sleep.

Kontrick, A. V., & Agarwal, G. (2024). Shift work, circadian misalignment, and healthcare worker well-being. Psychiatric Annals.

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

Morgenthaler, T. I., et al. (2007). Practice parameters for the clinical evaluation and treatment of circadian rhythm sleep disorders. Sleep.

North-West NHS Consultant Occupational Health Physicians Group. (2025). Consensus document on fatigue risk management in healthcare.

Patterson, P. D., et al. (2018). Evidence-based guidelines for fatigue risk management in emergency medical services. Prehospital Emergency Care.

Sack, R. L., et al. (2007). Circadian rhythm sleep disorders: Part I. Basic principles, shift work and jet lag disorder. Sleep.

Van Such, M., Lohr, R., Beckman, T., & Naessens, J. M. (2017). Extent of diagnostic agreement among medical referrals. Journal of Evaluation in Clinical Practice, 23(4), 870–874.

Wickwire, E. M., et al. (2017). Shift work and shift work sleep disorder: Clinical and organizational perspectives. Chest.

Working Time Society. (2019). A multi-level approach to managing occupational sleep-related fatigue. Industrial Health.

Full citation details (volume/issue/pages/DOIs) should be verified against research.caier.ai/science and completed before submission.