MEDBUG · CONFERENCE ABSTRACT

Development and validation of non-contact multi-modal sleep stage classification system

ASA 2026 · Accepted

Thi Cham Nguyen1, Andrew Vakulin1, Bastien Lechat1, Jack Manners1, Alison Teare1, Karen Reynolds2, Danny J. Eckert1, Duc Phuc Nguyen1* and Peter Catcheside1*.

*Flinders Health and Medical Research Institute Flagship Sleep Health Program, College of Medicine and Public Health, Flinders University, South Australia The Medical Device Research Institute, College of Science and Engineering, Flinders University, Bedford Park, South Australia, Australia. * PC and PN are co-senior authors*

Study objectives To develop and validate a mattress-sensor and deep learning-based sleep stage classification system to enable simplified multi-night sleep monitoring.

Methods A 100-million parameter deep learning model was first developed using a VGG1D encoder for feature extraction, and a 12-layer Transformer encoder to capture long-range temporal dependencies within pre-recorded raw polysomnography (PSG) signal recordings. A sleep stage classifier was then developed using 68 overnight recordings from 68 participants who underwent simultaneous expert scored PSG and unobtrusive ballistocardiographic mattress-sensor and bedside sensor device recordings. The classifier was then independently validated against expert-scored PSG data from 22 separate PSG recordings from 15 different participants using Cohen's Kappa, epoch-by-epoch classification accuracy, receiver operator characteristic (ROC) and Bland-Altman analyses.

Results Overall, four-stage classification accuracy was 65% (95% CI: 52–77%), with a Cohen's Kappa of 0.45 (95% CI: 0.44–0.46) and overall, ROC area-under-curve (AUC) of 0.86. Classification performance was highest for REM (sensitivity 76.3%, AUC 0.90) and wake (sensitivity 73.2%, AUC 0.89) and lower for deep (59.3%, AUC 0.86) and light (49.2%, AUC 0.71). Classifier versus PSG derived sleep metrics were strongly correlated with no significant bias for total sleep time (r=0.76, bias±1.96SD -17.6[-38.6 to 3.5] min, p=0.117), sleep efficiency (r=0.71, bias -3.7[-7.9 to 0.5] min, p=0.099), and sleep onset latency (r=0.78, bias 0.3 [-5.5 to 5.9] min, p=0.933).

Conclusions This approach provides a robust, high-performance and scalable method for multi-night sleep staging to help address problems of night-to-night variability in sleep measures not evaluated with conventional 1-night PSG.