www.nature.com/scientificdata

scientific data

OPEN

DATA DESCRIPTOR

Ear-EEG sleep monitoring data sets

Kaare Bjarke Mikkelsen[1]✉, Yousef Rezai Tabar[1], Laura Rævsbæk Birch[1], Simon Lind Kappel[1], Christian Bech Christensen[1], Lars Dalskov Mosgaard[2], Marit Otto[3], Martin Christian Hemmsen[2], Mike Lind Rank[2] & Preben Kidmose[1]

Here we present data from two studies, both of which had the purpose of investigating the potential of using electroencephalograms measured from the ear ('ear-EEG') for sleep monitoring in a home environment. In total, 320 nights were recorded. All nights were recorded with ear-EEG, and some were also recorded using scalp-EEG and/or wristworn actigraphy. All subjects were recorded multiple times. To our knowledge, this is the most extensive open access data set available for mobile EEG development, and possibly also the best open access dataset for studying repeated sleep monitoring on individuals. We describe the details of each data set, including data quality measures, and compare the sleep scoring performance to a previously published dataset.

Background & Summary

We here describe a collection of 2 data sets consisting of in total 320 sleep recordings, all performed using the ear-EEG method[1], on 30 healthy subjects. These constituent data sets have each first been described in Mikkelsen *et al.*[2] and Tabar *et al.*[3] and will be referred to as EESM19 and EESM23 (EESM meaning 'Ear-EEG Sleep Monitoring' and the number referring to the year of first publication). The unscored portion of EESM19 was first described in the later paper Mikkelsen *et al.*[4]. Due to rules regarding anonymization of the data sets, they have not been shared previously, hence this data descriptor to describe the formats and other details surrounding the two data sets.

The data was collected as part of an endeavor to develop the ear-EEG recording method[1,5] into a capable sleep monitoring method. To our knowledge, this collection is, at the time of writing, the largest open data set concerning mobile EEG monitoring (sleeping or otherwise), and the largest open collection of repeated EEG sleep recordings on the same subjects.

While there exist multiple different sleep monitoring solutions based on EEG, we believe that this data set will be useful for a broad selection of biomedical and sleep researchers. We base this not only on the quality of the data set itself, but also previous results indicating that many different recording setups can lead to similarly good sleep monitoring[6]. Thus, we hope that this data set may in general increase the interest in high quality mobile sleep monitoring.

Additionally, we also believe that this data will be useful for researchers working on other aspects of mobile brain monitoring. The data set contains long EEG recordings in uncontrolled environments, using recording techniques relevant for future mobile brain-computer-interface implementations.

In Table 1 is seen an overview of the two data sets. We see that they differ in both type of EEG recording equipment, size as well as which additional sensors were included.

Methods

Please see Table 1 for an overview of the differences between the two datasets, as described below:

For both data sets, healthy, adult research subjects were recruited through public channels and word-of-mouth. After obtaining informed consent, subjects were fitted for ear-EEG recordings. For EESM19, this consisted of taking molds of the outer ear (performed at the local audiological clinic), for EESM23, a selection of generic ear pieces were tried to find the best fit.

After fitting ear pieces had been prepared for the subject, two types of recordings took place: first combined recordings of ear-EEG and polysomnography (PSG) during either 4 (EESM19) or 2 (EESM23) nights. These recordings were not back-to-back, but scheduled according to equipment and time restrictions. After finishing the combined recordings, half of the subjects in EESM19 and all of the subjects in EESM23 underwent ear-EEG only recordings. For EESM19, this was 12 additional nights, for EESM23 it was 10.

[1]Department of Electrical and Computer Engineering, Aarhus University, Aarhus, Denmark. [2]T&W Engineering, Lynge, Denmark. [3]Aarhus University Hospital, Aarhus, Denmark. ✉e-mail: mikkelsen.kaare@ece.au.dk

	EESM19	EESM23
	 Individualized ear piece, **dry** IrO electrodes	 Generic ear piece, **dry** IrO electrodes
Number of ear-Electrodes	6+6	2+3
PSG and ear-EEG	20 subjects, 4 nights each	10 subjects, 2 nights each
Manual scoring	80 nights (twice)	20 nights
Male / Female	7 / 13	6 / 4
Mean age, range (years)	25.9, 23-36	27.4, 22-35
ear-EEG only	10 subjects, 12 nights each	10 subjects, 10 nights each
ear-EEG Amplifier	TMSi Mobita	In-house device
PSG amplifier	TMSi Mobita	TMSi Mobita
Ground electrode pos.	Neck	In-ear
Actigraphy	Yes	—
Comfort questionnaire	Yes	Yes
ASSR recordings	Yes	—
Calibration recording	Yes	Yes

Table 1. Overview of the two data sets, with the different ear piece designs clearly visible at the top.

After all nights, the subjects filled out questionnaires regarding their sleep quality and the comfort of the equipment, and during all EESM19 nights, the participants also wore wrist-worn actigraphs.

Below is given additional detail about the steps mentioned above:

- **Polysomnography (PSG):** Also known as a 'partial PSG', consisting of left and right EOG, three chin EMG electrodes, F3, F4, C3, C4, O1, O2, M1 and M2, placed according to the AASM standard[7]. EEG amplifier depends on the study. PSG setups were mounted in a laboratory by the investigators. All PSG recordings were manually scored by trained sleep technicians. In the case of EESM19, each PSG recording was scored twice, by different technicians from different sleep centers. All manual scorings are included in the data sets. The sleep scorers were instructed to only score epochs for which they felt reasonably certain of the correct stage - in other words to not shy away from using the 'unscored' option. This resulted in 5% and 3% unscored epochs in EESM19 and and EESM23, respectively. Most unscored epochs are localized to a few recordings.
- **ear-EEG:** EEG electrodes mounted on ear pieces to achieve in-ear EEG, as described in Looney *et al*, Kidmose *et al.* and Mikkelsen *et al.*[1,5,8]. In both studies, ear electrodes were made from iridium oxide (dry electrodes, see Kappel *et al.*[9]), while the number of recording electrodes in each ear varied from 2 to 6. Depending on the study, either individualized ear pieces, designed based on ear impressions, or generic ear pieces based on a fixed design were used. In EESM19 and EESM23, there are recordings with and without simultaneous PSG. In EESM19, the PSG and ear-EEG electrodes were connected to the same amplifier (a TMSi Mobita). For these recordings, the equipment was mounted in a laboratory. Otherwise (so, for PSG-free recordings in EESM19 and all recordings in EESM23), the ear-EEG ear pieces were put in shortly before the subject went to bed, by the study participants themselves according to instructions given by the investigators previously. In EESM19, the ear-EEG only recordings included a single EOG electrode (EOG-r) as well.
- **Actigraphy:** Wrist-mounted actigraphy was recorded using 'GENEactiv' wrist bands from Activinsights. Besides triaxial actigraphy, the wristbands also recorded temperature and ambient light.
- **Comfort questionnaire:** In both studies, the subjects were asked to answer a short questionnaire in the morning, describing their experience sleeping with the ear-EEG. The questions in the questionnaire concerned the experienced sleep quality, whether the night's sleep was different from the subject's usual sleep and the comfort of the ear-EEG device.

All EEG recordings (both ear-EEG and polysomnography) were performed using an average referencing scheme, and have been saved in the same format.

In Fig. 1 is shown examples of a person wearing both PSG and ear-EEG recording equipment, as well as the ear-EEG equipment itself.

Additional details about the recording types can be found either by consulting the README-files within each data set, or the original papers[2,3].

Fig. 1 Examples of the equipment used in the two studies. (**A**) person wearing partial PSG with ear-EEG, from EESM19. (**B**) Person wearing only ear-EEG with a single EOG electrode, from EESM19. (**C**) Dry-contact custom ear-piece. Written consent was obtained from the person in the photos.

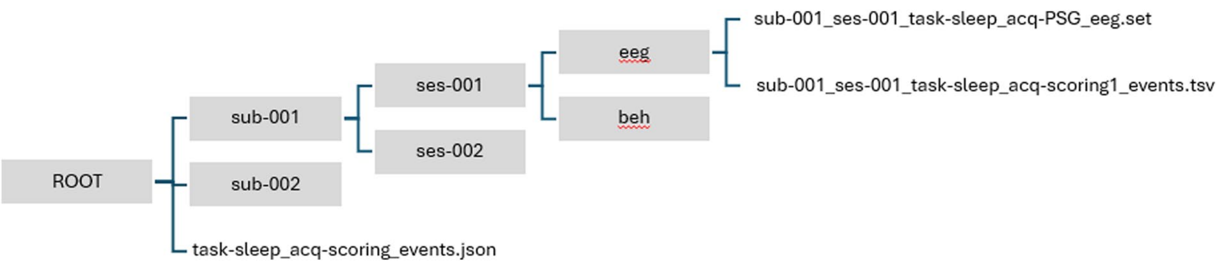

Fig. 2 Overview of how data sets are structured according to BIDS[13], taking EESM19 as an example. Subjects are top level, then sessions, then modalities. For EESM19, subject would go up to '020', and session can go up to '016'. json files for generic tsv types are placed on the top level, as shown.

All subjects gave written, informed consent. The studies were approved by the Central Denmark Region Committees on Biomedical Research Ethics as well as the Danish Medicines Agency. Case numbers are 1-10-72-413-17 and 2017111085 for EESM19, and 1-10-72-13-20 and 2020012619 for EESM23. Publication was not mentioned in the informed consent form. Because of this, prior to publication of the data, the local GDPR office of 'Region Midt' in Central Jutland considered the state of anonymization of the data set and judged it to be fully anonymized. Because of this, the GDPR office judged that publication of the data is not in violation of the GDPR.

Types of Recordings

The sessions in both studies are represented by multiple different recordings. The different types of recording are:

- **Calibration recordings:** To get an estimate for electrode quality, subjects watched a short video with various instructions for performing jaw clenches, opening and closing their eyes and performing horizontal eye movements. The video, with its instructions, is included in both EESM19 and EESM23 under 'stimuli'. An in-depth analysis of the responses, as recorded in EESM19, can be found in Mikkelsen *et al.*[10].
- **ASSR recording:** To obtain objective measures of electrode connection quality, 'auditory steady state responses' (ASSR) were recorded in the laboratory after EEG setups were mounted.
- **Sleep recording:** Sleep recordings took place in the subject's own home.

Study participants were primarily recruited among university students. No participants had known neurological or sleep disturbances.

Data Records

The two data sets are available at openneuro.org, with reference numbers 'ds005185' for EESM19[11] and 'ds005178' for EESM23[12].

All data sets have been formatted according to the Brain Imaging Data Set (BIDS) standard[13]. All EEG files are stored as EEGLAB .set files[14], all manual scoring is stored as .tsv event files, and all actigraphy recordings are stored as tsv.gz, in accordance with the BIDS standard. All .tsv files have accompanying .json files describibing each column. In following the standard, all files are sorted first into 'subject', then 'session' and finally 'modality' folders, and for redundancy, each file is named according to this convention. In Fig. 2 is shown a small slice

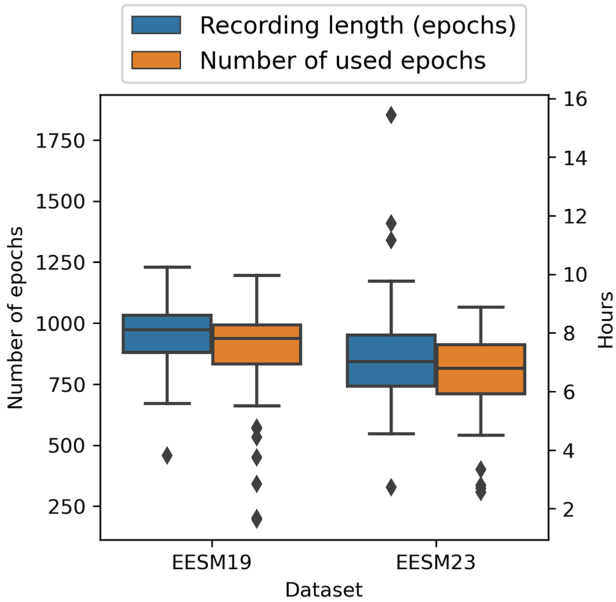

Fig. 3 Length of each scored recording (in 30 second epochs), as well as the number of epochs suitable for automatic scoring after data cleaning.

through EESM19, showing the location of the PSG recording for subject 001, session 001, together with one of the manual scorings. We recommend looking into the BIDS standard[13] for the full details.

Technical Validation

Methods. While many different analyses are possible with this data, we here focus on its validity for automatic sleep scoring, in particular using the ear-EEG data together with the manual scorings. In the following we demonstrate a full analysis pipeline, starting from the raw data as shared, finishing in a sleep scoring analysis. The code to reproduce the analysis, with comments, can be found at our dedicated gitlab repository[15].

There are two relevant steps in this processing: (1) data cleaning and (2) the sleep scoring itself. Data cleaning was done with in-house developed Python codes, performing a mixture of linear filtering and nonlinear detection and removal of transient artifacts. Largely, this cleaning pipeline is a standardization and translation (from Matlab to Python) of the procedures which were used in Mikkelsen *et al.*[2] and Tabar *et al.*[3]. It may here be important to note that the artifact detection was thoroughly inspected and controlled during the initial studies, but also that this processing pipeline is being applied 'after' data publication. Meaning that it is very easy for users of the data set to experiment with other preprocessing methods, completely replacing our suggested approach (indeed, we encourage people to investigate this).

Using the cleaned data, spectrograms are calculated and fed to a version of SeqSleepNet[16] pretrained on the SHHS dataset[17,18], using similar settings to what was used in Mikkelsen *et al.*[6]. This pretrained version is applied to the two data sets in the following manner: The SHHS-version of the network was first finetuned on EESM19 in a leave-one-subject out manner. Separately, a different version of the pretrained network was first fine-tuned on all subjects in EESM19, and then was further finetuned to EESM23, which is significantly smaller than EESM19 in terms of labeled recordings.

After training the models, we compare the outputs from both of the finetuned models to the manual scorings for the same data, by calculating Cohen's kappa[19] for each comparison. For comparison, we applied the 'EESM23' approach to the already published dataset EESM17[20,21] and included the results in the same figure. Epochs marked as 'unscored' by the manual scorer were disregarded in this analysis.

Note that we do not delve into the similar sleep scoring of the PSG recordings, but a very similar flow would work there, and most likely yield even better performance, as was shown in Mikkelsen *et al.*[6].

Results. In Fig. 3 is shown the distributions of recording sizes for the nightly recordings for both data sets. Distributions are shown for both before and after cleaning of the data. We see that the before / after distributions are quite similar for each data set, indicating a generally high data quality.

In Fig. 4 is shown the distributions of Cohen's kappa values. We see that all values are comfortably above the chance level (0), but we also see a marked difference between the data sets. EESM19 is the largest of the 3 data sets, and thus it is not entirely surprising that it is also the best performing. However, we note that the obtained performance for EESM23 is somewhat less than what was reported in Tabar *et al.*[3]. We have not been able to isolate the reason for this discrepancy, but note that the low average kappa in particular is driven by two recordings with kappa values below 0.4. We suspect that this may be related to the translation of the data cleaning pipeline from Matlab to Python. We hope future work with these open data sets can shed some light on this.

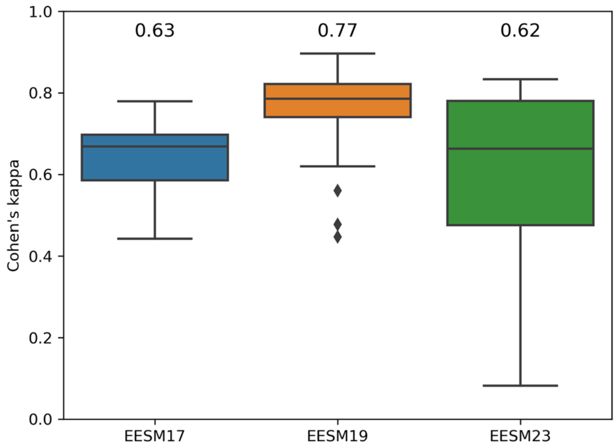

Fig. 4 Boxplots showing kappa distributions after fine tuning for the three data sets (EESM17 included for comparison). Above each box is written the average kappa. Note that chance level for the Cohen's kappa value is 0, and for healthy, young subjects, the interrater agreement using PSG data is around 0.8. Mikkelsen *et al.*[6] found it to be 0.82 for EESM19.

Usage Notes

Both data sets have been reformatted according to the BIDS standard[13]. This means that data is easily accessed using standard software and packages. In the accompanying analysis pipeline, all preprocessing is done with a combination of MNE[22], MNE-bids[23], numpy[24] and scipy[25].

All EEG data is saved in the .set format of the MATLAB package EEGLAB, making it very easy to work with from Matlab as well. In line with the BIDS specification, each data set has a detailed README file describing all peculiarities of the data set.

In addition to the sleep analysis pipeline, we have also included the code used for translating from source data to BIDS format, as well as the unit tests performed on the BIDS version. As there are certain peculiarities related to the specific hardware setups, which must be properly taken care of before using the data, we strongly recommend using the BIDS version of the dataset.

For completeness, we should point out that this pipeline is designed for the sleep data. For other types of data, such as the ASSR data or the calibration recordings, we suggest processing such as used in Kappel *et al.*[26] or Mikkelsen *et al.*[10], respectively.

The data has been used as benchmark data sets for developing automatic sleep scoring models[27,28], as well as simply investigating the information content of ear-EEG sleep recordings[29,30].

Additionally, these data can be used to investigate 'inter-intra' subject variation, both in simple evoked stimuli[10] and sleep features[31]. It may also be useful for investigating sleep 'fingerprints', both to monitor individual changes over time[32], as well as investigate methods to increase individual sleep scoring performance[33].

It may also be possible to use the data to investigate how to use calibration data for real-time updating of EEG models:in both datasets there are separate 'calibration' recordings, part of which made the basis for the analysis in Mikkelsen *et al.*[10]. These may correlate with individual variations among the participants.

Code availability

A complete processing pipeline for the sleep data, in Python, is found under 'code' in each data set.

Received: 6 June 2024; Accepted: 31 January 2025;
Published online: 19 February 2025

References

1. Mikkelsen, K. B., Kappel, S. L., Mandic, D. P. & Kidmose, P. EEG Recorded from the Ear: Characterizing the Ear-EEG Method. *Frontiers in Neuroscience* **9**, https://doi.org/10.3389/fnins.2015.00438 (2015).
2. Mikkelsen, K. B. *et al.* Accurate whole-night sleep monitoring with dry-contact ear-EEG. *Scientific Reports* **9**, 1–12, https://doi.org/10.1038/s41598-019-53115-3 (2019).
3. Tabar, Y. R. *et al.* At-home sleep monitoring using generic ear-EEG. *Frontiers in Neuroscience* **17** (2023).
4. Mikkelsen, K. B. *et al.* Self-applied ear-EEG for sleep monitoring at home. In *2022 44th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, 3135-3138, https://doi.org/10.1109/EMBC48229.2022.9871076, ISSN:2694-0604 (2022).
5. Looney, D. *et al.* The in-the-ear recording concept: User-centered and wearable brain monitoring. *IEEE Pulse* **3**, 32–42 (2012).
6. Mikkelsen, K. *et al.* Sleep monitoring using ear-centered setups: Investigating the influence from electrode configurations. IEEE transactions on bio-medical engineering PP, https://doi.org/10.1109/TBME.2021.3116274 (2021).
7. Berry, R. B. *et al.* AASM scoring manual updates for 2017 (version 2.4). *Journal of Clinical Sleep Medicine* **13**, 665–666, https://doi.org/10.5664/jcsm.6576 (2017).
8. Kidmose, P., Looney, D., Ungstrup, M., Lind, M. & Mandic, D. P. A study of evoked potentials from ear-EEG. *IEEE Transactions on Biomedical Engineering* **60**, 2824–2830 (2013).
9. Kappel, S. L., Rank, M. L., Toft, H. O., Andersen, M. & Kidmose, P. Dry-Contact Electrode Ear-EEG. *IEEE Transactions on Biomedical Engineering* **66**, 150–158, https://doi.org/10.1109/TBME.2018.2835778 (2019).

10. Mikkelsen, K. B., Tabar, Y. R., Christensen, C. B. & Kidmose, P. EEGs Vary Less Between Lab and Home Locations Than They Do Between People. *Frontiers in Computational Neuroscience* **15**, https://doi.org/10.3389/fncom.2021.565244 (2021).
11. Mikkelsen, K. B. *et al.* Ear-EEG Sleep Monitoring 2019 (EESM19) - OpenNeuro, https://doi.org/10.18112/openneuro.ds005185.v1.0.0.
12. Rezai Tabar, Y. *et al.* Ear-EEG Sleep Monitoring 2023 (EESM23) - OpenNeuro, https://doi.org/10.18112/openneuro.ds005178.v1.0.0.
13. Gorgolewski, K. J. *et al*. The brain imaging data structure, a format for organizing and describing outputs of neuroimaging experiments. *Scientific Data* **3**, 160044, https://doi.org/10.1038/sdata.2016.44 (2016).
14. Delorme, A. & Makeig, S. EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods* **134**, 9–21, https://doi.org/10.1016/j.jneumeth.2003.10.009 (2004).
15. AU207178. Eesm quality demo. Accessed: 2024-11-26 (2024).
16. Phan, H., Andreotti, F., Cooray, N., Chén, O. Y. & Vos, M. D. SeqSleepNet: End-to-End Hierarchical Recurrent Neural Network for Sequence-to-Sequence Automatic Sleep Staging. *IEEE Transactions on Neural Systems and Rehabilitation Engineering* **27**, 400–410, https://doi.org/10.1109/TNSRE.2019.2896659 (2019).
17. Zhang, G.-Q. *et al*. The National Sleep Research Resource: towards a sleep data commons. *Journal of the American Medical Informatics Association: JAMIA* **25**, 1351–1358, https://doi.org/10.1093/jamia/ocy064 (2018).
18. Quan, S. F. *et al*. The Sleep Heart Health Study: design, rationale, and methods. *Sleep* **20**, 1077–1085 (1997).
19. Cohen, J. A Coefficient of Agreement for Nominal Scales. *Educational and Psychological Measurement* **20**, 37–46, https://doi.org/10.1177/001316446002000104 (1960).
20. Mikkelsen, K. B., Villadsen, D. B., Otto, M. & Kidmose, P. Automatic sleep staging using ear-EEG. *Biomedical Engineering Online* **16** (2017).
21. Mikkelsen, K. B., Villadsen, D. B., Birch, L., Otto, M. & Kidmose, P. Ear-EEG Sleep Monitoring 2017 (EESM17), https://doi.org/10.18112/OPENNEURO.DS004348.V1.0.4 (2024).
22. Gramfort, A. *et al.* MEG and EEG data analysis with MNE-Python. *Frontiers in Neuroscience* **7** (2013).
23. Appelhoff, S. *et al*. MNE-BIDS: Organizing electrophysiological data into the BIDS format and facilitating their analysis. *Journal of Open Source Software* **4**, 1896, https://doi.org/10.21105/joss.01896 (2019).
24. Harris, C. R. *et al*. Array programming with NumPy. *Nature* **585**, 357–362, https://doi.org/10.1038/s41586-020-2649-2 (2020).
25. Virtanen, P. *et al*. SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nature Methods* **17**, 261–272, https://doi.org/10.1038/s41592-019-0686-2 (2020).
26. Kappel, S. L., Christensen, C. B., Mikkelsen, K. B. & Kidmose, P. Reference configurations for ear-EEG steady-state responses. *Conference proceedings: Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Annual Conference.* **2016**, 5689–5692, https://doi.org/10.1109/EMBC.2016.7592018 (2016).
27. Phan, H. *et al*. SleepTransformer: Automatic Sleep Staging With Interpretability and Uncertainty Quantification. *IEEE Transactions on Biomedical Engineering* **69**, 2456–2467, https://doi.org/10.1109/TBME.2022.3147187 (2022).
28. Phan, H. *et al.* L-SeqSleepNet: Whole-cycle Long Sequence Modelling for Automatic Sleep Staging. *IEEE Journal of Biomedical and Health Informatics1*. **10**, https://doi.org/10.1109/JBHI.2023.3303197 (2023).
29. Tabar, Y. R. *et al.* Ear-EEG for sleep assessment: a comparison with actigraphy and PSG. *Sleep and Breathing* https://doi.org/10.1007/s11325-020-02248-1 (2020).
30. Tabar, Y. R., Mikkelsen, K. B., Rank, M. L., Hemmsen, M. C. & Kidmose, P. Muscle activity detection during sleep by ear-EEG. In *2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, 1007–1010 (IEEE, 2020).
31. Kjaer, T. W., Rank, M. L., Hemmsen, M. C., Kidmose, P. & Mikkelsen, K. Repeated automatic sleep scoring based on ear-EEG is a valuable alternative to manually scored polysomnography. *PLOS Digital Health* **1**, e0000134, https://doi.org/10.1371/journal.pdig.0000134 (2022).
32. Hemmsen, M., Mikkelsen, K., Rank, M. & Kidmose, P. Long-term monitoring of trait-like characteristics of the sleep electroencephalogram using ear-EEG. *Sleep* **44**, A109–A109, https://doi.org/10.1093/sleep/zsab072.271 (2021).
33. Borup, K., Kidmose, P., Phan, H. & Mikkelsen, K. Automatic sleep scoring using patient-specific ensemble models and knowledge distillation for ear-EEG data. *Biomedical Signal Processing and Control* **81**, 104496, https://doi.org/10.1016/j.bspc.2022.104496 (2023).

Acknowledgements

Portions of this work were sponsored by the Innovation Fund Denmark, grant 7050-00007.

Author contributions

K.B.M. took part in design of both experiments, carried out all data collection related to EESM19, took part in designing preprocessing pipelines for both data sets, wrote this manuscript and performed BIDS translation for EESM19. Y.R.T. took part in designing EESM23, carried out all data collection related to EESM23 and took part in designing the preprocessing pipeline for EESM23. L.R.B. performed BIDS translation for EESM23. S.L.K. took part in designing and maintaining equipment used for EESM19 and EESM23. C.B.C. designed code and stimuli for the ASSR recordings in EESM19. L.D.M. designed part of the preprocessing pipeline for EESM23. M.O. was responsible for manual sleep scoring of all P.S.G. recordings. M.C.H. and M.L.R. took part in the design of EESM19 and EESM23. P.K. took part in the design of both experiments, produced the equipment, and was involved in the preprocessing of both data sets. All authors reviewed the manuscript.

Competing interests

LDM, MCH and MLR are or were employed in either T&W Engineering or the sister company Widex A/S. These companies are involved in the continued development of ear-EEG. The remaining authors report no conflicts of interest.

Additional information

Correspondence and requests for materials should be addressed to K.B.M.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.