The Chinese Clinical Sleep Database: An Innovative Database System Includes Large-Scale Clinical Data of Chinese Population

Ruichen Fang[1,2], Yihong Cheng[1,2], Fan Li[3], Yan Xu[1,2], Yuanhui Li[4], Xiang Liu[4], Simin Guo [iD][4], Yuling Wang[1,2], Jinnong Jiang[1,2], Dan Zhou[1,2], Bin Zhang [iD][1,2]

[1]Department of Psychiatry, Sleep Medicine Center, Nanfang Hospital, Southern Medical University, Guangzhou, People's Republic of China; [2]Key Laboratory of Mental Health of the Ministry of Education, Guangzhou, People's Republic of China; [3]School of Biomedical Engineering, Faculty of Medicine, Dalian University of Technology, Dalian, 116024, People's Republic of China; [4]Adai Technology (Beijing) Co., Ltd, Beijing, People's Republic of China

Correspondence: Bin Zhang, Department of Psychiatry, Sleep Medicine Center, Nanfang Hospital, No. 1838 North, Guangzhou Avenue, Guangzhou, 510515, People's Republic of China, Email zhang73bin@hotmail.com

Purpose: In this study, we established the Chinese Clinical Sleep Database (CCSD), aiming to provide a safe, scalable, and user-friendly database that includes high-quality clinical data from Chinese population to facilitate sleep research.

Material and Methods: We collect individual's demographic data, scales, anthropometric measurements, clinical diagnosis, and polysomnography (PSG) recordings from the routine medical process of sleep medicine centers using standardized procedures. The distributed cluster storage technology are utilized to store these data. The structured data are stored in a high-performance MySQL database, while the unstructured data are stored in an object storage service. And we have developed an online data platform to share and manage our data.

Results: The data collection has been conducted in three hospitals. In the preliminary stage of data collection (from October 18, 2022 to September 4, 2023), our database included a total of 1183 patients. Among them, 56.8% were male and their ages ranged from 3 to 88 years. These patients were diagnosed with various types of sleep disorders.

Conclusion: Since the CCSD's inception, it has demonstrated good stability, security, and scalability. As an public database, the CCSD also exhibits user-friendliness. The CCSD contains comprehensive clinical data, which can contribute to the advancement of the diagnosis and treatment strategies for sleep disorders, ultimately promoting sleep health.

Keywords: sleep medicine, methodology, database, data collection, collaboration tool

Introduction

Humans spend about one-third of their lives sleeping, and sleep is essential for maintaining good mental and physical health. Unhealthy sleep has been shown to be associated with a variety of health conditions, including obesity,[1] cardiovascular disease,[2] affective disorders,[3] and neurological disorders.[4] With the development of the economy and society, the incidence of sleep disorders increases. According to the survey data from the Chinese Sleep Research Association in 2021, there were over 300 million people in China who suffered from sleep disorders and related issues.[5] In light of the growing demand for better sleep health, it is essential to conduct in-depth research on sleep. Analyzing and mining large clinical data samples can help advance diagnosis and treatment strategies for sleep disorders and provide information for making public health policies.[6,7] In China, despite the gradual accumulation of clinical data in various sleep medicine centers, it has become evident that there are issues such as uneven data quality, diverse data formats, and inconvenient data management among different sleep medicine centers.[8] Establishing a standardized procedure for data collection and storage can enhance collaboration among different centers, thereby making progress in sleep research.

In order to provide resources for sleep research, public sleep databases, such as Sleep-EDF, ISRUC-sleep, and Sleep Heart Health Study (SHHS), have been established. Public sleep databases contain a vast amount of biophysiological signals and disease-related data. Researchers can use these data to perform secondary analyses. For example, by analyzing data from 1850 participants in the SHHS, Javaheri et al found that a decrease in N3 stage sleep was correlated to a higher prevalence of hypertension in both male and female individuals.[9] Azarbarzin et al also utilized data from SHHS and found that the hypoxic burden of sleep apnea could predict cardiovascular-related mortality.[10] However, the current public sleep databases mainly contain data from European and American populations. There are differences among different populations in terms of sleep structure, sleep quality, and prevalence of sleep disorders.[11] In order to understand the characteristics of sleep disorders and sleep-related diseases in Chinese population, it is crucial to establish a public sleep database for Chinese population.

In addition, data from public databases are commonly used for training automatic sleep analysis models. For example, Li et al used single-channel electroencephalogram (EEG) signals from SHHS and Sleep-EDF to train and validate an end-to-end deep learning-based automatic sleep staging model.[12] Prabha et al extracted features from single-channel abdominal respiratory effort signals sourced from the ISRUC-sleep database and constructed a model to automatically detect and classify sleep apnea events.[13] Because the detailed information about data collection and participants in the public databases is open to investigators, public databases can serve as benchmarks for comparing differences in algorithm performance.[14] However, due to the lack of validation with a large number of clinical populations, existing automatic sleep analysis models are rarely used in clinical practice.[15] Establishing a database with a large number of clinical samples may help developing models that meet clinical practice requirements and have high reliability.[16,17]

In this study, in collaboration with Adai Technology (Beijing) Ltd., Co, Beijing, China, we established a public sleep database - The Chinese Clinical Sleep Database (CCSD, psg.wangli-tech.com),[18] aiming to continuously collect and store the clinical data from routine clinical care and health monitoring processes through standardized procedures.

Materials and Methods
Population
The population of CCSD are the individuals presented to the sleep medicine centers. The inclusion criteria are as follows: (1) The individual undergo PSG monitoring (2) No gender or age-based limitations. Before PSG monitoring, individuals are required to sign an informed consent form. The content of the form includes informed consent for PSG, informed consent for synchronous video recording during PSG, and informed consent for making all clinical data (excluding personal private information such as name and phone number) publicly available for research purposes. If there is an occurrence of signal channel detachment or unforeseen interruption during PSG recording, the individual's data will not be included in the database. To ensure the privacy of each individual, the names and contact information are encrypted in the database. The study was approved by the Medical Ethics Committee of Nanfang Hospital, Southern Medical University (NFEC-2023-371).

The Procedure of Data Collection
In order to continuously obtain clinical data from sleep medicine centers, we have designed a set of workflow procedures. As shown in Figure 1, the clinicians make preliminary diagnoses, and arrange PSG for individuals. The nurses assist patients in completing questionnaires and measure their height, weight, neck circumference, and vital sighs. The sleep technicians conduct PSG recordings and score sleep stages and events adhering to the guidelines set forth by the American Academy of Sleep Medicine (AASM).[19] The clinicians then review the results of examinations and refine their diagnoses. The diagnosis of sleep disorders are made based on the International Classification of Sleep Disorders, third edition (ICSD-3).[20] The comorbidities of the patient, such as anxiety, depression, hypertension are also recorded. In the end, the clinicians upload the individual's diagnosis, questionnaire evaluation results, anthropometric measurements, and PSG report to the database. The biophysiological signal recordings, annotation files, audios, and videos of PSG are uploaded by a sleep technician.
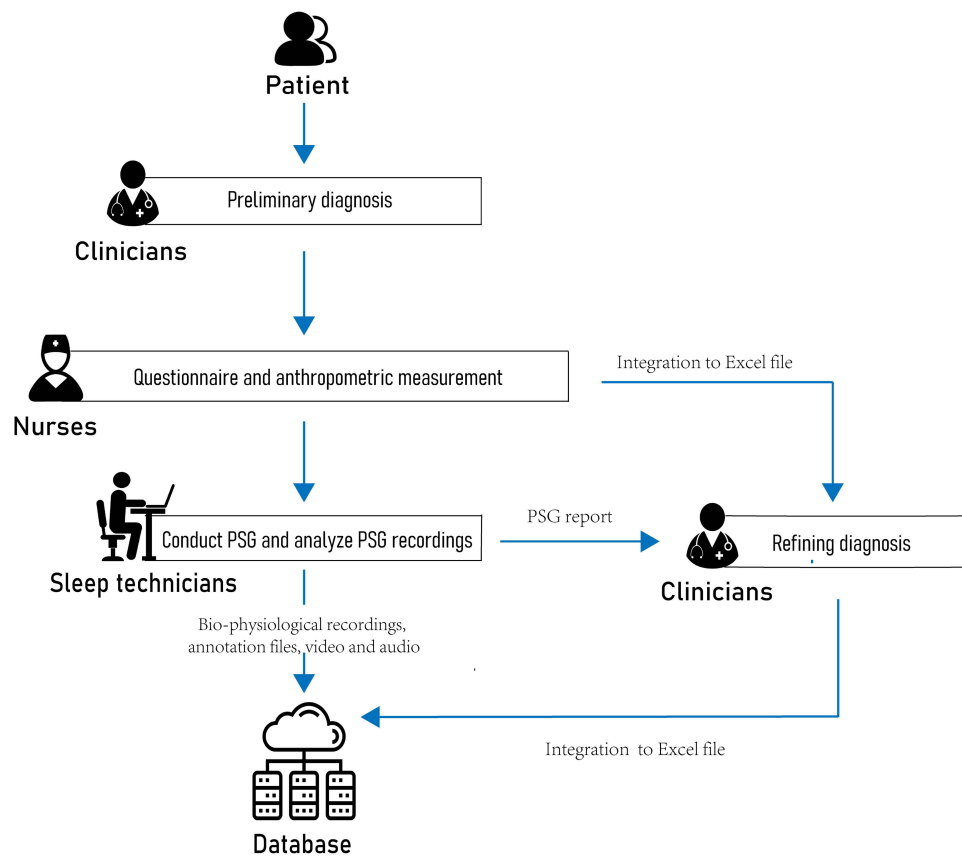
Figure 1 Workflow for continuous clinical data collection in sleep medicine centers.
Abbreviation: PSG, polysomnography.

Measurements

Questionnaires

A structured questionnaire is used to collect demographic data, including age, gender, ethnicity, residential address, education, marital status, occupation, employment status, frequency of night shift work, and medical history. The questionnaire also includes scales for evaluating various aspects of the individual's sleep and emotional well-being: Pittsburgh Sleep Quality Index (PSQI), Morning and Evening Questionnaire-5 (MEQ-5), Epworth Sleepiness Scale (ESS), and Hospital Anxiety and Depression Scale (HADS).[21–24] These scales provide information about the patient's sleep quality, chronotype, daytime sleepiness, as well as the levels of anxiety and depression.

Polysomnography

PSG recordings can be divided into diagnostic PSG, positive airway pressure (PAP) titration sleep study, and multiple sleep latency test (MSLT). Diagnostic PSG is used to diagnose or exclude sleep disorders. PAP titration sleep study is designed to determine the optimal pressure of PAP devices required to maintain the upper airway opening for obstructive sleep apnea (OSA) patients.[25] MSLT is conducted during the daytime after a diagnostic PSG, employing four or five naps to determine daytime sleepiness. It frequently serves as an authoritative examination for narcolepsy.[26] According to different diagnostic and therapeutic purposes, at least one of the abovementioned recordings is performed for an individual.

The channels and sampling rates of different PSG recordings are shown in Table 1. Diagnostic PSG consists of 6 channels of EEG, 2 channels of electrooculography (EOG), electromyography (EMG) of the mentalis muscle, electrocardiography (ECG), respiratory flow, nasal pressure, respiratory effort, oxygen saturation, snoring, body position, EMG of tibialis anterior muscle, and a time-synchronized video registration. The signal channels of the PAP titration sleep study are the same as those in diagnostic PSG, except that the respiratory airflow is measured using a PAP device.[27] The MSLT only includes 3 channels of EEG, 2 channels of EOG, chin EMG, and ECG.[26]

Table 1 Montage and Sampling Rate of Diagnostic PSG, PAP Titration, and MSLT

	Diagnostic PSG	**PAP Titration**	**MSLT**
EEG	6 channels: F4-M1, C4-M1, O2-M1, F3-M2, C3-M2 and O1-M2; Sampling rate: ≥ 500 Hz.	Same as the Diagnostic PSG; Sampling rate: ≥ 500 Hz.	3 channels: F4-M1 or F3-M2, C4-M1 or C3-M2, O2-M1 or O1-M2; Sampling rate:≥ 500 Hz.
EOG	2 channels: E1-M2 (E1:1 cm below and 1 cm lateral to the left outer canthus), E2-M2 (E2: 1 cm above and 1 cm lateral to the right outer canthus); Sampling rate: ≥ 500 Hz.	Same as the Diagnostic PSG; Sampling rate: ≥500 Hz.	Same as the Diagnostic PSG; Sampling rate: ≥ 500 Hz.
EMG	3 Chin EMG channels, bilateral anterior tibial EMG; Sampling rate: ≥ 500 Hz.	Same as the Diagnostic PSG; Sampling rate: ≥ 500 Hz.	3 Chin EMG channels; Sampling rates:≥ 500 Hz.
ECG	Sampling rate:≥ 500 Hz.	Sampling rate: ≥ 500 Hz.	Sampling rate: ≥ 500 Hz.
Respiratory flow	Nasal pressure transducer and oronasal thermal flow sensor; Sampling rate ≥ 100 Hz.	PAP devices flow; Sampling rate:≥ 100 Hz.	NA
Respiratory effort	Respiratory inductive plethysmography (RIP); Sampling rate ≥ 100 Hz.	Same as Diagnostic PSG; Sampling rate: ≥500 Hz.	NA
Oxygen saturation	Pulse oximetry; Sampling rate ≥ 25 Hz.	Same as Diagnostic PSG.	NA
Snoring	Condenser mic; Sampling rate ≥ 500 Hz.	Same as the Diagnostic PSG.	NA
Body position	Gravity-detecting electric sensor; Sampling rate: ≥ 1 Hz.	Same as the Diagnostic PSG.	NA
Infrared video	Camera and infrared light source.	Same as the Diagnostic PSG.	Same as the Diagnostic PSG.

Abbreviations: EEG, electroencephalography; EMG, electromyogram; EOG, electrooculogram; ECG, Electrocardiogram; NA, not applicable; PSG, polysomnography; PAP, positive airway pressure; MSLT, multiple sleep latency test.

Sleep and Related Events Scoring

Sleep scoring contains various tasks, including sleep staging, identification of arousal events, respiratory events, limb movements, and other sleep-related events. The manual scoring process encompasses sleep stages and respiratory events. Other sleep-related events are automatically identified by the data acquisition software and later reviewed and corrected by a sleep technician.

Sleep staging involves dividing an overnight into epochs of 30 seconds each and assigning each epoch to one of the five stages: non-rapid eye movement (NREM) sleep (including N1, N2, and N3), rapid eye movement (REM) sleep, and wake. A 5-minute time window is typically used to identify and classify respiratory events, including sleep apnea and hypopnea. Apnea is defined as a drop in the peak airflow signal of at least 90% of the baseline for at least 10 seconds. Hypopnea is defined as a reduction in the amplitude of the airflow signal by at least 30% of the baseline for a period longer than 10 seconds, accompanied by a decrease in blood oxygen saturation of ≥3% or arousal, or with a decrease in blood oxygen saturation of at least 4%.

Parameters of PSG Reports

The diagnostic PSG report includes parameters for sleep scoring, arousal events, respiratory events, sleep stage-related respiratory events, positional respiratory events, limb movement, and cardiac events, as shown in Table 2. The PAP titration sleep study report includes the same parameters as the diagnostic PSG report, with additional information about the PAP device used for titration, including brand, model, mode, initial pressure, 90% pressure, and leak rate. The MSLT report includes the sleep onset latency and REM onset latency for each nap, and the average sleep onset latency and REM onset latency for the five naps.

Table 2 Parameters Included in the Diagnostic PSG Report

Sleep Stage
- Lighting on/off time
- Total sleep time, Total recording time, Time in Bed
- Sleep onset latency, REM onset latency
- Time of wake after sleep onset, number of wake stages
- Sleep efficiency
- The duration of N1, N2, N3 and REM sleep
- The percentage of N1, N2, N3 and REM sleep in total sleep time

Arousal events
- Number and index of arousals during REM, NREM sleep and overnight sleep
- Number and index of respiratory event-related arousals during REM, NREM sleep and overnight sleep
- Number and index of limb movement arousals during REM, NREM sleep and overnight sleep
- Number and index of spontaneous arousal during REM, NREM sleep and overnight sleep

Respiratory events
- Average duration, number, and index of obstructive, mixed, and central sleep apnea
- Average duration, number, and index of sleep apnea
- Average duration, number, and index of hypopnea
- Apnea-hypopnea index
- Number and index of respiratory effort-related arousal
- Respiratory disturbance index
- Mean and minimum arterial oxygen saturation of sleep
- Number and index of oxygen desaturation of ≥3%

Respiratory events during REM and NREM sleep
- Respiratory events parameters during REM sleep
- Respiratory events parameters during NREM sleep

Respiratory events in various positions
- Respiratory events parameters in the supine position
- Respiratory events parameters in the left decubitus position
- Respiratory events parameters in the Right decubitus position

Limb movement events
- Number and Index of periodic limb movement during overnight sleep
- Number and Index of periodic limb movement during REM and NREM sleep

Cardiac events
- Average, lowest, and highest heart rate during sleep
- Number of bradycardia or tachycardia episodes
- Number of narrow or wide complex tachycardia episodes
- Number and longest duration of atrial fibrillation episodes

Abbreviations: PSG, polysomnography; REM, rapid eye movement; NREM, non-rapid eye movement.

Database Storage Architecture

The database employs a distributed cluster storage technology for its storage architecture (Figure 2). Structured data (questionnaires, anthropometric measurements, PSG report, etc.) are stored in the MySQL database. The unstructured data (biophysiological recordings, annotation files, videos, audios, etc.) are stored in an object storage service (OSS). The MySQL database storage has the characteristics of high performance, high availability, and easy expansion, which can construct efficient data structures and realize the storage and query of a large number of information. The OSS is a big data storage technology suitable for storing large files. It features automatically scalable storage space that can expand as the amount of data increases.
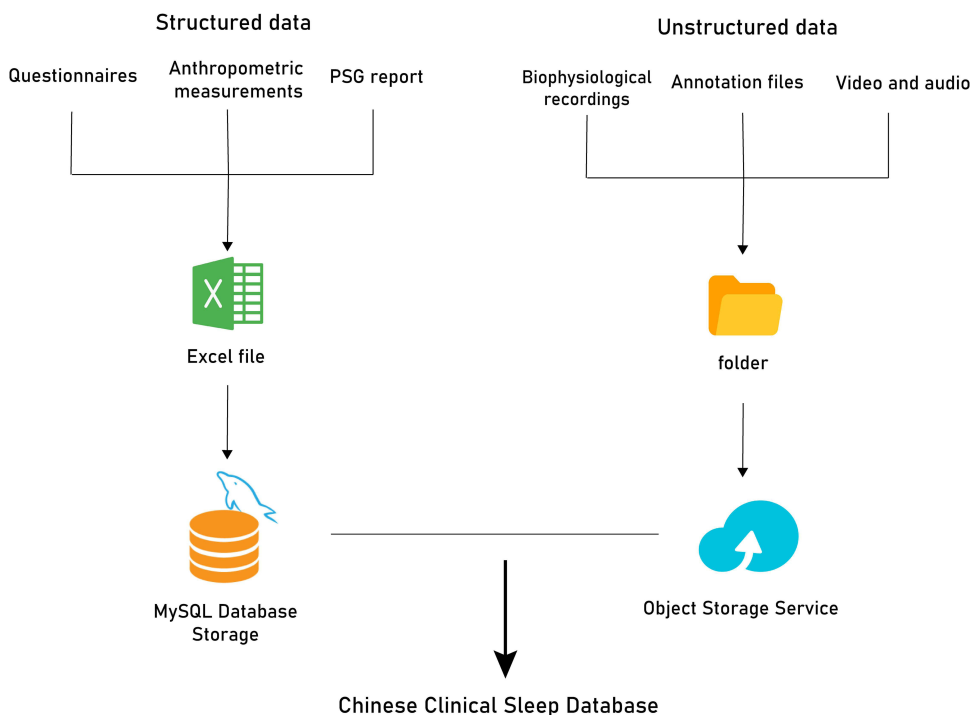
Figure 2 Storage architecture of the CCSD. MySQL database are designed for the storage of structured data, whereas object storage service is specifically for unstructured data. **Abbreviation**: PSG, polysomnography.

Data Platform

The data in our database can be accessed through a dedicated data platform. Users need to authenticate when they log on to the platform with two-factor authentication (2FA) in order to ensure the security of the data. Users can upload and download data through the platform. We use multithreading and resumable transfer technologies to improve the speed and stability of data transmission. The data platform offers fuzzy retrieval and keyword retrieval capabilities to help users quickly find the information they need.

Result

Currently, the CCSD is collecting data from Southern Medical University Nanfang Hospital, Taihe Hospital in Shiyan City, and Guizhou Second People's Hospital. In the preliminary stage of our data collection (from October 18, 2022 to September 4, 2023), a total of 1183 patients were included, with 56.8% of them being males. The age was 42.50 ± 15.44 years (mean ± SD), ranging from 3 to 88 years.

There were 1535 PSG recordings in our database, including 1214 diagnostic PSG recordings, 293 PAP titration sleep recordings and 28 MSLT recordings. For the diagnostic recordings, 983 patients had recordings for only one night, and 101 patients had recordings for 2 or 3 nights. For the PAP titration sleep study, 194 patients had undergone diagnostic PSG the night before. All patients involved in the MSLT had completed a diagnostic PSG on the previous night. Table 3 lists the detailed number of these recordings.

Table 3 PSG Recordings in the CCSD

Type of PSG	Number of Recordings	Details of Patients
Diagnostic PSG	1214	983 patients had recordings for one night; 101 had recordings for 2 or more nights
PAP titration sleep study	293	194 patients had a diagnostic PSG recording from the previous night
MSLT	28	All patients had a diagnostic PSG recording from the previous night

Abbreviations: PSG, polysomnography; PAP, positive airway pressure; MSLT, multiple sleep latency test; CCSD, Chinese Clinical Sleep Database.

Table 4 The Types of Sleep Disorders Covered in the CCSD

Diagnosis	Number of Cases	Percentage (%)
Sleep breath disorders	548	47.5
Insomnia	276	25.7
Sleep-related movement disorders	27	1.7
Narcolepsy	16	1.5
Parasomnias	12	1.0
Other sleep disorders	210	15.0
Comorbid sleep disorders	76	6.0
Normal population	18	1.7

Abbreviation: CCSD, Chinese Clinical Sleep Database.

Table 4 shows the types of sleep disorders that are currently included in our database. Sleep disorders are classified into seven major categories according to ICSD-3.[20] Our database includes sleep breath disorders, insomnia, sleep-related movement disorders, central disorder of hypersomnolence, parasomnias, and other types of sleep disorders.

Discussion

The CCSD is the first sleep database for Chinese population. We continuously collect clinical data from multiple centers using a standardized data collection method. We have successfully included over 1000 clinical samples, and the data in the database are still growing. Our plan for the future is to expand the number of cases to ten thousands in two years.

With innovative data storage design and efficient data management platform, our database exhibits the following characteristics: 1) Accessibility and usability: The data platform includes user-friendly web interface features for data entry, retrieval, management, sharing, and reporting. 2) Privacy and Security: The database uses state-of-The-art authentication and encryption systems to ensure privacy and security. The database includes a robust disaster recovery system, ensuring that the data is protected against unforeseen circumstances. Regular backups and redundant systems are in place to provide a fail-safe against data loss. 3) Scalability: The use of distributed cluster storage technology can help to store structured and unstructured data, ensuring a high availability and reliability of the data. The database is scalable to accommodate increasing volumes of data as more patients are included.

The CCSD is pooled from sleep medicine centers in grade-A tertiary hospitals, which have a wide range of patient sources. By combining data collection with the routine clinical work of these sleep medicine centers, our database has included various sleep disorders. As the number of patients increases, the data scale of our database will continue to expand. Our database can provide valuable resources for research on rare sleep disorders such as narcolepsy; it can be also used for research on comorbid sleep disorders and sleep disorders combined with other systemic diseases.

When collecting data for the CCSD, each sleep medicine center followed a standardized procedure and recorded PSG according to the guidance of the AASM. The standardized data collection procedure not only guarantees the integrity and reliability of the data, but also ensures comparability and consistency across different centers.[28] In addition, adopting a unified data format also helped strengthening the cooperation among various sleep medicine centers and promoting the data sharing and technological exchange. With the advancement of database construction, we hope to establish a nationwide sleep research system to delve into the sleep issues faced by different regions, age groups, and populations with different lifestyles. This will provide stronger support for improving people's sleep health.[8]

The utilization of distributed cluster storage technology enabled the inclusion of diverse unstructured data in our database, such as videos and audios. Video and audio recordings captured during sleep are necessary for diagnosing some sleep disorders, such as REM sleep behavior disorder (RBD). RBD is a type of parasomnia that is characterized by REM sleep without atonia (RWA), accompanied by dream-related speech and/or complex motor behaviors.[29] Videos and PSG recordings are the gold standard for diagnosing RBD.[30] Some sleep-related movement disorders, such as sleep bruxism, also require video and audio assistance for diagnosis.[31] Our database can provide resources for the research of these diseases.

PAP titration sleep study and MSLT are also common examinations in sleep medicine centers. However, the two types of PSG recordings are not currently available in the existing public database. Our database include both PAP titration sleep study

and MSLT recordings. The PAP titration study can be used to investigate the efficacy of continuous positive airway pressure therapy, and the MSLT recordings can be used to explore biomarkers for excessive daytime sleepiness.[32,33]

There are still some limitations in our current work. Firstly, due to the impact of the COVID-19 epidemic, the data collection in the early stage has been limited, resulting in a sample size smaller than we expected. Secondly, despite our strict control over the data collection procedure, there are still some individuals with missing data due to technical issues or lack of cooperation from the individuals. Lastly, it is currently not possible to directly conduct statistical analysis on the data platform.

Conclusion

The CCSD contains comprehensive clinical data. Since its inception, this database has demonstrated good stability, security, and scalability. As an public database, the CCSD also exhibits user-friendliness. The database can contribute to the advancement of the diagnosis and treatment strategies for sleep disorders, ultimately promoting sleep health.

In the future, we plan to recruit more institutions to participate in data collection. Additionally, we will conduct follow-ups with individuals in the database. We will also enhance the functionality of the data platform to provide investigators with greater convenience.

Data Sharing Statement

The data of CCSD can be accessed through *psg.wangli-tech.com*. In order to obtain the data, researchers need to contact the Department of Psychiatry (Sleep Medicine Center) of Nanfang Hospital and submit a research protocol with a clear scientific question. After internal review, researchers can be provided with the necessary account and password to access the database.

Acknowledgments

Disclosure

Yuanhui Li, Xiang Liu, and Simin Guo are affiliated with Adai Technology (Beijing) Co., Ltd. The authors report no other conflicts of interest in this work.

References

1. Chaput J, McHill AW, Cox RC, et al. The role of insufficient sleep and circadian misalignment in obesity. *Nat Rev Endocrinol*. 2023;19(2):82–97. doi:10.1038/s41574-022-00747-7
2. Huang T, Mariani S, Redline S. Sleep irregularity and risk of cardiovascular events: the multi-ethnic study of atherosclerosis. *J Am Coll Cardiol*. 2020;75(9):991–999. doi:10.1016/j.jacc.2019.12.054
3. Pearson O, Uglik-Marucha N, Miskowiak KW, et al. The relationship between sleep disturbance and cognitive impairment in mood disorders: a systematic review. *J Affect Disord*. 2023;327:207–216. doi:10.1016/j.jad.2023.01.114
4. Bishir M, Bhat A, Essa MM, et al. Sleep deprivation and neurological disorders. *Biomed Res Int*. 2020;2020:5764017. doi:10.1155/2020/5764017
5. Xin H, Suning L, Junxiang Y, et al. Research status of and recommendations for prevention and control of sleep disorders in China. *J Sichuan Univ*. 2023;54:226–230.
6. Benke K, Benke G. Artificial intelligence and big data in public health. *Int J Environ Res Public Health*. 2018;15(12):2796. doi:10.3390/ijerph15122796
7. Pastorino R, De Vito C, Migliara G, et al. Benefits and challenges of big data in healthcare: an overview of the European initiatives. *Eur J Public Health*. 2019;29(Supplement_3):23–27. doi:10.1093/eurpub/ckz168
8. Jiahui D, Xiaolin H, Xiaoxing L, et al. Past, present and future of sleep medicine in China. *Journal of Peking University*. 2023;55:567–573.
9. Javaheri S, Zhao YY, Punjabi NM, Quan SF, Gottlieb DJ, Redline S. Slow-wave sleep is associated with incident hypertension: the sleep heart health study. *Sleep*. 2018;41(1). doi:10.1093/sleep/zsx179
10. Azarbarzin A, Sands SA, Stone KL, et al. The hypoxic burden of sleep apnoea predicts cardiovascular disease-related mortality: the osteoporotic fractures in men study and the sleep heart health study. *Eur Heart J*. 2019;40(14):1149–1157. doi:10.1093/eurheartj/ehy624
11. Ancoli-Israel S. Sleep and ethnicity. *Behav Sleep Med*. 2010;8(4):191–193. doi:10.1080/15402002.2010.509186

12. Li F, Yan R, Mahini R, et al. End-to-end sleep staging using convolutional neural network in raw single-channel EEG. *Biomed Signal Process Control*. 2021:63. doi:10.1016/j.bspc.2020.102203

13. Prabha A, Yadav J, Rani A, Singh V. Respiratory effort signal based sleep apnea detection system using improved random forest classifier. *IETE J Res*. 2021. doi:10.1080/03772063.2021.1999864

14. Khalighi S, Sousa T, Santos JM, Nunes U. ISRUC-sleep: a comprehensive public dataset for sleep researchers. *Comp Metho Prog Bio*. 2016;124:180–192. doi:10.1016/j.cmpb.2015.10.013

15. Chen C, Wang B, Chen W, Han F. The application and prospects of artificial intelligence technology in the field of sleep medicine. *Natl Med J China*. 2021;101:1710–1716.

16. Nath S, Rahimy E, Kras A, Korot E. Toward safer ophthalmic artificial intelligence via distributed validation on real-world data. *Curr Opin Ophthalmol*. 2023;34(5):459–463. doi:10.1097/ICU.0000000000000986

17. Kim BJ, Kim BS, Mun JH, Lim C, Kim K. An accurate deep learning model for wheezing in children using real world data. *Sci Rep*. 2022;12(1):22465. doi:10.1038/s41598-022-25953-1

18. Chinese Clinical Sleep Database. Department of psychiatry, sleep medicine center, Nanfang hospital. China: Chinese Clinical Sleep Database; 2023. Available from: psg.wangli-tech.com. Accessed November 19, 2023.

19. Berry R, Quan S, Abreu A. *The AASM Manual for the Scoring of Sleep and Associated Events: Rules, Terminology and Technical Specifications*. 2.6. ed. Darien: American Academy of Sleep Medicine; 2023.

20. American Academy of Sleep Medicine. *International Classification of Sleep Disorders*. third edition ed. Darien: American Academy of Sleep Medicine; 2014.

21. Buysse DJ, Reynolds CFR, Monk TH, Berman SR, Kupfer DJ. The Pittsburgh sleep quality index: a new instrument for psychiatric practice and research. *Psychiatry Res*. 1989;28(2):193–213. doi:10.1016/0165-1781(89)90047-4

22. Weixia L, Aizezi MUYESE, Zhitao X, et al. Validity and reliability of the Chinese version of morningness/eveningness questionnaire-5 items (MEQ-5) in students of technical schools. *Chin Mental Health J*. 2016;30:406–412.

23. Johns MW. A new method for measuring daytime sleepiness: the Epworth sleepiness scale. *Sleep*. 1991;14(6):540–545. doi:10.1093/sleep/14.6.540

24. Zhenxiao S, Huaxue L, Linying J, et al. Reliability and validity of Hospital Anxiety and Depression Scale. *Chin J Clinicians*. 2017;11:198–201.

25. Kushida CA, Chediak A, Berry RB, et al. Clinical guidelines for the manual titration of positive airway pressure in patients with obstructive sleep apnea. *J Clin Sleep Med*. 2008;4(2):157–171.

26. Krahn LE, Arand DL, Avidan AY, et al. Recommended protocols for the multiple sleep latency test and maintenance of wakefulness test in adults: guidance from the American Academy of Sleep Medicine. *J Clin Sleep Med*. 2021;17(12):2489–2498. doi:10.5664/jcsm.9620

27. Patil SP, Ayappa IA, Caples SM, Kimoff RJ, Patel SR, Harrod CG. Treatment of adult obstructive sleep apnea with positive airway pressure: an American Academy of Sleep Medicine clinical practice guideline. *J Clin Sleep Med*. 2019;15(2):335–343. doi:10.5664/jcsm.7640

28. Bhatt A. Data quality - The foundation of real-world studies. *Perspect Clin Res*. 2023;14(2):92–94. doi:10.4103/picr.picr_12_23

29. Hu MT. REM sleep behavior disorder (RBD). *Neurobiol Dis*. 2020;143:104996. doi:10.1016/j.nbd.2020.104996

30. Antelmi E, Lippolis M, Biscarini F, Tinazzi M, Plazzi G. REM sleep behavior disorder: mimics and variants. *Sleep Med Rev*. 2021;60:101515. doi:10.1016/j.smrv.2021.101515

31. Manfredini D, Ahlberg J, Winocur E, Lobbezoo F. Management of sleep bruxism in adults: a qualitative systematic literature review. *J Oral Rehabil*. 2015;42(11):862–874. doi:10.1111/joor.12322

32. Brillante R, Cossa G, Liu PY, Laks L. Rapid eye movement and slow-wave sleep rebound after one night of continuous positive airway pressure for obstructive sleep apnoea. *Respirology*. 2012;17(3):547–553. doi:10.1111/j.1440-1843.2012.02147.x

33. Dietmann A, Gallino C, Wenz E, Mathis J, Bassetti CLA. Multiple sleep latency test and polysomnography in patients with central disorders of hypersomnolence. *Sleep Med*. 2021;79:6–10. doi:10.1016/j.sleep.2020.12.037