Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Feasibility of home sleep apnea testing with NightOwl and SleepImage in patients with type 2 diabetes: an external pilot study

Mathilde Hvidtfelt Kjeldsen
matkj@regionsjaelland.dk
Zealand University Hospital Køge
Line Troest Dolleris
Zealand University Hospital Køge
Asbjørn Kørvel-Hanquist
Zealand University Hospital Køge
Gitte Bjørn Hvilsom
Zealand University Hospital Køge
Mads Mose Jensen
Zealand University Hospital Køge
Troels Wienecke
Zealand University Hospital
Preben Homøe
Zealand University Hospital Køge

Research Article

Keywords: Denmark, Feasibility, Home sleep apnea testing, HSAT, NightOwl, Nox T3, Obstructive sleep apnea, OSA, SleepImage, Type 2 diabetes

Posted Date: November 11th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11430334/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: Competing interest reported. The present study and the planned main study were investigator-initiated. ResMed donated the NightOwl devices used in the study. Under the agreement with ResMed, the company is entitled to review material intended for publication. By

agreement with ResMed, the review of the present manuscript was limited to the methods section in order to verify the accuracy of the device description, and MyCardio LLC, the manufacturer of SleepImage, was offered the same opportunity although no corresponding agreement existed. Both companies suggested a few minor wording changes, which were adopted to improve clarity. No substantive content was changed, no information was removed, and submission was not delayed. Beyond this review, neither company had any role in the design of the study, the collection, analysis or interpretation of data, or in the writing of the manuscript. The authors declare no other competing interests.

Title

Feasibility of home sleep apnea testing with NightOwl and SleepImage in patients with type 2 diabetes: an external pilot study

Author information

No.	Full name	e-mail
1	Mathilde Hvidtfelt Kjeldsen* (MHK)	matkj@regionsjaelland.dk
2	Line Troest Dolleris (LTD)	linetp@regionsjaelland.dk
3	Asbjørn Kørvel-Hanquist (AKH)	asbp@regionsjaelland.dk
4	Gitte Bjørn Hvilsom (GBH)	gibj@regionsjaelland.dk
5	Mads Mose Jensen (MMJ)	madsmj@regionsjaelland.dk
6	Troels Wienecke (TW)	trw@regionsjaelland.dk
7	Preben Homøe (PH)	prho@regionsjaelland.dk

* Corresponding author

Abstract

Background

Obstructive sleep apnea (OSA) is more prevalent and substantially underdiagnosed in patients with type 2 diabetes (T2DM), whose regular follow-up offers an opportunity for OSA screening. The current diagnostic approach requires resources not available in that setting, whereas the less complex and less obtrusive

wearables NightOwl and SleepImage are promising alternatives. Their diagnostic accuracy must be established in patients with T2DM, in whom cardiac autonomic neuropathy may affect the signals acquired by both devices. This external pilot feasibility study aimed to assess device and study feasibility prior to a diagnostic accuracy study.

Methods

Adults with T2DM in Denmark were recruited through a tertiary diabetes outpatient clinic, general practice, and Facebook. In this single-arm, non-randomized study, participants underwent home sleep apnea testing on three consecutive nights with the two wearables NightOwl and SleepImage, with the cardio-respiratory monitor Nox T3 added on the third night as a comparator. Failed test nights were repeated. Device feasibility comprised the proportion of analyzable recordings per device and participants' experience with the two wearables, assessed by a questionnaire using visual analog scale (VAS) ranging from 0 to 100 and categorical items. Study feasibility comprised the inclusion rate, recruitment source, and measures of completion. All outcomes were analyzed descriptively.

Results

Forty-three adults were included. Median age was 61 years, and 51% were male. Most were recruited through Facebook (79%), and the inclusion rate was 3.9 participants per month. Median VAS scores were 83 for NightOwl and 92 for SleepImage. Analyzable recordings were obtained on 75% to 85% of intended test nights across all three devices. Sixty-seven percent completed all planned test nights after a median of three attempted nights, whereas non-completers attempted a median of six nights. Moderate-to-severe OSA was found in 56% of participants with an analyzable Nox T3 recording.

Conclusions

Participants reported a favorable experience with both NightOwl and SleepImage, and the main diagnostic accuracy study will proceed with active recruitment added. Its design should account for a technical success rate of roughly three in four recordings and a completion rate of two in three.

Trial registrations

Registered at ClinicalTrials.gov (NCT07778121) on 14 January 2026. Retrospectively registered during recruitment.

Keywords

Denmark; Feasibility; Home sleep apnea testing; HSAT; NightOwl; Nox T3; Obstructive sleep apnea; OSA; SleepImage; Type 2 diabetes

Key messages regarding feasibility

Existing uncertainties regarding the feasibility

This external pilot study precedes a diagnostic accuracy study of obstructive sleep apnea (OSA) comparing the wearables NightOwl and SleepImage with cardio-respiratory monitoring in patients with type 2 diabetes. Participants' experience with the wearables, the technical success rate of the three devices in this population, the recruitment rate, and adherence to the study procedure remained to be determined.

Key feasibility findings

Median ratings of the experience with NightOwl and SleepImage exceeded 80 on a 0 to 100 visual analog scale. Analyzable recordings were obtained on 75% to 85% of intended test nights for all three devices with overlapping confidence intervals. The solely passive recruitment strategy yielded the intended sample size, however with an inadequate inclusion rate. Two in three participants completed the full home sleep apnea testing, with the majority of completers needing no additional test nights, and 72% completed the night with all three devices. Among non-completers, the median number of test nights before declining further attempts was six (range: 3 to 7), indicating persistence.

Implications of the feasibility findings for the design of the main study

Participants' ratings of their experience with both NightOwl and SleepImage indicated that the devices would be acceptable to patients in a real-world setting. The main study will proceed with active recruitment added, more nights with all three devices worn to secure the primary outcome, identical criteria for an analyzable recording across devices, and a sample size calculation that accounts for the observed completion rate and, with caution, the observed distribution of OSA severity.

Background

Obstructive sleep apnea (OSA) is common and known to be more prevalent among patients with type 2 diabetes (T2DM) compared to the general population [1]. Untreated OSA is associated with poorer quality of life, which improves with treatment [2,3]. It is also linked to a higher comorbidity burden, with observational studies finding a reduction in long-term risk of major adverse health outcomes with treatment, while randomized controlled trials have failed to reproduce this effect [4]. Moderate-to-severe OSA was recently reported in 60% of patients with T2DM undergoing cardio-respiratory monitoring (CRM) at a Danish tertiary diabetes clinic [5], whereas only 10% of participants with diabetes, either type 1 or type 2, in a Danish general population study had OSA according to their medical records [6]. Although the two populations are not directly comparable, the gap between these prevalence estimates indicates that a substantial proportion of patients with T2DM have undiagnosed OSA, consistent with the recognized underdiagnosis of OSA in general [7]. Patients with T2DM already attend regular clinical follow-up at a diabetes clinic or with their general practitioner, providing an established contact from which systematic OSA screening could be initiated.

Systematic screening for OSA in patients with T2DM is currently limited by the low accuracy of screening questionnaires [5,8] and by the fact that the current diagnostic approach requires personnel with specialized training in sleep medicine and resources not normally available at diabetes clinics and in general practice. Diagnosis therefore depends largely on the patient's own initiative to seek sleep evaluation. In Denmark, the

most common approach to testing for OSA is CRM, typically performed by an otorhinolaryngologist [9]. The diagnostic setup requires extensive hardware, in-person instructions, wired download from the device, and specialized training in somnographic evaluation for interpretation. Newer wearables such as NightOwl™ and SleepImage® are promising alternatives that could be used in the follow-up setting for patients with T2DM, given their simple design and the lack of need for prior training for either patients or healthcare personnel. These devices record from a single finger-worn sensor, and healthcare personnel can access the acquired sleep data directly via cloud-based transfer. Their automatically scored results have been reported to achieve high diagnostic performance [10,11], thereby reducing the resources needed to evaluate the sleep data. Cloud-based data transfer further enables fewer visits compared to wired download of CRM recordings.

However, given the higher risk of cardiac autonomic neuropathy (CAN) among patients with T2DM [12], a complication that may affect the pulsatile signals acquired by NightOwl and SleepImage [13], the devices should be validated specifically in this population. The planned main study will therefore compare NightOwl and SleepImage with the CRM device Nox T3™ in patients with T2DM, with diagnostic agreement on a comparator night where all three devices are worn as the primary outcome. This pilot feasibility study was conducted to clarify uncertainties regarding participants' acceptance of the devices, the technical success rate of the recordings, the adequacy of the recruitment strategy, and the ability to complete the study procedure mirroring the subsequent main diagnostic accuracy study.

This study aimed to assess device and study feasibility. Device feasibility was defined as participants' experience with NightOwl and SleepImage, as well as the technical ability to obtain analyzable recordings from NightOwl, SleepImage, and the comparator CRM device Nox T3. Participants’ experience additionally addressed whether future clinical implementation would be realistic. Study feasibility was defined as the practical and logistical aspects of study progress, to identify issues requiring resolution or optimization before the main study. In addition, the distribution of OSA severity was described to inform the sample size calculation of the main study.

Methods

Study Design and Setting

A single-arm, non-randomized external pilot study was conducted, mirroring the design of the planned main study. Inspired by a similar study by Lildal et al., this study aimed to include 40 participants [14]. Reporting followed the CONSORT extension adapted to non-randomized pilot and feasibility studies [15,16].

The study site was located at the Department of Otorhinolaryngology, Zealand University Hospital (ZUH), Køge, Denmark, and enrolled participants between September 2025 and August 2026. One investigator (LTD) performed all visits from initial contact to the end of the study.

Participants

Participants were eligible if they were 18 years or older, diagnosed with T2DM, resided in Region Zealand, Denmark, and were able to speak and understand Danish. They were excluded if they had previously undergone surgery for OSA, had used treatment for OSA within the last month, had an implanted cardioverter defibrillator (ICD), or were clinically suspected of a sleep disorder other than OSA that might compromise the intended purpose of the devices. Written informed consent was obtained from all participants prior to study entry.

Recruitment materials included brief information about the study and eligibility criteria and were available as flyers at the diabetes clinic (Steno Diabetes Center Zealand) at ZUH Køge and in 18 general practices in Region Zealand. Posts were repeatedly published in local Facebook groups, and the study was also announced on the Danish Diabetes Association website. Recruitment was passive, meaning that interested individuals contacted the field investigator, who assessed eligibility.

Devices

NightOwl and SleepImage

NightOwl (ResMed, San Diego, California, United States) and SleepImage (MyCardio LLC, Denver, Colorado, United States) both comprise a finger-worn sensor, a smartphone app used to start and stop recordings via Bluetooth, and proprietary software. The NightOwl sensor (model: Mini) is a single-patient-use disposable sensor capable of recording up to ten nights of sleep data. It is shaped like an oval plate and attaches to the finger with adhesive tape. The SleepImage sensor (model: PO2) is rechargeable and reusable between participants. It is designed as a ring with a screen displaying current pulse rate and blood oxygen saturation. Both sensors acquire photoplethysmographic (PPG) signals, from which they derive oxygen saturation and pulse rate, as well as actigraphic signals. The two systems apply distinct proprietary signal-processing algorithms. NightOwl derives peripheral arterial tonometry (PAT) and other physiological parameters from PPG and accelerometry signals, whereas SleepImage uses PPG signals for sleep stage and state analysis from cardiopulmonary coupling (CPC) analysis and then further utilizes the PPG signal to derive breathing (plethysmography-derived respiration (PDR)) and oxygen metrics for evaluating sleep-disordered breathing events. Both devices generate a report for each test night, which is available in a secure cloud-based portal immediately after an ended recording. NightOwl does not generate a report if less than four hours of recording time is achieved, whereas SleepImage generates a report regardless of recording length but flags recordings shorter than four hours.

For the purpose of this feasibility study, an available report was considered analyzable. Assessment of the sleep measures in the reports falls within the scope of the main study.

Nox T3

The Nox T3 (Nox Medical, Reykjavík, Iceland), validated for OSA diagnostics [17], consists of a monitor attached with clips to the nightwear and connected to two respiratory inductance plethysmograph belts encircling the thorax and abdomen, and a nasal cannula. It connects via Bluetooth to a wrist-worn oximeter

(model: 3150, Nonin Medical Inc., Plymouth, Minnesota, United States) with a finger clip. The Nox T3 and the oximeter use disposable alkaline batteries with enough battery life for one to two nights, depending on the length of each recording. Signals recorded by the Nox T3 are downloaded via cable to Noxturnal software.

Sleep measures were based on the software's automated scoring, which has demonstrated good agreement with manual scoring [18]. Investigators LTD or MHK manually reviewed the recordings to assess signal quality and establish the start and end points of the analysis, while also removing artifacts (e.g., loss of signal), thereby replicating standard procedures at the study site. In accordance with local referral guidelines and international recommendations [19], recordings containing at least four hours of total sleep time (TST) and an overall signal quality of 90% were deemed analyzable.

Feasibility Questionnaire

A questionnaire developed by the authors for this study, inspired by the CACHET Unified Methodology for Assessment of Clinical Feasibility by Bardram [20], evaluated participants' experience with NightOwl and SleepImage. The questionnaire focused on the two devices separately and was organized into five domains: 1) Usability, 2) Sensor application and overnight fit, 3) Smartphone app, 4) Need for assistance, and 5) Health empowerment. It comprised 12 visual analog scale (VAS) items per device, distributed across all domains except 'Need for assistance', alongside categorical and free-text items. Three VAS items (items 6, 7, and 8) were inadvertently anchored in the opposite direction to the remaining items, with "completely agree" at 0 instead of 100. Item 6 was the last of six VAS items in the usability domain, and items 7 and 8 were the only VAS items in the sensor application and overnight fit domain. The questionnaire also included general items regarding experience with and positivity towards technology.

The original Danish and translated English versions of the questionnaire are provided in Supplementary material.

Supplementary table Feasibility questionnaire developed by the authors for this study. Items in part a) were presented for NightOwl first and then for SleepImage, with the device name replacing [the device]. Items in part b) concerned health

technology in general and were presented once.

*The condition for item 11 differed between the two devices, requiring "or" for NightOwl and "and" for SleepImage.

!Anchors were inadvertently reversed relative to the other VAS items.

VAS = visual analog scale.

a)

No.	Item	Response format
Usability		
1	Overall, I am satisfied with how easy [the device] was to use	VAS: Slider from completely disagree (0) to completely agree (100)
2	It was easy to learn how to use [the device]	VAS: Slider from completely disagree (0) to completely agree (100)
3	I had the information I needed to use [the device] correctly	VAS: Slider from completely disagree (0) to completely agree (100)
4	I used [the device] in the way I had been instructed to	VAS: Slider from completely disagree (0) to completely agree (100)
5	Using [the device] interfered minimally with my daily life	VAS: Slider from completely disagree (0) to completely agree (100)
6	My sleep was not affected by [the device]	VAS: Slider from completely agree (0) to completely disagree (100)!
Sensor application and overnight fit		

7	It was easy to apply [the device] to my finger	VAS: Slider from completely agree (0) to completely disagree (100)[!]
8	[the device] was correctly in place the following morning	VAS: Slider from completely agree (0) to completely disagree (100)[!]
9	Number of times during the first night I noticed that [the device] needed adjusting	0 times; 1 to 2 times; 3 to 4 times; 5 times or more; do not remember
10	Number of times during the second night I noticed that [the device] needed adjusting	0 times; 1 to 2 times; 3 to 4 times; 5 times or more; do not remember
11	Reasons that made adjustment of [the device] necessary *Condition: Asked only if item 9 and/or* 10 were not 0 times*	All that apply: itching; too tight; too loose; had fallen off; painful; other
12	Suggestions for improving the information and instructions on use of [the device]	Free text
Smartphone application		
13	The information in the app was easy to understand	VAS: Slider from completely disagree (0) to completely agree (100)
14	Overall, I am satisfied with how easy the app was to use	VAS: Slider from completely disagree (0) to completely agree (100)
Need for assistance		
15	I contacted the study staff for advice while at home	Yes; no
16	I should have sought telephone assistance to make [the device] work properly *Condition: Asked only if item 15 was answered no*	Yes; no.
17	Number of times I contacted the study staff for advice	1; 2; 3; 4 or more.

	Condition: Asked only if item 15 was answered yes	
18	I received help from a relative to make [the device] work	Yes; no
Health empowerment		
19	I am confident that [the device] measures correctly	VAS: Slider from completely disagree (0) to completely agree (100)
20	I feel that I have influence on my own health by using [the device] to be examined for sleep apnea	VAS: Slider from completely disagree (0) to completely agree (100)
Free comments		
21	Space to elaborate on the experience of using [the device]	Free text

Study procedure

Eligible participants were invited to the study site for a first visit comprising an interview on health history with focus on sleep, completion of three questionnaires (Berlin Questionnaire [21], STOP-Bang Questionnaire [22], both validated in Danish, and Epworth Sleepiness Scale [23]), self-reported height and weight, measurement of neck, waist and hip circumferences, an ear, nose and throat examination with flexible endoscopy, an examination for CAN with Vagus (Medicus Engineering ApS, Aarhus N, Denmark) [24] based on resting heart rate and heart rate variability during three simple maneuvers, and blood pressure measurement. Results for lipids (total cholesterol and triglycerides) and HbA1c were obtained from the medical record if available and less than three months old. Otherwise, new samples were drawn and analyzed at the regional laboratory following standard procedures. If testing for CAN failed with no obvious cause other than cardiac arrhythmia, participants were referred for an electrocardiogram (ECG). Finally,

participants received all three devices and a preconfigured smartphone to bring home, along with oral and written instructions for use.

Participants then performed home sleep apnea testing (HSAT) over three consecutive nights, wearing NightOwl and SleepImage during all three nights and the Nox T3 only during the third night. Participants completed the feasibility questionnaire before the comparator night with all three devices. Participants could contact investigator LTD by telephone for support at any point during the HSAT period. After completing the HSAT, the participants returned all devices at their final visit. Figure 1 outlines the procedure.

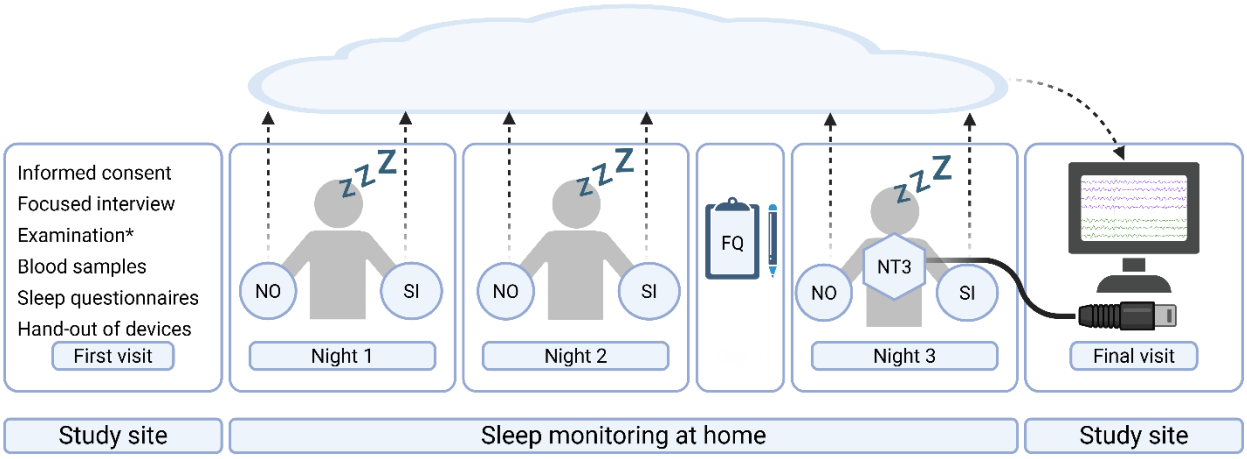


Figure 1 Overview of study procedures. The timeline runs from the first visit at the study site on the left, through three consecutive nights of home sleep apnea testing, to the final visit on the right. NightOwl and SleepImage were worn on all three nights, and Nox T3 was added on the third. The feasibility questionnaire was completed on day 3, after two nights of monitoring and before the night on which all three devices were worn. Any night was repeated with all scheduled devices if one or more recordings were non-analyzable. Dashed arrows indicate cloud-based data transfer, through which NightOwl and SleepImage recordings were available immediately after each night, whereas the cable indicates that Nox T3 recordings could be retrieved only by wired download at the study site.

NO = NightOwl, SI = SleepImage, NT3 = Nox T3, FQ = Feasibility questionnaire. *Examination comprised an ear, nose and throat examination with flexible endoscopy, blood pressure measurement, assessment of cardiac autonomic neuropathy, and measurement of neck, waist and hip circumference

Created in BioRender. Kjeldsen, M. (2026) https://BioRender.com/y3vmzrj

If one or more scheduled recordings were non-analyzable, the participant was instructed to repeat that night with all scheduled devices. Failed NightOwl and SleepImage recordings were identified the day after use, in which case the participant was contacted by telephone, whereas a failed Nox T3 recording was not apparent until the device was returned to the study site for download. If a participant reported an obvious cause of failure, for example that the Nox T3 had been dismounted during the night, the night was repeated with all three devices without returning the device for download first.

Study data were collected and managed using REDCap electronic data capture tools hosted at Region Zealand [25,26].

Outcomes and statistical analysis

All outcomes were displayed descriptively, with continuous measures presented as mean with standard deviation (SD) and median with either range or interquartile range (IQR), and categorical measures as counts and proportions.

Participants' experience with NightOwl and SleepImage, as reported in the feasibility questionnaire, was expressed as VAS scores by domain and overall and displayed as boxplots. Responses to the three reverse-anchored VAS items were excluded post hoc for both devices, comprising six responses per participant, as their distribution indicated inconsistent interpretation of the anchoring. The usability domain was therefore scored from five VAS items, and no VAS-based score was calculated for the sensor application and overnight fit domain. Responses to categorical items were also reported.

Analyzable recordings were reported per device relative to the number of intended test nights, accompanied by a post hoc sensitivity analysis applying the minimum criterion of TST $\geq$ 4 hours on all three devices, and not just the Nox T3. An intended test night is defined as a test night where the device was worn with the intention to record, and thus excludes nights where a scheduled device was never mounted. The Wilson method was applied to calculate 95% confidence intervals (CIs) [27].

Inclusion rate was calculated as the number of participants enrolled per month, and recruitment route relative to all enrolled participants. No active screening was performed. Completion of the HSAT was reported relative to all enrolled participants, together with the number of nights attempted and the duration from first visit to last test night, both reported separately for completers and for non-completers who declined further attempts.

Completeness of baseline data, as defined by data collected at the first visit, was reported for each variable. All baseline data intended for the main study were collected, although not analyzed in this study.

The prevalence of mild (5 ≤ apnea-hypopnea index (AHI) < 15), moderate (15 ≤ AHI <30), and severe (AHI ≥ 30) OSA was reported relative to all participants with a valid AHI from Nox T3, scored according to the 2012 scoring rules of the American Academy of Sleep Medicine (AASM) [28], with hypopneas requiring a ≥3% oxygen desaturation. The study was neither powered nor designed to estimate prevalence, and these figures are reported to inform the sample size calculation of the main study.

Missing data were reported and not imputed. For outcomes assessing completeness of the study procedure, missing data were included as a category in itself.

Analyses were performed in R version 4.6.1 [29] using the {tidyverse} packages [30]. Tables were produced with help from {gt} and {gtsummary} [31,32], and boxplot using {ggplot2} [33].

Results

Population

The study population comprised 43 participants, with a median age of 61 years (IQR: 52 to 67) and 51% male. The median time since diagnosis of T2DM (self-reported) was 8 years (IQR: 3 to 15), and 86% were managed by their general practitioner. All demographic characteristics are presented in Table 1.

252 Table 1 Baseline characteristics of the study population. Continuous variables are presented as median (interquartile range,

253 IQR) and categorical variables as counts and proportions.

254 *BMI was calculated from self-reported height and weight.

Characteristic	**N = 43**
Age in years	61 (52 to 67)
Male sex	22 (51%)
Body mass index in kg/m^2	30.2 (26.9 to 34.6)
Systolic blood pressure in mmHg	130 (124 to 141)
Diastolic blood pressure in mmHg	81 (75 to 88)
Cardiac autonomic neuropathy (CAN)	
CAN	9 (23%)
Borderline CAN	12 (31%)
Normal	18 (46%)
(Missing)	4
Setting for routine follow-up of diabetes	
General practitioner	37 (86%)
Tertiary diabetes clinic	1 (2%)
No follow-up	0 (0%)
Unknown	1 (2%)
Not reported	4 (9%)
Years since diagnosis of type 2 diabetes	8 (3 to 15)

255

Two participants did not obtain an analyzable Nox T3 recording. Among the remaining 41 participants, 40 had OSA, distributed as 17 (41%) with mild, 14 (34%) with moderate, and 9 (22%) with severe OSA.

Device feasibility

Five and six participants did not respond to any VAS or categorical items for NightOwl and SleepImage, respectively, leaving 38 and 37 participants contributing with complete responses to the device specific part of the questionnaire. The overall median score across all included VAS items was 83 (range: 54 to 100) in NightOwl and 92 (range: 27 to 100) in SleepImage. Across the three domains with retained VAS items, scores were slightly higher for SleepImage (median 96, 98, and 89 for usability, smartphone app and health empowerment, respectively) than for NightOwl (median 89, 92, and 84 in the same domains). VAS scores by domains and devices are visually displayed in Figure 2.

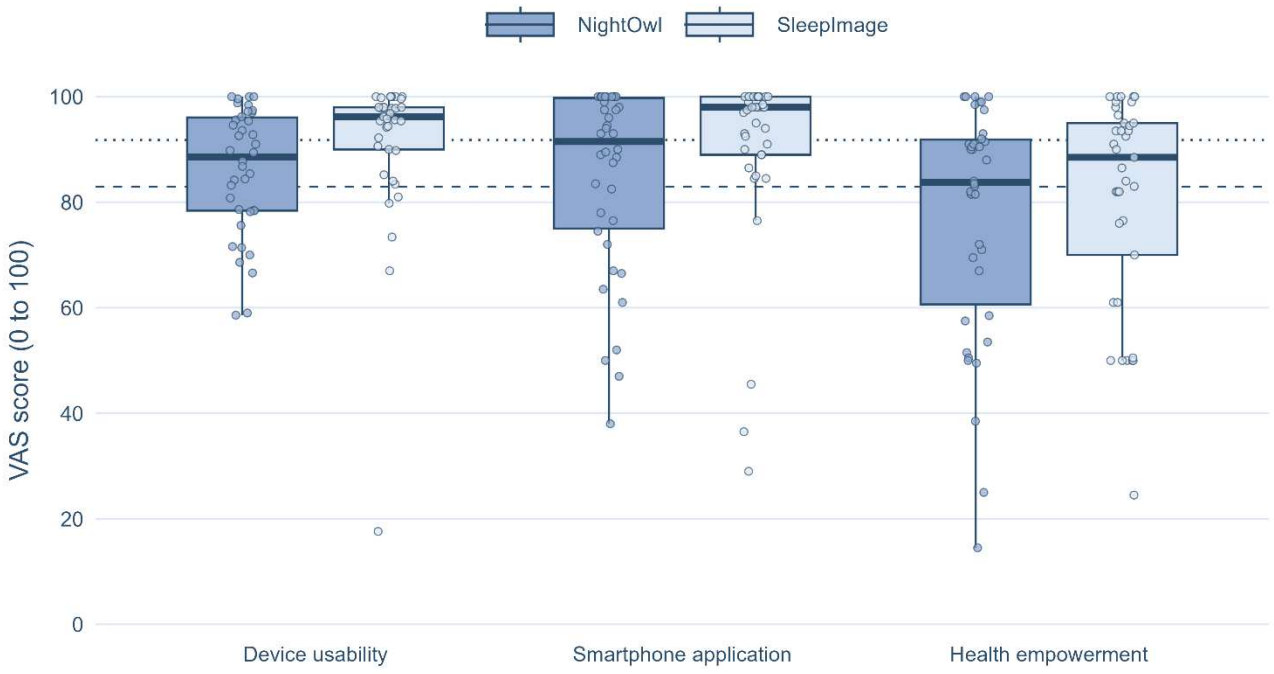


Figure 2 Visual analog scale (VAS) responses for NightOwl and SleepImage, grouped by domain. Scores range from 0 to 100, with higher scores indicating a more positive experience. Each circle represents one participant's response. The horizontal line within each box is the median, the box spans the interquartile range (IQR), and the whiskers extend to the most extreme observation within 1.5 times the IQR from the box. The two horizontal lines spanning the plot are the overall medians across all retained items,

shown dashed for NightOwl and dotted for SleepImage. Three VAS items (items 6, 7 and 8) were excluded post hoc and are therefore not included in the figure. Responses were available from 38 participants for NightOwl and 37 for SleepImage.

Most of the responders did not need to adjust the fit of either NightOwl or SleepImage during the test nights. The proportion reporting no need for adjustments increased from 66% and 76% on the first night to 82% and 84% on the second night for NightOwl and SleepImage, respectively. Five participants contacted the investigator during HSAT regarding NightOwl and three regarding SleepImage. Additionally, six participants received assistance from a relative in mounting and operating NightOwl, whereas only one participant received such help for SleepImage. Of the 36 participants who completed the questionnaire part regarding technology in general, 35 (97%) had a smartphone of their own, and all used it every day or every other day. The median response to the VAS item on positivity toward the use of health technology in general was 98 (range: 94 to 100).

NightOwl, SleepImage and Nox T3 yielded analyzable recordings in 83.5% (95% CI: 77.2 to 88.4), 85.4% (95% CI: 79.3 to 89.9) and 75.4% (95% CI: 63.3 to 84.5) of the intended test nights, respectively. The post hoc sensitivity analysis, applying a minimum of four hours of TST across all three devices rather than only the Nox T3, reduced the proportion for NightOwl to 80.0% (95% CI: 73.4 to 85.3) and SleepImage to 81.3% (95% CI: 74.8 to 86.4). Across all devices and analyses, the CIs overlapped (Table 2).

Table 2 Analyzable recordings per device under the predefined and post hoc applied criteria. Proportions are presented with 95% confidence intervals calculated by the Wilson method. Denominators differ between devices because NightOwl and SleepImage were worn on every test night, whereas Nox T3 was worn only on the night when all three devices were used, and because nights were repeated when a recording was non-analyzable.
Predefined criteria: Available reports for NightOwl and SleepImage. Nox T3, overall signal quality of at least 90% and total sleep time of at least four hours. Post hoc criteria: A minimum of total sleep time of at least four hours added to NightOwl and SleepImage.

Device	**Test nights**	**Predefined criteria**		**Post hoc criteria**	
		n	%	n	%
NightOwl	170	142	83.5 (77.2 to 88.4)	136	80.0 (73.4 to 85.3)

SleepImage	171	146	85.4 (79.3 to 89.9)	139	81.3 (74.8 to 86.4)
Nox T3	61	46	75.4 (63.3 to 84.5)	46	75.4 (63.3 to 84.5)

Study feasibility

After 11 months of enrollment, the intended sample size was reached, yielding an inclusion rate of approximately 3.9 participants per month. The most successful route of recruitment was through Facebook, providing 79% of all participants, followed by 14% through flyer material at general practices and the remainder through other or unknown routes.

Two in three (67%) obtained analyzable recordings from all scheduled devices in all three test nights. The median number of nights attempted among completers was 3 (range: 3 to 8), whereas the non-completers (33%) attempted a median of 6 (range: 3 to 7) test nights before declining further attempts. Median age was slightly higher among non-completers than completers. The median interval from first visit to last recording was 7 (range: 2 to 82) days among all participants. Overall, 72% of participants completed at least one night wearing all three devices (Table 3).

Table 3 Completion of home sleep apnea testing, number of nights attempted, and time spent in the study, presented overall and stratified by whether participants completed all three nights with all scheduled devices or declined further attempts. IQR = interquartile range, n = count, SD = standard deviation

	Overall N = 43	**Completed** N = 29 (67%)	**Declined further attempts** N = 14 (33%)
Age in years, median (IQR)	61.0 (52 to 67)	58 (49 to 65)	63.5 (56 to 70)
At least one night with all three devices, n (%)	31 (72%)	29 (100%)	2 (14%)

Nights attempted			
Mean (SD)	4.2 (1.5)	3.8 (1.2)	5.1 (1.6)
Median (range)	4 (3 to 8)	3 (3 to 8)	6 (3 to 7)
Completed within three nights, n (%)	17 (40%)	17 (59%)	0 (0%)
Days from first visit to last recording			
Mean (SD)	11.6 (14.2)	7.3 (7.3)	20.5 (20.1)
Median (range)	7 (2 to 82)	4 (2 to 33)	17 (2 to 82)

Complete data was collected for nearly all participants. CAN assessment failed in four participants, attributed to technical issues in one case and possible arrhythmia in the remaining three, of whom two completed the ECG. Blood samples were drawn from all participants where indicated.

Discussion

Summary of findings

In this external pilot study, participants rated both NightOwl and SleepImage favorably, and all three devices yielded analyzable recordings on at least three in four intended test nights. Recruitment, rather than acceptability or technical performance, was the main constraint on study progress. One in three participants did not complete the full HSAT, and among participants with an analyzable Nox T3 recording, 56% had moderate-to-severe OSA.

Device feasibility

To our knowledge, participants' experience with NightOwl and SleepImage has not previously been quantified. Caution is warranted when interpreting these estimates, since no threshold for acceptability was defined a priori. However, ratings above 80 (out of a maximum of 100) indicate that participants did not perceive either device as unacceptable, a minimum prerequisite for future implementation in routine care. They further provide a reference for future studies. Although comparing NightOwl and SleepImage was outside the study's scope and power, it is noteworthy that six participants required assistance from a relative to mount and operate NightOwl, compared with only one for SleepImage.

The need to adjust each device during the test night declined from the first to the second night, indicating an initial learning curve rather than a sustained difficulty.

A recent study reported analyzable recordings from NightOwl in 83.3% of planned test nights [34], thus in line with our findings, although the criteria for a failed recording were not specified. No similar estimate has been reported for SleepImage. For Nox T3, the present result of 75% of nights falls between two published estimates obtained under very different conditions. A feasibility study in children reported analyzable recordings in 53% of nights [14], whereas a validation study conducted in a sleep laboratory excluded only 7.5% of participants for poor signal quality [17]. The former differs from the present study in population, and the latter differs in setting where technologists were present. The problems reported in the pediatric study prompted us to ensure battery capacity before hand-out and to instruct participants to secure the nasal cannula and oximeter with adhesive tape, which may also have contributed to the higher success rate.

Study feasibility

This pilot study was planned with capacity to include and complete six participants per week, based on the assumption that the interval from first visit to last recording was less than five days. We found, however, that the median interval from first visit to last recording was seven days (range: 2 to 82) in the entire population. Even after accounting for this extra time per participant by reducing the expected participants to three per

week, capacity still exceeded the observed inclusion rate of 3.9 per month by more than threefold. It is thus recruitment, not site capacity, that constrained the pace of study progression.

The decision to rely solely on the more cost-effective [35] passive recruitment strategy was based on the considerable interest expressed when the study was presented in various T2DM fora before study start. However, this interest did not satisfactorily translate into enrolment, which reached less than a third of the available capacity.

Most participants were recruited through posts in local Facebook groups. Recruitment through social media has been reported to yield study populations reasonably close to the target population [36], but access presupposes internet use, and the participants were therefore likely more digitally literate than the wider T2DM population. Self-referral may additionally have favored particularly motivated participants. Notably, non-completers attempted more test nights before declining than completers required, which suggests that non-completion reflected difficulty with the procedure rather than unwillingness to continue.

Adding active recruitment in the following main study therefore involves a trade-off. Ratings of the devices may be lower, and the failure rate higher, if the result is a study population with less digital experience and less motivation to participate, but broader recruitment reduces the distance between the study population and the target population and thereby strengthens external validity.

Implications for the main study

Based on these findings, the main study should consider seven adjustments.

First, the inclusion rate should be increased by adding active recruitment, which may also improve external validity. Although recruitment materials were displayed at the tertiary diabetes clinic, only one participant was followed in that setting, and active recruitment there would address this underrepresentation of participants likely to have more severe T2DM and a higher burden of complications, including CAN. The size of the eligible population is unlikely to be a limiting factor, as more than 50,000 people in Region Zealand have been diagnosed with T2DM [37].

Second, the number of nights on which all three devices are worn should be increased. The primary outcome of the main study is the diagnostic agreement obtained on such a night, which only 72% of participants completed here. Three test nights were originally chosen to obtain an isolated rating of NightOwl and SleepImage before the Nox T3 was introduced, to buffer failed recordings and the first-night effect, to increase the number of devices to be managed gradually, and to permit assessment of internight variability.

Third, although the feasibility questionnaire was developed for the present study alone, the main study should retain some assessment of device feasibility, since a broader recruitment strategy may yield a population with different characteristics from the present one. Also, the main study should retain items on need for assistance to assess the burden before potential clinical implementation.

Fourth, the sample size calculation should account for the observed completion rate and, with caution, for the observed distribution of OSA severity.

Fifth, the return of Nox T3 should be coordinated so that, if possible, the recording can be downloaded and assessed while the participant is present, avoiding an additional visit when a night must be repeated. This need illustrates the practical advantage of the cloud-based transfer used by the two wearables.

Sixth, the same criteria for an analyzable recording should be applied to all three devices, in accordance with local instructions and AASM recommendations.

Seventh, because most participants (86%) had routine diabetes follow-up with their general practitioner, the original plan to obtain a broad severity spectrum of T2DM by equally including participants followed in primary and tertiary health care in the main study population should be abandoned.

Limitations

No progression criteria were specified in advance. The decision to proceed with the main study therefore rests on the investigators' overall judgment of the findings rather than on pre-specified thresholds, and the findings are used primarily to inform its design.

With 43 participants, the confidence intervals are correspondingly wide, but inferential testing was not intended in this descriptive pilot study. Moreover, recordings were counted per night, and because nights are clustered within participants, the confidence intervals for analyzable recordings are likely too narrow. The study was conducted at the only hospital-based otorhinolaryngology department in Region Zealand, which ensured local expertise and equipment but restricted participation to those living nearby or able to travel.

Three VAS items were inadvertently anchored in the opposite direction to the remainder, and their post hoc exclusion meant that no VAS-based score could be calculated for the sensor application and overnight fit domain, thus reducing the basis on which participants' experience is quantified.

The criteria for an analyzable recording were not identical across devices, which was only partially accommodated by the post hoc application of a minimum of four hours TST on all three devices and not just Nox T3, whereas the minimum criterion regarding signal quality for Nox T3 was not translated into the measures acquired by either NightOwl or SleepImage.

Passive recruitment may have introduced participation bias, inflating the findings of OSA among participants. The observed proportion with OSA is consistent with recent findings from a tertiary diabetes clinic [5], but both populations are selected, and the estimates should be applied to the sample size calculation with corresponding caution, particularly if routes of recruitment change. The predominance of participants recruited through Facebook may similarly have shifted the ratings of the devices upwards.

Finally, the future screening setting this work ultimately aims for involves a simple, single device, whereas participants here managed three, which paradoxically adds complexity to these initial steps of assessing the potential of a less complex diagnostic approach. Both NightOwl and SleepImage were retained because they differ in hardware, proprietary algorithm, and pricing model, so that a single study population yields two independent validations rather than one. Reducing the number of devices would probably improve completion, but at the cost of half the knowledge gained.

Conclusions

The newer wearables NightOwl and SleepImage were highly rated by participants and yielded analyzable recordings in proportions similar to those of the Nox T3. From a feasibility perspective, they are therefore promising, while their diagnostic agreement with the current approach remains to be established. The passive recruitment strategy relying on self-referral yielded an inclusion rate of almost four participants per month, well below this study's capacity and expectation. Adding active recruitment to the strategy might improve study feasibility and external validity. Two in three completed the study procedure. Non-completers attempted more nights before declining further attempts than completers, indicating that non-completion reflected difficulty with the procedure rather than unwillingness to continue.

These findings support proceeding to the main diagnostic accuracy study with the modifications described above. The study also provides reference estimates and identifies practical challenges that may inform similar studies.

List of abbreviations

AASM	American Academy of Sleep Medicine
AHI	Apnea-hypopnea index
CAN	Cardiac autonomic neuropathy
CI	Confidence interval
CPC	Cardiopulmonary coupling
CRM	Cardio-respiratory monitor
ECG	Electrocardiogram
HbA1c	Glycated hemoglobin
HSAT	Home sleep apnea test
ICD	Implanted cardioverter defibrillator

IQR	Interquartile range
OSA	Obstructive sleep apnea
PAT	Peripheral arterial tonometry
PPG	Photoplethysmography
SD	Standard deviation
T2DM	Type 2 diabetes
TST	Total sleep time
VAS	Visual analog scale
ZUH	Zealand University Hospital

Declarations

Ethics approval and consent to participate

The study was approved by the Danish Medical Research Ethics Committees (2403689) and registered with the Region Zealand record of processing activities related to scientific research projects (p-2024-16547). All participants provided written informed consent before inclusion.

Consent for publication

Not applicable.

Availability of data and materials

The datasets generated and analyzed during the current study are not publicly available due to the General Data Protection Regulation (GDPR) and the Danish Data Protection Act.

Competing interests

The present study and the planned main study were investigator-initiated. ResMed donated the NightOwl devices used in the study. Under the agreement with ResMed, the company is entitled to review material intended for publication. By agreement with ResMed, the review of the present manuscript was limited to the methods section in order to verify the accuracy of the device description, and MyCardio LLC, the manufacturer of SleepImage, was offered the same opportunity although no corresponding agreement existed. Both companies suggested a few minor wording changes, which were adopted to improve clarity. No substantive content was changed, no information was removed, and submission was not delayed. Beyond this review, neither company had any role in the design of the study, the collection, analysis or interpretation of data, or in the writing of the manuscript. The authors declare no other competing interests.

Funding

The Novo Nordisk Foundation and Steno Diabetes Center Zealand financially supported this study, and neither influenced study design, data analysis, or reporting of results. The University of Copenhagen partially covered the article processing charge. ResMed donated the NightOwl devices.

Author's contributions

MHK obtained regulatory and ethical approvals, planned the practical conduct of the study, performed the statistical analysis, and drafted the manuscript.

LTD assessed eligibility, enrolled all participants, conducted all study visits and participant contact, and entered all study data into REDCap.

MMJ conceived the idea of evaluating newer HSAT wearables for screening purposes and contributed practical expertise on the devices.

AKH planned the sleep evaluation so that participants were assessed in line with that of other patients with suspected OSA at the Department of Otorhinolaryngology, Zealand University Hospital, Køge, and was consulted on clinical questions throughout.

GBH, TW and PH contributed to the design and conduct of the study.

All authors revised the manuscript and approved the final version.

Acknowledgements

The authors thank the general practices and the local diabetes clinic for making the recruitment materials available to their patients. They also extend their gratitude to the health care personnel at the sleep clinic at ZUH Køge for their practical support.

Authors' information (optional)

Not applicable.

References

[1] Reutrakul S, Mokhlesi B. Obstructive Sleep Apnea and Diabetes. Chest 2017;152:1070–86. https://doi.org/10.1016/j.chest.2017.05.009.

[2] Baldwin CM, Griffith KA, Nieto FJ, O'Connor GT, Walsleben JA, Redline S. The Association of Sleep-Disordered Breathing and Sleep Symptoms with Quality of Life in the Sleep Heart Health Study. Sleep 2001;24:96–105. https://doi.org/10.1093/sleep/24.1.96.

[3] McEvoy RD, Antic NA, Heeley E, Luo Y, Ou Q, Zhang X, et al. CPAP for Prevention of Cardiovascular Events in Obstructive Sleep Apnea. N Engl J Med 2016;375:919–31. https://doi.org/10.1056/NEJMoa1606599.

[4] Gottlieb DJ, Punjabi NM. Diagnosis and Management of Obstructive Sleep Apnea: A Review. JAMA 2020;323:1389. https://doi.org/10.1001/jama.2020.3514.

[5] Asyaei P, Johansen MS, Jakobsen C, Stokbro K, Jensen S, Henriksen JE, et al. Obstructive sleep apnea in type 2 diabetes: prevalence and limitations of the Berlin questionnaire and Epworth sleepiness scale in identifying high-risk patients. Sleep Breath 2025;29:353. https://doi.org/10.1007/s11325-025-03533-7.

[6] Kjeldsen MH, Hvilsom GB, Kørvel-Hanquist A, Wienecke T, Homøe P. Epidemiology of obstructive sleep apnea in a Danish rural population – cross-sectional data from the Lolland-Falster Health Study. Eur Arch Otorhinolaryngol 2026. https://doi.org/10.1007/s00405-026-10363-0.

[7] Young T, Evans L, Finn L, Palta M. Estimation of the Clinically Diagnosed Proportion of Sleep Apnea Syndrome in Middle-aged Men and Women. Sleep 1997;20:705–6. https://doi.org/10.1093/sleep/20.9.705.

[8] Westlake K, Polak J. Screening for Obstructive Sleep Apnea in Type 2 Diabetes Patients – Questionnaires Are Not Good Enough. Front Endocrinol 2016;7. https://doi.org/10.3389/fendo.2016.00124.

[9] Hägerström EG, Kiær EK, Ovesen T, Von Buchwald C, Bille J, Petersen KB, et al. Chronic obstructive sleep apnoea. Ugeskr Læger 2025:1–9. https://doi.org/10.61409/V04250291.

[10] Massie F, Mendes De Almeida D, Dreesen P, Thijs I, Vranken J, Klerkx S. An Evaluation of the NightOwl Home Sleep Apnea Testing System. Journal of Clinical Sleep Medicine 2018;14:1791–6. https://doi.org/10.5664/jcsm.7398.

[11] Lu M, Brenzinger L, Rosenblum L, Salanitro M, Fietze I, Glos M, et al. Comparative study of the SleepImage ring device and polysomnography for diagnosing obstructive sleep apnea. Biomed Eng Lett 2023;13:343–52. https://doi.org/10.1007/s13534-023-00304-9.

[12] Davis TME, Tan E, Davis WA. Prevalence and prognostic significance of cardiac autonomic neuropathy in community-based people with type 2 diabetes: the Fremantle Diabetes Study Phase II. Cardiovasc Diabetol 2024;23:102. https://doi.org/10.1186/s12933-024-02185-3.

[13] Ryals S, Chiang A, Schutte-Rodin S, Chandrakantan A, Verma N, Holfinger S, et al. Photoplethysmography-new applications for an old technology: a sleep technology review. J Clin Sleep Med 2023;19:189–95. https://doi.org/10.5664/jcsm.10300.

[14] Lildal TK, Bertelsen JB, Ovesen T. Feasibility of conducting type III home sleep apnoea test in children. Acta Otolaryngol 2021;141:707–13. https://doi.org/10.1080/00016489.2021.1930152.

[15] Eldridge SM, Chan CL, Campbell MJ, Bond CM, Hopewell S, Thabane L, et al. CONSORT 2010 statement: extension to randomised pilot and feasibility trials. Pilot Feasibility Stud 2016;2:64. https://doi.org/10.1186/s40814-016-0105-8.

[16] Lancaster GA, Thabane L. Guidelines for reporting non-randomised pilot and feasibility studies. Pilot Feasibility Stud 2019;5. https://doi.org/10.1186/s40814-019-0499-1.

[17] Cairns A, Wickwire E, Schaefer E, Nyanjom D. A pilot validation study for the NOX T3(TM) portable monitor for the detection of OSA. Sleep Breath 2014;18:609–14. https://doi.org/10.1007/s11325-013-0924-2.

[18] Kristiansen S, Traaen GM, Øverland B, Plagemann T, Gullestad L, Akre H, et al. Comparing manual and automatic scoring of sleep monitoring data from portable polygraphy. Journal of Sleep Research 2021;30:e13036. https://doi.org/10.1111/jsr.13036.

[19] Kapur VK, Auckley DH, Chowdhuri S, Kuhlmann DC, Mehra R, Ramar K, et al. Clinical Practice Guideline for Diagnostic Testing for Adult Obstructive Sleep Apnea: An American Academy of Sleep Medicine Clinical Practice Guideline. J Clin Sleep Med 2017;13:479–504. https://doi.org/10.5664/jcsm.6506.

[20] Bardram JE. CACHET Unified Methodology for Assessment of Clinical Feasibility. Copenhagen, Denmark: Copenhagen Center for Health Technology; 2018.

[21] Netzer NC, Stoohs RA, Netzer CM, Clark K, Strohl KP. Using the Berlin Questionnaire to identify patients at risk for the sleep apnea syndrome. Ann Intern Med 1999;131:485. https://doi.org/10.7326/0003-4819-131-7-199910050-00002.

[22] Chung F, Yegneswaran B, Liao P, Chung SA, Vairavanathan S, Islam S, et al. STOP Questionnaire: A Tool to Screen Patients for Obstructive Sleep Apnea. Anesthesiology 2008;108:812–21. https://doi.org/10.1097/ALN.0b013e31816d83e4.

[23] Johns MW. A New Method for Measuring Daytime Sleepiness: The Epworth Sleepiness Scale. Sleep 1991;14:540–5. https://doi.org/10.1093/sleep/14.6.540.

[24] Ejskjaer N, Fleischer J, Fleischer J, Jacobsen PE, Poulsen PL, Nygaard H. A Pocket-size Device to Detect Autonomic Neuropathy. J Diabetes Sci Technol 2008;2:692–6. https://doi.org/10.1177/193229680800200421.

[25] Harris PA, Taylor R, Minor BL, Elliott V, Fernandez M, O'Neal L, et al. The REDCap consortium: Building an international community of software platform partners. Journal of Biomedical Informatics 2019;95:103208. https://doi.org/10.1016/j.jbi.2019.103208.

[26] Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)—A metadata-driven methodology and workflow process for providing translational research informatics support. Journal of Biomedical Informatics 2009;42:377–81. https://doi.org/10.1016/j.jbi.2008.08.010.

[27] Wilson EB. Probable Inference, the Law of Succession, and Statistical Inference. Journal of the American Statistical Association 1927;22:209–12. https://doi.org/10.1080/01621459.1927.10502953.

[28] Berry RB, Budhiraja R, Gottlieb DJ, Gozal D, Iber C, Kapur VK, et al. Rules for Scoring Respiratory Events in Sleep: Update of the 2007 AASM Manual for the Scoring of Sleep and Associated Events: Deliberations of the Sleep Apnea Definitions Task Force of the American Academy of Sleep Medicine. Journal of Clinical Sleep Medicine 2012;08:597–619. https://doi.org/10.5664/jcsm.2172.

[29] R Core Team. R: A Language and Environment for Statistical Computing 2021.

[30] Wickham H, Averick M, Bryan J, Chang W, McGowan L, François R, et al. Welcome to the Tidyverse. JOSS 2019;4:1686. https://doi.org/10.21105/joss.01686.

[31] Iannone R, Cheng J, Schloerke B, Haughton S, Hughes E, Lauer A, et al. gt: Easily Create Presentation-Ready Display Tables 2026.

[32] Sjoberg D D, Whiting K, Curry M, Lavery J A, Larmarange J. Reproducible Summary Tables with the gtsummary Package. The R Journal 2021;13:570. https://doi.org/10.32614/RJ-2021-053.

[33] Wickham H. ggplot2: elegant graphics for data analysis. Second edition. Cham: Springer international publishing; 2016.

[34] Hashiba M, Al-Alak A, Nielsen SK, Kayser L, Nouhravesh N, Elmegaard M, et al. Feasibility of a decentralised trial of sleep apnoea screening in patients with atrial fibrillation. Dan Med J 2025:1–10. https://doi.org/10.61409/A12240892.

[35] Estabrooks P, You W, Hedrick V, Reinholt M, Dohm E, Zoellner J. A pragmatic examination of active and passive recruitment methods to improve the reach of community lifestyle programs: The Talking Health Trial. Int J Behav Nutr Phys Act 2017;14:7. https://doi.org/10.1186/s12966-017-0462-6.

[36] Whitaker C, Stevelink S, Fear N. The Use of Facebook in Recruiting Participants for Health Research Purposes: A Systematic Review. J Med Internet Res 2017;19:e290. https://doi.org/10.2196/jmir.7071.

[37] Diabetestal n.d. https://www.diabetestal.nu/ (accessed September 8, 2026).