Obstructive Sleep Apnea Syndrome in Company Workers: Development of a Two-Step Screening Strategy with a New Questionnaire

ARTICLE *in* JOURNAL OF CLINICAL SLEEP MEDICINE: JCSM: OFFICIAL PUBLICATION OF THE AMERICAN ACADEMY OF SLEEP MEDICINE · OCTOBER 2015

Impact Factor: 3.05

6 AUTHORS, INCLUDING:

Michiel Eijsvogel

Medisch Spectrum Twente

22 PUBLICATIONS **144** CITATIONS

SEE PROFILE

Winfried J. Randerath

Krankenhaus Bethanien gGmbH, Solingen

287 PUBLICATIONS **1,311** CITATIONS

SEE PROFILE

Johan Verbraecken

Universitair Ziekenhuis Antwerpen

194 PUBLICATIONS **1,481** CITATIONS

SEE PROFILE

TITLE PAGE

SUBMISSION NUMBER
TJC-15-00511

MANUSCRIPT TITLE
Obstructive Sleep Apnea Syndrome in company workers: development of a two-step screening strategy with a new questionnaire

SUBTITLE
Two-step community screening for obstructive sleep apnea syndrome

RUNNING HEAD
Two-step community screening for OSAS

AUTHORS, HIGHEST ACADEMIC DEGREES AND INSTITUTIONAL AFFILIATIONS
MSc Michiel M Eijsvogel
Department of Pulmonology and Sleep Medicine, Medisch Spectrum Twente Hospital, Enschede, The Netherlands.

MSc Sytske Wiegersma
MIRA-Institute for Biomedical Technology & Technical Medicine, University of Twente, Enschede, The Netherlands.

PhD Winfried Randerath
Institute for Pneumology at the university Witten/Herdecke, Clinic for Pneumology and Allergology, Centre of Sleep Medicine and Respiratory Care, Bethanien Hospital, Solingen, Germany.

PhD Johan Verbraecken
Department of Pulmonary Medicine and Multidisciplinary Sleep Disorders Centre, Antwerp University Hospital and University of Antwerp, Belgium.

MSc Esther Wegter-Hilbers
Occupational health service provider HumanCapitalCare B.V., Son, The Netherlands.

PhD Job van der Palen
Department of Epidemiology, Medisch Spectrum Twente Hospital, Enschede.
Department of Research Methodology, Measurement and Data Analysis, University of Twente, Enschede, The Netherlands.

INSTITUTION AT WHICH STUDY WAS PERFORMED
Medisch Spectrum Twente Hospital, Pulmonary and Sleep department
Haaksbergerstraat 55. 7513 ER Enschede, The Netherlands.

ACKNOWLEDGEMENTS

We are indebted to Piet-Hein van Mechelen (chairman Dutch Sleep Apnea organisation (http://www.apneuvereniging.nl) for practical study coordination and support. We thank Reindert P.J. van Steenwijk, Pulmonary and Sleep Physician, Academic Medical Centrum, Amsterdam, for critical reviewing the manuscript.

DISCLOSURE STATEMENT

Home sleep recording by the Netherlands Slaap Instituut (http://nederlandsslaapinstituut.nl) and data-collection by HumanCapitalCare (www.humancapitalcare.nl) was sponsored by Philips NV, Eindhoven, the Netherlands.
The Department of Research Methodology, Measurement and Data Analysis, of the University of Twente, Enschede, the Netherlands received a grant of Philips NV, Eindhoven, the Netherlands for support in performing the statistical analyses. Philips NV Netherland had no access to employee data, study results, and was not involved in manuscript preparation or writing.
None of the authors: Michiel Eijsvogel, Sytske Wiegersma, Winfried Randerath, Johan Verbraecken, Esther Wegter-Hilbers, Job van der Palen had any conflict of interest.

OFF-LABEL OR INVESTIGATIONAL USE

None.

CORRESPONDING AUTHORS ADDRESS, PHONE, FAX AND E-MAIL ADDRESS

Michiel M. Eijsvogel.
Medisch Spectrum Twente Hospital; PO 50000 7500 KA Enschede, the Netherlands
Phone: 0031-6-22251667 or 0031-53-4281434 or 0031-53-4872000 (Hospital)
Email address: m.eijsvogel@mst.nl or eijsvogel@online.nl

ABSTRACT

Study Objectives: To develop and evaluate a screening questionnaire and a two-step screening strategy for Obstructive Sleep Apnea Syndrome (OSAS) in healthy workers.

Design: Cross-sectional study

Setting and participants: A total of 1861 employees comprising healthy blue and white-collar workers in two representative plants in the Netherlands from a worldwide consumer electronic company were approached to participate.

Interventions: Employees were invited to complete various sleep questionnaires, and undergo nasal flow recording and home polysomnography on two separate nights.

Measurements and Results: Of the 1861 employees 249 gave informed consent and from 176 (70.7%) all nasal flow and polysomnography data were available. OSAS was diagnosed in 65 (36.9%). A combination of age, absence of insomnia, witnessed breathing stops and 3-way scoring of the Berlin and STOPBANG questionnaires best predicted OSAS. Factor analysis identified a six-factor structure of the resulting new questionnaire: snoring, snoring severity, tiredness, witnessed apneas, sleep quality and daytime well-being. Subsequently, some questions were removed, and the remaining questions were used to construct a new questionnaire. A scoring algorithm, computing individual probabilities of OSAS as high, intermediate or low risk, was developed. Subsequently, the intermediate risk group was split into low and high probability for OSAS, based on nasal flow recording. This two-step approach showed a sensitivity of 63.1%, and a specificity of 90.1%. The latter is important for low prevalence populations.

Conclusion: A two-step screening strategy with a new questionnaire and subsequent nasal flow recording is a promising way to screen for OSAS in a healthy worker population.

Keywords: Sleep apnea syndrome, screening, questionnaire, home-recording, polysomnography

Trial registration: Development and validation of a screening instrument for Obstructive Sleep Apnea Syndrome in healthy workers. Netherlands Trial Register (www.trailregister.nl), number: NTR2675.

INTRODUCTION

Obstructive sleep apnea syndrome (OSAS) is a prevalent and treatable disease with often impaired daytime performance and increased cardiovascular and metabolic risks Recently, increasing awareness for these consequences is reflected in a growing interest in screening for OSAS[1-3]. Screening for OSAS can be of importance in a hospital setting (preoperative patients), primary care, work environment and is specific groups as in commercial drivers[2]. In the hospital setting, the pre-test prevalence for any disease is higher than in primary care and the primary goal of hospital based screening is mostly to 'rule in' the disease[4]. In community screening for diseases with no fatal outcome, the goal is to 'rule out' the disease; due to the low prevalence, the huge group of subjects with no disease should be correctly identified without further costly and demanding testing; a high specificity and high NPV is than needed. The price to be paid is often a reduced sensitivity, so accepting missed cases. Also a confirmation test (in OSAS a sleep study) in subjects with positive test results should be performed. The argument to do community screening it as follows: it is better to screen realistically and diagnose at least a relevant proportion of patients out of a huge group of healthy people, than not to screen at all. In our study we focused on this community type of screening.

Screening can be costly, especially in the case of low prevalence rates, and only low-burden methods will result in a sufficient response rate[2,5]. In a systematic review on screening questionnaires for OSAS, only 20 out of 4105 studies were considered valid for inclusion, due to different study population (sleep clinic or hospital), methodological aspects (no polysomnography (PSG) as gold standard, no clearly defined Apnea hypopnea index (AHI) and not sufficient data to draw a two by two table)[5]. Finally, only ten of them were analyzed. From these only three studies were done in the general population using two different questionnaires[5]. The most used questionnaire in the general population and in primary care is the Berlin questionnaire (BerlinQ) with five PSG validated studies published.

Using an AHI ≥ 5 as a definition for OSAS, sensitivity varies between 36% and 86%, whereas the specificity varies between 70%-84%[6-9]. In a meta-analysis looking at studies on screening for OSAS in a preoperative setting, improved accuracy was noted when questionnaires, body measurements, morphometric data and oximetry were combined[10]. Most questionnaires result in a dichotomous answer: high or low probability for OSAS; questionnaires with three possible outcomes (high – intermediate – low probability) allow application of an extra test for the intermediate result.

In this so-called two-step strategy, Gurubhagavatula et al. performed additional oximetry in the intermediate group among commercial drivers, resulting in a sensitivity of 74% and high specificity of 89%[11].

As part of their social responsibility, Philips NV Netherlands planned to screen their workers in The Netherlands (13.500) for OSAS. Given the large numbers, the paucity of literature in occupational population based screening and the financial consequences a screening strategy was developed, using a combination of existing questionnaires. In order to validate this strategy, factor analysis was performed, together with PSG as the gold standard, in a sample of employees. In advance, it was decided to develop a two-step strategy including nasal flow (NFlow) recording.

METHODS AND MEASUREMENTS

Study Design

After obtaining informed consent, a set of diagnostic questionnaires were completed by all participants, while NFlow recording and home-based PSG (as gold standard) were performed in all employees as well, without knowledge of the outcome of the applied questionnaires and recordings. The Medical Ethical Committee Twente at Enschede, The Netherlands, approved this study, which is registered at Netherlands Trial Register (www.trailregister.nl) number NTR2675.

Study population and sample size

A total of 1861 employees comprising healthy blue and white-collar workers in two representative plants from Philips were approached through the regular Philips communication channels (newsletter, e-mail, website). Subsequently, workers received a letter at home explaining the aim of the study (to find a proper screening tool for OSAS, which is prevalent and often under diagnosed), the study outline (online questionnaires and two separate nights of sleep recording at home) and the possible benefit (stating that for most employees the investigation will not have any advantages, since only a minority will have OSAS). A reminder was sent out one week later. After returning the consent form by post, these employees got a personal username and password and were invited to fill in the online questionnaires.

Questionnaires

All questionnaires were translated into Dutch by the Philips Translation Services, with forward and backward translation by two different translators, and were administered as a computer-based questionnaire. Factor analysis was performed for the BerlinQ, STOP Questionnaire (STOPQ) and Athens Insomnia Scale (AIS) separately, to assess their validity in the Dutch language, and will be reported elsewhere.

The *Berlin questionnaire (BerlinQ)* has 11 dichotomous and polytomous items on snoring, wake time, sleepiness, blood pressure, and obesity; and is organized in three different categories. Subjects scoring positive on two or more of these categories are considered at high-risk and subjects who score positive on none or only on one of these categories are considered at low-risk for OSAS. The first category is positive when two or more positive responses are obtained on items 1-5, the second category is positive when two or more positive responses are given on items 6-8, and the third category is positive with at least one positive response on items 10-11 is obtained[6].

The *STOPBANG questionnaire (STOPBANGQ)* has eight dichotomous items, which can be divided into two parts[12]. The first part or *STOP questionnaire* consists of four items: snoring, tiredness, observed apneas, and blood pressure. The second part (BANG) consists of also four items: BMI, age, neck circumference, and gender. Subjects are considered at high-risk when three or more of the eight items are scored positive and at low risk for OSAS with a score of less than three positive items[16].

The *Athens Insomnia Scale (AIS)* can be used to assess insomnia and can be useful to distinguish between different causes of sleepiness[14]. It consists of eight questions. The questions can be scored using a four-point scale, ranging from zero (‘no problem at all’) to

three ('very serious problem'). The total score of the AIS-8 ranges from 0-24. Subjects with a score of six or higher are considered to be at high-risk of having insomnia. We used the very high cut off score of ten to exclude severe insomnia from our study population (see definition OSAS). For an estimated prevalence of 10%, the negative predictive value is 94%[14,15].
The *Epworth sleepiness scale (ESS)* is an eight point self-completed questionnaire to assess the tendency to fall asleep in eight different daytimes conditions[16], with a score of 0-3 pro item and a total scoring range of 0-24, with a cut off value of ≥ 10.
The *Pittsburgh Sleep Quality Index* is a self-rated questionnaire, which assesses sleep quality over the last month, and can distinguish between good and poor sleepers[17]. Nineteen questions generate five subgroups of information: sleep quality, sleep latency, sleep duration, sleep efficiency and sleep disturbances and two additional subgroups consisting of use of sleep medication and daytime functioning. These seven groups are each scored by severity (0-3), leading to a total score range of 0-21. The normal cut-off value is > 5.

Nasal flow recording
All participating employees were asked to undergo one night of home NFlow recording (RU-Sleeping; Philips-Respironics)[3,18]. Nasal-flow is measured by pressure transduction with a nasal cannula connected to a small recording unit. Respiratory events are scored by the device when the peak-to peak nasal pressure waveform initially fall at least below 50% of the baseline value and remained below at least 75% of the baseline value for a minimum of 10 s. If a given hour had three "not valid" periods (at least 12 min without valid airflow), then that hour of the recording was discarded, marked as "ERR" and did not count toward the respiratory event index. At the end of the workday, the portable device with nasal cannula was provided with verbal and written instruction. The next morning, the device was collected at the workplace and read out. For a valid recording, a minimum of four hours recording without 'error' reading was needed.

Polysomnography
Home PSG was performed with the Alice PDx (Philips-Respironics) and further analyzed with Polysmith (Neurotronics, Gainesville, Florida) software. At the end of the workday, trained nurses supervised by a registered polysomnographic technologist applied the PSG electrodes and sensors. EEG was carried out with F4-M1, C4-M1 and O2-M1 derivations, together with nasal airflow (cannula), thoracic and abdominal movements (based on respiratory induction plethysmography, chin EMG, vertical and horizontal eye movements, heart rate, and finger oximetry were measured as well. Sensor choice, settings and scoring were performed according to the AASM 2007 rules[19]. The alternative hypopnea definition (≥ 50% nasal flow amplitude drop with ≥ 3% O_2 desaturation or arousal) was applied. Polysomnographic technologists from the Nederlands Slaap Instituut (http://nederlandsslaapinstituut.nl) hooked up the patients, performed the home PSG and analyzed the data.

OSAS definition
The applied definition of OSAS was based on AHI, symptoms and the absence of severe insomnia (AIS > 10). An AHI ≥ 15 was defined as OSAS and an AHI 5-14 was only defined as OSAS, if severe insomnia was absent and symptoms (daytime sleepiness or at least 2 minor symptoms) were present as described in the AASM and the Dutch OSAS guideline[20,21]. Information about symptoms were obtained from the questionnaires (Table 1)

Table 1. OSAS symptoms and related questions in questionnaires.

Additional measurements

Medical history (concerning previously diagnosed OSAS, cardiovascular disease, , hypertension), and anthropometric measurements like neck circumference, blood pressure, body weight and height, and age were obtained during a 20 minute one-to-one consult by occupational nurses from the health advisory company, HumanCapitalCare (Eindhoven, The Netherlands). No specific resting time protocol was used for blood pressure measurement.

STATISTICAL ANALYSIS

Development of questionnaire and two-step screening strategy

Based on factor analyzes and multivariate logistic regression a set of questions was selected that best predicted the presence or absence of OSAS in a healthy workers population. Many of these questions were derived from the BerlinQ and the STOPBANGQ. In contrast to routine two-way outcome scoring, these questionnaires were also scored leading to a three-way outcome: high, intermediate and low risk for OSAS. For the BerlinQ categories and STOPBANGQ items the following scoring was used respectively for low, intermediate and high risk: 0, 1-2, 3 categories and ≤1, 2-4, ≥5 positive answers.
The questions selected formed the so-called Philips questionnaire (PhilipsQ). By multivariate logistic regression analysis, a scoring algorithm was constructed to compute individual probabilities of OSAS. These probabilities were divided into a high, intermediate and low risk for OSAS. For employees ending up in the intermediate risk category according to this PhilipsQ, results from the NFlow measurement with a cut off NFlow value of 15 events/hour were used to finally classify these employees into either high or low risk for OSAS.

Factor analysis

Factor analysis was based on item response theory (IRT) instead of Classical Test Theory (CTT), which has several advantages over CTT[22,23]. IRT enables to conduct factor analysis on variables that are not continuous, an aspect often ignored when factor analysis is applied [13,24-26]. Furthermore, IRT can deal with missing data and scales that contain items with different response formats, which is an important advantage, since the PhilipsQ contains both dichotomous and polytomous items and consists of several introductory and follow-up items that result in many missing data[22,23,27]. Finally, IRT offers detailed insight in the psychometric characteristics[27].
Factor analysis was performed in two steps; first of all, the factor structures of the three questionnaires used in the PhilipsQ (the BerlinQ, STOPQ and AIS) were analyzed. Secondly, factor analysis was performed on the PhilipsQ to assess if the observed responses could be explained by an underlying structure and to assess the psychometric properties of the items. The exploratory factor analysis was conducted with oblique cf-parsimax rotation, which minimizes variable and factor complexity and results in a simple, easily interpretable factor structure, using

Mplus software (version 5.2, 2007, Muthén & Muthen, Los Angeles, USA)[28]. For the BerlinQ, the AIS and the PhilipsQ, Samejima's graded response model for ordered categorical (ordinal) variables was fitted[29]. For the STOPQ, a two-parameter logistic model for binary variables was used. Item parameters were estimated using a mean-adjusted weighted least squares estimator[30]. Based on the table of critical values of Stevens, factor loadings greater than 0.364 were considered significant[31]. Model fit was assessed based on Chi-Square statistics, the Root Mean Square Error of Approximation (RMSEA), the Comparative Fit Index (CFI) and the Tucker Lewis Index (TLI), using recommendations for adequate fit given by Hu and Bentler[32]. Good model fit is considered as empirical evidence that an underlying structure is present, beneath the observed responses on the items[23].

RESULTS

Baseline data

Of the invited 1861 employees 249 gave informed consent. Finally, 241 employees (12.9%) completed all questionnaires. Due to further refusal to participate or prior diagnosis of OSAS 196 appointments for PSG and NFlow recording on two separate night were made; 196 home PSGs were performed, of whom 186 were valid (six battery failures (in the first group of employees) and four PSGs showed incomplete data due to a memory bug. Of the 196 appointments for NFlow three employees did not show up. Eight employees had incomplete data (< 4 hours), hence 185 flow-measurements were valid for analysis. In 176 employees, complete data from questionnaires, physical measurements, NFlow recordings and PSGs were obtained for analysis (figure1).

Figure1. Study flow scheme

Eighty-six (48.8%) employees did have an AHI ≥5, while 65 employees (36.9%) were diagnosed as OSAS (Table 2). Significant differences between no OSAS and OSAS were found for age, BMI, neck circumference, (AHI), ODI, mean SO_2, arousal index, WASO, sleep efficiency, NREM1, NREM3, REM, NFlow index, STOPQ high risk and STOPBANGQ high risk. In employees with an AHI ≥ 5 or with OSAS the STOPQ and the STOPBANGQ, but not de BerlinQ, was significant more positive (high risk for OSAS) compared to employees with respectively an AHI< 5 or without OSAS (Table 2).

Table 2. Morphometric, PSG and questionnaires characteristics of all employees categorized by AHI or OSAS.

Descriptive statistics questionnaire

Comparison of fit indices indicated that a seven-factor model had a significantly ($p < 0.01$) better fit than models with fewer factors: χ^2 (98) = 114.62, CFI = 0.999, TLI = 0.998, RMSEA = 0.027. However, this model did not result in an interpretable factor solution. Therefore, the six-factor model, also demonstrating a good fit (χ^2 (114) = 159.33, CFI = 0.998, TLI = 0.996,

RMSEA = 0.041), was which resulted in a more interpretable solution. The factor solution for the PhilipsQ is shown in table 3. Factor 1 contains only one item, which represents 'snoring'. Factor 2 consists of four items and seems to represent 'snoring severity'. Factor 3 represents 'tiredness', reflected by three items on tiredness and one item on daytime well-being. Factor 4, 'observed apneas', contains two items on observed apneas and one item with regard to the frequency of nodding off while driving. Five items reflect factor 5, 'sleep quality'. Factor 6, 'daytime well-being', consists of six items with regard to tiredness and functioning during the daytime and an additional item on blood pressure. The item nodding off while driving did not load on any of the factors.

Table 3. Factor loadings of Philips Questionnaire items

Logistic regression analysis

Multivariate logistic regression analysis identified a number of variables as predictors for OSAS, shown in Table 4. As can be seen we used a novel, three-way scoring for the BerlinQ and STOPBANGQ as the normal two-way scoring (not shown) was less predictive than the tree -way scoring.

Table 4. Variables predicting OSAS

A new questionnaire (PhilipsQ) was constructed out of these best predictive questions and questionnaires (see supplement).

With the predictive factors an algorithm can be written to compute the individual estimated probability (P) that a subject will have OSAS: $P(OSAS) = 1 / (1+e^{-X})$, where X = -6.322 + 2.828 x AIS + 2.745 x STOPBANGQ High + 0.965 x STOPBANGQ Intermediate + 4.640 x BerlinQ High + 1.584 x BerlinQ Intermediate + 0.790 x Age + 0.810 x BerlinQ Q-5.

Exploratory logistic regression analyzes, testing different cut off values of OSAS probabilities, led to the final cut off values of 35% for low probability of OSAS and 55% for high probability of OSAS. In the second step, for employees with a probability between 35 and 55% the NFlow index with a cut-off of ≥15 breathing stops (>10 seconds) per hour was added to finally classify these employees into high or low risk as well. The decision to use these cut-off values was based on our preference to have primarily a high specificity and negative predictive value for screening purposes; the test characteristics of the low, high and intermediate + NFlow steps of the Philips questionnaire and final two-step strategy is presented in table 5. As can been seen, the sensitivity and specificity of the two-step strategy is 63.1% and 90.1% respectively. The AUC of the ROC curve is 0.81 (CI 0.75-0.88).

Table 5. Test characteristics Philips questionnaire and two-step strategy.

To translate these test characteristics into outcome in screening populations with different prevalences, the negative (NPV) and positive predictive values (PPV) were calculated (Table 6). To read this table an example for 20% prevalence is given: in a screening population with

an estimated prevalence of OSAS of 20%, 29.1% will need NFlow recording for the intermediate outcome of the PhilipsQ; finally 20.5% will have a high probability for OSAS and will be referred to a sleep clinic, whereas OSAS will be confirmed in about 2/3 (61.5%) of the referred persons. In those with a negative outcome of the two-step strategy, 90.7% should be correctly identified as negative for OSAS.

Table 6. Outcome two-step screening strategy according to prevalence of OSAS

A current practice is to use the BerlinQ or STOPBANGQ questionnaire also in a low prevalence setting as in community screening and refer the cases with high-risk outcome for further testing. In this study we used these questionnaires as well. The sensitivity and specificity of the BerlinQ was 46.2% and 69.4% and for the STOPBANGQ 80.0% and 55.0%, respectively. The prevalence in our sample (due to self-selection) was high: 36.9%. However, for testing in a low prevalence setting in for instance 1000 subjects and prevalence of 10%, the PPV is very low for both questionnaires (Berlin and STOPBANG), due to many false positives compared to the Philips two-step strategy (Table 7). A higher negative predictive value (NPV) and a small reduction in true positive rate (3.7% for the Philips two-step strategy, compared to 4.6% and 5.1% for the Berlin and STOPBANG questionnaire, respectively) is also observed. All true and false positives should be referred for diagnostic sleep studies.

Table 7. Estimated outcome of three screening strategies

DISCUSSION

Our study, investigating a two-step strategy to detect OSAS in a healthy workers population demonstrated a sensitivity of 63.1%, and specificity of 90.1%. Most questionnaires used to screen for OSAS in low-prevalence groups use a high versus low risk outcome and have a limited sensitivity and specificity[6-9]. It has been shown that the limited sensitivity and specificity can be improved with three-way scoring (high, low, intermediate risk of OSAS)[11]. This allows application of an extra test for subjects ending up in the intermediate risk group. The new PhilipsQ was constructed from the most predictive questions out of a number of established OSAS and sleep questionnaires.
The high odds ratio of the AIS scoring is due to the used definition of OSAS, in which employees with an AIS score >=10 were excluded as having OSAS. Only seven employees in the three-way scoring of the BerlinQ where at high risk (three positive categories) explaining the very high odds ratio of the variable 'BerlinQ High vs. Low'. In the multivariate analysis, age was almost contributing statistically (0.052), and we decided to retain this variable in the overall formula to predict the individual probability of OSAS. The BerlinQ question five about breathing stops was not contributing statistically significantly, but left in the prediction as a separate variable since it was shown before as a predictive question in literature and guidelines[6,12].
Factor analysis of the 21 questions in the PhilipsQ showed a six-factor structure as best interpretable, consisting of the following factors: snoring, snoring severity, tiredness, observed apneas, sleep quality and daytime well-being. The question 'nodding off while

driving' (PhilipsQ question eight) did not load on any factor and therefore, it was not included in the final Philips questionnaire. It was shown before that 'nodding off while driving' is a conflicting question. Netzer et al. reported a reduction of the Cronbach's α from 0.86 to 0.63 due to this question in category II of the BerlinQ[6]. The related question 'nodding off frequency' (PhilipsQ question nine) was therefore removed as well.

We constructed a scoring algorithm to compute individual probabilities of OSAS. These probabilities were divided into a high, intermediate and low risk. The intermediate group could be split into low and high probability for OSAS, based on NFLow with a cutoff of ≥ 15. This two-step strategy has a sensitivity of 63.1%, and specificity of 90.1%. Depending on the prevalence of OSAS in screening populations, a NPV of at least 90.7% for a prevalence ranging from 5-20% can be anticipated. For the much smaller group of subjects with a positive test, the positive predictive value (PPV) varies from 25-61%; these subjects with a positive test results should than be referred for a diagnostic sleep study. The PhilipsQ is constructed of body measurements (Body mass index (BMI), age, neck circumference, gender) and questions from the BerlinQ, STOPQ and AIS. These questionnaires are translated into many languages, which enables its implementation outside the Dutch language. Internal consistency for the different questionnaires used within the PhilipsQ has been published in the literature. For the BerlinQ, the Cronbach's α for the category I questions (PhilipsQ questions 1-5) is 0.92, while 0.63 for category II questions (PhilipsQ questions 6-9) (or 0.86 when leaving out PhilipsQ questions 8)[6]. Leaving out question eight in the free available PhilipsQ was therefore done after this study. For the STOPQ (PhilipsQ questions 10-13) and AIS (PhilipsQ questions 14-21), Cronbach's α is 0.92 and 0.89, respectively[12,14].

In line with our study, Cowan et al. found no discriminating properties in the normal (two-way) scoring of the BerlinQ, in contrast to the STOPQ and STOPBANGQ[33]. Several prediction models for OSAS, combining morphometric data have been published in patients referred to a sleep clinic. The Multivariate Apnea Risk (MAP) and the Sleep Apnea Clinical Score (SACS) are such examples[34,35]. Evaluating these and other models for OSAS (AHI≥10) in referred patients showed a sensitivity range of 76-96% and a specificity range of 13-54%[38]. These authors concluded that clinical prediction models are not sufficiently accurate to discriminate between patients with and without OSAS[36].Takegami et al. tested a four variable prediction model (gender, blood pressure BMI, snoring) for OSAS (AHI≥15) in employees during a health check, and reported a sensitivity of 93% and a specificity of 66% [37]. In this study, the study population was obscured with referred patients, and oximetry and type-III portable monitoring was used instead of PSG. In a dataset from the Sleep Heart Health Study, Caffo et al. recently published about (modern ensemble learning) algorithms for predicting OSAS with a RDI ≥ 7. They found BMI, age, snoring frequency, waist circumference and snoring loudness as the variables with the largest impact on prediction performance[38]. Except for or waist circumference (not measured) the same variables in our predicting algorithm where included.

However, Caffo et al. did not investigate a two-step strategy, probably explaining the moderate specificity of 70%. On the other hand, Gurubhagavatula et al. applied such two-step screening strategy based on the multivariable apnea prediction index and oximetry in commercial drivers, which resulted in a sensitivity of 74% and a specificity of 89% for OSAS (AHI≥5)[11]. These results were quite similar to our data. This somewhat higher sensitivity could be explained by using an AHI>5 instead of our OSAS definition or as the authors stated, possibly due to regression towards the mean (due to strategy validation in the same

development cohort). In a primary care setting, Chai-Coetzer et al. developed a two-step strategy to select OSAS, with a four-item questionnaire (waist circumference, breathing stops (BerlinQ, Q-5), snoring and age) combined with oximetry to select OSAS (AHI≥30)[39]. They reported a sensitivity of 88% and specificity of 82% in the validation group. Their OSAS50 screening questionnaire had a two-way result with a scoring range of 0-10 instead of an algorithm[39].

A limitation of our study was that only 12.9% of the invited employees gave informed consent. However, a low participation rate and self-selection are inherent to community based studies in healthy employees, especially when there is a substantial study burden to the participants (five questionnaires and two nights of sleep monitoring). Employees with sleep related complaints were probably more willing to participate, explaining the extremely high prevalence of OSAS (36.9%). On the up-side, this self-selection resulted in a sufficient number of employees with OSAS that allowed a reasonable assessment of sensitivity and specificity of our two-step strategy. Our study is not unique in this respect, as another study in commercial drivers reported the same limitations; complete datasets (questionnaires, PSG and oximetry data) were obtained in 9.4% with an observed prevalence for OSAS of 28%[11]. Another limitation, and also related to self-selection is that NPV and PPV could not be determined accurately in a healthy workers population, since these parameters depend on the true prevalence of OSAS in the population studied. Therefore, we estimated NPV and PPV for several hypothetical prevalences. Finally, another drawback is the lack of a validation group. Instead of splitting our group, we decided to use all the employees for the estimation sample to obtain the best predictive statistics. A new validation study will be set up in the future.

One could argue about excluding severe insomnia in our definition of OSAS. Nevertheless, insomnia is one of the major sleep disorders with daytime symptoms that can result in false positive test results in OSAS questionnaires, especially in the presence of benign snoring, what occurs in 30-50% of the adult population[6]. On the other hand, co-existence of OSAS and insomnia and its possible interaction is subject of recent interest[40,41]. However, for screening for OSAS in a large population, we preferred to avoid this group of persons with overt insomnia especially in mild OSAS. For this purpose, we used the AIS, which has a simple scale constructed just as the ESS. Instead of the widely used cut-off value of six, we used the cut-off value of ten, with a NPV and PPV of respectively 94% and 88%[15]. This resulted in 13 out of 176 employees with overt insomnia according to the AIS (mean score 12.6 ± 1) Four employees with insomnia did indeed have an AHI ≥ 5. Two of them had an AHI >15 and these would be missed in applying the Philips two-step strategy.

We used nasal flow, instead of oximetry, simply because Philips NV funded this study and could easily provide a large number of Respironics nasal flow recorders (RU-sleeping, Philips-Respironics, Murrysville, USA). Three different single NFlow devices (SleepCheck, Flow Wizard, RU-sleeping) and one NFlow -oximetry combined device (ApneaLink) have been tested against PSG, all with high agreement for OSAS[18,39,42,43]. In two studies, oximetry and NFlow recording did have equivalent accuracy for diagnosing OSAS in the home setting[43,44]. Chai-Coetzer et al. reported less signal loss (3%) (with oximetry) compared to NFlow (9%)[39]. We found 8 out of 184 (4.3%) invalid NFlow recordings. The RU-sleeping flow recorder, used in this study, is a small recorder, which can even be sent by post and applied by the subject after reading the simple instruction.

The sensitivity of 63.1% found in our study can perhaps be interpreted as moderate. On the other hand, in a healthy almost asymptomatic population, this is probably the best to get.

Almost 2/3 of the true OSAS employees could be identified with this relatively simple two-step strategy. Due to large numbers of subjects without having the disease in this setting, the primary goal in screening is often to correctly exclude subjects without the disease. Hence, a test strategy with a high specificity (here 90%), and therefore relatively few false positives, and if possible, with a high negative predictive value, is essential.
The cost savings of our two-step strategy will be substantial for community screening compared with a quick questionnaire only; for the STOPBANGQ and BerlinQ this amounts to a difference of 324 (405-81) and 256 (337-81) diagnostic sleep studies (PSG or portable monitoring in clinic or at home), respectively, minus the extra costs for the cheap single channel nasal flow recording at home in the intermediate group of the Philips questionnaire which have to be done in 27.2% or 272 subjects in the given example in 1000 subjects with 10% OSAS prevalence (Table 6 and 7).
In conclusion, a two-step screening strategy with the Philips questionnaire and subsequent nasal flow recording is a promising way to screen for OSAS in a large group of healthy workers.

ACKNOWLEDGEMENTS

We are indebted to Piet-Hein van Mechelen (chairman Dutch Sleep Apnea organization (http://www.apneuvereniging.nl) for practical study coordination and support. We thank Reindert P.J. van Steenwijk, Pulmonary and Sleep Physician, Academic Medical Centrum, Amsterdam, for critical reviewing the manuscript.

DISCLOSURE STATEMENT

Home sleep recording by the Netherlands Slaap Instituut (http://nederlandsslaapinstituut.nl) and data-collection by HumanCapitalCare (www.humancapitalcare.nl) was sponsored by Philips NV, Eindhoven, the Netherlands.
The Department of Research Methodology, Measurement and Data Analysis, of the University of Twente, Enschede, the Netherlands received a grant of Philips NV, Eindhoven, the Netherlands for support in performing the statistical analyses. Philips NV Netherland had no access to employee data or study results, and was not involved in manuscript preparation or writing. None of the authors had any conflict of interest.

REFERENCES

1 Young T, Palta M, Dempsey J, Skatrud J, Weber S, Badr S. The occurrence of sleep-disordered breathing among middle-aged adults. N Engl J Med 1993;328:1230-5.

2 Baumel MJ, Maislin G, Pack AI. Population and occupational screening for obstructive sleep apnea: are we there yet? Am J Respir Crit Care Med 1997;155:9-14.

3 Seetho IW, Wilding JP. Screening for obstructive sleep apnoea in obesity and diabetes--potential for future approaches. Eur J Clin Invest 2013;43:640-55.

4 Boyko EJ. Ruling out or ruling in disease with the most sensitive or specific diagnostic test: short cut or wrong turn? Med Decis Making 1994;14:175-9.

5 Abrishami A, Khajehdehi A, Chung F. A systematic review of screening questionnaires for obstructive sleep apnea. Can J Anaesth 2010;57:423-38.

6 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;131:485-91.

7 Hrubos-Strom H, Randby A, Namtvedt SK et al. A Norwegian population-based study on the risk and prevalence of obstructive sleep apnea. The Akershus Sleep Apnea Project (ASAP). J Sleep Res 2011;20:162-70.

8 Kang K, Park KS, Kim JE et al. Usefulness of the Berlin Questionnaire to identify patients at high risk for obstructive sleep apnea: a population-based door-to-door study. Sleep Breath 2013;17:803-10

9 Simpson L, Hillman DR, Cooper MN et al. High prevalence of undiagnosed obstructive sleep apnoea in the general population and methods for screening for representative controls. Sleep Breath 2013;17:967-73.

10 Ramachandran SK, Josephs LA. A meta-analysis of clinical screening tests for obstructive sleep apnea. Anesthesiology 2009;110:928-39.

11 Gurubhagavatula I, Maislin G, Nkwuo JE, Pack AI. Occupational screening for obstructive sleep apnea in commercial drivers. Am J Respir Crit Care Med 2004;170:371-6.

12 Chung F, Yegneswaran B, Liao P et al. STOP questionnaire: a tool to screen patients for obstructive sleep apnea. Anesthesiology 2008;108:812-21.

13 Gilley WF, Uhlig GE. Factor analysis and ordinal data. Education 1993;114:258-64.

14 Soldatos CR, Dikeos DG, Paparrigopoulos TJ. Athens Insomnia Scale: validation of an instrument based on ICD-10 criteria. J Psychosom Res 2000;48:555-60.

15 Soldatos CR, Dikeos DG, Paparrigopoulos TJ. The diagnostic validity of the Athens Insomnia Scale. J Psychosom Res 2003;55:263-7.

16 Johns MW. Reliability and factor analysis of the Epworth Sleepiness Scale. Sleep 1992;15:376-81.

17 Buysse DJ, Reynolds CF, III, Monk TH, Berman SR, Kupfer DJ. The Pittsburgh Sleep Quality Index: a new instrument for psychiatric practice and research. Psychiatry Res 1989;28:193-213.

18 Grover SS, Pittman SD. Automated detection of sleep disordered breathing using a nasal pressure monitoring device. Sleep Breath 2008;12:339-45.

19 The AASM Manual for the Scoring of Sleep and Associated Events: Rules, Terminology and Technical Specifications. Westchester, Illinois: American Academy of Sleep Medicine, 2011.

20 Sleep-related breathing disorders in adults: recommendations for syndrome definition and measurement techniques in clinical research. The Report of an American Academy of Sleep Medicine Task Force. Sleep 1999;22:667-89.
21 Eijsvogel M, Schimsheimer R, Vos P, de Vries N, Raats C, Trooster R. CBO Richtlijn. Diagnostiek en behandeling van het obstructief slaapapneu syndroom bij volwassenen. van Zuiden. 2009.
22 Hays RD, Morales LS, Reise SP. Item response theory and health outcomes measurement in the 21st century. Med Care 2000;38:II28-42.
23 van den Berg SM, Glas CA, Boomsma DI. Variance decomposition using an IRT measurement model. Behav Genet 2007;37:604-16.
24 Bartholomew DJ, Steele F, Moustaki I, Galbraith JI. Analysis of multivariate social science data. Boca Raton, Florida: Chapman & Hall/CRC, 2008.
25 Jöreskog KG, Moustaki I. Factor analysis of ordinal variables: A comparison of three approaches. Multivariate Behavioral Research 2001;36:347-87.
26 Vermunt JK, Magidson J. Factor analysis with categorical indicators: A comparison between traditional and letent class approaches. In: Van der Ark A, Croon MA, Sijtsma K. New developments in categorical data analysis for the social and behavioural sciences. Mahwah: Lawrence Erlbaum Ass., 2005:41-62.
27 Siemons L, Ten Klooster PM, Taal E et al. Validating the 28-Tender Joint Count Using Item Response Theory. J Rheumatol 2011;38:2557-64.
28 Sass DA, Schmitt TA. A comparative investigation of rotation criteria within exploratory factor analysis. Multivariate Behavioral Research 2010;45:73-103.
29 Samejima F. Estimation of latent ability using a response pattern of graded scores. Psychometrika Monograph Supplement 1969;34:1-100.
30 Computationally efficient estimation of multilevel high-dimensional latent variable models. American Statistical Association; 2007.
31 Stevens JP. Applied multivariate statistics for the social sciences. Hillsdale, NJ: Erlbaum, 1980.
32 Hu L, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. Structural Equation Modeling: A Multidisciplinary Journal 1999;6:1-55.
33 Cowan DC, Allardice G, Macfarlane D et al. Predicting sleep disordered breathing in outpatients with suspected OSA. BMJ Open 2014;4:e004519.
34 Flemons WW, Whitelaw WA, Brant R, Remmers JE. Likelihood ratios for a sleep apnea clinical prediction rule. Am J Respir Crit Care Med 1994;150:1279-85.
35 Maislin G, Pack AI, Kribbs NB et al. A survey screen for prediction of apnea. Sleep 1995;18:158-66.
36 Rowley JA, Aboussouan LS, Badr MS. The use of clinical prediction formulas in the evaluation of obstructive sleep apnea. Sleep 2000;23:929-38.
37 Takegami M, Hayashino Y, Chin K et al. Simple four-variable screening tool for identification of patients with sleep-disordered breathing. Sleep 2009;32:939-48.
38 Caffo B, Diener-West M, Punjabi NM, Samet J. A novel approach to prediction of mild obstructive sleep disordered breathing in a population-based sample: the Sleep Heart Health Study. Sleep 2010;33:1641-8.
39 Chai-Coetzer CL, Antic NA, Rowland LS et al. A simplified model of screening questionnaire and home monitoring for obstructive sleep apnoea in primary care. Thorax 2011;66:213-9.

40 Beneto A, Gomez-Siurana E, Rubio-Sanchez P. Comorbidity between sleep apnea and insomnia. Sleep Med Rev 2009;13:287-93.
41 Luyster FS, Buysse DJ, Strollo PJ, Jr. Comorbid insomnia and obstructive sleep apnea: challenges for clinical practice and research. J Clin Sleep Med 2010;6:196-204.
42 Rofail LM, Wong KK, Unger G, Marks GB, Grunstein RR. The utility of single-channel nasal airflow pressure transducer in the diagnosis of OSA at home. Sleep 2010;33:1097-1105.
43 de Almeida FR, Ayas NT, Otsuka R et al. Nasal pressure recordings to detect obstructive sleep apnea. Sleep Breath 2006;10:62-9.
44 Rofail LM, Wong KK, Unger G, Marks GB, Grunstein RR. Comparison between a single-channel nasal airflow device and oximetry for the diagnosis of obstructive sleep apnea. Sleep 2010;33:1106-14.

Table 1. OSAS symptoms and related questions in questionnaires

	Symptoms		*Source*	*Detail*
One	*Major symptom:*			
	Excessive Daytime Sleepiness (EDS)	or	ESS > 10	Total score ESS
		or	Berlin-Q Q-8	Have you ever nodded off or fallen asleep while driving a vehicle? (yes, ≥1-2 times/week)
		or	AIS Q-8	Sleepiness during the day 0: None 1: Mild 2: Considerable 3: Intense (≥1)
or	≥ *2 of the following minor symptoms:*			
	Choking or gasping during sleep	or	N / A	
	Recurrent awakenings from sleep	or	AIS Q-2	Awakenings during the night. 0: No problem 1: Minor problem 2: Considerable problem 3: Serious problem or did not sleep at all (≥2)
	unrefreshed sleep	or	Berlin-Q Q-6	How often do you feel tired or fatigued after your sleep (≥3-4 time/week)
	Daytime fatigue	or	STOP-Q Q-2	Do you often feel tired, fatigued, or sleepy during the daytime? Yes/No
		or	Berlin-Q Q-7	During your wake time, do you feel tired, fatigued or not up to par ? (≥3-4 time/week)
	Impaired concentration	or	N/A	
and	*Absence of severe Insomnia*		AIS ≤ 10	Total score AIS

OSAS symptoms according to Chicago criteria[19]: EDS or ≥2 of the minor symptoms, not otherwise explained (severe insomnia was excluded). N/A: not applicable. ESS=Epworth Sleepiness Scale; Berlin-Q=Berlin Questionnaire; AIS=Athens insomnia Scale; STOP-Q= Stop questionnaire; Q=Question. See methods for OSAS definition.

Table 2. Morphometric, PSG and questionnaires characteristics of all employees categorized by AHI or OSAS.

		AHI <5 n 90	AHI≥5 n 86	No-OSAS n 111	OSAS N 65
Gender, Male	(n)	61 (67.8 %)	74 (86.0 %)**	80 (72.1 %)	55 (84.6 %)
Age	(year)	44.1 ±9.2	50.5 ± 8.0 ***	45.2 ± 9.2	50.6 ± 8.2 ***
BMI	(index)	26.2 ± 4.8	27.6 ± 3.6 *	26.4 ± 4.5	27.8 ± 3.8 *
RR Systolic	(mmHg)	137.2 ±15.6	139.9 ±16.1	137.5 ± 15.3	140.3 ± 16.9
RR diastolic	(mmHg)	82.2 ± 9.6	84.5 ± 8.4	80.0 ± 9.3	84.5 ± 8.8
Neck circumf.	(cm)	37.9 ± 3.2	39.7 ± 2.5 ***	38.2 ± 3.1	39.7 ± 2.5 ***
AHI	(index)	1.9 (0.0-4.9)	10.5 (8.0-21.2)***	3.4 (0.7-4.5)	13.6 (8.8-23.4)***
ODI 3%	(index)	1.5 (0.4-3.1)	5.7 (2.0-10.7)***	1.7 (0.4-3.9)	5.8 (2.4-14.1)***
Mean SO2%	(number)	96.0 (95.0-96.0)	95.0 (94.0-96.0)***	96.0 (95.0-96.0)	95.0 (94.0-96.0)***
RERAindex	(index)	0.0 (0.0-0.0)	0.0 (0.0-0.0)	0.0 (0.0-0.0)	0.0 (0.0-0.3)
Arousalindex	(index)	0.6 (0.2-2.3)	6.7 (2.9-13.5)***	1.1 (0.3-3.3)	7.0 (2.9-14.8)***
TST	(hours)	6.8 (6.3-7.3)	6.5 (6.1-7.3)	6.6 (6.2-7.3)	6.5 (6.1-7.4)
Sleeplatency	(min)	5.0 (2.0-12.0)	5.0 (3.0-10.0)	6.0 (2.0-12.0)	5.0 (2.5-10.0)
WASO	(min)	26.0 (15.8-45.0)	31.5 (20.8-57.3)	26 (15.0-96.0)	34 (22.0-59.6)*
Sleepefficiency	(%)	94.0 (91.0-96.0)	93.0 (86.0-95.3)*	94.0 (90-96)	92.0 (86-95)*
NREM 1 / TST	(%)	12.5 (8.0-16.0)	17.0 (11.0-21.3)***	14.0 (9.0-17.0)	18.0 (11.0-22.5)**
NREM II / TST	(%)	49.0 (44.0-54.3)	48.0 (43.0-54.0)	49.0 (44.0-54.0)	48.0 (43.0-54.0)
NREM III /TST	(%)	13.5 (8.8-18.3)	11.0 (7.0-16.0)*	13.0 (8.0-18.0)	11.0 (6.0-16.0)*
REM / TST	(%)	16.0 (13.0-20.0)	15.0 (11.0-18.0)**	16.0 (12.0-19.0)	14.0 (11.5-17.0)*
Nasal flow	(index)	7.2 (3.5-12.1)	16.0 (9.6-23.7)***	8.0 (3.8-14.2)	16.8 (8.4-24.0)***
ESS	(score)	5.0 (2.0-7.0)	5.0 (3.0-7.0)	4.0 (2.0-7.0)	6.0 (3.0-8.0)*
ESS ≥ 10	(n)	9 (10.0 %)	7 (8.1 %)	9 (8.1 %)	7 (10.8 %)
AIS >10	(n)	9 (10.0 %)	4 (4.7 %)	11 (9.9 %)	2 (3.1 %)
PSQI > 5	(n)	34 (37.8 %)	25 (29.8 %)[c]	36 (32.4 %)	23 (36.5 %)[c]
Berlin-Q high-risk [b]	(n)	29 (32.2 %)	35 (40.7 %)	34 (30.6 %)	30 (46.2 %)
Stop-Q high-risk [b]	(n)	33 (36.7 %)	47 (54.7 %)*	38 (34.2 %)	42 (64.6 %) ***
StopBang-Q high-risk [b]	(n)	41 (45.6 %)	61 (70.9 %)**	50 (45.0 %)	52 (80.0 %) ***

In all 176 subjects, morphometric measurements, PSG, separate single nasal flow recording and different questionnaires were taken. Subjects are categorized according to AHI (AHI: <5 or ≥5) and OSAS (no OSAS or OSAS); OSAS is defined as AHI 5-14 + symptoms (derived from questionnaires) and without insomnia (AHI<10) or an AHI>15 without restriction (see also methods); n = number; [b] Q = questionnaire; [c] two missing PSQI questionnaires; data expressed as numbers with percentage, mean ± SD or data expressed as median with interquartile range; * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$ for differences between AHI: <5 versus ≥5 and for no-OSAS versus OSAS.

Table 3. Factor loadings of Philips Questionnaire items

Philips Questionnaire	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6
	Snoring	Snoring severity	Tiredness	Observed apneas	Sleep quality	Daytime well-being
Q-1 Snoring (BerlinQ)	1.574					
Q-2 Snoring loudness (BerlinQ)		0.849				
Q-3 Snoring frequency (BerlinQ)		0.433				
Q-4 Snoring bothered others (BerlinQ)		0.578				
Q-5 Observed apneas (BerlinQ)				1.018		
Q-6 Tiredness after sleep (BerlinQ)			0.817			
Q-7 Tiredness during wake time (BerlinQ)			0.681			0.377
Q-8 Nodding off while driving (BerlinQ)	---	---	---	---		---
Q-9 Nodding off frequency (BerlinQ)				0.444		
Q-10 High blood pressure (BerlinQ/StopQ)						0.400
Q-11 Snoring (StopQ)		0.899				
Q-12 Tiredness (StopQ)			0.436			0.540
Q-13 Observed apneas (StopQ)				0.839		
Q-14 Sleep induction (AIS)					0.515	
Q-15 Awakenings nighttime (AIS)					0.693	
Q-16 Early final awakening (AIS)					0.695	
Q-17 Total sleep duration (AIS)					0.671	
Q-18 Overall sleep quality (AIS)					0.584	
Q-19 Daytime well-being (AIS)			0.432			0.638
Q-20 Daytime functioning (AIS)						0.770
Q-21 Daytime sleepiness (AIS)						0.597

Factor loadings > 0.364 Q-number = questions number Philips questionnaire; BerlinQ = Berlin questionnaire; StopQ = Stop questionnaire; AIS = Athens insomnia scale.

Table 4. Variables predicting OSAS

	Odds Ratio (CI)	Sig.	Coefficient[#]
AIS scoring <10 vs. >=10	16.9 (1.7-165)	0.015	2.828
STOPBANG-Q High vs. Low	15.57 (2.5-95)	0.003	2.745
STOPBANG-Q Intermediate vs. Low	2.62 (0.8- 8.6)	0.112	0.965
Berlin-Q High vs. Low	103 (4.0-2697)	0.005	4.640
Berlin-Q Intermediate vs. Low	4.88 (1.0-23.0)	0.045	1.584
Age >45 vs. <=45	2.20 (1.0-4.9)	0.052	0.790
Berlin-Q Q-5 (Breathing stops) Yes vs. No	2.25 (0.6-7.9)	0.207	0.810
Constant	0.002		-6.322

CI= confidence interval; # derived coefficients for formula Philips-Q to compute individual risk for OSAS. AIS=Athens Insomnia Scale; STOPBANG-Q= STOPBANG questionnaire with scoring points low (≤ 1), intermediate (2-4) and high (≥5) risk for OSAS Berlin-Q= Berlin questionnaire with scoring points low (0), intermediate (1-2) and high (3) risk for OSAS.

Table 5. Test characteristics Philips questionnaire and two-step strategy.

	Sensitivity	Specificity
Low cut off only (35%)	78.5% (51/65)	70.3% (78/111)
High cut off only (55%)	33.8% (22/65)	95.5% (106/111)
35-55% + NFlow recording (15)[#]	65.5% (19/29)	78.6% (22/28)
Final two-step screening strategy	63.1% (41/65)	90.1% (100/111)

[#] Employees in the intermediate cut off group after results with NFlow (Nasal Flow) recording (cut-off value: index of ≥ 15).

Table 6. Outcome two-step screening strategy according to prevalence of OSAS.

Prevalence	NPV [#]	PPV [#]	NFlow % [+]	Sleep clinic % [+]
5%	97.9%	25.1%	26.2%	12.6%
7.5%	96.8%	34.1%	26.7%	13.9%
10%	95.6%	41.5%	27.2%	15.2%
15%	93.3%	52.9%	28.1%	17.9%
20%	90.7%	61.5%	29.1%	20.5%

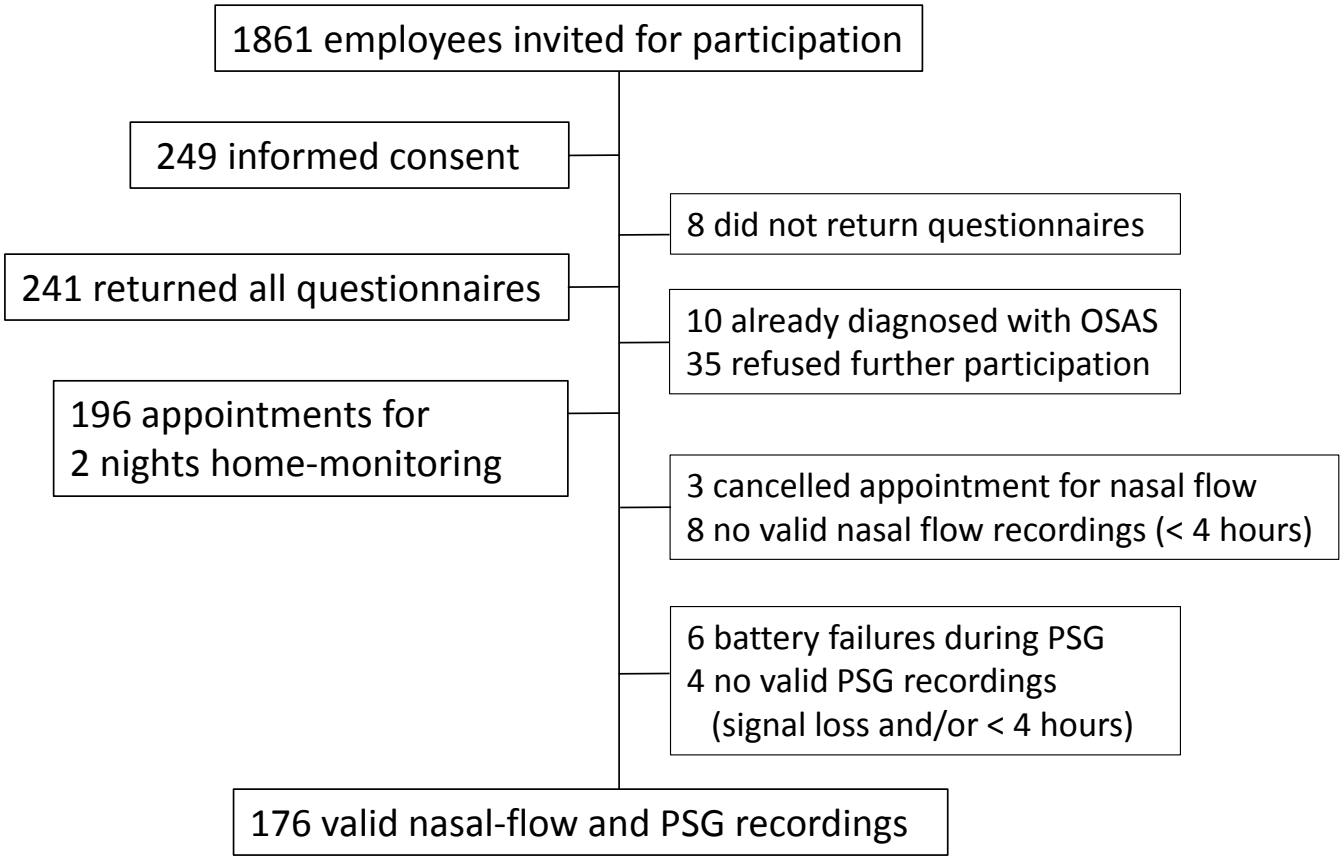
[#] negative (NPV) and positive predicting value (PPV).

[+] percentage of subjects invited for NFlow (nasal flow) recording or percentage of subjects referred to sleep clinic for P(S)G.

Table 7. Estimated outcome of three screening strategies

	NPV	TP	PPV	FP
BerlinQ	91.2 %	46/1000	12.0 %	337/1000
STOPBANGQ	91.0 %	51/1000	11.2 %	405/1000
Two-step strategy with PhilipsQ	95.6 %	37/1000	41.5 %	81/1000

Estimated outcome in 1000 subjects with 10% prevalence of obstructive sleep apnea syndrome.
NPV= negative predictive value; PPV= positive predictive value; TP= true positives; FP= false positives;
BerlinQ = Berlin Questionnaire; STOPBANGQ= STOPBANG questionnaire; PhilipsQ= Philips questionnaire;
Two-step strategy = PhilipsQ with only additional nasal flow recording for the intermediate risk result.


1861 employees invited for participation
249 informed consent
8 did not return questionnaires
241 returned all questionnaires
10 already diagnosed with OSAS
35 refused further participation
196 appointments for
2 nights home-monitoring
3 cancelled appointment for nasal flow
8 no valid nasal flow recordings (< 4 hours)
6 battery failures during PSG
4 no valid PSG recordings
(signal loss and/or < 4 hours)
176 valid nasal-flow and PSG recordings