Pacemaker guided screening for severe sleep apnea, a possible option for patients with atrial fibrillation: A systematic review and meta-analysis

Peer-reviewed author version

Wyckmans, Martin; Tukanov, Eldar; Winters, Robbe; Stinissen, Robin; VERMEULEN, Helene; DENDALE, Paul & DESTEGHE, Lien (2021) Pacemaker guided screening for severe sleep apnea, a possible option for patients with atrial fibrillation: A systematic review and meta-analysis. In: Pacing and Clinical Electrophysiology, 44 (8) , p. 1421-1431.

DOI: 10.1111/pace.14256
Handle: http://hdl.handle.net/1942/34417

Pacemaker guided screening for severe sleep apnea, a possible option for patients with atrial fibrillation: A systematic review and meta-analysis

Martin Wyckmans[a], BSc, Eldar Tukanov[a], BSc, Robbe Winters[a], BSc, Robin Stinissen[a], BSc, Helene Vermeulen[b], MSc, Paul Dendale[a,c,] MD PhD, Lien Desteghe*, [a,c,d,e] MSc PhD

a: Faculty of Medicine and Life Sciences, Hasselt University, Martelarenlaan 42, 3500 Hasselt, Belgium

b: Interuniversity Institute for Biostatistics and statistical Bioinformatics, Data Science Institute, Hasselt University, Martelarenlaan 42, 3500 Hasselt, Belgium

c: Heart Center Hasselt, Jessa Hospital, Stadsomvaart 11, 3500 Hasselt, Belgium

d: Research Group Cardiovascular Diseases, University of Antwerp, Prinsstraat 13, 2000 Antwerp, Belgium

e: Cardiology Department, Antwerp University Hospital, Drie Eikenstraat 655, 2650 Edegem, Belgium

*Address for correspondence:

Lien Desteghe

Hasselt University

Heart Centre, Jessa Hospital

Stadsomvaart 11

3500 Hasselt

Belgium

Phone: 32-11-33 70 88

Fax: 32-11-33 70 38

Email: liendesteghe@gmail.com | lien.desteghe@uantwerpen.be | lien.desteghe@uhasselt.be

Funding : none

Disclosures : none

Brief title: CIED for sleep apnea screening

Word count: 4262 words (excluding abstract, references, figures and tables)

Keywords: atrial fibrillation -- obstructive sleep apnea -- implantable defibrillator -- artificial pacemaker -- polysomnography -- ambulatory electrocardiography

Abstract

Introduction

Obstructive sleep apnea is often underdiagnosed in atrial fibrillation (AF) patients although it is an important risk factor. A systematic review and meta-analysis was performed to assess which techniques cardiac implantable electronic devices (CIED) and Holter monitors use to screen for sleep apnea (SA), and to evaluate if these are suitable for AF patients from a diagnostic accuracy perspective.

Methods

The search was conducted in accordance with the PRISMA-guidelines. PICO was defined as (P) patients with AF, (I) Holter monitors or CIED suitable for screening for SA, (C) overnight polysomnography (PSG), (O) positive screening with subsequent positive polysomnographic diagnosis of SA. Optimal index test cut-off points corresponding to reference test cut-off for severe SA (PSG-AHI≥30) were compared. Meta-analysis was conducted for the diagnostic odds ratio (DOR), with forest plot and ROC-curve for summary DOR.

Results

A total of 5 prospective cohort studies (n=192) were included in the systematic review of which 4 studies (n=132) were included in the meta-analysis. All included studies use transthoracic impedance measurement as a screening parameter. No studies evaluating Holter monitors were included. The population consisted of patients indicated for pacemaker implantation. The summary DOR was 27,14 (8,83; 83,37), AUC was 0,8689 (0,6872; 0,9456) and Q* was 0,8390 (0,7482; 0,9013).

Conclusion

At optimal pacemaker-cut-off, pacemaker-guided screening for severe SA in patients with AF can be an effective triage tool for clinical practice. Further studies with larger sample sizes are needed to strengthen the evidence for this conclusion.

Introduction

Atrial fibrillation (AF) is the most common type of cardiac arrhythmia with an estimated prevalence of 1 to 2% in the general population.[1] Although this condition is associated with other comorbidities, several studies suggest that AF is independently associated with an increased mortality, further highlighting the importance of early diagnosis and treatment.[2] Several modifiable risk factors have been identified for the development and progression of AF, including obstructive sleep apnea (OSA).[3] Both pathologies share risk factors, such as obesity and hypertension, and are closely related to the presence of various cardiovascular diseases.[4] Therefore, current literature strongly suggests that a significant relation exists between AF and the presence of OSA.[5]

Obstructive sleep apnea is the most common type of sleep disordered breathing, with an estimated prevalence of 3 to 7% in the general population.[6] If untreated, severe OSA is associated with increased cardiovascular mortality of any cause.[7] In a population with AF, OSA is prevalent in up to 62% of the population, further highlighting the importance of the interplay between both pathologies.[8] Early diagnosis and appropriate treatment of OSA is linked with a decrease in recurrence of AF and other arrhythmias, making it an important factor for the treatment of these cardiovascular diseases.[9] Although OSA is very common, it remains underdiagnosed in patients with AF as those patients often do not report the typical OSA-related symptoms, while further increasing the severity of arrhythmias.[10] Available screening options or diagnostic tools for OSA include questionnaires, scoring systems and Home Sleep Apnea Testing (HSAT).[8,31] Effectiveness of these screening tools, however, has yet to be confirmed in an AF population, while the current golden standard for OSA diagnosis remains polysomnography (PSG) during an overnight stay at the sleep clinic.[11]

Cardiac implantable electronic devices (CIED) may also be used in the diagnosis and management of arrhythmias such as AF.[12] Some CIED have incorporated sleep apnea (SA)-

detection methods making use of algorithms. These devices could be used as a fast and practical method for early detection of SA in patients with AF. The aim of this review is to determine the added value and accuracy of Holter monitors and CIED in patients with AF for detection of SA. The aim of the meta-analysis is to evaluate the accuracy of transthoracic impedance measurement to screen for severe SA.

Methods

Eligibility criteria

The PICO was defined as (i) patients with AF for population; (ii) Holter monitors or CIED suitable for screening for SA; (iii) standard practice of screening for SA, being overnight PSG, as control and (iv) positive screening with subsequent positive polysomnographic diagnosis of SA with an apnea-hypopnea index (AHI)≥15 as outcome. Only published, full-length articles from 2000 and onwards were considered. Language restrictions were set for English and Dutch.

Inclusion criteria were: the focus of the study is screening for SA; the index test is compared to the reference test; the reference test in the study is overnight PSG. Exclusion criteria during abstract screening were: diagnosis of SA defined as PSG-AHI<15; exclusion of patients based on a previous diagnosis of SA; article screens for central sleep apnea (CSA) only; and on-going studies. Full-text article exclusion was conducted based on the following criteria: definition of SA was AHI<15, focus of the study is not diagnosis or screening for SA, no validation with PSG, sample size includes <10% patients with AF (to ensure that the device could be used in patients with AF).

Data search

Articles were gathered from PubMed, Embase and Cochrane. The search was concluded in February 2020 in accordance with the PRISMA-guidelines.[13]

The following keywords were used ‘sleep apn(o)ea’, ‘atrial fibrillation’, ‘arrhythmia’ and ‘dysrhythmia’ and the Medical Subject Headings (MeSH) terms ‘sleep apnea syndrome’, ‘sleep disordered breathing’ and ‘Epworth Sleepiness Scale’. The search was expanded by identifying synonyms or closely-related words. References of included articles were searched manually to identify any missing articles. The full search strategy can be found in supplementary table 1. Four reviewers (RS, ET, RW and MW) independently assessed titles and abstracts for inclusion of relevant references, followed by screening of full-text articles. Each article was assessed by at least two of these researchers. Publications were included in this review if Holter monitors and/or CIED detecting SA, were evaluated in patients diagnosed with AF. Publications were excluded in case of case reports, review articles, letters to the editor, animal studies, poster abstracts and articles published before 2000.

Data collection

A checklist of data points was made. The checklist included author and year of publication, data collection period, percentage of the population with AF, index test, reference test, index test comparison points for PSG-AHI≥15 and PSG-AHI≥30, statistical measurements (i.e. sensitivity, specificity, positive predictive value and negative predictive value) of comparison points PSG-AHI≥15 and PSG-AHI≥30, study details, sample size, main findings of the study, funding, country where the study was performed and cut-off point of index test recommended by the article.

Risk of bias and applicability

Risk of bias and concerns of applicability in individual studies were evaluated using the QUADAS-2 tool.[14] Studies were graded within four domains for risk of bias; (i) methods used by researchers for patient selection, (ii) administration and interpretation of index test, (iii) administration and interpretation of reference standard, and (iv) flow and timing of study conduct. The first three of

these domains were also evaluated for concerns of applicability. These domains are graded as 'high risk', 'low risk' or 'unclear' by four reviewers (RS, ET, RW, MW).

Statistical analysis

Results of individual studies were presented at a common cut-off point for the reference test, namely PSG-AHI≥30, which corresponds to a diagnosis of severe SA. This cut-off point was chosen instead of PSG-AHI≥15, as statistical reporting at PSG-AHI≥15 by the included studies was insufficient for meta-analysis. In case the respiratory disturbance index (RDI) was used, the RDI corresponded to the AHI.

2x2 contingency tables were extracted from the original studies. From these contingency tables, sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV) were calculated, including their respective confidence intervals at 5% significance level. Calculation of confidence intervals was done by the Wilson score interval for binomial proportion with continuity correction.[15] Diagnostic odds ratio (DOR) for individual studies was calculated, with a Haldane correction applied in case one of the cells in the contingency table was equal to zero.[16]

Meta-analysis based on summary data was conducted using DOR pooling, to circumvent the threshold effect.[17] One study, *Shalaby et al. (2006)*[21], was removed from the meta-analysis due to insufficient reporting of statistical results at optimal test performance. Studies were compared at an equal cut-off point for the reference test and the optimal cut-off point reported for the screening tool. Analysis of heterogeneity was performed by Cochran Q test, I^2 and H^2. As these tests have low power in meta-analyses with a small amount of studies, individual studies were evaluated qualitatively based on study design and investigated population to determine whether fixed-effect or random-effects is more appropriate. Due to the similarity of the included studies, a fixed-effect model was deemed appropriate. The Mantel-Haenszel estimator was used in StatsDirect to calculate the weighted average of the DOR's observed in the primary studies.[18,19]

Weighting of individual studies was based on the amount of false results relative to the individual study population.[19] The DOR's of individual studies and the pooled DOR were visualised using a forest plot. A symmetric ROC-curve with confidence bands was generated from the pooled DOR using MS Excel. Individual study ROC-curves were generated, with the optimal cut-off points visualised as reported in the included articles. In case a Haldane correction was performed for the meta-analysis, both the reported and corrected optimal cut-off points were visualised.

From the symmetric ROC-curve, the Area Under the Curve (AUC) and the optimum Q*, i.e. the point where the sensitivity and specificity are equal, were extracted.

Deeks' test for publication bias evaluation was not used, as this test has a low power in meta-analyses with a small amount of studies and limited individual sample sizes.[20] Therefore, publication bias was evaluated on a qualitative basis for six factors as proposed by Ioannidis (2005).[28]

Results

Search results and study characteristics

The search for studies assessing CIED and Holter monitors to detect SA amongst patients with AF yielded 250 results (Figure 1). After removal of duplicates, 199 publications were screened on title and abstract. A total of 24 articles were screened on full text, five of which met the inclusion criteria: *Shalaby et al. (2006)*[21], *Defaye et al. (2014)*[22], *Barbieri et al. (2018)*[23], *Chen et al. (2019)*[24] and *Defaye et al. (2019)*[25]. A summary of study characteristics can be found in supplementary table 2.

Three studies were found that use the ApneaScan algorithm; *Barbieri et al. (2018)*, *Chen et al. (2019)* and *Defaye et al. (2019)*. The ApneaScan algorithm continuously measures transthoracic impedance by sending a low-voltage signal from the lead and the implantable cardioverter defibrillator (ICD) scan. Changes in impedance are used to create a waveform that is used for

respiration counting. Devices with this algorithm are enabled when the patient is expected to sleep. At night, the algorithm detects apnea-hypopnea events by assessing whether an amplitude threshold is reached within a time interval.

This amplitude threshold baseline value is calculated using the average amplitude between inhalation and expiration and is continually updated for small changes in the respiratory cycle. If the amplitude threshold is not exceeded for a time interval longer than the duration threshold, an apnea-hypopnea event is recorded. Generally, a duration threshold of 10 seconds is used. An apnea event is defined as two consecutive large breaths, where the time interval between breaths exceeds the duration threshold. A hypopnea event is also defined as two consecutive large breaths where the time interval between breaths exceeds the duration threshold, but where the time between the two consecutive large breaths also includes small breaths that do not exceed the amplitude threshold.[23-25] Based on these results, an average amount of apnea-hypopnea events per hour is calculated, also called the Apnea-Hypopnea-Index (AHI). Pacemaker results are abbreviated as PM-AHI and polysomnography results are abbreviated as PSG-AHI.

Other algorithms use a comparable technique to the ApneaScan.[26] For example, SAM-RDI is the respiratory disturbance index evaluated by the sleep apnea monitoring algorithm used in Defaye et al. (2014).[22] This SAM-RDI is similar to the PM-AHI.

Results of individual studies

The five articles included were dated from 2006 to 2019 (supplementary table 2). All articles investigated pacemakers equipped with transthoracic impedance sensors to screen for SA using algorithms. Sample size varied from 21 to 60. No articles that were included, used Holter monitors to screen for SA. Three of the five studies enrolled patients indicated for a pacemaker before the pacemaker was implanted. Shalaby et al. (2006) included patients who had a pacemaker implanted before enrolment and Barbieri et al. (2018) included heart failure patients receiving cardiac resynchronization therapy (CRT) upgrading. Statistical results for the recommended

pacemaker cut-offs, compared with cut-offs of PSG-AHI≥15 and PSG-AHI≥30, were calculated. Statistical results could not be calculated for *Shalaby et al. (2006)* for the recommended cut-off point of PM-AHI≥39 at PSG-AHI≥30. *Defaye et al. (2014)* recommend a cut-off point of SAM-RDI≥20, with sensitivity 88,89%, specificity 84,62%, PPV 88,89%, NPV 84,62% and DOR 44,00. *Barbieri et al. (2018)* recommend a cut-off point of PM-AHI≥32, with sensitivity 83,33%, specificity 60,00%, PPV 45,45%, NPV 90,00% and DOR 7,50. *Chen et al. (2019)* recommend a cut-off point of PM-RDI≥41, with sensitivity 81,82%, specificity 88,64%, PPV 64,29%, NPV 95,12% and DOR 35,10. *Defaye et al. (2019)* recommend a cut-off point of PM-AHI≥30, with sensitivity 100%, specificity 76,47%, PPV 66,67%, NPV 100% and DOR 51,00. The statistical results of the recommended cut-off points at PSG-AHI≥30 are graphically illustrated in Figure 2.

As the recommended pacemaker cut-off based on the cut-off PSG-AHI≥30 for the reference test was not reported with statistical values by *Shalaby et al. (2006)*, the study was removed from the meta-analysis. Only two studies included recommended PM-AHI cut-offs based on a PSG-AHI≥15 cut-off point, namely *Shalaby et al. (2006)* and *Chen et al. (2019)*. *Shalaby et al. (2006)* recommend a cut-off point of PM-AHI≥37, without reporting of statistical values. In *Chen et al. (2019),* the best cut-off value reported for a PSG-AHI≥15 cut-off point was PM-RDI≥26 with sensitivity 100%, specificity 70.6%, PPV 67.7% and NPV 100%.

Meta-analysis

A meta-analysis was conducted using four studies. Statistical heterogeneity tests showed no significant heterogeneity (Cochran Q = 1,56561 (with p-value = 0,6672); $H^2 = 1$; $I^2 = 0\%$ [0%; 67,9%]). A qualitative evaluation of the study protocol and the population of the included studies was performed. The investigated population consists of patients with an indication for pacemaker implantation with a similar age and body mass index distribution. All patients were included consecutively. Intervention and control were performed similarly, with pacemaker estimation of

AHI-score and PSG-measurement done concurrently in all studies. Three of the four pacemakers used the same algorithm for AHI-score, with the fourth having similar parameters in their algorithm, making the results comparable. Based on this evaluation, it was concluded that heterogeneity was negligible and fixed-effect model was appropriate for the meta-analysis. Results for the summary DOR were calculated using a fixed-effect model based on the Mantel-Haenszel estimator and were visualised using a forest plot (Figure 3).

Summary DOR was 27,14 (8,83; 83,37). AUC was 0,8689 (0,6872; 0,9456) and Q* was 0,8390 (0,7482; 0,9013). The mean of the Q* for the four studies were used to visualise the individual studies on the symmetric ROC curve (Figure 4). Three of the studies were noted to be above the mean, with one outlier *(Barbieri et al. 2018)* below the lower confidence bound of the ROC curve. Symmetric ROC-curves were generated for the individual studies and statistical parameters for optimal screening cut-off were plotted as reported by the original articles (Figure 5). Additionally, the statistical parameters for Defaye et al. 2019 were recalculated using the Haldane correction, giving a sensitivity of 94.44% and a specificity of 75%.

Risk of bias in individual studies

The results of the risk of bias using the QUADAS-2 tool are summarised in Figure 6.[14] One study *(Shalaby et al. 2006)* had no consecutive or random sample of patients enrolled, and the selection of patients could therefore introduce bias. Two studies *(Barbieri et al. 2018, Chen et al. 2019)* had not blinded the interpreters of the index test results from the results of the reference standard, and the index test could therefore introduce bias.

In three studies *(Defaye et al. 2014, Chen et al. 2019, Defaye et al. 2019)*, not all patients were included in the analyses, while in two studies *(Defaye et al. 2014, Chen et al. 2019)*, not all patients received the reference standard, resulting in a high risk of bias in flow and timing. Due to this patient drop-out, the characteristics of these patients were lost, altering the results of the

population that was investigated. Reasons for patient drop-out were patients not wanting to undergo both index test and reference standard, and malfunctioning of either transthoracic impedance sensors or reference standard.

Applicability

Patient selection and reference standard were noted to have low concerns of applicability across all five articles (Figure 7). Concerns regarding the index test were rated as high for *Barbieri et al. (2018)* due to inclusion of CSA, whereas these concerns were rated low for the four remaining articles.

Publication bias

Due to the small amount of included studies and limited individual sample sizes, Deeks' test would have a low power to detect publication bias.[27] Therefore, publication bias was evaluated on a qualitative basis for six factors as proposed by Ioannidis (2005).[28] Of note, the studies conducted have a small average sample size (n=33). This means the individual studies have less power. A publication bias towards positive results is deemed likely. No explicit conflicts of interest were mentioned in the articles described above. Sponsoring and other support in the various articles is described in supplementary table 2.

Discussion

This review shows that transthoracic impedance measurement of pacemakers can be a promising technique to screen for SA. Nevertheless, considerations should be made to interpret the results effectively.

Early detection of OSA in a population with AF remains an underestimated necessity.[29] Currently, the golden standard for detection and diagnosis of OSA is a polysomnographic sleep study. Usage of such an in-hospital overnight sleep study remains challenging due to high costs and long waiting lists.[30] The introduction of ambulatory screening tools could refer patients more efficiently, if the tools used have a proper balance of sensitivity and specificity, and improve current referral practices in a healthcare setting. The literature search was restricted to CIED and Holter monitors, but other screening tools could possibly be effective for OSA screening in a population with AF, such as the Berlin Questionnaire.[31]

A notable finding was that only one article, *Gonçalves et al. (2019)*, compared a population composed of patients with AF to a population without AF.[32] However, only new diagnoses of AF were included, which could influence the results. For this reason, and for not meeting the inclusion criteria, the article was not included in the systematic review. Current literature suggests that more extensive studies are needed for an adequate screening tool specifically meant for a population with AF, as its prevalence is increasing.[33]

It should also be pointed out that, even though there are articles published that used Holter monitors for screening for sleep apnea, none of those found met the eligibility criteria. An example is the study of Hsu et al. (2020), which evaluated the use of a single-lead ECG patch to screen for sleep apnea, but this study excluded patients with persistent AF.[34]

Screening tools should focus on the detection of at least moderate SA (PSG-AHI≥15), as treatment mostly starts above this threshold. Only two studies included recommended PM-AHI cut-offs based on a PSG-AHI≥15 cut-off point, namely *Shalaby et al. (2006)* and *Chen et al. (2019)*.[21,24] It is noteworthy that the included articles are generally pointed towards patients with severe SA (PSG-AHI≥30). In this case, parts of the population that could also benefit from SA treatment may be missed. During the literature search, there was also a study that included mild SA (PSG-AHI≥5) as their focus group.[32] However, this is a group of patients that do not all necessarily require treatment and therefore this study was excluded from the review. Thus, as treatment is recommended for patients with moderate SA, screening for the PSG-AHI≥15 subgroup would be more useful. Deciding on a common reference test threshold to report in articles is recommended, to better evaluate the screening tools.

Optimal cut-off points for screening tools, when evaluated in a scientific setting, that use a continuous measure like AHI are based on the highest sum of sensitivity and specificity. Deviating from the optimal cut-off for increased accuracy of one statistical measurement will inevitably result in a decreased accuracy of the other statistical measurement.[35]

The symmetric ROC curve shows that Q*, i.e. the point where sensitivity and specificity are equal, is 0,8390 (0,7482; 0,9013). Importantly, the meta-analysis cannot make definite conclusions on the sensitivity and specificity of the test, only on the diagnostic accuracy according to the pooled DOR. Caution should be used with the interpretation of these results.

As a theoretical example using the summary DOR as a reference, at 90% specificity, estimated sensitivity is 75,10%. In clinical practice, this means that more missed diagnoses would occur and less people would be unnecessarily referred. On the other hand, when looking at 90% sensitivity, estimated specificity is 75,10%. Reciprocally, this results in less missed diagnoses and an increase in unnecessary referrals. In this case, priority can be given to a high sensitivity, as a proper screening tool aims not to miss diagnoses.

In the meta-analysis, individual study ROC-curves were generated and the reported optimal cut-off points were plotted (Figure 5). For Defaye et al. 2019, the reported cut-off point did not fit into the ROC-curve, while the cut-off point with the Haldane correction did. It may be possible that the statistical parameters corresponding to the optimal cut-off in the original article are an incidental result, not representing the true efficacy of the device.

The studies included mentioned occurrences of missing data. Most notable was *Barbieri et al. (2018)* in which 20 of the 41 recordings failed to yield measurements.[23] The article attributes this to the high amount of included patients with CSA. Relative to the number of patients with either OSA or CSA, the percentage of failures (50%) was equal for both the OSA and CSA subgroups. This makes their explanation unlikely as a cause of failed recordings.

Importantly, of the recordings that did yield measurements in the aforementioned study, CSA patients were still represented more numerously compared to other studies analysed in the review. Of the 21 valid recordings, 10 were of patients with CSA, being 47.6% of total valid recordings. For comparison, *Chen et al. (2019)* had one patient (1,8%) with CSA, *Defaye et al. (2019)* had seven patients (28%) with CSA. *Defaye et al. (2014)* did not differentiate between CSA and OSA. *Barbieri et al. (2018)* further state that AP scan® measurements commonly overestimate sleep apnea, which causes a lower specificity.[23] A possible explanation for this overestimation is that CSA events are overestimated by the transthoracic impedance sensors. *Luo et al. (2009)* showed that recording of the chest and abdominal movement by respiratory inductance plethysmography in conventional PSG overestimated the frequency of CSA events.[36] This overestimation has since been adjusted during PSG by measurement of oesophageal pressure (Pes) and diaphragm electromyogram (EMGdi).[36] As the pacemaker algorithm uses chest movement for screening and cannot use Pes or EMGdi for correction, the overestimation

could possibly be extrapolated to the pacemakers investigated. An overestimation in CSA-events by pacemakers might include patients in the category of severe SA before diagnosis is confirmed via PSG. When these patients undergo PSG-study, they do not receive the diagnosis of severe SA and therefore are classified as a false-positive, decreasing specificity of the pacemaker algorithm. This might be an explanation why *Barbieri et al. (2018)* has a considerably worse specificity. As such, overestimation of AHI in patients with CSA is possible when using CIED based on transthoracic impedance.

When using these devices, it is important to evaluate the goal of implementation in standard practice. Substitution of PSG is not feasible, because PSG is still the golden standard.[11] Instead, using this technology as a form of triage (i.e. first testing) is possible. The effectiveness of a triage system, and whether it should be used in clinical practice, is best evaluated using PPV and NPV. To reduce the amount of unneeded referral, preference goes to a high PPV. To reduce the amount of missed diagnoses, preference goes to a high NPV. For example, in *Defaye et al. (2019)*, the reported PPV and NPV would result in 33,30% of referrals to be unneeded, but no missed diagnoses would occur.[25]

The pacemakers evaluated did not have screening for SA as the primary goal in daily clinical practice.[22-25] Using these devices to screen for SA incurs no immediate costs for healthcare providers or patients. It could even be stated that this is a manner to decrease costs, as this data is readily available and can lead to better screening and treatment of SA in a patient population with already a cardiovascular history. This, in turn, may lead to a better quality of life and possibly less cardiovascular events, as treatment of SA in these patients will decrease or prevent risk factors associated with SA, such as AF and stroke.[9]

There are no specific obstacles for the usage of the screening algorithms in CIED devices to screen for SA. Referral for PSG to validate the screening, does have known obstacles. Long

waiting lists and fear for the procedure can make patients defer or decline further testing by in-hospital sleep testing.[30] These factors should not decrease the feasibility of the devices for screening purposes. Using the devices will increase the number of people, both true positives and false positives, referred to the sleep clinic for PSG. However, use of effective screening tools allows for less missed diagnoses, which should be the primary goal.

Limitations of the review method

Only articles in the English and Dutch language were reviewed. Embase, Medline and Cochrane were the only databases consulted. No studies were identified that evaluate the device-types in a population of exclusively AF, as was the goal of the original search strategy. This meant an inclusion criterion had to be adjusted. Studies describing a population with >10% patients with AF were taken into account, to be sure that devices investigated were suitable for patients with AF. This makes it possible that studies which investigate devices suitable for patients with AF, but do not have >10% patients with AF in their study, were missed. No definite conclusions can be made concerning the sensitivity and specificity of the transthoracic impedance measurement in pacemakers as a screening tool, since the analysis of the DOR only allows statements to be made on diagnostic accuracy as a whole.

Conclusion

Screening for SA by transthoracic impedance sensors in CIED is a promising technique. At optimal cut-off, pacemakers can be useful to screen for severe SA and serve as a triage-tool. However, more studies with a larger group of patients and evaluated at a common cut-off for the index and reference test are needed to strengthen the evidence, as results are not conclusive enough. At the moment, over-reliance on these devices for SA testing is not recommended and

decisions should be clinically correlated. Studies composed exclusively of patients with AF will allow a better evaluation of this technique in this select population.

Potential conflicts of interest

None of the other authors has any personal conflicts of interest.

Acknowledgements

This study is part of Limburg Clinical Research Center, supported by the foundation Limburg Sterk Merk, province of Limburg, Flemish government, Hasselt University, Ziekenhuis Oost-Limburg and Jessa Hospital.

Data availability statement

The data supporting this systematic review and meta-analyses are from previously reported studies and datasets, which have been cited within the text.

References

1. Camm AJ, Kirchhof P, Lip GY, et al. Guidelines for the management of atrial fibrillation: the Task Force for the Management of Atrial Fibrillation of the European Society of Cardiology (ESC). Eur Hear J. 2010;31(19):2369-2429. doi:10.1093/eurheartj/ehq278
2. Benjamin Emelia J, Wolf Philip A, D'Agostino Ralph B, Silbershatz H, Kannel William B, Levy D. Impact of Atrial Fibrillation on the Risk of Death. Circulation. 1998;98(10):946-952. doi:10.1161/01.CIR.98.10.946
3. T. JC, Samuel WL, S. AJ, et al. 2014 AHA/ACC/HRS Guideline for the Management of Patients With Atrial Fibrillation: Executive Summary. Circulation. 2014;130(23):2071-2104. doi:10.1161/CIR.0000000000000040
4. Peppard PE, Young T, Palta M, Skatrud J. Prospective Study of the Association between Sleep-Disordered Breathing and Hypertension. N Engl J Med. 2000;342(19):1378-1384. doi:10.1056/NEJM200005113421901
5. Stevenson IH, Teichtahl H, Cunnington D, Ciavarella S, Gordon I, Kalman JM. Prevalence of sleep disordered breathing in paroxysmal and persistent atrial fibrillation patients with normal left ventricular function. Eur Heart J. 2008;29(13):1662-1669. doi:10.1093/eurheartj/ehn214
6. Punjabi NM. The epidemiology of adult obstructive sleep apnea. Proc Am Thorac Soc. 2008;5(2):136-143. doi:10.1513/pats.200709-155MG
7. Young T, Finn L, Peppard PE, et al. Sleep disordered breathing and mortality: eighteen-year follow-up of the Wisconsin sleep cohort. Sleep. 2008;31(8):1071-1078.
8. Desteghe L, Hendriks JML, McEvoy RD, et al. The why, when and how to test for obstructive sleep apnea in patients with atrial fibrillation. Clin Res Cardiol. 2018;107(8):617-631. doi:10.1007/s00392-018-1248-9
9. Fein AS, Shvilkin A, Shah D, et al. Treatment of Obstructive Sleep Apnea Reduces the

Risk of Atrial Fibrillation Recurrence After Catheter Ablation. J Am Coll Cardiol. 2013;62(4):300-305. doi:https://doi.org/10.1016/j.jacc.2013.03.052

10. Broughton C, Piracha S, McGibbon K, Phan T, Allen MB. 15Is sleep apnoea unrecognised in patients with atrial fibrillation? EP Eur. 2018;20(suppl_4):iv7-iv8. doi:10.1093/europace/euy198.008
11. Kapur VK, Auckley DH, Chowdhuri S, 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 JCSM Off Publ Am Acad Sleep Med. 2017;13(3):479-504. doi:10.5664/jcsm.6506
12. Karthekeyan B, Nandipati S, Thangavel P, Vakamudi M, Kamalakkannan G, Kodali R. Pacemakers and Implantable Cardioverter Defibrillator. Int J Cardiol Lipidol Res. 2014;1:33-42. doi:10.15379/2410-2822.2014.01.01.04
13. Liberati A, Altman DG, Tetzlaff J, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. Ann Intern Med. 2009;151(4):W65-94. doi:10.7326/0003-4819-151-4-200908180-00136
14. Whiting PF, Rutjes AWS, Westwood ME, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. Ann Intern Med. 2011;155(8):529-536. doi:10.7326/0003-4819-155-8-201110180-00009
15. Wilson EB. Probable Inference, the Law of Succession, and Statistical Inference. J Am Stat Assoc. 1927;22(158):209-212. doi:10.1080/01621459.1927.10502953
16. Higgins JP, Deeks JJ, Altman DG. Chapter 16: Special topics in statistics. In: Higgins JP, Green S, eds. Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0. The Cochrane Collaboration; 2011. http://srdta.cochrane.org/.
17. Macaskill P, Gatsonis C, Deeks J, Harbord R, Takwoingi Y. Chapter 10: Analysing and Presenting Results. In: Deeks J, Bossuyt P, Gatsonis C, eds. Cochrane Handbook for

Systematic Reviews of Diagnostic Test Accuracy Version 1.0. ; 2010. http://srdta.cochrane.org/.

18. Freemantle N. StatsDirect—Statistical Software for Medical Research in the 21st Century. BMJ. 2000;321(7275):1536. doi:10.1136/bmj.321.7275.1536

19. Mantel N, Haenszel W. Statistical aspects of the analysis of data from retrospective studies of disease. J Natl Cancer Inst. 1959;22(4):719-748.

20. Bossuyt P, Davenport C, Deeks J, Hyde C, Leeflang M, Scholten R. Chapter 11: Interpreting results and drawing conclusions. In: Deeks J, Bossuyt P, Gatsonis C, eds. Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy Version 0.9. The Cochrane Collaboration; 2013. http://srdta.cochrane.org/.

21. Shalaby A, Atwood C, Hansen C, et al. Feasibility of automated detection of advanced sleep disordered breathing utilising an implantable pacemaker ventilation sensor. Pacing Clin Electrophysiol. 2006;29(10):1036-1043. doi:10.1111/j.1540-8159.2006.00496.x

22. Defaye P, de la Cruz I, Martí-Almor J, et al. A pacemaker transthoracic impedance sensor with an advanced algorithm to identify severe sleep apnea: The DREAM European study. Heart Rhythm. 2014;11(5):842-848. doi:10.1016/j.hrthm.2014.02.011

23. Barbieri F, Dichtl W, Heidbreder A, et al. Sleep apnea detection by a cardiac resynchronisation device integrated thoracic impedance sensor: A validation study against the gold standard polysomnography. PLoS One. 2018;13(4):e0195573-e0195573. doi:10.1371/journal.pone.0195573

24. Chen R, Chen K, Dai Y, Zhang S. Utility of transthoracic impedance and novel algorithm for sleep apnea screening in pacemaker patient. Sleep Breath. 2019;23(3):741-746. doi:10.1007/s11325-018-1755-y

25. Defaye P, Mendelson M, Tamisier R, et al. Validation of an apnea and hypopnea detection algorithm implemented in implantable cardioverter defibrillators. The AIRLESS study. Sci Rep. 2019;9(1):9597. doi:10.1038/s41598-019-45255-3

26. D'Onofrio A, La Rovere MT, Emdin M, et al. Implantable cardioverter-defibrillator-computed respiratory disturbance index accurately identifies severe sleep apnea: The DASAP-HF study. Hear Rhythm. 2018;15(2):211-217. doi:10.1016/j.hrthm.2017.09.038
27. Debray TPA, Moons KGM, Riley RD. Detecting small-study effects and funnel plot asymmetry in meta-analysis of survival data: A comparison of new and existing tests. Res Synth Methods. 2018;9(1):41-50. doi:10.1002/jrsm.1266
28. Ioannidis JPA. Why most published research findings are false. PLoS Med. 2005;2(8):e124-e124. doi:10.1371/journal.pmed.0020124
29. S. GA, Gregg P, M. CS, et al. Association of Atrial Fibrillation and Obstructive Sleep Apnea. Circulation. 2004;110(4):364-367. doi:10.1161/01.CIR.0000136587.68725.8E
30. Escourrou P, Luriau S, Rehel M, Nédelcoux H, Lanoë JL. Needs and costs of sleep monitoring. Stud Health Technol Inform. 2000;78:69-85.
31. 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(7):485-491. doi:10.7326/0003-4819-131-7-199910050-00002
32. Gonçalves IS, Agostinho JR, Silva G, et al. Accuracy and utility of a pacemaker respiratory monitoring algorithm for the detection of obstructive sleep apnea in patients with atrial fibrillation. Sleep Med. 2019;61:88-94. doi:10.1016/j.sleep.2019.01.051
33. Peppard PE, Young T, Barnet JH, Palta M, Hagen EW, Hla KM. Increased prevalence of sleep-disordered breathing in adults. Am J Epidemiol. 2013;177(9):1006-1014. doi:10.1093/aje/kws342
34. Hsu Y-S, Chen T-Y, Wu D, Lin C-M, Juang J-N, Liu W-T. Screening of obstructive sleep apnea in patients who snore using a patch-type device with electrocardiogram and 3-axis accelerometer. J Clin sleep Med JCSM Off Publ Am Acad Sleep Med. 2020;16(7):1149-1160. doi:10.5664/jcsm.8462
35. Youden WJ. Index for rating diagnostic tests. Cancer. 1950;3(1):32-35. doi:10.1002/1097-

0142(1950)3:1<32::aid-cncr2820030106>3.0.co;2-3

36. Luo Y-M, Tang J, Jolley C, et al. Distinguishing obstructive from central sleep apnea events: diaphragm electromyogram and esophageal pressure compared. Chest. 2009;135(5):1133-1141. doi:10.1378/chest.08-1695

Figure legends

Figure 1: PRISMA Flowchart of the selection process.

Figure 2: Statistical results of pacemakers evaluated at PSG-AHI≥30.

Figure 3: Diagnostic Odds Ratio Forest Plot in a Fixed Effect model.

Figure 4: Symmetric ROC-curve of summary-DOR with confidence bounds. The red curve is the lower confidence bound, the blue curve is the symmetric ROC curve according to the summary DOR and the green curve is the upper confidence bound. The diagonal purple line intersects the Q*-points for the three curves. X-axis represents (1-specificity) and y-axis represents sensitivity. The included studies were positioned based on their Q*-point.

Figure 5: Symmetric ROC-curve of individual study DOR with statistical parameters at optimal cut-off. The x-axis represents (1-specificity) and the y-axis represents (sensitivity). The diagonal black line intersects the curves at their respective Q*, as shown in Figure 4.

Figure 6: Studies with low, high or unclear RISK of BIAS.

Figure 7: Studies with low, high or unclear CONCERNS of APPLICABILITY.