Sleep Medicine 88 (2021) 116–133

Contents lists available at ScienceDirect

Sleep Medicine

journal homepage: www.elsevier.com/locate/sleep

Review Article

Prediction of the obstruction sites in the upper airway in sleep-disordered breathing based on snoring sound parameters: a systematic review

Zhengfei Huang [a, b, *], Ghizlane Aarab [a], Madeline J.L. Ravesloot [c], Ning Zhou [a, d], Pien F.N. Bosschieter [c], Maurits K.A. van Selms [a], Chantal den Haan [e], Nico de Vries [a, c, f], Frank Lobbezoo [a], Antonius A.J. Hilgevoord [b]

[a] Department of Orofacial Pain and Dysfunction, Academic Center for Dentistry Amsterdam (ACTA), University of Amsterdam and Vrije Universiteit Amsterdam, Amsterdam, the Netherlands
[b] Department of Clinical Neurophysiology, OLVG, Amsterdam, the Netherlands
[c] Department of Otorhinolaryngology - Head and Neck Surgery, OLVG, Amsterdam, the Netherlands
[d] Department of Oral and Maxillofacial Surgery, Amsterdam UMC Location AMC and Academic Centre for Dentistry Amsterdam (ACTA), University of Amsterdam, Amsterdam, the Netherlands
[e] Medical Library, Department of Research and Education, OLVG, Amsterdam, the Netherlands
[f] Department of Otorhinolaryngology - Head and Neck Surgery, Antwerp University Hospital (UZA), Antwerp, Belgium

ARTICLE INFO

Article history:
Received 15 July 2021
Received in revised form
16 September 2021
Accepted 12 October 2021
Available online 21 October 2021

Keywords:
Snoring sound
Upper airway
Frequency
Prediction model
Machine learning

ABSTRACT

Background: Identification of the obstruction site in the upper airway may help in treatment selection for patients with sleep-disordered breathing. Because of limitations of existing techniques, there is a continuous search for more feasible methods. Snoring sound parameters were hypothesized to be potential predictors of the obstruction site. Therefore, this review aims to i) investigate the association between snoring sound parameters and the obstruction sites; and ii) analyze the methodology of reported prediction models of the obstruction sites.
Methods: The literature search was conducted in PubMed, Embase.com, CENTRAL, Web of Science, and Scopus in collaboration with a medical librarian. Studies were eligible if they investigated the associations between snoring sound parameters and the obstruction sites, and/or reported prediction models of the obstruction sites based on snoring sound.
Results: Of the 1016 retrieved references, 28 eligible studies were included. It was found that the characteristic frequency components generated from lower-level obstructions of the upper airway were higher than those generated from upper-level obstructions. Prediction models were built mainly based on snoring sound parameters in frequency domain. The reported accuracies ranged from 60.4% to 92.2%.
Conclusions: Available evidence points toward associations between the snoring sound parameters in the frequency domain and the obstruction sites in the upper airway. It is promising to build a prediction model of the obstruction sites based on snoring sound parameters and participant characteristics, but so far snoring sound analysis does not seem to be a viable diagnostic modality for treatment selection.

1. Introduction

Obstructive sleep apnea (OSA) is a common sleep-related breathing disorder; the reported prevalence of OSA in the general adult population ranges from 9% to 38% [1]. OSA is characterized by repeated partial or complete obstruction of the upper airway during sleep, which may lead to oxygen desaturation, respiratory arousals, and non-restorative sleep due to frequent awakenings [2]. Because of patient's complaints and potential health risks, the

* Corresponding author. Department of Orofacial Pain and Dysfunction, Academic Centre for Dentistry Amsterdam (ACTA), Gustav Mahlerlaan 3004, 1081 LA Amsterdam, the Netherlands.
E-mail address: zf.huang@acta.nl (Z. Huang).

https://doi.org/10.1016/j.sleep.2021.10.015

management of OSA has drawn more attention in the field of sleep medicine [3]. It has been suggested that identification of the level(s), degree, and direction/configuration of the obstruction site(s) in the upper airway is essential in the diagnostic work-up of OSA patients and the treatment decision-making process. For instance, patients with complete concentric collapse on palatal level (CCCp) or lateral oropharyngeal collapse need a significantly higher continuous positive airway pressure (CPAP) level [4]; prescription of oral appliance (OA) therapy makes more sense in case of obstruction at the tongue base than in case of CCCp and complete lateral oropharyngeal collapse [5]; the indication of upper airway surgery is largely based on anatomically correctable features; and CCCp is considered an absolute contraindication for upper airway stimulation [6,7].

Nowadays, nine modalities are frequently used to investigate possible upper airway obstructions. Among these, drug-induced sleep endoscopy (DISE) is a unique and dynamic technique. During drug-induced sleep, the upper airway is observed using a flexible endoscope, and information on upper airway obstruction can be obtained [8]. However, there are some limitations to DISE, such as the influence of drug selection on DISE findings [9] and potential inaccuracy of DISE results due to the difference between natural sleep and drug-induced sleep [10]. Alternatively, sleep videofluoroscopy (SVF) is a localization technique that combines fluoroscopy and video recording, enabling direct visualization of dynamic airway change [11]. However, because SVF can only provide two-dimensional lateral views, it cannot explain lateral movement of the upper airway [12]. In addition, SVF is also performed during drug-induced sleep, which, as mentioned above, may lead to inaccuracy of results due to the difference between natural sleep and drug-induced sleep. Airway pressure measurement is a system by which the obstruction site in the upper airway can be determined during natural sleep. In this system, a pressure transducer catheter with sensors is inserted through a naris into the esophagus. The absence of pressure deflection indicates the presence of obstruction [13]. However, this system can only indicate the lowest obstructive site; higher possible obstruction sites are left out. In addition, this system gives no information about abnormal anatomic structures in the obstruction site, such as enlarged tonsils [14]. Airflow shape is reported to be associated with the obstruction sites as well. According to the contours of the airflow reference patterns, different obstruction sites can be identified [15]. However, this method can only be used to identify single-level obstruction and is not applicable for patients with multilevel obstructions [15]. Dynamic magnetic resonance imaging (MRI) is sometimes used for identification of the obstruction sites in the upper airway during either wakefulness or sleep. This method, however, creates noise that may keep patients awake, is costly, and may cause claustrophobic effects [16]. Müller's maneuver, simulated snoring, lateral cephalometry, and modified Mallampati classification can all be used to evaluate soft tissues and/or skeletal anatomies while patients are awake and to identify the obstruction sites in the upper airway. However, previous studies found that results obtained during wakefulness show poor agreement with results observed during sleep [17,18]. Because of these drawbacks of all of these diagnostic methods, there is a continuous search for less invasive and more feasible methods to identify the levels, degree, and configuration of obstructions in the upper airway.

According to the definition of the International Classification of Sleep Disorders – Third Edition (ICSD-3), snoring is characterized as "audible vibrations of the upper airway during respiration in sleep" [19]. Snoring generally occurs during upper airway collapse and the structures/regions in the upper airway may contribute to the collapse (Fig. 1). The classical Starling resistor model is commonly used to explain the mechanism of snoring, hypopnea, and apnea [20]. In this model (Fig. 2), the pharynx is considered as a collapsible conduit, which is mounted between a rigid upstream segment (the nasal cavity) and a rigid downstream segment (the trachea). The pressure that is applied to the pharynx is from the tissues surrounding it and, in this model, the surrounding tissues are considered as a chamber where the air pressure (tissue pressure; P_t) can be controlled. The patency of the collapsible conduit is determined by its transmural pressure. Previous studies on the classical Starling resistor model suggested that, during inspiration, once downstream pressure (P_d) drops below P_t (eg, due to the excessive inspiratory effort or increased P_t), then an obstruction site develops and flow limitation occurs [21,22]. In addition, it was

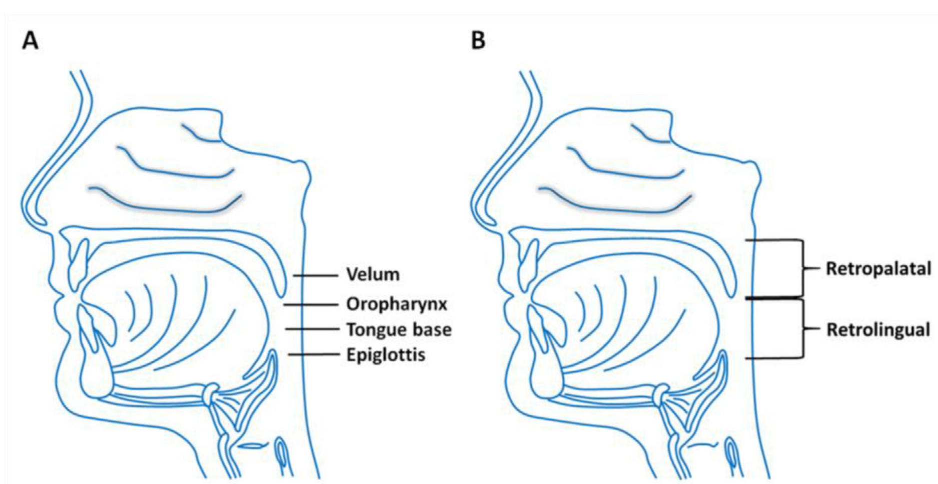

Fig. 1. (1.5 column; using color for online version only) Schematic visualization of the upper airway. Velum, oropharynx, tongue base, and epiglottis are commonly involved in the obstruction of the upper airway, either individually or in every possible combination (A). According to the region-based classification, an obstruction can also be classified as retropalatal obstruction or retrolingual obstruction (B). (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

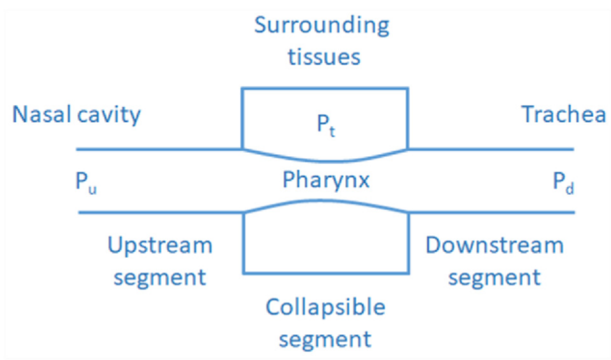

Fig. 2. (Single column; using color for online version only) Schematic visualization of the classical Starling resistor model. The horizontal conduit with a collapsible segment represents the airway. The upstream segment represents the nasal cavity and the downstream segment represents the trachea. P_u is the pressure in the upstream segment and P_d is the pressure in the downstream segment. The collapsible segment represents the pharynx and P_t is the pressure that is applied to the pharynx, which is from the tissues surrounding it. In this model, the surrounding tissues of the pharynx are considered as a chamber where the air pressure (P_t) can be controlled. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

found that the obstruction site prevents the reduction in upstream pressure (P_u; usually equals to atmospheric pressure) that is caused by the reduction in P_d. Therefore, P_u remains constant [21,22]. In this case, the inspiratory flow is independent of P_d and is determined by the difference between P_u and P_t (ie, $P_u - P_t$), since P_t replaces P_d as the effective downstream pressure to inspiratory flow [21,22]. In the classical Starling resistor model, snoring is found to occur during flow limitation (ie, $P_u > P_t > P_d$) [21,23,24]. Once P_u also drops below P_t (eg, due to increased P_t), apnea occurs due to the cessation of airflow (ie, $P_t > P_u > P_d$) [21]. However, inconsistent findings were observed between studies on the classical Starling resistor model and studies on OSA patients, suggesting that the classical Starling resistor model still fails to capture the full complexity of airway collapse [20,22]. Another theorem that may explain the mechanism of snoring is Bernoulli's principle, but it is somewhat inaccurate as Bernoulli's principle applies to changes in flow and pressure through a fixed, rigid constriction. Bernoulli's principle in snoring was thoroughly discussed in a previous study [25].

Given the noisy nature, various criteria have been employed in previous studies to identify snoring sound [26,27]. In the last few years, more and more studies suggest that snoring sound carries information about the upper airway and sleep, eg, the presence of OSA and sleep stage [28,29]. The informative nature of snoring sound has been reported in OSA-related studies. Acar et al. [30] reported that the peak frequency (Hz) of snoring sound increased with the increase of the severity of OSA. Azarbarzin and Moussavi [31] distinguished OSA snorers from non-OSA snorers based on the features of snoring sound, eg, average power and zero crossing rate, and they reported an accuracy of 96.4%. In addition, a positive linear association between the intensity of snoring sound and the severity of excessive daytime sleepiness in non-OSA patients was reported [32].

In general, sound is analyzed in three domains, which are the temporal domain, the intensity domain, and the frequency domain. Analyses in the temporal domain address the variation of sound over time. Sound intensity is defined as the sound power per unit area, of which the most common unit to measure is decibel (dB). In digital audio recordings, the amplitudes of sound waves represent the sound intensity. For research purposes, analyses in the temporal and intensity domains cannot provide enough information of a sound signal. This suggests the importance of frequency analyses. In the frequency domain, a complicated signal in the temporal domain can be decomposed into different frequency components to obtain its frequency spectrum using Fourier transform (Fig. 3). Similar to speech, the acoustic characteristics of snoring sound are determined by the anatomy of the upper airway [33]. The tension and length of the vocal cords affect the sound's vibration frequency in speaking. Similarly, the dynamic properties of the upper airway walls will influence the vibrations produced in snoring. The chamber-like anatomy enables the upper airway to act as a resonance system, and different frequencies of the original sound waves are either attenuated or amplified in this “resonance chamber”. Based on the “source-filter theory” [34,35], the hypothesis was proposed that snoring sound parameters may be potential predictors of the obstruction sites in the upper airway [36].

To the best of our knowledge, the feasibility of predicting the obstruction sites in the upper airway based on snoring sound parameters is not yet fully explored. Therefore, the available literature was reviewed with the aim to i) investigate the association between snoring sound parameters and the obstruction sites determined by objective methods (eg, DISE, dynamic MRI) during sleep; and ii) analyze the methodology and the participant characteristics in the studies reporting prediction models of the obstruction sites as to provide suggestions for further study.

2. Methods

2.1. Search strategy

This systematic review is reported in accordance with the PRISMA statement (Preferred Reporting Items for Systematic Reviews and Meta-Analyses; www.prisma-statement.org). The protocol for this review was registered at PROSPERO (ID: CRD42020179769). To identify all relevant publications, a systematic search was conducted in the bibliographic databases PubMed, Embase.com, CENTRAL, Web of Science, and Scopus from inception up to June 23, 2021, in collaboration with a medical information specialist (CH), who provided help to formulate a comprehensive search strategy. The search was performed without date and language restrictions. The full search strategies for all databases can be found in Appendix 1 of the Supplementary Material.

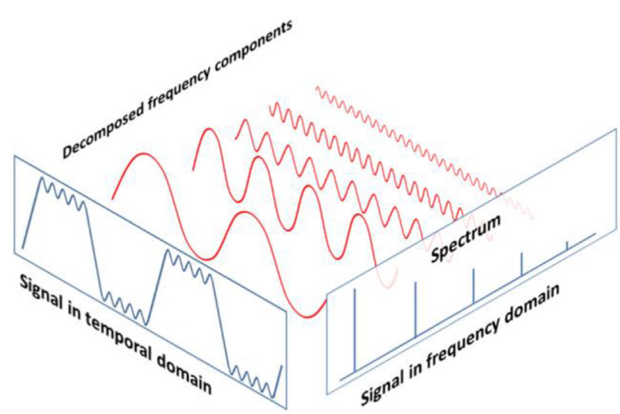

Fig. 3. (Single column; using color for online version only) Schematic visualization of Fourier transform. The Fourier transform is an algorithm, by which a complicated signal in temporal domain (left panel) can be decomposed into sinusoidal oscillations at distinct frequencies (“Decomposed frequency components”), with each sinusoidal oscillation having its own amplitude. A spectrum (right panel) is the projection of the amplitudes and frequencies of the sinusoidal oscillations on a plane. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

2.2. Inclusion and exclusion criteria

Inclusion criteria:

- Studies on adult OSA patients and/or snorers;
- Randomized Clinical Trial (RCT) studies, cross-sectional studies, case–control studies, cohort studies, or conference abstracts (if necessary, authors are approached for full-text);
- Studies that either report the association(s) between snoring sound parameters and the obstruction sites determined by objective methods (eg, DISE, dynamic MRI) during sleep, or report prediction model(s) of the obstruction sites based on snoring sound parameters.

Exclusion criteria:

- Studies on children or animals;
- No full-text available;
- Reviews, editorials, case reports.

2.3. Study selection

Two reviewers (ZH and NZ) independently screened all titles and abstracts for eligibility. Differences in judgement were resolved through a consensus procedure, guided by a third reviewer (AH). In the full-text review, which included the decision of inclusion/exclusion of studies, quality assessment, and data extraction, two reviewers (ZH and NZ) took different responsibilities. ZH performed all the procedures and NZ checked the results. Differences in judgement were resolved through a consensus procedure guided by AH.

2.4. Quality assessment and data extraction

During the full-text review of the articles, it was found that all the included studies on the association between snoring sound parameters and the obstruction sites were cross-sectional studies. Hence, the Appraisal tool for Cross-Sectional Studies (AXIS) was used to assess the methodological quality of these studies, in terms of introduction, methods, results, discussion, and other [37]. Because AXIS does not provide cut-off points for various quality grades, according to a previous study [38], studies with a total score higher than 15 were considered to be of high quality, those of 10–15 were considered to be of moderate quality, while those scoring less than 10 were considered to be of poor quality. Studies that were considered as being of "poor quality" were excluded.

For studies that reported prediction models, the Prediction model Risk Of Bias ASsessment Tool (PROBAST) was used to assess both the risk of bias and the applicability of the models [39]. The tool consists of four domains (participants, predictors, outcome, and analysis) and has 20 questions that facilitate reaching overall judgement of risk of bias and applicability. A prediction model is considered "low risk of bias" and/or "low concerns regarding applicability", if risk of bias and/or applicability are evaluated as "low" in all domains. A prediction model is considered "high risk of bias" and/or "high concerns regarding applicability", if risk of bias and/or applicability are assessed as "high" in one or more domain. A prediction model is considered "unclear risk of bias" and/or "unclear concerns regarding applicability", if risk of bias and/or applicability are assessed as "unclear" in one or more domain and the other domains are "low".

As for data extraction, the following information was extracted from the studies reporting the association(s) between snoring sound parameters and the obstruction sites: (i) country, (ii) sample size (N), (iii) population, (iv) aim(s), (v) recording of snoring sound, (vi) snoring sound parameter(s), (vii) method of identification of the obstruction sites (viii) main finding(s), and (ix) conclusion(s). In addition to these data, the name of the prediction model was also extracted for the studies reporting prediction model of the obstruction sites.

3. Results

3.1. Study selection

The literature search generated a total of 2256 references: 395 in PubMed, 561 in Embase.com, 28 in CENTRAL, 646 in Scopus, and 626 in Web of science. After deduplication, 1016 references remained. The flow chart of the search and selection process is presented in Appendix 2 of the Supplementary Material. Of the 1016 articles obtained from databases and other sources (hand searching; n = 1), 28 articles were found eligible for inclusion in this study.

3.2. Quality assessment

The results of the quality assessments can be found in Appendix 3 of the Supplementary Material. Of the 28 included studies, 12 studies reported association(s) between snoring sound parameters and the obstruction sites in the upper airway. Among these studies, four [41,44,45,49] were ranked as being of high quality, the other eight studies [36,40,42,43,46–48,50] were ranked as being of moderate quality, and no study was excluded because of poor methodological quality. In short, among the twenty questions that are included in AXIS, eleven studies did not report sample size calculation; eleven studies did not report the response rate; eleven studies did not report any information about non-responders; and ten studies did not report what method was used to address non-responders. These were the main reasons why the majority of the studies were classified as being of moderate quality for the purpose of this review.

The other 16 studies reported prediction models of the obstruction sites based on snoring sound parameters. Among these studies, six [51,52,58,63,64] were identified as being of low risk of bias and low concern regarding applicability, while the other ten studies [53–57,59–62,66] were identified as being of low concern regarding applicability but high risk of bias due to the small sample size.

3.3. Study characteristics

All included articles were in English. The full-texts of two German articles and one Korean article were screened, but the German articles were excluded due to reporting "other aspect of snoring" while the Korean article was excluded as it was a review article. Sample size in the included studies ranged from 9 to 219. The characteristics of included studies reporting the association between snoring sound parameters and the obstruction sites are shown in Table 1, the characteristics of included studies reporting prediction model of the obstruction sites are shown in Table 2.

Of the 16 studies reporting prediction models, one study [53] built a prediction model using conventional statistical methods and reported overall accuracy. The other 15 studies built machine learning models (see below). Of the 15 studies, four studies [60–62,64] reported overall accuracy and eleven studies used unweighted average recall (UAR) to evaluate the accuracy of the machine learning models due to the unbalanced distribution of obstruction sites. The UAR was defined as [58]:

Table 1
Characteristics of included studies reporting the association between snoring sound parameters and the obstruction sites.

Study	Country	N	Population	Aim(s)	Recording of snoring sound	Snoring sound parameter(s)	Identification of the obstruction site(s)	Main finding(s)	Conclusion(s)
Agrawal et al., 2002 [40]	UK	16	Primary snorer	To use observations during induced snoring to make correlations between sound frequency and the site of upper airway narrowing	Contact condenser microphone (between the nasal ala and corner of the mouth)	DF, center frequency, and power ratio (750 Hz)	DISE (VOTE)	The center frequencies of palatal, tonsillar, tongue-base, and epiglottic snores were 391 Hz, 445 Hz, 1094 Hz, and 442 Hz, respectively The power ratio of palatal, tonsillar, tongue-base, and epiglottic snores were 7, 12, 0.2, and 53, respectively The combination of palate and tongue snoring showed both low and high frequency components, with a center frequency of 404 Hz	Induced snores contain a higher frequency component of sound, suggesting that DISE may not accurately reflect natural snoring
Chang et al., 2014 [41]	Taiwan	9	Patients with excessive snoring and sleep-disordered breathing	To synchronize dynamic MRI signals and acoustic measurements aiming to valuate dynamic upper airway obstruction during sleep	Build-in microphone of MRI (5–20 mm from patient's mouth)	Sound intensity (amplitude), soft tissue vibration time, and collapse index	Dynamic MRI (retropalatal, retroglossal, or combined)	There was a significant correlation between the collapse index and soft tissue vibration time ($P < 0.03$) The collapse index and soft tissue vibration time were significantly different between pure retropalatal and combined snoring ($P < 0.0001$)	The pilot study proves that synchronized MRI and acoustic recordings can characterize the sites of airway obstruction during sleep
Gurpinar et al., 2020 [42]	Turkey	55	Candidates for OSA surgical intervention	To integrate the physical findings of the OSAS patients during DISE with the snoring sound analysis of the susceptible	Smartphone (above the shoulder of each subject)	Mean frequency, DF, and F0	DISE (retropalatal, retrolingual, or multilevel)	The mean frequency, F0, and DF of retrolingual snoring are significantly higher than those of retropalatal snoring ($P = 0.001$) The mean frequency, F0, and DF of multilevel snoring are higher than those of retropalatal snoring, but the differences are not significant ($P = 0.419$) The mean frequency, F0, and DF of multilevel snoring are significantly lower than those of retrolingual snoring ($P = 0.025$)	To determine the site of obstruction using the sound analysis of snores is a cost-effective method which can be done during DISE
Herzog et al., 2014 [43]	Germany	41	Male patients with suspected sleep-disordered breathing	To evaluate the characteristics of different types of snoring sounds in order to create algorithms for acoustic analyses of snoring sounds	An external condenser microphone (30 cm above patient's mouth)	Sound pressure level (dBA), center frequency, and psychoacoustic parameters (loudness, sharpness, roughness, fluctuation strength)	DISE (the first snoring episode after an apneic event, tonsillar snoring, velar snoring, or obstructive velar snoring)	Velum snoring showed lower center frequency than the other snoring patterns ($P < 0.05$) The sound pressure level of obstructive velar snoring was higher than the other three snoring patterns ($P < 0.05$)	Nocturnal snoring might be differentiated by psychoacoustic algorithms

								Velar snoring showed lower loudness than the other snoring patterns ($P < 0.05$) Tonsil snoring showed higher sharpness than the other snoring patterns ($P < 0.05$)	
Koo et al., 2016 [44]	South Korea	32	Male OSA patients	To determine the spectrographic pattern, sound intensity (dB), F0, F1, F2, and F3 of snoring sounds caused by different obstruction sites	Smartphone above the patient's shoulder	Sound intensity (dB), spectrographic pattern, F0, F1, F2, and F3	DISE (retropalatal or retrolingual)	In spectrograph, retropalatal obstruction tended to have sharp regular peaks, while retrolingual obstruction showed gradual onset and decay There was no significant difference in the intensity of the snoring sounds between retropalatal and retrolingual level obstructions ($P = 0.8$) The F1 and F2 of retrolingual level obstruction were significantly higher than those of retropalatal obstruction ($P < 0.05$)	The analysis of formants will be a useful screening test for the prediction of obstruction sites in the upper airway
Lee et al., 2016 [45]	Taiwan	36	OSAHS patients	To examine associations between acoustic parameters of whole night snoring sounds during natural sleep and obstruction sites (multi-level and other levels) defined by DISE	Non-contact microphone (100 cm above patient's head) was used to record snoring sounds simultaneously with PSG	Snoring index (events/hour), intensity (dB; max, mean), mean frequency, and DF	DISE (VOTE)	Participants with more low frequency snoring events ($\geq$28 events/h) were 7.86 times more likely to have complete velopharynx obstruction than participants with less low frequency snoring events (<28 events/h; $P = 0.010$) Participants with a higher mean frequency ($\geq$1220 Hz) were 1.29 times more likely to have lateral oropharyngeal wall obstructions than participants with a lower mean frequency (<1220 Hz; $P = 0.013$) Participants with a higher peak frequency ($\geq$1775 Hz) were 4.80 times more likely to have complete tongue base obstructions than participants with a lower peak frequency (<1775 Hz; $P = 0.033$) Participants with a higher mean intensity ($\geq$66 dB) at frequency of 301–850 Hz were 1.14 times more likely to have epiglottitis obstructions than	Snoring sound analysis may be helpful in determining obstruction sites

(continued on next page)

Table 1 (*continued*)

Study	Country	N	Population	Aim(s)	Recording of snoring sound	Snoring sound parameter(s)	Identification of the obstruction site(s)	Main finding(s)	Conclusion(s)
								participants with a lower mean intensity (<66 dB; p = 0.020)	
Miyazaki et al., 1998 [46]	Japan	75	Patients with sleep related respiratory disorders	To find a way of predicting the site of airway obstruction by analyzing the snoring sound acoustically	Handheld condenser microphone	F0	Monitoring the pressure in the upper airway (soft palate, tonsil/tongue base, combined type, or larynx type)	Average F0 was 102.8 ± 34.9 Hz in the soft palate snoring, 331.7 ± 144.8 Hz in tonsil/tongue base snoring, 115.7 ± 58.9 Hz in the combined snoring, and 249.4 ± 79.7 Hz in larynx snoring	Acoustic analysis of snoring sound is useful as a screening method to determine the obstruction site in the upper airway
Osborne et al., 1999 [47]	UK	11	Patients who underwent DISE	To differentiate palatal snoring from non-palatal snoring	DAT recorder	Crest factor	DISE (exclusive soft palate or exclusive lower segment)	Crest factor over 2.7 were palatal and below 2.7 were lower segment	Crest factor is a reliable indicator of palatal snoring
Quinn et al., 1996 [36]	UK	10	Subjects known to suffer from heavy snoring but not obstructive sleep apnea	To study palatal snoring during DISE	Contact microphone (upper lip)	Center frequency	DISE (soft palate or tongue base)	Palatal snoring showed a mean center frequency of less than 420 Hz and a mean standard deviation of less 370 Hz Tongue base snoring showed a mean center frequency of higher than 650 Hz and a mean standard deviation of higher than 430 Hz	Palatal and tongue base snorers can be differentiated from each other using the criteria of center frequency greater or less than 500 Hz and frequency deviation greater or less than 400 Hz
Saunders et al., 2004 [48]	UK	35	Patients undergoing DISE with AHI less than 40	To establish whether acoustic analysis of snoring sound could replace DISE in a clinical setting	Non-contact microphone (30 cm above patient's mouth)	Center frequency	DISE (soft palate, tongue base, or mixed type)	Central frequency of 90 Hz could be used to differentiate palatal snoring and tongue base snoring The calculated irregularity of the sound showed no correlation with the type of snoring	Acoustic analysis may be helpful to differentiate palatal snoring tongue base snoring, but it is unlikely to replace DISE
Won et al., 2012 [49]	South Korea	90	Patients complaining of snoring and/or sleep apnea	To evaluate the acoustic characteristics of snoring according to obstruction site determined by SVF	Contact microphone at parasternal area	F1, F2, intensity (not specified), pitch, jitter, and shimmer	SVF (soft palate or pharyngeal lateral wall, which included tonsil, tongue base, epiglottis, and a combination of these structures)	The pitch and F1 of tongue base and epiglottis obstruction were significantly higher than soft palate group (pitch: $P = 0.007$ and $P = 0.001$, respectively; F1: $P = 0.019$ and $P = 0.047$, respectively) The pitch and F1 of both soft palate + tongue base group and soft palate + epiglottis group were significantly higher than those of soft palate group (pitch: $P = 0.007$ and $P = 0.007$, respectively; F1: $P = 0.04$ and $P = 0.008$, respectively) No significant difference was found among the	Pitch and F1 differed according to site of upper airway obstruction

								different groups in jitter and shimmer (*P > 0.05*)	
Xu et al., 2010 [50]	China	30	Male OSAS patients	To investigate the correlation between the sound spectrum of snoring sound and the site of obstruction in patients with OSAS	Non-contact microphone (20 cm above patient's mouth)	DF, center frequency, and the proportion of power in low, middle, and high frequency band (<800 Hz; 800–2000 Hz; >2000 Hz)	Monitoring the pressure in the upper airway (upper level: above the free margin of the soft palate; lower level: below the free margin of the soft palate)	The mean of DF, central frequencies, and proportions of energy from 800 Hz to 2000 Hz and above 2000 Hz of the first snoring sounds after lower-level obstructive apneas were higher than those after upper-level obstructive apneas	It might be possible to use sound spectral analysis as a supplementary or screening method to determine the obstructive site in OSAS

Note. DAT = digital audio tape; dB = decibel; dBA = A-weighted decibel; DF = dominant frequency (also known as peak frequency); DISE = drug-induced sleep endoscopy; F0 = fundamental frequency; F1 = the first formant; F2 = the second formant; F3 = the third formant; MFCCs = Mel-frequency cepstral coefficients; MRI = magnetic resonance imaging; OSA = obstructive sleep apnea; OSAHS = obstructive sleep apnea-hypopnea syndrome; OSAS = obstructive sleep apnea syndrome; PSG = polysomnography; SVF = sleep videofluoroscopy; VOTE = velum, oropharyngeal lateral walls, tongue base, and epiglottis classification.

$$UAR = \frac{\sum_{\lambda=1}^{\Lambda} Recall_{\lambda}}{\Lambda}$$

where Λ is the number of classes, and $Recall_{\lambda}$ is the class-specific recall, ie, the ratio of instances of class λ that is classified correctly of the λ -th class.

Machine learning is a subset of artificial intelligence, which focuses on making predictions using a computer. Based on the different sample data, different machine learning approaches can be used. For example, to analyze sample data with labelled outputs, supervised learning can be used [67]; to analyze sample data without labelled outputs and to find structures in the data by the machine learning model itself, unsupervised learning can be used [68]. In addition, different machine learning classifiers can be used for different applications. For example, deep neural networks (DNN) and Gaussian mixture models (GMM) are often used for automatic speech recognition [69,70], while support vector machine (SVM) is commonly used for classification and regression analysis [71].

3.3.1. Participant characteristics

Of the 28 included studies (Tables 1 and 2), six studies [43,44,50,55–57] only included male patients. Among these studies, only Xu et al. [50] stated that female patients were excluded as it is not clear whether gender is a factor that influences snoring sound parameters. The other studies did not specify why they only included male patients. It also needs to be noted that, of the included studies, 13 studies [42,44,45,50–52,58,60–65] only included OSA patients, eight studies [49,53–57,59,66] included both OSA patients and primary snorers, three studies [41,43,46] included patients with sleep-disordered breathing, two studies [36,40] only included primary snorers, and two studies [47,48] included patients who underwent DISE.

3.3.2. Methods for the recording of the snoring sound

Of the 28 included studies (Tables 1 and 2), 27 studies digitally recorded snoring sound and saved recordings on a computer, ie, the snoring sound captured by a microphone was digitized and saved on a computer in an audio file format. One study [47] digitally recorded snoring sound and saved recordings on a digital audio tape (DAT), which in appearance is similar to a cassette tape but can record signal digitally. As for the microphone used in these studies, 12 studies [41–45,48,50,53,60–62,66] used a non-contact microphone, and the position of the microphone varied from approximately 5 mm away from patient's mouth to the ceiling of the room. Three studies used a contact microphone; two of these studies [36,40] placed the microphone around the mouth, while the other study [49] placed the microphone at the parasternal area. In the other 13 studies, the position of the microphone was not mentioned.

3.3.3. Methods for the identification of the snoring sound

As for the identification of snoring events (Tables 1 and 2), 16 studies [36,41,43,48–50,53–57,59–62,66] manually selected snoring events from recordings. Six machine learning studies [51,52,58,63–65] used a publicly-available snoring sound corpus (Munich Passau Snore Sound Corpus [MPSSC]), which was developed for a sub-challenge in the INTERSPEECH 2017 Computational Paralinguistics Challenge [77]. In order to prepare the snoring sound corpus, an algorithm was used to identify audio events in sound recordings that were recorded during DISE. Then, an experienced researcher listened to all sound episodes and classified them manually as either snoring sound or other sound. Gurpinar et al. [42] and Lee et al. [45] employed an automatic detection

Table 2
Characteristics of included studies reporting prediction models of the obstruction sites.

Study	Country	N	Population	Aim(s)	Recording of snoring sounds	Snoring sound parameter(s)	Identification of the obstruction site(s)	Model (classifier)	Main finding(s)	Conclusion
Amiriparian et al., 2017 [51]	Germany	219	Publicly available snore sound corpus (MPSSC) based on OSA patients	To assess the possibility to classify snoring sound based on the combination of spectrograms and pre-trained CNN	Headset microphone, stand-mounted microphone, handheld microphone, and fixed microphone on forehead	Deep spectrum features for all combinations of CNN-descriptors and spectrogram color maps (jet: varying from blue to green to red; gray: varying from black to grey to white; viridis: varying from blue to green to yellow)	DISE (VOTE)	SVM	A UAR of 67% was achieved based on the combination of CNN descriptor AlexNet *fc*7 and viridis color map	Using the deep spectrum feature extraction method and linear SVM as a classifier, it was able to outperform the baseline for the sub-challenge
Demir et al., 2018 [52]	Turkey	219	Publicly available snore sound corpus (MPSSC) based on OSA patients	To investigate the use of low-level image texture features in classification of snore sounds	Headset microphone, stand-mounted microphone, handheld microphone, and fixed microphone on forehead	LBP and HOG	DISE (VOTE)	SVM	A UAR of 72.6% was obtained by combining LBP Features and HOG Extraction	Low level image texture features (LBP and HOG) are reliable for snoring sound classification
Peng et al., 2017 [53]	China	74	Patients diagnosed with OSA or primary snoring	To explore the possibility of using the acoustic parameters to differentiate the sources of snoring sounds	Non-contact microphone (30 cm above patient's mouth)	F0, F1, F2, and F3	DISE (VOTE; palatal snoring, mix type of palatal and lateral wall snoring, lateral snoring)	AUC, NPV, PPV, and overall accuracy	The overall accuracy of F0 was 60.8% The overall accuracy of F2 was 62.4%	F0 might be used to distinguish palatal snoring sound from non-palatal snoring sound. F2 is less sensitive than F0
Qian et al., 2016 [54]	Germany	24	Subjects diagnosed with primary snoring or OSA	To compare wavelet feature set with some frequently-used acoustic features	Headset microphone	Wavelet features (variance, length, and entropy of wavelet coefficients) and frequently-used features (F0, F1, F2, F3, MFCCs, power ratio [800 Hz], and crest factor)	DISE (VOTE)	SVM	A UAR of 71.2% was achieved using wavelet feature set, which outperformed other frequently-used feature sets	The classifier model based on wavelet feature set outperformed the classifier model based on conventional feature sets
Qian et al., 2016 [55]	Germany	40	Male patients diagnosed with OSA or primary snoring	To systematically compare different acoustic features, and classifiers for their performance in the classification of the excitation location of snore sounds	Headset microphone and handheld microphone	Crest factor, F0, F1, F2, F3, SFFs, SER, MFCCs, power ratio (800 Hz), EMDF, and WEF	DISE (VOTE)	K-NN, LDA, FNN, SVM, RF, ELM, and KELM	The best classification performance is achieved by a combination of all feature sets with RF classifier (UAR of 78%)	The results show that multi-feature analysis is a promising means to help identifying the anatomical mechanisms of snore sound
Qian et al., 2017 [56]	Germany	40	Male patients diagnosed with OSA or primary snoring	To evaluate two kinds of wavelet features for their performance when classifying snore sounds	Headset microphone and handheld microphone	WTE and WPTE	DISE (VOTE)	SVM, k-NN, LDA, RF, ELM, KELM, MLP, and DNN	WTE features achieved the highest UAR of 60.4% using SVM	The state-of-the-art machine learning methods including were not as efficient when compared to conventional classifiers (e. g., SVM)
Qian et al., 2018 [57]	Germany	40	Male patients diagnosed with OSA or primary snoring	To compare the performance of each kind of feature set	Headset microphone and handheld microphone	Crest factor, SERs, F1, F2, F3, MFCCs, power ratio (800 Hz), WTE, WPTE, WEF	DISE (VOTE)	SVM, K-NN, LDA, RF, ELM, KELM, MLP, and DNN	The highest UAR of 72.8% was obtained by the combination of SERs feature set and DNN classifier	There are no significant differences between varied classifiers when fed with a same feature set

										The combination of SERs feature set and DNN classifier can present the best performance
Qian et al., 2019 [58]	Germany	219	Publicly available snore sound corpus (MPSSC) based on OSA patients	To test a novel method based on multiresolution WT and BoAW approach	Headset microphone, stand-mounted microphone, handheld microphone, and fixed microphone on forehead	Wavelet features (WTE and WPTE) and conventional features (F1, F2, F3, MFCCs, SERs, and SFFs)	DISE (VOTE)	Naïve Bayes	The UAR based on wavelet features was 69.4%	The machine listening model based on wavelet features outperformed the model based on conventional snoring sound parameters
Schmitt et al., 2016 [59]	Germany	24	Subjects diagnosed with primary snoring or OSA	To improve classifier model	Headset microphone	MFCCs, F1, F2, F3, and wavelet features (energy, variance, waveform length, and entropy of wavelet coefficients)	DISE (VOTE)	SVM	The combination of MFCCs, formants, and wavelet-based features achieved an UAR of 79.5%	The combination of MFCCs, formants, and wavelet-based features are suitable for the classification of snore sounds
Sebastian et al., 2019 [60]	Australia	13	OSA patients	To use snoring data to classify three sites (lateral wall, palate, and tongue base) of obstruction in the upper airway.	Microphone placed on the ceiling above the patient's bed during PSG	MFCCs based feature set	Airflow shape (lateral wall, palate, and tongue base)	GMM	The model achieved an overall accuracy of 78.9%	Acoustic properties of snoring sound are likely helpful in identifying different obstruction sites
Sebastian et al., 2020 [61]	Australia	45	OSA patients	To automatically classify OSA patients into four categories based on the predominant site-of-collapse with snore data	Microphone placed on the ceiling above the patient's bed during PSG	MFCCs and Chroma features	Airflow shape (palate, lateral wall, tongue base, and multi-level)	LDA	The model achieved an overall accuracy of 65%	The snore signal analysis can be used to identify the predominant site-of-collapse in OSA patient
Sebastian et al., 2020 [62]	Australia	58	OSA patients	to automatically classify OSA patients into four categories based on the predominant site-of-collapse of the hypopnoea events of a night's recording using snore data	Microphone placed on the ceiling above the patient's bed during PSG	F0, MFCC, spectral entropy	Airflow shape (palate, lateral wall, tongue base, and multi-level)	LDA	The model achieved an overall accuracy of 62%	The results demonstrate that the audio signal recorded during sleep can successfully identify the site-of-collapse in the upper airway
Sun et al., 2020 [63]	China	219	Publicly available snore sound corpus (MPSSC) based on OSA patients	To develop an algorithm that distinguishes VOTE snoring	Headset microphone, stand-mounted microphone, handheld microphone, and fixed microphone on forehead	MFCC of the trend of the spectrum	DISE (VOTE)	SVM	By using the MFCC of the trend of the spectrum, the proposed approach achieves a UAR of 87.5%	The MFCC of the trend of the spectrum is a promising feature for capturing the characteristics of snoring
Sun et al., 2021 [64]	China	219	Publicly available snore sound corpus (MPSSC) based on OSA patients	To develop an algorithm to improve the accuracy of locating an obstructive site by capturing the information of the state embedded in snoring	Headset microphone, stand-mounted microphone, handheld microphone, and fixed microphone on forehead	Transformed signal and MFCC	DISE (VOTE)	SVM	The model yielded an overall accuracy of 92.2%	The characteristics of snores are related to the state of the upper airway
Vesperini et al., 2018 [65]	Italy	219	Publicly available snore sound corpus (MPSSC) based on OSA patients	To identify the type of snoring among four target classes representing the snore	Headset microphone, stand-mounted microphone, handheld microphone, and fixed	SCAT	DISE (VOTE)	DNN	A UAR of 67.7% was obtained by combining SCAT and DNN model	The DNN based classifier in combination with SCAT is an effective means to

(continued on next page)

Table 2 (*continued*)

Study	Country	N	Population	Aim(s)	Recording of snoring sounds	Snoring sound parameter(s)	Identification of the obstruction site(s)	Model (classifier)	Main finding(s)	Conclusion
				sounds' excitation location	microphone on forehead					identify the snore sounds' excitation location
Zhang et al., 2020 [66]	China	76	Subjects diagnosed with primary snoring or OSA	To classify the vibration patterns of the snoring source	Non-contact microphone (30 cm above patient's mouth)	Compressed HOG	DISE (V, E, O, V + T/E, V + O, VOTE)	SVM	The model achieved an accuracy of 89.8% for recognizing the seven vibration patterns of the snoring source based on the sound recordings	It was demonstrated that the noncontact acoustic analysis is a promising alternative to determine the snoring source of patients

Note. BoAW = bag-of-audio-word; CNN = convolutional neural network; DISE = drug-induced sleep endoscopy; DNN = Deep Neural Networks; ELM = Extreme Learning Machines; EMDF = Empirical Mode Decomposition-Based Features; F0 = fundamental frequency; F1 = the first formant; F2 = the second formant; F3 = the third formant; FNN = Feedforward Neural Networks; GMM = Gaussian mixture model; HOG = Histogram of Oriented Gradients; KELM = Kernel-ELM; K-NN = K-Nearest Neighbors; LBP = Local Binary Pattern; LDA = Linear Discriminant Analysis; LLDs = low-level descriptors; MFCCs = Mel-frequency cepstral coefficients; MLP = Multilayer Perceptrons; MPSSC = Munich Passau Snore Sound Corpus; OSA = obstructive sleep apnea; PSG = polysomnography; RF = Random Forests; SERs = subband energy ratios; SFFs = spectral frequency features, which include peak frequency and center frequency; SCAT = deep scattering spectrum; SVM = support vector machine; UAR = Unweighted Average Recall; VOTE = velum, oropharyngeal lateral walls, tongue base, and epiglottis classification; WEF = wavelet energy features, which include WPTE and WTE; WPTE = wavelet packet transform energy; WT = wavelet transformation; WTE = wavelet transform energy.

algorithm to extract snoring sounds. Osborne et al. [47] only stated that snoring events were confirmed on DISE video. The other three studies [40,44,46] did not mention based on which criterion the snoring sounds were identified.

3.3.4. Snoring sound parameters

The acoustic characteristics of sound can be divided into three domains, which are the temporal, intensity, and frequency domains. Of the included studies, various snoring sound parameters in all domains were used (Table 1, Table 2). The explanations of all snoring sound parameters are shown in Table 3 and Table 4. It was noted that five studies included snoring sound intensity (sound pressure level), but different units were used. Koo et al. [44] and Lee et al. [45] used decibel (dB) as unit to describe sound intensity, whereas Herzog et al. [43] used A-weighted decibel (dBA) for sound intensity. A-weighting is characterized by attenuation of low frequency (<1000 Hz) sounds as to imitate sound intensity perceived by the human ear [78]. Won et al. [49] did not specify which unit they used to describe sound intensity. Chang et al. [41] used the amplitude of the sound wave as a representation of sound intensity. In addition, two studies [43,49] also used psychoacoustic parameters, which is a branch of psychophysics involving the scientific study of sound perception and audiology, ie, how humans perceive various sounds [72,73] (Table 3). The psychoacoustic parameters also can be divided into intensity domain (loudness and shimmer; Table 3) and frequency domain (jitter, pitch, and sharpness; Table 3).

3.3.5. Methods and criteria for the identification of the obstruction sites

Of the 28 included studies (Table 1, Table 2), seven studies used different methods (dynamic MRI [41], monitoring pressure [46,50], SVF [49], and airflow shape [60–62]) and different criteria to locate the obstruction sites. The other 21 studies used DISE to identify the obstruction sites. Of the 21 DISE studies, seven studies [36,42–44,47,48,66] used different criteria to classify the obstruction sites; 14 studies used the VOTE classification [40,45,51–59,63–65]. By using the VOTE classification, not only the level (velum, oropharyngeal lateral wall, tongue base, and epiglottis), but also the degree (no obstruction [collapse less than 50%], partial obstruction [collapse between 50% and 75%], and complete obstruction [collapse more than 75%]) and configuration (anteroposterior [A-P], lateral, and concentric) of the obstruction sites can be assessed [79].

As for the degree of the obstruction, 18 studies [36,40,41,46–50,53–57,59–62,66] did not indicate the degree of the obstruction. Seven studies [42,51,52,58,63–65] only included patients with partial obstruction in the upper airway. Koo et al. [44] only included obstruction sites with collapse more than 75%. Lee et al. [45] included both partial (77–99%) and complete (100%) obstruction sites. In addition to velar snoring and tonsillar snoring, Herzog et al. [43] also included obstructive velar snoring and snoring terminating an apnea event. No study distinguished the configuration of the obstruction site. It needs to be noted that all 28 studies included snorers whose obstruction site(s) is (are) also the excitation/vibration site(s) of snoring sound. Of the 28 studies (Tables 1 and 2), 17 studies [36,43,44,47,50–52,54–60,63–65] only included snorers with single-level obstruction, while the other eleven studies [40–42,45,46,48,49,53,61,62,66] included snorers with both single-level obstruction and multilevel obstruction, ie, snoring sound was generated by the vibration of more than one site in the upper airway.

Another point that needs to be noted is that, of the 28 included studies, all snoring sounds were recorded simultaneously with the identification of the obstruction sites in the upper airway, except for the studies from Lee et al. [45] and Sebastian et al. [60–62], in

Table 3
Explanations of the snoring sound parameters that are used in studies reporting the association between snoring sound parameters and the obstruction sites.

Domain	Parameter	Explanation
Temporal	Duration of soft tissue vibration [41]	The duration of soft tissue vibration obtained by measuring the duration of snoring sound
Intensity	Crest factor [47]	Crest factor indicates how extreme the peak of a sound wave is. Crest factor of 1 indicates no peaks and higher crest factors indicate peaks. Crest factor is the peak amplitude of the waveform divided by the root mean square (RMS) value of the waveform $Crest\ factor = \frac{\lvert Amp_{Peak}\rvert}{Amp_{RMS}}$ $\lvert Amp_{Peak}\rvert =$ the absolute peak amplitude $Amp_{RMS} =$ the RMS of amplitudes
Intensity	Sound pressure level (SPL; also known as sound intensity) [43]	A logarithmic measurement of the effective pressure of a sound relative to a reference value $dBSPL = 20 \times \log_{10}\left(\frac{Sound\ pressure_{RMS}}{Sound\ pressure_{ref}}\right)$ $Sound\ pressure_{RMS} =$ the RMS of sound pressure $Sound\ pressure_{ref} =$ the reference sound pressure
Frequency	Center frequency [40]	A measure of a central frequency between the upper and lower cutoff frequencies. It is usually defined as the mean of the lower cutoff frequency and the upper cutoff frequency of a band-pass system or a band-stop system $Center\ ferquency = \frac{\int_0^{Fs/2} fP(f)df}{\int_0^{Fs/2} P(f)df}$ $P(f) =$ spectral power density $Fs =$ sampling frequency $df =$ the standard deviation for the frequency range
Frequency	Dominant frequency (DF; also known as peak frequency) [42]	The frequency that carries the most energy, ie, the frequency with the largest amplitude on a spectrum
Frequency	Fundamental frequency (F0) [44]	The lowest frequency of a waveform. In mathematics, the F0 of a signal is the greatest common divisor (GCD) of all the frequency components contained in a signal
Frequency	Formants (F1, F2, F3) [44]	A group of frequencies amplified by a resonator, ie, three specific frequencies of snoring sound that are amplified by resonating in the upper airway $H_{(z)} = \frac{1}{1 - \sum_{q=1}^{p} \alpha_q z^{-q}}$ $\alpha_q\ (q = 1, 2, 3, \ldots, p) =$ the parameters of linear predictive coding (LPC), one of the most powerful speech analysis techniques
Frequency	Mel-frequency cepstral coefficients (MFCCs) [55]	MFCCs is a set of coefficients that are derived from a spectrum, which can better represent human's perception to sound, because the Mel scale approximates the human auditory system's response MFCCs are commonly derived as follows: 1. Take the Fourier transform of (a windowed excerpt of) a signal 2. Map the powers of the spectrum obtained above onto the Mel scale, using triangular overlapping windows 3. Take the logs of the powers at each of the Mel frequencies 4. Take the discrete cosine transform of the list of Mel log powers, as if it were a signal 5. The MFCCs are the amplitudes of the resulting spectrum
Frequency	Power ratio [57]	The relative amount of sound emanating below and above a certain frequency $Power\ ratio_{800} = \log_{10}\frac{\sum_{f_i=0}^{800}(\lvert Amp_{f_i}\rvert)^2}{\sum_{f_i=800}^{f_c}(\lvert Amp_{f_i}\rvert)^2}$ $Power\ ratio_{800} =$ the power ratio at the frequency of 800 Hz $f_c =$ the cut-off frequency of the snoring sound spectrum $\lvert Amp_{f_i}\rvert =$ the absolute amplitude of the spectrum at the frequency of f_i
-	Collapse index [41]	The smallest area percent of the axial airway during snoring, obtained in dynamic MRI report
-	Snoring index [45]	The number of snore events per hour during sleep (events/hour)
-	Psychoacoustic parameters [72,73]	Psychoacoustics is the branch of psychophysics involving the scientific study of sound perception and audiology, ie, how humans perceive various sounds
Intensity	- Loudness	It is the counterpart of sound pressure level (see below) in psychoacoustics, measured in phon. It is the subjective perception of sound pressure by humans $N = \int_0^{24Bark} N'dz$ $N =$ loudness $N' =$ specific loudness $dz =$ differential
Intensity	- Shimmer	A measurement for the instability of intensity
Frequency	- Jitter	A measurement of the instability of frequency
Frequency	- Pitch	Pitch is a perceptual property of sounds that allows their ordering on a frequency-related scale
Frequency	- Sharpness	It is a measurement of the high frequency content of a sound, measured in acum. The bigger the proportion of high frequencies, the "sharper" the sound $S = 0.11\frac{\int_0^{24Bark} N'g(z)zdz}{\int_0^{24Bark} N'dz}$ $S =$ Sharpness $g(z) =$ critical band rate

Table 4
Explanations of the snoring sound parameters that are used in studies reporting prediction models of the obstruction sites.

Domain	Parameter	Explanation
Temporal-frequency	Histogram of Oriented Gradients (HOG) and Local Binary Pattern (LBP) [74]	LBP is a type of visual descriptor used for classification in computer vision, and HOG is a feature descriptor used in computer vision and image processing for the purpose of object detection. The combination of LBP and HOG can extract distinctive features from the spectrogram images of snoring sound and improve the detection performance
Temporal-frequency	Chroma features [75]	Chroma features are the representation of a music audio. The entire spectrum of the music audio is projected onto 12 bins, representing 12 different semitones (or chroma) of the musical octave
Temporal-frequency	Deep scattering spectrum (SCAT) [76]	An efficient representation of an audio signal based on the scattering transform, which is a new time-frequency signal processing tool
Temporal-frequency	Deep spectrum features [51]	Visual representations for audio data, ie, plots of spectrograms or chromagrams, that are drawn from DeepSpectrum, which is a Python toolkit
Frequency	Empirical Mode Decomposition-Based Features (EMDF) [55]	Empirical Mode Decomposition (EMD) is an adaptive time-space analysis method suitable for processing series that are non-stationary and non-linear. EMDF include the subband EMD energy ratio and the entropy of the subband EMD energy ratio $EMD_{ratio}(k) = E_k/E$ $H_{EMD_{ratio}} = -\sum_{k=1} EMD_{ratio}(k)\log_{10}(EMD_{ratio}(k))$ EMD_{ratio} = subband EMD energy ratio $H_{EMD_{ratio}}$ = entropy of the subband EMD energy ratio E_k = the energy (sum of squares) of the k -th level intrinsic mode functions (IMFs) decomposed by EMD from the snoring sound E = the total energy of the whole snoring sound within EMD
Frequency	Spectral entropy [62]	The spectral entropy represents the spectral power distribution of a signal and measures how sinusoidal the signal is
Frequency	Subband energy ratio (SER) [57]	SER describes the relative energy distribution in subbands of the snoring sound spectrum $SER_{1000}(j) = \frac{\sum_{f_i=1000j-1}^{1000j}(\lvert Amp_{f_i}\rvert)^2}{\sum_{f_i=0}^{f_c}(\lvert Amp_{f_i}\rvert)^2}$ $SER_{1000}(j)$ = 1000 Hz SER feature set j = 1, 2, 3, …, 8. f_c = the cut-off frequency of the snoring sound spectrum $\lvert Amp_{f_i}\rvert$ = the absolute amplitude of the spectrum at the frequency of f_i
Frequency	Wavelet transform (WT) [54–58]	WT is a tool for the analysis of signals in frequency domain. Different with the fixed window in Fourier transform, WT has flexible window, by which a higher time resolution can be provided for high frequency components and a higher frequency resolution can be provided for low frequency components
Frequency	- Wavelet packet transform energy (WPTE)	Wavelet packet transform (WPT) is a kind of WT, by which both the high frequency component and low frequency component of a signal can be decomposed The WPTE is calculated based on WPT coefficients (see below for the information about wavelet coefficients): $E_{\phi_{l,m}} = \sqrt{\frac{\sum_n (W_{l,m,n})^2}{N_m}}, n = 0, 1, 2, \ldots, 2^{l-1}$ $W_{l,m}$ = WPT coefficients obtained from the signal at the subspace $V_{l,m}$ N_m = the number of wavelet coefficients in the m -th decomposition level
Frequency	- Wavelet transform energy (WTE)	The WTE is calculated based on WT coefficients: $\widehat{E}_{\widehat{\phi}_l} = \frac{(\widehat{W}_l)^2}{\sum_{l=1}^{L_{max}}(\widehat{W}_l)^2}$ $\widehat{W}_l$ = coefficients generated by WT at the l-th decomposition level L_{max} is the maximum level for wavelet decomposition
Frequency	- Wavelet coefficients [54–58]	Wavelet analysis is a measurement of similarity between the selected wavelet and the original function. The wavelet coefficients indicate how close the function is to the wavelet at each decomposition level

which the snoring sounds were recorded during polysomnography (PSG).

3.4. Association between snoring sound parameters and the obstruction site

3.4.1. Parameters in temporal domain

Of the included studies, only one study, which was classified as a high-quality study, reported the association between the snoring sound parameters in temporal domain and the obstruction site in the upper airway. Chang et al. [41] used dynamic MRI to identify the obstruction site and found that there was a significant correlation between collapse index (ie, the smallest percent area of the axial airway during snoring; Table 3) and soft tissue vibration time ($P < 0.03$), and that the collapse index and soft tissue vibration time were significantly different between pure retropalatal and combined snoring ($P < 0.0001$). In summary, the available studies are not indicative of association between the obstruction sites in the upper airway on the one hand and the snoring sound parameters in temporal domain on the other hand.

3.4.2. Parameters in intensity domain

Four studies reported findings of the association between snoring sound intensity and obstruction sites identified using DISE. Lee et al. [45] found that participants with a higher mean intensity (≥66 dB) at a frequency of 301–850 Hz were 1.14 times more likely to have epiglottic obstruction than participants with a lower mean intensity (<66 dB; $P = 0.02$). However, Koo et al. [44] demonstrated that there was no significant difference in the intensity of the snoring sounds between retropalatal and retrolingual level obstructions ($P = 0.8$). Two moderate-quality studies from Herzog et al. [43] and Osborne et al. [47] also reported the associations between snoring sound intensity and the obstruction sites. Herzog et al. [43] found that the sound pressure level of obstructive velar snoring was higher than those of velar snoring and tonsillar snoring ($P < 0.05$), and that velar snoring showed lower loudness (Table 3) than that of tonsil snoring ($P < 0.05$). Osborne et al. [47] reported that a crest factor of 2.7 can be used to differentiate palatal snoring from snoring generated from lower segments in the upper airway. In summary, the available studies are not indicative of association between the obstruction sites in the upper airway on the one hand and the snoring sound parameters in intensity domain on the other hand.

3.4.3. Parameters in frequency domain

The high-quality studies from Lee et al. [45] and Koo et al. [44] also reported the associations between the characteristic frequencies of snoring sound and the obstruction sites. Lee et al. [45] found that participants with more low-frequency snoring events were more likely to have velopharyngeal obstruction than participants with less low-frequency snoring events ($P = 0.010$). In addition, participants with a higher mean frequency were more likely to have lateral oropharyngeal wall obstructions than participants with a lower mean frequency ($P = 0.013$). Koo et al. [44] reported that the first and second formants (F1 and F2; Table 3) of retrolingual level obstruction were significantly higher than those of retropalatal obstruction ($P < 0.05$), suggesting that the characteristic frequencies of the snoring sound generated from the lower part of the upper airway was higher than those of the snoring sound generated from the upper part of the upper airway. This finding was supported by the SVF study from Won et al. [49], which was identified as high-quality study as well. In that study, it was found that the pitch and F1 of tongue base and epiglottis obstruction were significantly higher than those of the soft palate group ($P < 0.05$), while there was no significant difference in jitter and shimmer (Table 3) among different groups ($P > 0.05$). Moderate-quality studies also confirmed this finding. The DISE study from Gurpinar et al. [42] found that the mean frequency, fundamental frequency (F0; Table 3), and dominant frequency (DF; Table 3) of retrolingual snoring were significantly higher than those of retropalatal snoring ($P = 0.001$). Xu et al. [50] identified obstruction site(s) by monitoring the pressure in the upper airway and reported that the mean of DF and center frequency (Table 3) of the first snoring sounds after lower-level obstructive apneas were higher than those after upper-level obstructive apneas ($P < 0.05$). The DISE study from Herzog et al. [43] found that velum snoring had lower center frequency and sharpness (Table 3) than tonsillar snoring ($P < 0.05$). In addition, both DISE studies from Quinn et al. [36] and Saunders et al. [48] reported the lower center frequency of palatal snoring than that of tongue base snoring. The only deviating result was reported in another DISE study from Agrawal et al. [40], in which the center frequencies of palatal, tonsillar, tongue-base, and epiglottic snores were reported to be 391 Hz, 445 Hz, 1094 Hz, and 442 Hz, respectively.

Another finding was that multilevel obstruction that consisted of the upper-level obstruction and lower-level obstruction in the upper airway, which were found to generate snoring sounds with relatively low characteristic frequencies and relatively high characteristic frequencies, respectively, generated snoring sounds with medium characteristic frequencies. Results from all DISE studies, the SVF study, and the upper airway pressure study supported this finding. The high-quality SVF study from Won et al. [49] found that the pitch and F1 (Table 3) of both the soft palate + tongue base group and the soft palate + epiglottis group were significantly higher than those of the pure soft palate group ($P < 0.05$). The moderate-quality DISE study from Gurpinar et al. [42] found that the mean frequency, F0, and DF of multilevel snoring are significantly lower than those of retrolingual snoring ($P = 0.025$) and are higher than those of retropalatal snoring, respectively, but the differences with retropalatal snoring are not significant ($P = 0.419$). In addition, another DISE study [40] reported that the DF and center frequency were 137 Hz (range: 105–189 Hz) and 391 Hz (range: 253–1027 Hz) in palatal snoring, 1243 Hz (range: 1215–1277 Hz) and 1094 Hz (range: 1059–1200 Hz) in tongue base snoring, and 190 Hz (range: 115–223 Hz) and 404 Hz (312–605 Hz) in mixed palate and tongue snoring. The moderate-quality upper airway pressure study from Miyazaki et al. [46] reported that the average F0 was 102.8 ± 34.9 Hz in soft palate snoring, 331.7 ± 144.8 Hz in tonsil/tongue base snoring, and 115.7 ± 58.9 Hz in mixed snoring.

In summary, available evidence points toward associations between the snoring sound parameters in frequency domain and the obstruction sites in the upper airway. It is found that the characteristic frequencies (center frequency, DF, F1, F2, mean frequency, and pitch; Table 3) of the snoring sound generated from the lower level of the upper airway are higher than those of the snoring sound generated from the upper level of the upper airway, and that multilevel obstruction that consists of the upper-level obstruction and lower-level obstruction in the upper airway generates snoring sounds with medium characteristic frequencies (center frequency, DF, F0, F1, mean frequency, and pitch; Table 3).

3.5. Prediction model of the obstruction site based on snoring sound parameters

In the present review, in addition to the above-mentioned studies on the associations between snoring sound and obstruction site(s), 16 prediction model studies were included as well. Of the 16 studies, 13 studies used DISE to identify the obstruction sites, while three studies from Sebastian et al. [60–62] identified the obstruction sites based on the airflow shape. In prediction model studies, the accuracy of the prediction model is usually reported. Peng et al. [53] built a prediction model based on F0 and formants (F1, F2, F3; Table 3) using conventional statistical methods. The authors found that F0 and F2 can be used to differentiate palatal snoring from mixed type (palatal snoring + lateral wall snoring) and pure lateral wall snoring. The overall accuracy of F0 and F2 were 60.8% and 62.4%, respectively. The three studies from Sebastian et al. [60–62] built machine learning models based on Mel-frequency cepstral coefficients (MFCCs; Table 3), MFCCs + Chroma features (Table 3, Table 4), MFCCs + F0 + spectral entropy (Table 3, Table 4), respectively, as to classify soft palate, lateral wall, and tongue base snoring. The authors reported that the respective overall accuracies of the three models were 78.9%, 65%, and 62%. Vesperini et al. [65], Sun et al. [63,64], and Zhang et al. [66] built prediction models based on spectrum-related characteristics (Table 4) and reported an UAR of 67.7% [65], an UAR of 87.5% [63], an overall accuracy of 92.2% [59], an accuracy of 89.8% [66], respectively.

The other eight studies were performed by a same team. The authors extensively tested the performance of machine learning models with different classifiers and snoring sound parameters; the reported UAR ranged from 60.4% [56] to 79.5% [59]. The snoring sound parameters in these studies included conventional acoustic parameters [54,58] (eg, crest factor, F0, MFCCs; Table 3), newly used snoring sound parameters [56,57] (eg, wavelet transform energy [WTE], wavelet packet transform energy [WPTE]; Table 4), and spectrum-related characteristics in other machine learning models [51,52] (eg, deep spectrum features, Local Binary Pattern [LBP], Histogram of Oriented Gradients [HOG]; Table 4). The machine learning models included conventional classifiers [52,54] (eg, support vector machine [SVM]) and the state-of-the-art classifiers

[56,57] (eg, Deep Neural Networks [DNN]). The authors concluded that there was no significant difference between different classifiers when fed with a same feature set [57]. In addition, it was found that the classifier model based on wavelet features outperformed the model based on conventional snoring sound parameters [58], and that multi-frequency feature (MFCCs, formants, and wavelet-based features; Table 3, Table 4) analysis was a promising means to help identifying the anatomical mechanisms of snoring sound [59].

In summary, the reported accuracies of prediction models are promising. For machine learning models, there was no significant difference between different classifiers when fed with a same feature set, and multi-feature analysis showed better performance in the prediction of the obstruction sites.

4. Discussion

Previous studies reported the informative nature of snoring sound and hypothesized that snoring sound parameters are potential predictors of the obstruction sites in the upper airway. Because of the limitations of the existing diagnostic modalities (see Introduction), researchers and clinicians try to predict the obstruction sites in the upper airway by analyzing snoring sound parameters. This method can provide information on the anatomy of the upper airway during natural sleep. In addition, it is more affordable and patient-friendly than the majority of the existing modalities, since it only needs a snoring sound recording during sleep; no drugs or other instruments are needed. In theory, snoring sound analysis seems a feasible method to predict the obstruction sites, but its feasibility has not been investigated. Therefore, the present review aimed to i) investigate the association between snoring sound parameters and the obstruction sites determined by objective methods, eg, DISE or dynamic MRI; and ii) analyze the methodology and the participant characteristics of the studies reporting prediction models for the obstruction sites as to provide suggestions for further study.

4.1. Association between the characteristic frequencies of snoring sound and the obstruction site

One of the main findings of this review is that the characteristic frequencies (center frequency, DF, F1, F2, mean frequency, and pitch; Table 3) of the snoring sound generated from the lower level of the upper airway were higher than those of the snoring sound generated from the upper level of the upper airway. This finding is partially supported by another systematic review, which reported that snoring sound that originated from the epiglottis level had a higher pitch than palatal snoring [80]. Another finding is that multilevel obstruction that consisted of the upper-level obstruction and lower-level obstruction in the upper airway generated snoring sounds with medium frequencies (center frequency, DF, F0, F1, mean frequency, and pitch; Table 3). It has been known that the dimensions of a tube have important effect on the resonance frequencies [81], where a smaller and narrower tube will lead to higher resonance frequencies. The influence of the dimensions of the upper airway on breathing sound has been reported as well. Rembold and Suratt found that children who generated high-frequency inspiration sounds had a significantly narrower upper airway ($P = 0.02$) [82].

4.2. Association between snoring sound intensity and the obstruction site

Results from the included studies are not indicative of an association between snoring sound intensity and the obstruction sites. This may be due to the differences caused by different recording systems used in the studies. In addition, Saha et al. [83] reported that the only factor that may positively and significantly correlate to snoring sound intensity was the degree of the upper airway narrowing. This was reported by Chang et al. [41] as well. In another study from Ng et al. [84], the authors investigated the impact of changing the cross-sectional areas of the pharynx and oral cavity on snoring sound, and found that narrowing the pharyngeal airway consistently increased the amplitude of F1. Taken together, these results suggest that the effect of the degree of the upper airway obstruction on snoring sound should be taken into consideration in further studies.

4.3. Influence of participant characteristics on snoring sound

A point that needs to be noted is that, in the prediction model studies, the authors included snoring sound parameters, but did not include participant characteristics, which would have helped to improve the prediction model. The influence of participant characteristics on snoring sound has been reported. Azarbarzin and Moussavi [85] found that gender, body mass index (BMI), height, and apnea hypopnea index (AHI) were the parameters that can influence snoring sound significantly ($P < 0.05$). This finding was also supported by other studies [86–89]. Among these studies, the positive association between snoring sound intensity and AHI was thoroughly discussed [86,87]. Wilson et al. [88] reported that the mean snoring sound intensity was significantly higher for men than for women ($P < 0.05$), indicating an influence of gender on snoring sound. BMI was reported to be positively associated with snoring sound intensity in studies from Peng et al. [87] and Wilson et al. [88]. Ng et al. [89] reported the positive association between neck circumference and the loudness, annoyance, and roughness of snoring sound. The positive association between height and snoring sound is a finding of interest, and we hypothesize that this association may attribute to the association between the length of upper airway and snoring sound. Such association was reported previously by Saha et al. [83], viz., the length of the upper airway was inversely correlated with the resonant frequencies of snoring sound ($P < 0.05$). The above-mentioned evidence suggests that including participant characteristics as independent factors may be a good way to improve the performance of prediction models.

4.4. Methods for recording and identification of the snoring sound

It needs to be noted that, to date, there are no standard recording and analysis procedures for snoring sound. The application of different recording and analysis systems will definitely lead to difficulties in comparing results between studies. As a first step of snoring sound-related studies, sufficient attention should be given to the quality of the snoring sound recording, since it will undoubtedly influence the corresponding analysis results. Among all the technical details in the recording of snoring sound, the position of the microphone may be the most important one. In a previous study, Herzog et al. [90] found that the microphone position influenced the frequency components of snoring sound. Specifically, non-contact microphones had a wider frequency range than contact microphones, which presents a decreased sensitivity to the frequency components above 1000 Hz. The authors also suggested that a contact microphone was a useful screening device (eg, in polygraphy devices), whereas a non-contact microphone was the better choice for a natural analysis of snoring sounds. Another study from Azarbarzin et al. [91] reported that the snoring sounds recorded using a non-contact microphone were not as characteristic as those recorded over the trachea, and the tracheal snoring sounds showed better performance in clustering snoring sounds into two groups based on their acoustic parameters.

Due to the partially contradictory results, no conclusion can be drawn about which kind of microphone is more suitable to make recordings for snoring sound analyses. However, given the desired comfortableness of the sleeper, a non-contact microphone may be a better choice. In addition, snoring sounds recorded by a non-contact microphone close to snorer can register the snoring sounds as they are perceived by snorers' bed partners. This may be helpful in investigating the disturbance caused by the snoring sounds to snorers' bed partners. Also, for a sound recording with a contact microphone, one has to take into consideration the sound conduction effect of subcutaneous tissue, especially in an OSA population with markedly varying degrees of obesity [92]. Probably, these are the reasons why the majority of the included studies performed snoring sound recordings using non-contact microphones. For non-contact microphones, it is important that the generated snoring sounds go directly to the microphone so that the microphone can receive as much signals as possible, and that the reflection of snoring sound by walls and furniture should be minimized as to obtain the original snoring sound, which means that the non-contact microphone should be placed in free field (direct field), where the reflection of sound is negligible [93]. Therefore, the non-contact microphone should point to the sleeper's mouth and the distance between the non-contact microphone and patient's mouth should be shorter than that between the non-contact microphone and the nearest wall or furniture that may lead to the reflection of sound.

After the recording and pre-processing of snoring sound, another hurdle one may be confronted with is the identification of snoring events. Normally, researchers set an acoustic criterion by which the snoring events are identified. Unfortunately, this criterion is highly variable and subjective due to the lack of standardization and the existence of various recording systems. The variable and subjective identification procedure leads to a variety in the field of snoring sound analysis, and to difficulties in comparing the various studies, which consequently hinders the development of an evidence-based snoring sound-related study. In previous studies, the employed acoustic criterion for the identification of snoring events varied from 36 dB to 76 dB and from 40.5 dBA to 60 dBA [26,27,94,95] with a large variation of the distance between microphone and patient's mouth. Leto et al. [26] set a sound intensity of 36 dB as the criterion of detecting snoring sounds. This level was based on the results of their pilot studies, as well as on the suggestion from World Health Organization (WHO), viz., indoor continuous sound pressure level above 30 dB should be avoided during sleep [96]. Rohrmeier et al. [27] used a sound intensity of 46.6 dBA to differentiate snoring sound from breathing sound, and obtained an optimum sensitivity and specificity of 76.9% and 78.8%, respectively. Blumen et al. [95] identified snoring events based on both the frequency and the intensity of snoring sound; the authors defined snoring sound as sound of which the frequency band was within 20–200 Hz, and the sound intensity was higher than 76 dB. It is undoubtedly better to identify snoring events based on snoring sound parameters in both intensity and frequency domains than only based on snoring sound parameters in intensity domain. However, it seems that a frequency band of 20–200 Hz is too narrow to include all kinds of snoring sounds, sounds of which the frequencies can range from less than 200 Hz to higher than 1000 Hz [46,50].

4.5. Limitations

This review inevitably has some limitations. Firstly, all 28 studies only included snorers whose obstruction site(s) is (are) also the excitation/vibration site(s) of snoring sound. However, it is a clinical reality that lots of snorers, especially snoring OSA patients, have multilevel obstruction and usually only one obstruction site generates snoring sounds while the other obstruction site(s) does (do) not generate snoring sounds. To the best of the authors' knowledge, no study has been performed on the association between snoring sound parameters and all obstruction sites (both generating and not generating snoring sound). It would be clinically meaningful to investigate the predictive role of snoring sound on multilevel obstruction and to explore the possibility of predicting all obstruction sites in the upper airway rather than only the excitation site. Secondly, given the fact that there is no gold standard to investigate the upper airway during sleep and to identify the obstruction site, all studies using DISE, airflow shape, dynamic MRI, SVF, and upper airway pressure monitoring were included in this review. This prevents the potential bias caused by the use of a single reference test but leads to the difficulty in comparing the results between studies. A positive side of including multiple reference tests is that the main findings in the present review were supported by studies using all the above-mentioned modalities, strengthening the reliability of the findings. Lastly, as mentioned above, the studies and outcomes were too heterogeneous (different populations, different snoring sound parameters, different methods for the identification of the obstruction sites, etc.) to conduct a meta-analysis for further results comparison.

5. Conclusion

In this systematic review, available evidence points toward association between the snoring sound parameters in frequency domain and the obstruction sites in the upper airway. More studies are needed to investigate the associations between the obstruction sites on the one hand and the snoring sound parameters in temporal and intensity domains on the other hand. In addition, it is promising to build a prediction model of the obstruction sites based on snoring sound parameters and participant characteristics, but so far snoring sound analysis does not seem to be a viable diagnostic modality for treatment selection.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Acknowledgement

None.

Conflict of interest

None.

The ICMJE Uniform Disclosure Form for Potential Conflicts of Interest associated with this article can be viewed by clicking on the following link: https://doi.org/10.1016/j.sleep.2021.10.015.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.sleep.2021.10.015.

References

[1] Senaratna C, Perret J, Lodge C, et al. Prevalence of obstructive sleep apnea in the general population: a systematic review. Sleep Med Rev 2017;34:70–81. https://doi.org/10.1016/j.smrv.2016.07.002.

[2] Jordan A, McSharry D, Malhotra A. Adult obstructive sleep apnoea. Lancet 2014;383:736–47. https://doi.org/10.1016/S0140-6736(13)60734-5.
[3] Schwartz M, Acosta L, Hung Y, et al. Effects of CPAP and mandibular advancement device treatment in obstructive sleep apnea patients: a systematic review and meta-analysis. Sleep Breath 2018;22:555–68. https://doi.org/10.1007/s11325-017-1590-6.
[4] Lan M, Hsu Y, Lan M, et al. The predictive value of drug-induced sleep endoscopy for CPAP titration in OSA patients. Sleep Breath 2018;22:949–54. https://doi.org/10.1007/s11325-017-1600-8.
[5] Op de Beeck S, Dieltjens M, Verbruggen A, et al. Phenotypic labelling using drug-induced sleep endoscopy improves patient selection for mandibular advancement device outcome: a prospective study. J Clin Sleep Med 2019;15:1089–99. https://doi.org/10.5664/jcsm.7796.
[6] Strollo P, Soose R, Maurer J, et al., STAR Trial Group. Upper-airway stimulation for obstructive sleep apnea. N Engl J Med 2014;370:139–49. https://doi.org/10.1056/NEJMoa1308659.
[7] Vanderveken O, Maurer J, Hohenhorst W, et al. Evaluation of drug-induced sleep endoscopy as a patient selection tool for implanted upper airway stimulation for obstructive sleep apnea. J Clin Sleep Med 2013;9:433–8. https://doi.org/10.5664/jcsm.2658.
[8] Chong KB, De Vito A, Vicini C. Drug-induced sleep endoscopy in treatment options selection. Sleep Med Clin 2019;14:33–40. https://doi.org/10.1016/j.jsmc.2018.11.001.
[9] Viana A, Zhao C, Rosa T, et al. The effect of sedating agents on drug-induced sleep endoscopy findings. Laryngoscope 2019;129:506–13. https://doi.org/10.1002/lary.27298.
[10] Chang ET, Certal V, Song SA, et al. Dexmedetomidine versus propofol during drug-induced sleep endoscopy and sedation: a systematic review. Sleep Breath 2017;21:727–35. https://doi.org/10.1007/s11325-017-1465-x.
[11] Lee CH, Hong SL, Rhee CS, et al. Analysis of upper airway obstruction by sleep videofluoroscopy in obstructive sleep apnea: a large population-based study. Laryngoscope 2012;122:237–41.
[12] Kim DK, Lee WH, Lee CH, et al. Interrater reliability of sleep videofluoroscopy for airway obstruction in obstructive sleep apnea. Laryngoscope 2014;124:1267–71. https://doi.org/10.1002/lary.24509.
[13] Hudgel DW. Variable site of airway narrowing among obstructive sleep apnea patients. J Appl Physiol 1986;61:1403–9.
[14] Han D, Ye J, Wang J, et al. Determining the site of airway obstruction in obstructive sleep apnea with airway pressure measurements during sleep. Laryngoscope 2002;112:2081–5. https://doi.org/10.1097/00005537-200211000-00032.
[15] Genta P, Sands S, Butler J, et al. Airflow shape is associated with the pharyngeal structure causing OSA. Chest 2017;152:537–46. https://doi.org/10.1016/j.chest.2017.06.017. 10.1016/j.chest.2017.06.017.
[16] Razek A. Diagnostic role of magnetic resonance imaging in obstructive sleep apnea syndrome. J Comput Assist Tomogr 2015;39:565–71. https://doi.org/10.1097/RCT.0000000000000243.
[17] Zerpa VZ, Llatas MC, Porras GA, et al. Drug-induced sedation endoscopy versus clinical exploration for the diagnosis of severe upper airway obstruction in OSAHS patients. Sleep Breath 2015;19:1367–72. https://doi.org/10.1007/s11325-015-1266-z.
[18] Herzog M, Kellner P, Plößl S, et al. Drug-induced sleep endoscopy and simulated snoring in patients with sleep-disordered breathing: agreement of anatomic changes in the upper airway. Eur Arch Oto-Rhino-Laryngol 2015;272:2541–50. https://doi.org/10.1007/s00405-015-3559-z.
[19] American Academy of Sleep Medicine. International classifi cation of sleep disorders. 3rd ed. Darien, IL: American Academy of Sleep Medicine; 2014.
[20] Owens RL, Edwards BA, Sands SA, et al. The classical Starling resistor model often does not predict inspiratory airflow patterns in the human upper airway. J Appl Physiol 2014;116:1105–12. https://doi.org/10.1152/japplphysiol.00853.2013.
[21] Schwartz AR, Smith PL. CrossTalk proposal: the human upper airway does behave like a Starling resistor during sleep. J Physiol 2013;591:2229–32. https://doi.org/10.1113/jphysiol.2012.250654.
[22] Wellman A, Genta PR, Owens RL, et al. Test of the Starling resistor model in the human upper airway during sleep. J Appl Physiol 1985;117:1478–85. https://doi.org/10.1152/japplphysiol.00259.2014.
[23] Gold AR, Schwartz AR. The pharyngeal critical pressure. The whys and hows of using nasal continuous positive airway pressure diagnostically. Chest 1996;110:1077–88. https://doi.org/10.1378/chest.110.4.1077.
[24] Le TB, Moghaddam MG, Woodson BT, et al. Airflow limitation in a collapsible model of the human pharynx: physical mechanisms studied with fluid-structure interaction simulations and experiments. Phys Rep 2019;7:e14099. https://doi.org/10.14814/phy2.14099.
[25] Fajdiga I. Snoring imaging: could Bernoulli explain it all? Chest 2005;128:896–901. https://doi.org/10.1378/chest.128.2.896.
[26] Ieto V, Kayamori F, Montes M, et al. Effects of oropharyngeal exercises on snoring: a randomized trial. Chest 2015;148:683–91. https://doi.org/10.1378/chest.14-2953.
[27] Rohrmeier C, Herzog M, Ettl T, et al. Distinguishing snoring sounds from breath sounds: a straightforward matter? Sleep Breath 2014;18:169–76. https://doi.org/10.1007/s11325-013-0866-8.
[28] Herath D, Abeyratne U, Hukins C, et al. An investigation of critical frequency sub-bands of snoring sounds for OSA diagnosis. Conf Proc IEEE Eng Med Biol Soc 2019;2019:2568–71.
[29] Xue B, Deng B, Hong H, et al. Non-contact sleep stage detection using canonical correlation analysis of respiratory sound. IEEE J Biomed Health Inform 2020;24:614–25. https://doi.org/10.1109/JBHI.2019.2910566.
[30] Acar M, Yazıcı D, Bayar Muluk N, et al. Is there a relationship between snoring sound intensity and frequency and OSAS severity? Ann Otol Rhinol Laryngol 2016;125:31–6.
[31] Azarbarzin A, Moussavi Z. Snoring sounds variability as a signature of obstructive sleep apnea. Med Eng Phys 2013;35:479–85.
[32] Kalchiem-Dekel O, Westreich R, Regev A, et al. Snoring intensity and excessive daytime sleepiness in subjects without obstructive sleep apnea. Laryngoscope 2016;126:1696. 1670.
[33] Pevernagie D, Aarts RM, De Meyer M. The acoustics of snoring. Sleep Med Rev 2010;14:131–44.
[34] Owens F. Signal processing of speech. 1993.
[35] Childers D, Diaz J. Speech processing and synthesis toolboxes. J Acoust Soc Am 2000;108:1972–5.
[36] Quinn S, Huang L, Ellis P, et al. The differentiation of snoring mechanisms using sound analysis. Clin Otolaryngol Allied Sci 1996;21:119–23. https://doi.org/10.1111/j.1365-2273.1996.tb01313.x.
[37] Downes M, Brennan M, Williams H, et al. Development of a critical appraisal tool to assess the quality of cross-sectional studies (AXIS). BMJ Open 2016;6:e011458.
[38] Ow N, Mayo NE. Health-related quality of life scores of typically developing children and adolescents around the world: a meta-analysis with meta-regression. Qual Life Res 2020;29:2311–32. https://doi.org/10.1007/s11136-020-02519-0.
[39] Wolff R, Moons K, Riley R, et al., PROBAST Group. PROBAST: a tool to assess the risk of bias and applicability of prediction model studies. Ann Intern Med 2019;170:51–8. https://doi.org/10.7326/M18-1376.
[40] Agrawal S, Stone P, McGuinness K, et al. Sound frequency analysis and the site of snoring in natural and induced sleep. Clin Otolaryngol Allied Sci 2002;27:162–6. https://doi.org/10.1046/j.1365-2273.2002.00554.x.
[41] Chang Y, Huon L, Pham V, et al. There a relationship between snoring sound intensity and frequency and OSAS. Physiol Meas 2014;35:2501–12. https://doi.org/10.1088/0967-3334/35/12/2501.
[42] Gürpınar B, Saltürk Z, Kumral T, et al. Analysis of snoring to determine the site of obstruction in obstructive sleep apnea syndrome. Sleep Breath 2020. https://doi.org/10.1007/s11325-020-02252-5. 10.1007/s11325-020-02252-5.
[43] Herzog M, Plößl S, Glien A, et al. Evaluation of acoustic characteristics of snoring sounds obtained during drug-induced sleep endoscopy. Sleep Breath 2015;19:1011–9. https://doi.org/10.1007/s11325-014-1085-7.
[44] Koo S, Kwon S, Kim Y, et al. Acoustic analysis of snoring sounds recorded with a smartphone according to obstruction site in OSAS patients. Eur Arch Oto-Rhino-Laryngol 2017;274:1735–40. https://doi.org/10.1007/s00405-016-4335-4.
[45] Lee L, Lo Y, Yu J, et al. Snoring sounds predict obstruction sites and surgical response in patients with obstructive sleep apnea hypopnea syndrome. Sci Rep 2016;6:30629. https://doi.org/10.1038/srep30629.
[46] Miyazaki S, Itasaka, Ishikawa K, et al. Acoustic analysis of snoring and the site of airway obstruction in sleep related respiratory disorders. Acta Otolaryngol Suppl 1998;537:47–51.
[47] Osborne J, Osman E, Hill P, et al. A new acoustic method of differentiating palatal from non-palatal snoring. Clin Otolaryngol Allied Sci 1999;24:130–3. https://doi.org/10.1046/j.1365-2273.
[48] Saunders N, Tassone P, Wood G, et al. Is acoustic analysis of snoring an alternative to sleep nasendoscopy? Clin Otolaryngol Allied Sci 2004;29:242–6. https://doi.org/10.1111/j.1365-2273.2004.00800.x.
[49] Won T, Kim S, Lee W, et al. Acoustic characteristics of snoring according to obstruction site determined by sleep videofluoroscopy. Acta Otolaryngol 2012;132(Suppl 1):S13–20. https://doi.org/10.3109/00016489.2012.660733.
[50] Xu H, Huang W, Yu L, et al. Sound spectral analysis of snoring sound and site of obstruction in obstructive sleep apnea syndrome. Acta Otolaryngol 2010;130:1175–9. https://doi.org/10.3109/00016481003694774.
[51] Amiriparian S, Gerczuk M, Ottl S, et al. Snore sound classification using image-based deep spectrum features. INTERSPEECH 2017:3512–6.
[52] Demir F, Sengur A, Cummins N, et al. Low level texture features for snore sound discrimination. Conf Proc IEEE Eng Med Biol Soc 2018;2018:413–6. https://doi.org/10.1109/EMBC.2018.8512459.
[53] Peng H, Xu H, Xu Z, et al. Acoustic analysis of snoring sounds originating from different sources determined by drug-induced sleep endoscopy. Acta Otolaryngol 2017;137:872–6. https://doi.org/10.1080/00016489.2017.1293291.
[54] Qian K, Janott C, Zhang Z, et al. Wavelet features for classification of vote snore sounds. In: 2016 IEEE international conference on acoustics, speech and signal processing (ICASSP); 2016. p. 221–5.
[55] Qian K, Janott C, Pandit V, et al. Classification of the excitation location of snore sounds in the upper airway by acoustic multifeature analysis. IEEE Trans Biomed Eng 2017;64:1731–41. https://doi.org/10.1109/TBME.2016.2619675.
[56] Qian K, Janott C, Deng J, et al. Snore sound recognition: on wavelets and classifiers from deep nets to kernels. Conf Proc IEEE Eng Med Biol Soc 2017;2017:3737–40. https://doi.org/10.1109/EMBC.2017.8037669.
[57] Qian K, Janott C, Zhang Z, et al. Teaching machines on snoring: a benchmark on computer audition for snore sound excitation localisation. Arch Acoust Q 2018;43:465–75.
[58] Qian K, Schmitt M, Janott C, et al. A bag of wavelet features for snore sound classification. Ann Biomed Eng 2019;47:1000–11. https://doi.org/10.1007/s10439-019-02217-0.

[59] Schmitt M, Janott C, Pandit V, et al. A bag-of-audio-words approach for snore sounds' excitation localisation. Speech communication; 12. In: ITG symposium; 2016.
[60] Sebastian A, Cistulli P, Cohen G, et al. A preliminary study of the automatic classification of the site of airway collapse in OSA patients using snoring signals. Conf Proc IEEE Eng Med Biol Soc 2019;2019:1592–5. https://doi.org/10.1109/EMBC.2019.8856788.
[61] Sebastian a, Cistulli p, Cohen g, et al. Identifying the predominant site of upper airway collapse in obstructive sleep apnoea patients using snore signals. Annu Int Conf IEEE Eng Med Biol Soc 2020;2020:2728–31. https://doi.org/10.1109/EMBC44109.2020.9175626.
[62] Sebastian A, Cistulli P, Cohen G, et al. Automated identification of the predominant site of upper airway collapse in obstructive sleep apnoea patients using snore signal. Physiol Meas 2020;41:095005. https://doi.org/10.1088/1361-6579/abaa33.
[63] Sun J, Hu X, Chen C, et al. Amplitude spectrum trend-based feature for excitation location classification from snore sounds. Physiol Meas 2020;41:085006. https://doi.org/10.1088/1361-6579/abaa34.
[64] Sun J, Hu X, Peng S, et al. Automatic classification of excitation location of snoring sounds. J Clin Sleep Med 2021;17:1031–8. https://doi.org/10.5664/jcsm.9094.
[65] Vesperini F, Galli A, Gabrielli L, et al. Snore sounds excitation localization by using scattering transform and deep neural networks. In: 2018 international joint conference on neural networks (IJCNN); 2018.
[66] Zhang Y, Zhao Z, Xu H, et al. Vibration pattern recognition using a compressed histogram of oriented gradients for snoring source analysis. Bio Med Mater Eng 2020;31:143–55. https://doi.org/10.3233/BME-201086.
[67] Russell S, Norvig P. Artificial intelligence: a modern approach. 3rd ed. 2010.
[68] Bousquet O, von Luxburg U, Raetsch G. Advanced lectures on machine learning. 2004.
[69] Goehring T, Bolner F, Monaghan J, et al. Speech enhancement based on neural networks improves speech intelligibility in noise for cochlear implant users. Hear Res 2017;344:183–94. https://doi.org/10.1016/j.heares.2016.11.012.
[70] Montalvão J, Istrate D, Boudy J, et al. Sound event detection in remote health care - small learning datasets and over constrained Gaussian mixture models. Conf Proc IEEE Eng Med Biol Soc 2010;2010:1146–9. https://doi.org/10.1109/IEMBS.2010.5627149.
[71] Ben-Hur A, Horn D, Siegelmann H, et al. Support vector clustering. J Mach Learn Res 2002:125–37.
[72] Rohrmeier C, Herzog M, Haubner F, et al. The annoyance of snoring and psychoacoustic parameters: a step towards an objective measurement. Eur Arch Oto-Rhino-Laryngol 2012;269:1537–43. https://doi.org/10.1007/s00405-011-1878-2.
[73] Eberhard Z, Hugo F. Psychoacoustics-facts and models. 2nd ed. 1999.
[74] Wang X, Han T, Yan S. An HOG-LBP human detector with partial occlusion handling. In: 2009 IEEE 12th international conference on computer vision; 2009. p. 32–7.
[75] Jensen J, Christensen M, Ellis D, et al. A tempo-insensitive distance measure for cover song identification based on chroma features. In: 2008 IEEE international conference on acoustics, speech and signal processing; 2008. p. 2209–12. https://doi.org/10.1109/ICASSP.2008.4518083.
[76] Anden J, Mallat S. Deep scattering spectrum. IEEE Trans Signal Process 2014;62:4114–28.
[77] Janott C, Schmitt M, Zhang Y, et al. Snoring classified: the munich-passau snore sound corpus. Comput Biol Med 2018;94:106–18. https://doi.org/10.1016/j.compbiomed.2018.01.007.
[78] Pierre R, Maguire D. The impact of A-weighting sound pressure level measurements during the evaluation of noise exposure. Engineering; 2004.
[79] Kezirian E, Hohenhorst W, de Vries N. Drug-induced sleep endoscopy: the VOTE classification. Eur Arch Oto-Rhino-Laryngol 2011;268:1233–6. https://doi.org/10.1007/s00405-011-1633-8.
[80] Torre C, Camacho M, Liu S, et al. Epiglottis collapse in adult obstructive sleep apnea: a systematic review. Laryngoscope 2016;126:515–23. https://doi.org/10.1002/lary.25589.
[81] Wolfe J, Garnier M, Smith J. Vocal tract resonances in speech, singing, and playing musical instruments. HFSP J 2009;3:6–23. https://doi.org/10.2976/1.2998482.
[82] Rembold C, Suratt P. An upper airway resonator model of high-frequency inspiratory sounds in children with sleep-disordered breathing. J Appl Physiol 1985;98:1855–61. https://doi.org/10.1152/japplphysiol.01231.2004.
[83] Saha S, Bradley T, Taheri M, et al. A subject-specific acoustic model of the upper airway for snoring sounds generation. Sci Rep 2016 23;6:25730. https://doi.org/10.1038/srep25730.
[84] Ng A, Koh T, Baey E, et al. Role of upper airway dimensions in snore production: acoustical and perceptual findings. Ann Biomed Eng 2009;37:1807–17. https://doi.org/10.1007/s10439-009-9745-7.
[85] Ali Azarbarzin, Moussavi Zahra. Do anthropometric parameters change the characteristics of snoring sound? Conf Proc IEEE Eng Med Biol Soc 2011;2011:1749–52. https://doi.org/10.1109/IEMBS.2011.6090500.
[86] Acar M, Yazıcı D, Muluk N, et al. Is there a relationship between snoring sound intensity and frequency and OSAS severity? Ann Otol Rhinol Laryngol 2016;125:31–6. https://doi.org/10.1177/0003489415595640.
[87] Peng H, Xu H, Gao Z, et al. Acoustic analysis of overnight consecutive snoring sounds by sound pressure levels. Acta Otolaryngol 2015;135:747–53. https://doi.org/10.3109/00016489.2015.1027414.
[88] Wilson K, Stoohs R, Mulrooney T, et al. The snoring spectrum: acoustic assessment of snoring sound intensity in 1,139 individuals undergoing polysomnography. Chest 1999;115:762–70. https://doi.org/10.1378/chest.115.3.762.
[89] Ng A, Koh T. Using psychoacoustics of snoring sounds to screen for obstructive sleep apnea. Annu Int Conf IEEE Eng Med Biol Soc 2008;2008:1647–50. https://doi.org/10.1109/IEMBS.2008.4649490.
[90] Herzog M, Kühnel T, Bremert T, et al. The impact of the microphone position on the frequency analysis of snoring sounds. Eur Arch Oto-Rhino-Laryngol 2009;266:1315–22. https://doi.org/10.1007/s00405-008-0858-7.
[91] Azarbarzin A, Moussavi Z. A comparison between recording sites of snoring sounds in relation to upper airway obstruction. Conf Proc IEEE Eng Med Biol Soc 2012;2012:4246–9. https://doi.org/10.1109/EMBC.2012.6346904.
[92] Chordekar S, Adelman C, Sohmer H, et al. Soft tissue conduction as a possible contributor to the limited attenuation provided by hearing protection devices. Noise Health 2016;18:274–9. https://doi.org/10.4103/1463-1741.192476.
[93] Kuttruff H. Room acoustics. Elsevier; 1991.
[94] Goswami U, Black A, Krohn B, et al. Smartphone-based delivery of oropharyngeal exercises for treatment of snoring: a randomized controlled trial. Sleep Breath 2019;23:243–50. https://doi.org/10.1007/s11325-018-1690-y.
[95] Blumen M, Quera-Salva M, Vaugier I, et al. Is snoring intensity responsible for the sleep partner's poor quality of sleep? Sleep Breath 2012;16:903–7. https://doi.org/10.1007/s11325-011-0554-5.
[96] Guidelines for community noise. World Health Organization website. (http://www.who.int/docstore/peh/noise/Commnoise4.htm).