Estimation of the Apnea-Hypopnea Index from Epoch-based Classification of the ECG using Modulations of QRS Area and Respiratory Myogram Interference

Christoph Maier[1,2], Heinrich Wenz[1], Hartmut Dickhaus[1]

[1]Heidelberg University, Heidelberg, Germany
[2]Heilbronn University, Heilbronn, Germany

Abstract

Most approaches for ECG-based detection of sleep apnea classify fixed-duration epochs (here: 60s) in a binary way with respect to the presence or absence of respiratory events. Identifying time-locked modulations in QRS-amplitude and respiratory myogram interference with a single feature, this paper presents a new method to transform such results into an estimate of the apnea-hypopnea-index (AHI). The basic idea is to translate the local period duration of the event-related quasi-periodic oscillations into a weighting factor for each epoch's result. We present a binary and a ternary strategy (including a borderline class) for identification of patients with an AHI ≥ 15. Results on a large (N = 140) and representative clinical sample indicate 100% sensitivity and 86.4% accuracy for the binary strategy. The ternary strategy achieves 95.1% accuracy for a subset of 83.6% of the recordings. The remaining 16.4% of the sample are classified as 'boderline cases'. Additional validation on the independent Physionet Apnea ECG Database (N = 69) resulted in perfect separation of the apnea group from the control cases. We conclude that it is possible to provide a physiologically meaningful estimate of the AHI from epoch-based classifications of the ECG: It promises robust detection of apnea patients in various clinical scenarios such as screening of Holter ECGs.

1. Introduction

Sleep-related breathing disorders (SRBD) represent a medical condition with severe social and cardiovascular consequences. They imply a significant reduction of quality of life. Despite their high prevalence, diagnosis of SRBD is typically delayed, and the estimated number of unreported cases is as high as 85% [1]. In this context, application of the Holter-ECG as a screening tool for SRBD that facilitates earlier detection appears highly promising. Research on this topic has been well established within the Computers in Cardiology community since more than 10 years [2].

The clinical standard that quantifies the severity of SRBD is the apnea-hypopnea index (AHI). It is defined as the average number of respiratory events per hour of sleep. An AHI ≥ 15 is considered to be critical independent of an association with daytime symptoms [3].

Most ECG-based approaches do not provide an estimate of the AHI. Rather, they perform a classification of fixed-duration epochs for the presence or absence of respiratory events [4, 5, 6, 7]. If an AHI is estimated, this is often based on a sample-regression of the number of apnea-positive epochs on the true AHI [7]. From a physiological point of view, this is unsatisfactory since it practically means that the estimated AHI value for a specific record varies dependent on the AHI values of the remaining records in the sample.

Our paper presents a new, simple and physiologically motivated method to transform the results of an epoch-based classification scheme into an AHI estimate. Results of a screening application for severe cases (AHI ≥ 15) are presented using two large, independent data sets. Finally, a ternary decision strategy is suggested that increases classification accuracy and specificity by identifying cases with borderline constellation.

2. Material and Methods

A sample of 140 overnight Holter ECG registrations (Mortara H12+, 8 channels, 1 kHz/channel) was recorded from 121 different patients with suspected SBAS at the Sleep Medicine Center of the Thoraxklinik Heidelberg. Respiratory events were extracted from synchronized polysomnograms (Alice 4) and mapped to adjacent epochs with one minute time resolution. They served as reference for epoch-based apnea detection. With the only exception of persisting supraventricular arrhythmia, we did not exclude co-morbidity or confounding medication. 70 of the recordings had an AHI ≥ 15 and 40 had an AHI ≤ 5.

As an independent data set we included the Physionet Apnea ECG Database [8]. It comprises 70 overnight ECG records (1 channel, 100 Hz) with minute-by-minute apnea annotations, coming from 40 severe apnea cases (group

A), 20 healthy controls (group C) and 10 borderline cases (group B). We only used 69 out of the 70 recordings. Record b04 was excluded owing to its bad signal quality.

After 50 Hz notch-filtering, QRS detection and ectopic beat classification, a first beat-to-beat respiratory surrogate signal was derived by approximating the area under the QRS complexes in windows of 120 ms width. Essentially, this corresponds to the method described in [9]. A second respiratory surrogate was based on quantifying the amount of myogram interference in segments that extend from 80 to 430 ms after each QRS fiducial point. This region of the ECG is typically free of high-frequency activity of cardiac origin. Therefore, the residual standard deviation after 60 Hz high-pass filtering was taken as an estimate of myogram activity. The QRS area and the myogram interference series were interpolated by means of cubic splines and equidistantly resampled at a rate of 3 Hz. We refer to those as Q series and M series in the remainder of the paper. Finally, we extracted an envelope-related signal from the M series (eM). Details on the calculations are given in [10, 11].

Fig. 1 shows exemplary time-courses of these signals.

It is of fundamental importance for our approach, that clinically relevant cases of SRBD typically exhibit repetitive sequences of respiratory events. As fig. 1 demonstrates, these elicit time-locked quasi-periodic low-frequency modulations in each of the series Q, M, and eM. We have recently developed a method that quantifies the joint occurrence of such oscillations in these three series (after band-pass filtering) for detection of SRBD in epochs of T_{ep} = 60s duration [10, 11]. Its output is a binary statement $c(n) \in \{0, 1\}$, where $c(n) = 1$ indicates the presence of events in epoch n. It is based on the selection of a candidate prototype event pattern with a fixed duration of 30s for each epoch. The candidate is initially located as the dominant extremum in the band-pass filtered Q-series. If an epoch truly contains respiratory events, the modulations are expected to occur coincidently in all series. Therefore, in the same temporal position as determined in Q, candidates in eM and M are extracted after band-pass filtering. In each of the three series, the recurrence of the candidate pattern in the epoch's local vicinity (5 min) is probed separately by means of normalized cross-correlation. Finally, the three correlation functions are multiplied element-wise [10]. Again, in case of respiratory events, the maxima should coincide resulting in their relative augmentation in the product correlation function. The final classification feature consists in the sum of all product correlation values that exceed a pre-specified threshold. ROC-analysis resulted in an epoch-based sensitivity of 0.855 and a specificity of 0.86 for the clinical sample [10].

Our new approach for AHI estimation rests upon the results of epoch-based approach as sketched above. Its basic idea is to assess the number of respiratory events in each epoch n and use it as a weight factor $w(n)$ for the epoch's classification result $c(n)$. The sum over all epochs provides the total number of events which is finally normalized to the record duration (in hours) to obtain the AHI estimate:

$$AHI^* \quad = \quad \frac{1}{T_{rec}^{[h]}} \sum_{n=0}^{N-1} w(n) \cdot c(n)$$

Obviously, epochs with $c(n) = 0$ do not contribute to the AHI^*. If an epoch n is scored $c(n) = 1$ we can assume the existence of a low-frequency oscillatory pattern (fig. 1) in its vicinity, where each period corresponds to one respiratory event. We estimate the inter-event duration by doubling the average zero-crossing distance $\overline{T_{zc}}(n)$of the band-pass-filtered QRSA-series (fig. 1) in a local window of 3 min. The reciprocal of this value, normalized to the epoch duration T_{ep}, is used as an estimate of the number of respiratory events $w(n)$ within epoch n (please note that this may be a fractional number).

$$w(n) \quad = \quad \frac{T_{ep}}{2 \cdot \overline{T_{zc}}(n)}$$

The epoch classification was based on a receiver operating characteristics (ROC) analysis [10]. Two different ROC-thresholds were considered: the first (Θ_a) maximizing accuracy, the second (Θ_b) balancing sensitivity and specificity. Θ_b relates to the point on the ROC curve with the smallest Euclidean distance to the point of perfect separation.

Both thresholds were then used to separately estimate an AHI^*. This provides a threshold-dependent screening decision $S(\Theta)$ according to

$$S(\theta) \quad = \quad \begin{cases} A^+ & if \quad AHI^*(\theta) \geq 15 \\ A^- & if \quad AHI^*(\theta) < 15 \end{cases}$$

A^+ indicates the presence, A^- the absence of significant SRBD. We considered two screening strategies. The first is a binary strategy S_B based on Θ_b that assigns a definite decision to each record according to

$$S_B \quad = \quad S(\theta_b)$$

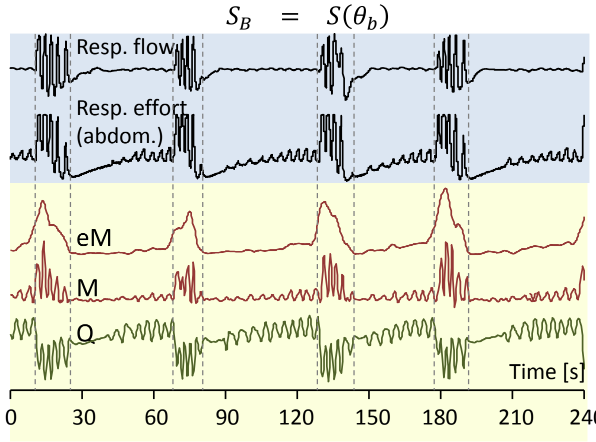

Figure 1. Upper traces: time-courses of respiratory flow and respiratory effort during repetitive mixed apneas. Lower traces: re-sampled respiratory surrogate signals QRS area (Q), myogram interference (M) and envelope of myogram interference (eM).

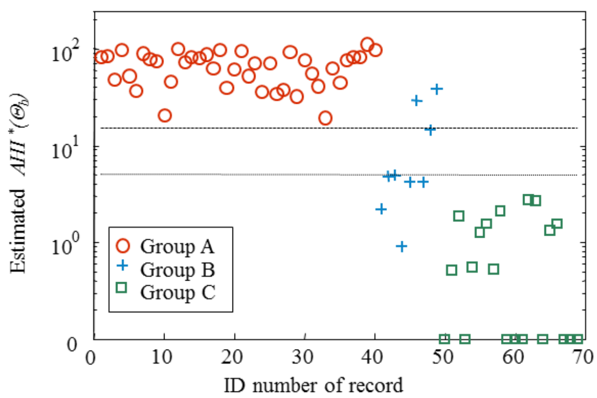

Figure 2. Result of the AHI estimation for the records of the Physionet Apnea ECG Database. The different symbols refer to the subject groups A, B, and C. Horizontal lines indicate AHI values of 5 and 15.

The second is a ternary strategy S_T which aims to increase specificity by additionally identifying records with borderline constellation *B*. This situation is considered to be present if the screening decisions for the two thresholds do not coincide:

$$S_T = \begin{cases} A^+ & \text{for} \quad S(\theta_a) = S(\theta_b) = A^+ \\ A^- & for \quad S(\theta_a) = S(\theta_b) = A^- \\ B & for \quad S(\theta_a) \neq S(\theta_b) \end{cases}$$

3. Results

Table 1 shows the results for the binary classification strategy S_B. All cases with AHI $\geq$ 15 are recognized correctly resulting in a sensitivity of 1. However there are 19 false positive cases, i.e. the specificity is 0.729 and the accuracy 0.864 (kappa: 0.73).

The results for the ternary strategy S_T are given in table 2. A total of 23/140 records (16.4%) is classified as 'borderline', i.e. only for 117/140 recordings (83.6%) a definitive statement is made. Within that reduced sample, still all positive cases are detected correctly (61/61, sensitivity 1.0) and the number of false positives reduces to only 5/56 (specificity 0.911, accuracy 0.951, kappa 0.91). It should be noted that 3 out of the 5 false positives show an AHI $\geq$ 10.

Fig. 2 shows the result of the AHI^* estimation using Θ_b in 69/70 records of the PADB. The values clearly reflect the group structure and achieve perfect separation

Table 1. Confusion matrix for the binary classification strategy SB in the clinical data sample.

	AHI ≥ 15	AHI < 15	
$S_B = A^+$	70	19	89
$S_B = A^-$	0	51	51
	70	70	140

Table 2. Confusion matrix for the ternary classification strategy S_T in the clinical data sample.

	AHI ≥ 15	AHI < 15	
$S_T = A^+$	61	5	66
$S_T = A^-$	0	51	51
$S_T = B$	9	14	23
	70	70	140

4. Discussion and conclusion

This paper demonstrates that it is possible to provide a physiologically meaningful estimate of the AHI from epoch-based classifications of the ECG without explicit delineation of single respiratory events. The basic idea is that, in clinically relevant cases, the self-sustaining nature of the patho-mechanism elicits quasi-periodic modulations in the ECG-derived respiratory surrogate signals. Their period is directly related to the repetition-rate of the events. Epoch-based estimates of these periods can therefore be transformed to weighting-factors that approximate the local number of events. Accumulation over all epochs that contain events results in an estimate of the AHI.

The results show that this AHI can be used to detect patients in need for treatment with perfect sensitivity. Despite of its good accuracy of 0.864 the specificity of the binary strategy (0.729) appears probably too low for a screening application in a general Holter-ECG population. Too many false-positive detections would result in unreasonable allocation of health-care resources. Therefore, the binary strategy seems rather suitable as a means to exclude SRBDs when there is already a suspect for SRDB. In contrast, the ternary strategy provides a strong improvement in specificity (0.911) at the expense of a slightly reduced population for which a definitive conclusion is made. The introduction of the borderline class 'B' may be considered as an additional assertion of the individual accuracy of a specific record's decision. For the 83.4% of records for which a statement is made, the accuracy is now 0.955. Still, no patients are missed - the perfect negative predictivity of 1.0 holds for the entire 100% of the sample. This constellation appears appropriate for a general screening of Holter-ECGs. Patients with borderline ('B') outcome should undergo proper anamnesis including questionnaires, and alternative diagnostic procedures like nocturnal oximetry or ambulant polygraphy when indicated.

Detailed review of misclassified records has revealed that the comparatively low specificity of the binary strategy relates to a tendency to overestimate AHI values under certain conditions. One is the problem of respiratory events that occur in periods which are scored as 'wake' in the PSG. By definition, these are not counted in the clinical AHI score [3]. Second, very short repetitive events that do not reach the required minimum duration of 10s may still elicit pronounced ECG modulations. Owing to their short period duration such events produce a high weighting factor which exacerbates the problem. The appropriateness to exclude both conditions from the PSG-AHI may be questioned, but that discussion would definitely go beyond the scope of this paper.

Compared to other methods employing regression analysis, our approach is clearly motivated by physiology and depends solely on properties of the one record that is classified. Regression analysis entirely disregards that the duration of single events may vary between 10s and 90s. This potentially results in vastly different AHI values for the same number of epochs spent in disordered breathing.

With respect to detection accuracy, Heneghan et al. report 81.5% for AHI ≥ 15 in a strongly selected sample of 92 subjects [7]. Khandoker et al. report 100% accuracy, however with a black-box approach on a selected sample of only 16 subjects [12]. In contrast, our approach is entirely transparent and traceable from a physiological point of view. Our database is comparatively large and representative in the sense that it does not exclude subjects with typical comorbidity or medication. Still, our method achieves very good accuracy in conjunction with excellent robustness and generality. This is documented by the independent results on the Physionet Apnea ECG Database (fig. 2) which indicate perfect separation for groups A and C. It should be noted that the entire database (learning and tests set) is included here.

We conclude that it is possible to provide a physiologically meaningful estimate of the AHI from epoch-based classifications of the ECG: It promises robust detection of apnea patients in various clinical scenarios such as screening of Holter ECGs.

References

[1] Somers VK, White DP, Amin R, Abraham WT, Costa F, Culebras A, et al. Sleep apnea and cardiovascular disease: an American Heart Association/American College of Cardiology Foundation Scientific Statement from the American Heart Association Council for High Blood Pressure Research Professional Education Committee, Council on Clinical Cardiology, Stroke Council, and Council on Cardiovascular Nursing. J Am Coll Cardiol. 2008 Aug;52(8):686–717. Available from: http://-dx.doi.org/10.1016/j.jacc.2008.05.002.

[2] Moody GB, Mark RG, Goldberger A, Penzel T. Stimulating rapid research advances via focused competition: the Computers in Cardiology Challenge 2000. Computers in Cardiology 2000; 207–210.

[3] Iber C, Ancoli-Israel S, Chesson AL, for the American Academy of Sleep Medicine SFQ. The AASM-Manual for the Scoring of Sleep and Associated Events: Rules, Terminology and Technical Specifications. Westchester, Illinois; 2007.

[4] Penzel T, McNames J, Murray A, de Chazal P, Moody G, Raymond B. Systematic comparison of different algorithms for apnoea detection based on electrocardiogram recordings. Med Biol Eng Comput. 2002;40:402–407.

[5] de Chazal P, Heneghan C, Sheridan E, Reilly R, Nolan P, O'Malley M. Automated processing of the single-lead electrocardiogram for the detection of obstructive sleep apnoea. IEEE Trans Biomed Eng. 2003;50:686–696. Available from: http://dx.doi.org/10.1109/-TBME.2003.812203.

[6] Mendez MO, Bianchi AM, Matteucci M, Cerutti S, Penzel T. Sleep apnea screening by autoregressive models from a single ECG lead. IEEE Trans Biomed Eng. 2009;56(12):2838–2850. Available from: http://dx.doi.org/-10.1109/TBME.2009.2029563.

[7] Heneghan C, de Chazal P, Ryan S, Chua CP, Doherty L, Boyle P, et al. Electrocardiogram recording as a screening tool for sleep disordered breathing. J Clin Sleep Med. 2008;4(3):223–228.

[8] Goldberger AL, Amaral LA, Glass L, Hausdorff JM, Ivanov PC, Mark RG, et al. PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals. Circulation 2000; 101(23):E215–E220.

[9] Moody GB, Mark RG, Zoccola A, , Mantero S. Derivation of Respiratory Signals from Multi-lead ECGs. In: Murray A, editor. Computers in Cardiology. vol. 12. Washington, DC:: IEEE Computer Society Press; 1985. p. 113–116.

[10] Maier C, Wenz H, Dickhaus H. Robust Detection of Sleep Apnea from Holter ECGs by Joint Assessment of Modulations in QRS Amplitude and Respiratory Myogram Interference. In: Cerutti S, Dickhaus H, Mainardi L, Yana K, editors. Proceedings VII International Workshop on Biosignal Interpretation BSI2012. http://www.biomed.polimi.it/BSI2012/BSIarticles/00710069.pdf; 2012. .

[11] Maier C, Wenz H, Dickhaus H. Steps toward subject-specific classification in ECG-based detection of sleep apnea. Physiol Meas. 2011 Nov;32(11):1807–1819. Available from: http://dx.doi.org/10.1088/0967-3334/32/-11/S07.

[12] Khandoker AH, Gubbi J, Palaniswami M. Automated scoring of obstructive sleep apnea and hypopnea events using short-term electrocardiogram recordings. IEEE Trans Inf Technol Biomed. 2009 Nov;13(6):1057–1067. Available from: http://dx.doi.org/10.1109/-TITB.2009.2031639.

Address for correspondence.

Christoph Maier
Heidelberg University
Institute for Medical Biometry and Informatics
Im Neuenheimer Feld 305
69120 Heidelberg, Germany
christoph.maier@med.uni-heidelberg.de