Reprinted from: Progress In Sleep Apnea Research, ed. R. Ferber, *published NovaScience*, 2007 1

Automated screening for sleep apnoea in the home environment

Philip de Chazal, Conor Heneghan, Redmond Shouldice and Chern-Pin Chua

Abstract— We present a simple method for detection of sleep apnoea in the home environment. The method identifies episodes of disordered breathing and provides an estimated apnoea hyponoea index (AHI) by recording and automatic analysis of nighttime ECG and oximetry recordings. At the heart of the device is a pattern recognition system that identifies episodes of apnoea by analysing the heart variability, an ECG-derived respiration signal and blood oximetry values. The method has been tested on 125 subjects who have undergone a full overnight polysomnogram study. Results show that the method a) correctly annotates 89% of epochs as either normal or sleep disordered breathing and b) it separates control subjects from subjects with clinically significant sleep apnoea with a specificity and sensitivity exceeding 93%. These results demonstrate that home-based screening for sleep apnoea is a viable alternative to hospital-based tests with the added benefit of low cost and minimal waiting times.

Automated screening for sleep apnoea in the home environment

Philip de Chazal, Conor Heneghan, Redmond Shouldice and Chern-Pin Chua

Abstract— We present a simple method for detection of sleep apnoea in the home environment. The method identifies episodes of disordered breathing and provides an estimated apnoea hyponoea index (AHI) by recording and automatic analysis of nighttime ECG and oximetry recordings. At the heart of the device is a pattern recognition system that identifies episodes of apnoea by analysing the heart variability, an ECG-derived respiration signal and blood oximetry values.

The method has been tested on 125 subjects who have undergone a full overnight polysomnogram study. Results show that the method a) correctly annotates 89% of epochs as either normal or sleep disordered breathing and b) it separates control subjects from subjects with clinically significant sleep apnoea with a specificity and sensitivity exceeding 93%.

These results demonstrate that home-based screening for sleep apnoea is a viable alternative to hospital-based tests with the added benefit of low cost and minimal waiting times.

I. INTRODUCTION

Sleep apnoea is a cardiorespiratory disorder characterised by brief interruptions of breathing during sleep. Typical sleep patterns of a sufferer involve heavy snoring interspersed with obstruction of the upper airway, leading to waking and gasping for breath. Often the sufferer has no recollection of the sleep interruptions that can occur hundreds of times in a night. The primary health implications of sleep apnoea are its impact on the cardiovascular system (increased levels of hypertension, coronary arterial disease, arrhythmias), increased accident levels due to sleepiness, and quality of life issues. Obstructive sleep apnoea (OSA) is not a rare condition. It occurs in 2% to 4% of middle-aged adults (Young 1993) and in 1% to 3% of preschool children (Gislason 1995). However, despite the fact that apnoea has such health and quality of life implications, there is a surprisingly low public and medical awareness of the illness. Of the 10 to 20 million sufferers in the U.S. with moderate-to-severe sleep apnoea, it is estimated that only 10 to 15% have been diagnosed (Young 1997).

By definition, an apnoea is a cessation of airflow through the upper airway for a period of 10 seconds or longer, and is typically associated with a fall in SaO2 termed a desaturation. A hypopnoea is a reduction in airflow to less than 50% of normal airflow that leads to a desaturation of 3% or an electrocortical arousal (AASM 1999). Apnoeas and hypopnoeas are classified into three types:

- obstructive, in which respiratory effort is present, but the upper airway is partially or completely blocked
- central, in which the upper airway is open, but respiratory effort is absent or reduced, and
- mixed, in which both central and obstructive aspects are present. A typical mixed apnoea may show a period of central apnoea for several seconds, during which the upper airway occludes, followed by increased respiratory effort against the obstruction.

Most apneic events are terminated by recovery breaths, frequently (though not always) accompanied by an electrocortical arousal which is visible in an EEG recording.

Overnight polysomnography (PSG), a laboratory-based sleep study, is regarded as the gold standard for the diagnosis of sleep apnoea. It is widely agreed that PSG is a thorough and reliable test. However, it also receives its share of criticism. Firstly, PSG is inconvenient since it requires the patient to stay in hospital for one night. Secondly, it is an expensive process. This high cost is due to the need for the study to take place in a hospital setting, the requirement to have a sleep technician in attendance overnight, and the need to manually "score" the resultant measurements. Thirdly, many sleep centers worldwide are currently operating at full capacity and PSG usually suffers from a low availability reflected in up to 6 month waiting lists for testing. Therefore there is considerable interest in the development of reliable low-cost techniques for identification of subjects with sleep apnoea, particularly in those systems which can be reliably used in a home environment.

A joint study group drawn from the Amercian Thoracic Society, the American Academy of sleep Medicine, and the American College of Chest Physicians carried out a meta-analysis of reports on portable monitoring devices for assessment of sleep apnoea (Chesson 2003). Three categories of portable monitoring (PM) devices were reviewed with regard to their accuracy in distinguishing subjects with an apnoea-hypopnoea index (AHI) of greater or less than 15 in attended and unattended settings. Type 2 devices are defined as having a minimum of seven channels, including EEG, EOG, chin EMG, ECG or heart rate, airflow, respiratory effort, oxygen saturation, Type 3 devices should have a minimum of four channels, including ventilation or airflow (at least two channels of respiratory movement, or respiratory movement and airflow), heart rate or ECG and oxygen saturation) and Type 4 devices would measure only one or two parameters. Their conclusions were quite cautious in that they found insufficient evidence to recommend the use widespread use of portable monitoring devices, though they did find that Type 3 devices could be used in an attended setting to increase or to decrease the probability that a

Manuscript started July 2006

patient has an apnoea-hypopnoea index greater than 15. However, they did emphasize the need for continued study of portable monitoring technologies, with the collection of large well-controlled data sets, and more general consideration of the clinical data beyond simple use of an AHI threshold.

Despite the lack of clear evidence supporting unattended studies using devices with a small number of channels, there has been considerable exploration of such technologies, with the most widely explored low-cost option being unattended overnight ambulatory oximetry. However, there have been mixed reports in the literature on the efficacy of this technique. Overnight oximetry can be useful if it shows a pattern of cyclic desaturation, but in practice the limiting factor appears to be the negative predictive value (NPV) of the method. For example, Chiner at al. (Chiner 1999) quoted an NPV ranging from 38-48%. Conversely, it has also been suggested that the sensitivity of oximetry might be limited by the fact that some obstructive events may not lead to an obvious desaturation, i.e. the definition of an obstructive hypopnoea typically includes a desaturation or microarousal whereas an obstructive apnoea is defined purely by a cessation of flow for more than 10 seconds and not necessarily a desaturation (Series 2002). A further confounding factor is that commercial oximeters have different averaging times (averaging is used in oximeters to overcome limitations due to intermittent data loss due to poor perfusion). Davila et al. conducted a systematic study on the influence of averaging parameters and noted that sensitivity is lower and specificity higher when longer oximeter averaging periods (e.g. 12 seconds) were considered; this reversed when using shorter averaging length (e.g. 3 or 6 seconds) (Davila 2002). Oximeters are also subject to artifact due to motion and poor perfusion, which can lead to significant loss of data. For example, Yamashiro et al. used nocturnal oximetry as a screener for OSA reject 10% of subjects from the study due to poor SpO2 signal quality (Yamashiro 1995).

Assessment based on oximetry can range from simple detection of desaturations (defined as a decrease of 3% of 4% from a baseline level, over a certain period of time) to more complex measures such as CT80 or CT90 (the percentage of time where the SpO2 value is below 80% of 90% respectively, or derived measures such as the oximetry delta index (Magalang 2003).

While we do not intend to give an exhaustive survey of reports on the use of nocturnal oximetry in apnoea screening, it is useful to report on a number of studies, if only to conclude that the reported clinical performances are highly variable, and depend on the chosen clinical population, the parameters of oximetry measurement, and the statistical measures derived from the oximetry measurements. In particular, some studies report that the limiting factor is specificity, whereas others indicate that limited sensitivity is observed. For example, Series et al. found a sensitivity of 85% and specificity of 93% in distinguishing congestive heart failure patients with sleep related breathing disorders. (Series 2005). However, a drawback of oximetry was that it was unable to distinguish obstructive from central events reliably in these subjects (Series 2005). In an adult population, Zamarron et al. used spectral analysis of the oximetry signal to screen for OSA and found a sensitivity of 78% and a specificity of 89% (Zamarron 1999).

Conversely, Brouillette et al. performed oximetry in a group of children with suspected OSAS and compared it with simultaneous full polysomnography (Brouillette 2000). Patients with complex medical conditions were excluded. Compared with polysomnography, they found a PPV of 97% and an NPV of 47%, indicating that oximetry was useful when results were positive. However, patients with negative results of oximetry required full polysomnography for definitive diagnosis. False-positive results were found in patients with mild coexistent medical problems, such as obesity and asthma, suggesting that this technique is useful only in otherwise healthy children. Levy et al. reported on a study of 301 adults and found a Se of 98% with a Sp of 46% in distinguishing AHI>15 (Levy 1996).

As an alternative to oximetry as a screening tool, the observation of changes in heart rate associated with apneic events has long been suggested as a possible technique for simple identification of subjects with sleep apnoea syndrome (Guilleminault 1984). Following the initial work of Guilleminault et al. in this area, several researchers have proposed techniques for using ECG-based analysis for sleep apnoea screening (Dingli 2003, Roche 1999, Roche 2003, Stein 2003). The physiological rationale for this approach is as follows. Apneic events are typically (though not always) associated with a bradycardia, which is followed by an abrupt tachycardia simultaneously with recovery breaths. Therefore the tachogram of RR intervals will show a characteristic sawtooth pattern with a duration approximately corresponding to the duration of the apnoea plus recovery breaths (15-20 s). This time-domain fluctuation has been termed a Cyclical Variation in Heart Rate. Figure 1 illustrates some typical fluctuations in heart rate seen in a subject with severe OSA. Alternatively, frequency domain analysis of the RR interval series, will reveal augmented energies at the corresponding very-low-frequencies (VLF) – typically 0.02-0.05 Hz. In addition, a surface ECG also contains direct information about respiratory effort, since its amplitude is typically modulated by movement of the ribcage. This concept is shown in Figure 2, where we show how modulation of the ECG amplitude is directly related to measured ribcage movement.

We have previously reported on an algorithm (de Chazal 2003) for identifying patients with obstructive sleep apnoea by recognizing both autonomic and respiratory effort changes in the ECG associated with apneic events. This algorithm was developed using a single channel of modified lead V2 ECG recording extracted from polysomnograms comprising the Philipps University database (Penzel 2000) and arose out of our participation in the Computers in Cardiology Conference Challenge in 2000 (Moody 2000, Penzel 2002). The algorithm carries out an automated classification of an ECG recording. It divides a recording into one-minute epochs, and then estimates the probability of each epoch being from an “apneic” minute or a “normal respiration” minute. The per-epoch classifications are then combined to form an overall classification of the recording in terms of average minutes/hour of OSA. This minutes/hour measure is then mapped to an estimated AHI by using a linear transformation. The estimated AHI was used to classify subjects into three classes: moderate-to-severe OSA with AHI>15, mild OSA with 5<AHI<15, and controls with AHI<5. The results of this algorithm on an independent test set of 35 ECG recordings was that it correctly distinguished all moderate-to-severe cases from control subjects. The correlation between the estimated AHI and the PSG determined AHI was 0.9.

Given that both the ECG and oximetry have been suggested as useful screening tools for unattended use, it seems reasonable to investigate the simultaneous use of oximetry and the ECG signal for both predicting AHI and indicating the presence or otherwise of sleep disordered breathing. The literature suggests that simultaneously acquired "gold standard" polysomnography (PSG) is necessary to justify any quoted results, so in this paper we will consider the analysis of ECG and oximetry signals as a subset of a complete polysomnogram recording, allowing direct comparison of the proposed screening algorithm with PSG results. This study extends our previous work through the addition of oximetry signal and by testing of the system on a database of 125 subjects. The paper outlines a method for providing a reliable diagnostic measure of OSA based on measurement of the electrocardiogram and oximetry signals.

II. DATABASE

A. Subjects:

The dataset of 125 subjects was made up of two cohorts. The first cohort comprised 28 subjects randomly recruited over a six-month period (September 2002 to February 2003) from patients referred to the Sleep Disorders Clinic at St. Vincent's University Hospital (Dublin, Ireland). These subjects were referred to the clinic for evaluation of suspected OSAS. Subjects were over 18 years of age, had no known cardiac disease, autonomic dysfunction, and were not on medication known to interfere with heart rate, such as beta-blockers, digoxin, or calcium receptor anatagonists.

The second cohort comprised of two groups. The first group of 64 subjects was recruited from consecutive males attending the same Sleep Disorders Clinic for evaluation of suspected OSAS (from March 2004 to September 2005), who were free from other medical disorders and not commenced on regular medication. The second group comprised 33 healthy male control subjects recruited from the general population, who were matched according to age and BMI to the first group.

The protocol was approved by the Hospital's Ethics Committee, and all subjects provided written, informed consent.

TABLE 1: DEMOGRAPHIC DATA

n=125	**Value**
Age (years)	44 ± 8
Male:Female (n)	121 : 4
Body Mass Index (kg/m2)	33 ± 6
Apnoea-Hypopnoea Index (no/hr)	29 ± 29
Epworth Sleepiness Score	12 ± 6

TABLE 2: APNOEA SEVERITY

Apnoea Severity	**No. of subjects**
Not clinically significant (AHI < 5)	30
Mild (AHI 5 to 15)	28
Moderate (AHI 15 to 30)	19
Severe (AHI > 30)	48

B. Sleep Studies:

Overnight polysomnography was performed using the Jaeger-Toennies system (Erich Jaeger GmbH, Hoechberg, Germany). EEG (C4/A1, C3/A2), bilateral EOG, submental EMG and ECG (modified lead V2) were recorded using surface electrodes. Respiration was measured through oronasal flow (thermistor) and thoracic and abdominal movements (uncalibrated inductance plethysmography). Oxygen saturation was measured using finger pulse oximetry. Snoring was recorded using a surface microphone attached above the sternal notch, and body position was also monitored. All studies were performed in the sleep laboratory and supervised throughout by an experienced sleep technologist.

C. Expert Annotation

Following completion of the study, sleep staging was performed using full polysomnography by a single experienced sleep technologist. The scorer also produced an annotated respiratory event list which provides onset times, and durations of sleep-disordered breathing events including obstructive, mixed, and central apnoeas and hypopnoeas, and periodic breathing episodes.

Obstructive events were distinguished from central events by the presence or absence of paradoxical thoracic and abdominal movements during apnoeas or hypopnoeas. The sleep scorer was blind to the output of the automated analysis system. Subjects were also asked to complete the Epworth Sleepiness Score questionnaire (Johns 1991).

III. METHODS

We have used a combination of physiological knowledge of the manifestation of apnoea in the ECG and the oximetry signals and black-box pattern recognition methods to design three systems for estimating the AHI from the ECG and oximetry signals. The three systems are

- an ECG only system;
- an oximetry only system; and
- a combined ECG and oximetry system.

The third system is a unification of the first two systems and is shown in Figure 1. The three systems process the ECG and/or oximetry signals and provide two outputs. The first output is an epoch-by-epoch sequence of annotations of "normal" or "SDB" (sleep disordered breathing). The second output provides an estimated apnoea-hypopnea index (AHI) and is derived on the basis of the epoch-by-epoch annotations.

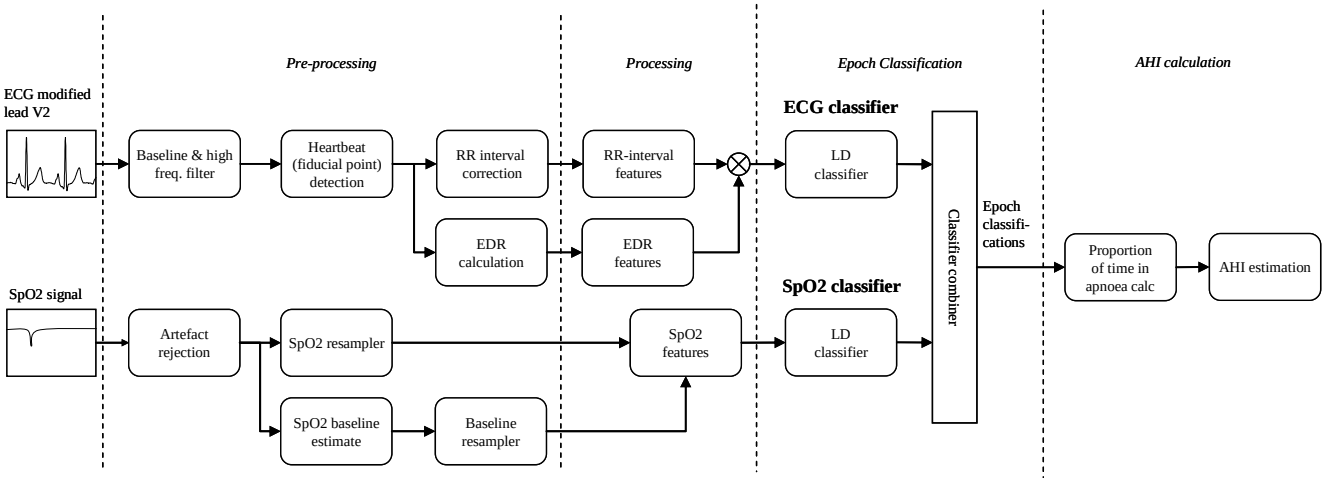

Figure 1: Simultaneous ECG and oximetry system for identifying epochs of apnoea and estimating AHI from overnight recordings.

A. Overall System Design

This study has adopted pattern recognition approach using supervised learning to obtain the three systems. Classifier methods were based on linear discriminants (Duda 2001) and our selected signal representations.

At its most fundamental level the system processes an epoch of data and allocates the epoch to either "normal" or "apnoea". To combine this with supervised learning a first step was to map the event based annotations of the expert determined respiration to epoch-based annotations and this is discussed below.

We considered various representations of the ECG and SpO2 signals using different features. Previous studies using ECG had shown that features based on the timing of QRS complexes (Guilleminault 1984, Penzel 1990, Hilton 1999, Roche 1999), and the amplitude of the ECG (Moody 1985, Moody 1986, Travaglini 1998) might be useful for apnoea identification. Both types of features were considered in this study

For oximetry systems, previous studies have used temporal features such as the percentage time below a certain level, the sum of the differences between successive readings (delta index), the number of dips in oxygen saturation per hour, and frequency based features such as the spectral peak in the SpO2 spectrum and pulse rate periodogram in the range 30-70 seconds. (Gyulay 1993, Zamarron 1999, Oeverland 2002, Golpe 1999). In this work, however, temporal SpO2 features only were considered for classification.

Classifier performance was determined using cross-validation using the available ECG and oximetry data.

The three systems (ECG-only, oximetry-only, ECG-oximetry) have separate stages. The four stages are a pre-processing, a processing, an epoch classification and finally an AHI calculation stage. Each of these stages is described below.

B. Annotations

There were six possible respiratory event annotations relating to apnoea by the scorer. These annotations were

- obstructive apnoea (OA),
- obstructive hypopnoea (OH),
- central apnoea (CA),

- central hypopnoea (CH),
- mixed apnoea (MA), and
- mixed hypopnoea (MH).

In addition, periodic breathing events were annotated but were not included in the calculation of the AHI index.
The apnoea hypopnoea index was determined by summing the number of respiratory events from the six above and dividing by the number of hours of sleeptime.

1) Mapping event-based to epoch-based annotations
The expert annotations provided with this study are event-based i.e the start and finish of an annotation correspond to the start and finish of the respiration event. As the systems presented here are epoch-based the first step was to map the expert annotations to epoch-based annotations. To achieve this the annotation time sequence was divided into epochs and the annotation of each epoch was assigned to a category as follows:

- Determine the duration of all the normal events in a classification epoch (Normal events are all events which are not apnoea events(see next bullet))
- Determine the duration of all apnoea events in an a classification epoch (Apnoea events = OA, CA, MA, OH, OH, MH, and CH)
- If the duration of the apnoea events exceeds 5 seconds then the epoch-based label is sleep disorder breathing (SDB)
- otherwise the epoch is labelled normal.

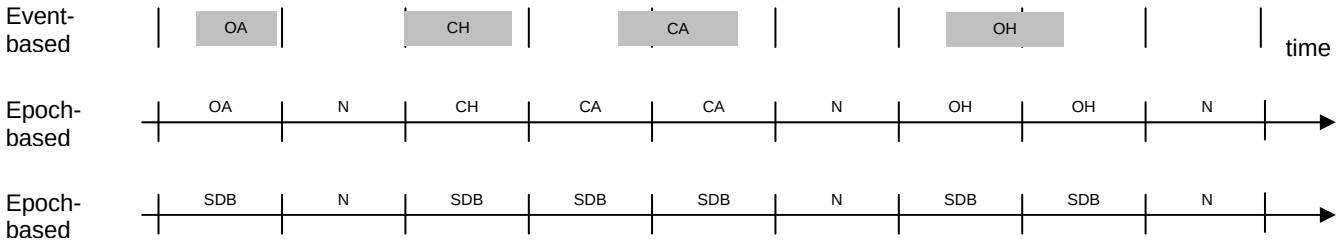

Figure 2: An example of the mapping of the event based annotations of the scorer to epoch based annotations required for development of the systems.

C. Preprocessing Stage

1) Filtering
A bandpass filter (0.5-40Hz) was used to remove unwanted baseline wander and high frequency interference.

2) Heartbeat Detection
We have implemented our own version of a QRS detector which determines QRS peaks using fuzzy classification of two ECG parameters. The first parameter is the normalized absolute amplitude of the ECG and the second parameter is a measure derived from the 3 points of inflection in the QRS. The detection performance of the system has been validated on the MIT-BIH arrhythmia database (Mark 1997). It detects over 99% of QRSs with a false detection rate of less than 1%.

3) RR-interval correction
RR-intervals were defined as the interval between successive QRS detection points. Due to poor signal quality and errors in the automatically generated QRS detections, the RR-interval sequences generated from both sets of QRS detection times contained physiologically unreasonable times. A first preprocessing step prior to calculating the ECG features was to calculate a corrected RR-interval sequence where all intervals were physiologically reasonable. The following automatic algorithm was developed for this purpose.
Suspect RR-intervals were found by applying a median filter of width five to the sequence of RR-intervals. This provided a robust estimate of the expected value for each RR-interval. Significant variations from this expected value led to it being flagged as a suspect RR-interval. Suspect RR-intervals could be due to either spurious QRS detections, or missed QRS complexes.
Spurious QRS detections were found by comparing the sum of adjacent RR-intervals with the robust RR-interval estimate. If this sum was numerically closer to the robust estimate than either of the individual RR-intervals then a spurious detection was deemed to be present. The two RR-intervals were merged to form a single RR-interval.
Conversely, we determined heuristically that if an RR-interval was a factor of 1.8 times or greater than the robust estimate then it was probable that one or more QRS complexes were missed. To estimate (interpolate) the times of the missing QRS complexes the RR-interval was divided by the sequence of integers 2,3,4,… until it best matched the robust estimate of the RR-interval. The single RR-interval was then subdivided by the appropriate integer to form a series of new detections.

4) ECG-Derived Respiratory (EDR) Signal
During the breathing cycle, the body-surface ECG is influenced by electrode motion relative to the heart and by changes in thoracic electrical impedance as the lungs fill and empty with air. The effect is most obviously seen as a slow modulation of the ECG amplitude at the same frequency as the breathing cycle (Moody 1985, Moody 1986, Travaglini 1998). To access this signal the original ECG signal was filtered with two median filters to remove the baseline wander. The original

ECG signal was processed with a median filter of 200 ms width to removed QRS complexes and P waves. The resulting signal was then processed with a median filter of 600ms width to remove T waves. The signal resulting from the second filter operation contained the baseline of the ECG signal, which was then subtracted from the original signal to produce the baseline corrected ECG signal.
A sample point of an ECG-derived respiratory signal (EDR) was then obtained by calculating the area enclosed by the baseline corrected ECG in the region 50 ms either side of the QRS detection point. An example of an EDR signal is shown in Figure 3

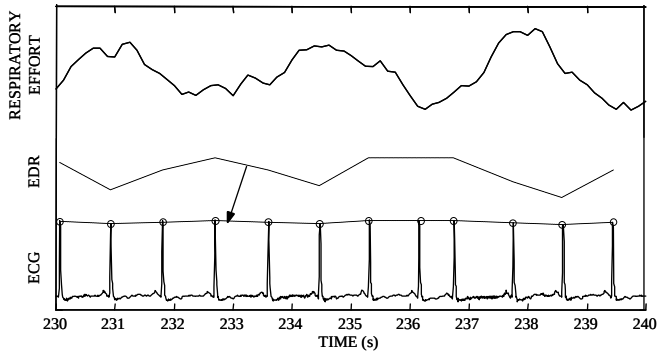

Figure 3: Estimating an ECG derived respiration signal (EDR) from the ECG.

5) Oximetry
The first step in SpO2 preprocessing was to remove obvious artefact. This was achieved by marking all changes of oxygen saturation between consecutive sampling intervals of greater than 4% per second as artefact. In addition, all SpO2 values of less than 65% or greater than 100.1% saturation were marked as artefact.

The next step is to produce an estimate of the running 5 minute average of the SpO2 signal. The artefact tagged samples were ignored in the running mean average calculation.

The final step in the preprocessing of the SpO2 signal was to resample both the original SpO2 signal and estimated baseline version at 0.1Hz.

D. Processing Stage

The preprocessing steps outlined above resulted in discrete index sequences of the RR-intervals, an EDR signal, an artefact tagged SpO2 signal and a SpO2 baseline signal. Based on these, a large set of features that could potentially be used for classification were considered. Features were generated for one-minute segments overlapped by 30 seconds. The features considered in this study were:
Interval-based power spectral density of the RR-intervals (DeBoer 1984),
Heart rate variability parameters (Teich 2000, Hilton 1999, Task 1996),
The power spectral density of the EDR signal, and
Oximetry time domain features
It is worth noting that none of the measures listed above consider the morphology of the ECG. It is implicitly assumed that the processes leading to apnoea occur at a location external to the heart and thus do not directly affect the generated cardiac potentials.

1) ECG-based features

a) Inter-beat intervals

The first step in calculating features from the RR intervals was to assess the quality of the RR interval sequence. If the average calculated heart rate was below 30 bpm or greater than 180 bpm or if four or more RR intervals were interpolated, then the RR intervals were considered artefact and RR features were not calculated for the epoch. If RR quality was OK then features relating based on frequency and time domain calculations were calculate.

An interval-based RR-interval power spectral decomposition (PSD) was calculated in the following way. A sequence of RR-intervals was associated with each one-minute segment. The index for this sequence was beat number, not time. The mean RR-interval for that segment was removed from each value, to yield a zero-mean sequence. The sequence was zero-padded to length 256, and the fast Fourier transform (FFT) was taken of the entire sequence. The magnitudes of the FFT coefficients were squared to yield a periodogram estimate of the PSD, which had high variance. Averaging of four adjacent frequency bins yielded a 64-point PSD estimate of which only the first 32 points were used as features (due to the symmetry of the upper and lower PSD point estimates). The *x*-axis has units of cycles/interval.

- Time domain features used included:

- the first five serial correlation coefficients,
- the log of the standard deviation of the RR intervals,
- the log of the standard deviation of the change in RR intervals (delta RR), and
- the log of the mean epoch RR interval.

b) EDR

An EDR quality estimate was produced by removing the mean and dividing by the standard deviation. A 100 point moving median filter is applied to the resulting EDR signal and saved as an EDR_{median} signal. The 3% to 97% sorted range of the EDR is calculated (a measure of spread). Differences between the original EDR and EDR_{median} greater than 1.8 times the calculated spread were considered to be artefact in the EDR signal.

If no artefact was detected in the EDR signal then the EDR signal was normalised by subtracting the mean and dividing by the standard deviation for each epoch of values. Features were obtained from the EDR power spectral density were the spectrum was calculated in an identical fashion to the RR-interval PSD except that the normalised EDR values were inputs into the transformation. The spectral variable was also defined as cycles/interval.

2) Oximetry-based features

If no artefact was detected in the oximetry signal in an epoch then the following temporal SpO2 saturation features were calculated for each 1 minute epoch, using the resampled signal:

1. The mean SpO2 value
2. The minimum SpO2 value
3. The number of SpO2 values of less than 92% saturation
4. The square root of the 5% to 95% spread in sorted SpO2 values
5. The mean of the absolute differences between successive SpO2 samples

In addition, the following two features were calculated using both the resampled SpO2 and estimated SpO2 baseline signals:

6. Count the number of times SpO2 and SpO2 estimated baseline are within +2.9% of each other during each epoch
7. Count the number of times SpO2 and SpO2 estimated baseline are within –2.9% of each other during each epoch

E. Epoch Classification Stage

Classifier models based on linear discriminants (LD) were utilised throughout this study. The model parameters were determined using the training data using 'plug-in' maximum likelihood estimates.
For linear discriminants the likelihood (*L*) function is defined as (Ripley 1996)

$$L = \sum_{k=1}^{c} \sum_{n=1}^{N_k} \log\left(f_k\left(\mathbf{x}_{kn}, \boldsymbol{\mu}_k, \boldsymbol{\Sigma}\right)\right) \qquad (1)$$

where the number of classes is *c*; the number of training examples in class *k* is N_k; and $f_k\left(\mathbf{x}_{kn}, \boldsymbol{\mu}_k, \boldsymbol{\Sigma}\right)$ is the value of the Gaussian distribution with mean $\boldsymbol{\mu}_k$ and common covariance $\boldsymbol{\Sigma}$ evaluated at training example $\mathbf{x}_{kn}$. The training process determines the parameter values of $\boldsymbol{\mu}_k$ and $\boldsymbol{\Sigma}$ that maximises the value of *L*.
This is maximized when the mean vectors are defined as

$$\boldsymbol{\mu}_k = \sum_{n=1}^{N_k} \mathbf{x}_{kn} \Big/ N_k \qquad (3)$$

and the covariance matrix is defined as

$$\boldsymbol{\Sigma} = \left(\sum_{k=1}^{c} \sum_{n=1}^{N_k} \left(\mathbf{x}_{kn} - \boldsymbol{\mu}_k\right)\left(\mathbf{x}_{kn} - \boldsymbol{\mu}_k\right)^T \right) \Big/ N . \qquad (4)$$

After determining the $\boldsymbol{\mu}_k$'s and $\boldsymbol{\Sigma}$ from the training data, a feature vector **x** is classified by assuming values for the prior probabilities π_k and calculated the estimated posterior probabilities, $P\left(k \mid \mathbf{x}\right)$ for the *k*th class using

$$P\left(k \mid \mathbf{x}\right) = \frac{\exp\left(y_k\right)}{\sum_{l=1}^{c} \exp\left(y_l\right)} \qquad (5)$$

where $y_k = -\frac{1}{2}\boldsymbol{\mu}_k^T \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu}_k + \boldsymbol{\mu}_k^T \boldsymbol{\Sigma}^{-1} \mathbf{x} + \log\left(\pi_k\right)$.
The prior probability of the classes were set equal and the final classification of a single feature set system was obtained by choosing the class with the highest posterior probability estimate from (5).

1) Combining Classifiers

To obtain a classification based on processing information from multiple feature sets simultaneously, the posterior probabilities obtained from each feature set were combined across the separate classifier outputs. Assuming the outputs from *M* classifiers are to be combined, the final posterior probability output $\overline{P}(k \mid \mathbf{x})$ was calculated from the individual classifier outputs $P_m(k \mid \mathbf{x})$ using the unweighted Bayesian addition integration scheme (Bloch 1996):

$$\overline{P}(k \mid \mathbf{x}) = \sum_{m=1}^{M} P_m(k \mid \mathbf{x}) \Big/ \sum_{l=1}^{c} \sum_{m=1}^{M} P_m(l \mid \mathbf{x}) \qquad (6).$$

As before the final classification is obtained by choosing the class with the highest posterior probability estimate. By using information from all available streams more efficient use of the available classification information was made.

F. AHI Calculation Stage

An estimated AHI was derived by first determining the average posterior probability value of the apnoea class and using this as input into a linear equation. The average posterior probability value was calculated as the average of the epoch-by-epoch posterior probability values. The coefficients of the linear fit were determined by plotting the square root of the average posterior probability value of apnoea versus the square root of the actual AHI and then using linear regression to find the slope of the line of best fit. The generally accepted range of AHI threshold between normal and clinically significant apnoea is an AHI value of 5 to 15. In clinical practice AHIs can range from 0 to 100 (or more). As the AHI range of 5 to 15 was very important we used the square root transformation to emphasise the contribution of the low AHI cases to the linear fit.

G. Performance Measures

1) Epoch-based

A two classifier model was trained that discriminated between normal and any type of sleep disordered breathing. Each epoch was thus labelled 'Normal' or 'SDB' by the system and the expert. Each epoch label by the system and expert was compared and the outcome determined as one of the following

- True positive (TP): an epoch is labelled as SDB by the expert and labeled as SDB by the system
- True negative (TN): an epoch is labelled as Normal by the expert and labeled as Normal by the system
- False positive (FP): an epoch is labelled as Normal by the expert and labeled as SDB by the system
- False negative (FN): an epoch is labelled as SDB by the expert and labeled as Normal by the system

Counts of these outcomes over all the epochs were made and the two way confusion matrix formed as shown in Table 3.

TABLE 3: TWO WAY CONFUSION MATRIX FOR EPOCH-BASED PERFORMANCE ASSESSMENT

		Actual	
Predicted		**Normal**	**SDB**
	Normal	TN	FN
	SDB	FP	TP

Using the table 3 the following performance measures calculated:

- Specificity = TN/(TN+FP)
- Sensitivity = TP/(TP+FN)
- Positive Predictivity = TP/(TP+FP)
- Negative Predictivity = TN/(TN+FN)
- Accuracy = (TN+TP)/(TP+TN+FN+FP)

2) AHI estimation

The second set of performance measures examined the performance of the system in separating normal subjects from 'apnoea' subjects on the basis of AHI. Firstly the clinically determined AHIs were used to determine the clinical classification using a predetermined AHI thresholds. Each record was thus categorised as 'normal', 'borderline', and 'apnoea' on basis of clinically determined AHI. Secondly, the estimated AHI was determined and thresholded as above to categorise each record as 'normal' or 'apnoea' on basis of predicted AHI.

With each record label by the system and expert a comparison of the labels was made and the outcome determined as one of the following

- True positive (TP): a record is labelled as Apnoea by the expert and labeled as Apnoea by the system
- True negative (TN): an epoch is labelled as Normal by the expert and labeled as Normal by the system
- False positive (FP): an epoch is labelled as Normal by the expert and labeled as Apnoea by the system
- False negative (FN): an epoch is labelled as Apnoea by the expert and labeled as Normal by the system

Counts of these outcomes over all records were made and the confusion matrix formed as shown in Table 4.

TABLE 4: CONFUSION MATRIX FOR PER-SUBJECT AHI ASSESSMENT

		Clinical AHI		
Predicted AHI		Normal (AHI_C <=5)	Borderline (5<AHI_C<15)	Apnoea (AHI_C >=15)
	Normal (AHI_P<=10)	TN	-	FN
	Apnoea (AHI_P >10)	FP	-	TP

Where AHI_P is predicted AHI value and AHI_C is clinical AHI value.
Using the Table 4 the following performance measures calculated:

- Specificity = TN/(TN+FP)
- Sensitivity = TP/(TP+FN)
- Positive Predictivity (PP) = TP/(TP+FP)
- Negative Predictivity (NP) = TN/(TN+FN)
- Accuracy = (TN+TP)/(TP+TN+FN+FP)

Note that 28 subjects were classified clinically as borderline in this study.

H. Classifier Performance Estimation

When developing a classifier it is important to be able to estimate the expected performance of the classifier on data not used in training. The available data must be divided into independent training and test sets. There are a number of schemes for achieving this and the most suitable for the size of data set used in this study, is *n*-fold cross validation (XV) (Kohavi 1995, Bishop 1995). This scheme randomly divides the available data into *n* approximately equal size and mutually exclusive "folds". For an *n*-fold XV run, *n* classifiers are trained with a different fold used each time as the test set, while the other *n*-1 folds are used for the training data. The choice of *n* influences the ratio of data used for training/testing. Cross validation estimates are generally pessimistically biased, as training is performed using a subsample of the available data.
A factor in this study was that while epochs of data were independent across records they were not independent within a record. To eliminate the bias that could potentially arise from having epochs from the same record in the training and testing data we applied the XV scheme to the records and not to epochs. In other words all epochs from one record were treated as a unit.
In this study we used a special case of XV (the leave-one-out XV scheme) where all but one available examples are used for training and one example used for testing. To achieve this, in turn, each recording was reserved as the test record and the remaining 125 records used as the training data. As we were processing 125 recordings to complete a run of XV, 125 classifiers were trained and tested.

IV. RESULTS AND DISCUSSION

A. Epoch-based classification

We first report on the performance of the three systems in correctly identifying epochs as being either normal or SDB. The performance is given in Table 5 in terms of specificity, sensitivity, negative and positive predictive values, accuracy and Cohen's kappa coefficient. While all three systems provide good performance, the oximetry-only system provides the highest specificity on a per-epoch basis. We note that the limiting factor for the oximetry-only system is its sensitivity, and we ascribe this to the fact that hypoponeas may lead to relatively modest desaturations. The trade-off in per-epoch performance between sensitivity and specificity will also depend upon the mapping from the event-based annotations to the epoch labels.

TABLE 5: EPOCH-BASED CLASSIFICATION RESULTS FOR THE ECG AND OXIMTERY DATA STREAMS.

	Specificity	**Sensitivity**	**NP**	**PP**	**Accuracy**	**Kappa**
ECG	78.3	78.9	91.1	56.9	78.5	0.51
Oximetry	94.8	71.4	90.4	83.0	88.7	0.69
ECG-oximetry	94.3	73.4	90.7	82.4	88.8	0.70

B. Per-Subject AHI Estimation

Table 6 provides the more clinically useful performance figures on a per-subject basis rather than on a per-epoch basis. All three systems show excellent performance in terms of sensitivity, and the oximetry and combined systems also have high specificity. We also note that one advantage of the combined system is that records which could not previously be reliably classified due to excess noise in the ECG or oximetry trace can be reliably scored.

TABLE 6: AHI-BASED CLASSIFICATION RESULTS FOR THE ECG AND OXIMTERY DATA STREAMS. THESE FIGURES DEMONSTRATE THE PERFORMANCE OF THE SYSTEM IN SEPARATING THE NORMAL CASES (I.E., AHI <5) FROM THOSE WITH CLINICALLY SIGNIFICANT APNOEA (AHI>15). THE NC COLUMN INDICATES THE NUMBER OF RECORDS THAT WERE NOT CLASSIFIED DUE TO POOR SIGNAL QUALITY.

	TN	FN	FP	TP	NC	Sensitivity	Specificity	Accuracy
ECG	20	11	5	58	2	92.1	64.5	83.0
Oximetry	28	2	4	59	3	93.7	93.3	93.5
ECG-oximetry	29	2	4	61	0	93.8	93.5	93.8

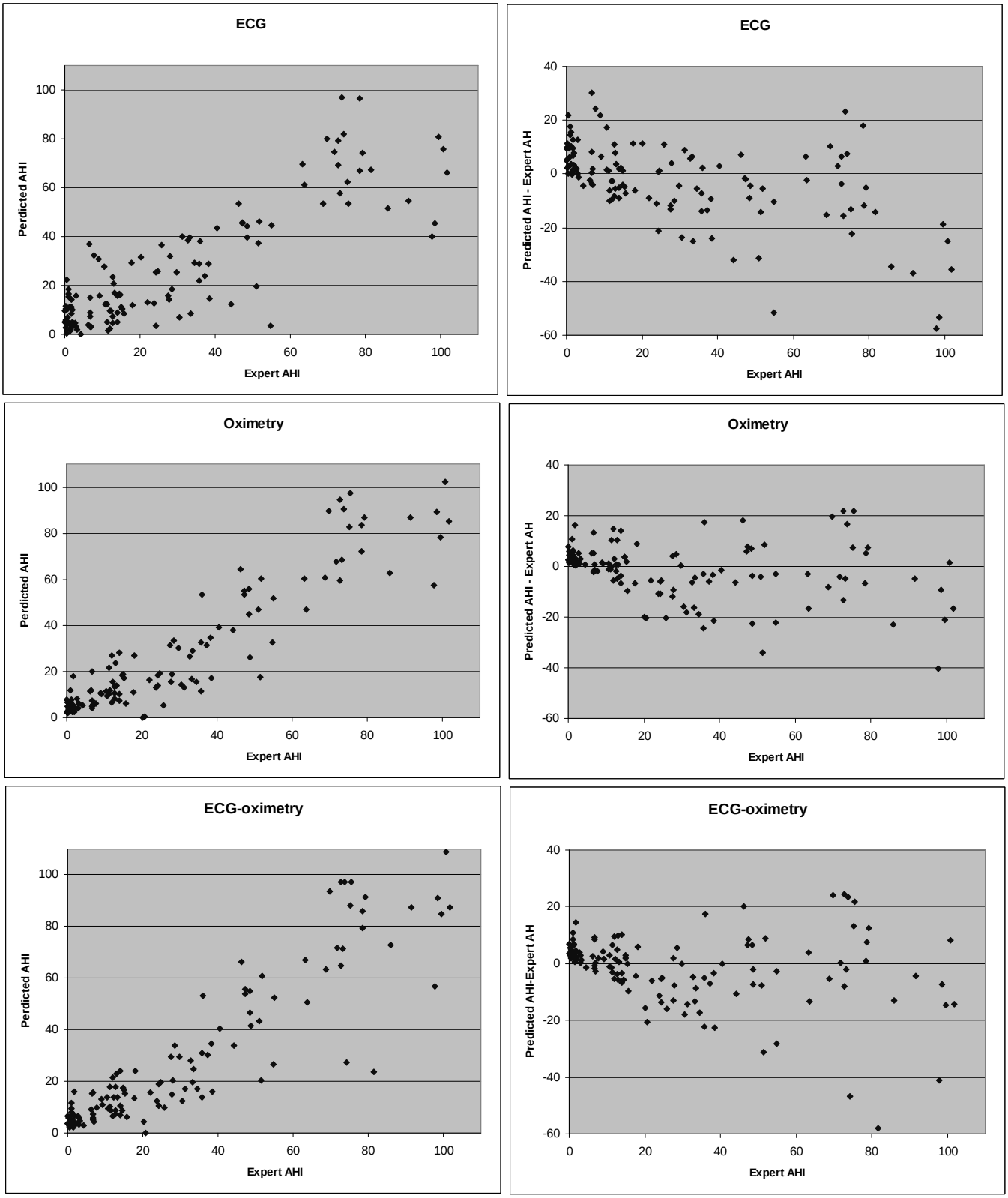

Figure 4: AHI plots for ECG only, Oximetry-only and ECG-oximetry systems. Left column) AHI estimated from the three systems versus AHI determined from the full polysomnogram. Right colum) Bland-Altman plot of the three systems and PSG AHI.

Figure 4 reveal that the oximetry-only and ECG-oximetry system more accurately estimated the PSG AHI value than the ECG-only system. On average all systems slightly underestimated the AHI. The bias of the ECG-only system was an underestimation of AHI by 2.7 events/hour, the bias of the oximetry-only was 1.5 events/hour and for the ECG-oximetry system the bias was 1.6 events/hour.

It is instructive to more carefully consider (for each classifier system) the 96 non-borderline cases (i.e., AHI≤5 or AHI≥15) in this study in order to determine the confounding factors in the analysis

1) ECG

Using the ECG-only system, sixteen classification errors were made with a sensitivity of 92% and a specificity of 65% in separating normal subjects from those with clinically significant apnoea. Two subjects could not be analysed due to poor signal quality. There were eleven false positives and five false negatives. All false negatives had a proportion of central hypopnea or central apnea events. A possible explanation of the failure of the ECG-only system to successfully screen these subjects is that the hypopnoeas did not trigger a sufficiently obvious bradycardia/tachycardia pattern to be detected by the system. The most significant error of the ECG-only system was the estimation of AHI of 3.3 for a subject with an AHI of 54.8. The false-positive results indicate that causes other than SDB can contribute to changes in heart rate during sleep. For example, we can postulate that periodic limb movements may induce associated cardiac events. It is more likely however that sub-clinical events (i.e., apneas which do not reach the ten-second duration definition, or hypopneas which do not meet the criterion of a 50% reduction in amplitude) cause a corresponding cardiac change, which will not be reflected in a contribution to that person's AHI score.

This study indicates that automated analysis of Holter recordings may have clinical utility in identifying patients likely to have obstructive sleep apnoea. A prime advantage of this technique is that it comes at no additional cost for subjects undergoing Holter monitoring as part of a normal cardiology work-up.

The sensitivity and specificity of our technique are comparable to those reported elsewhere in the literature for ECG-based screening. Roche et al. have presented several reports on automated recognition of subjects with obstructive sleep apnoea syndrome (OSAS) using analysis of heart rate variability (Roche 1999), inter-beat interval times (Roche 2002), and wavelet-based analysis of the RR interval series (Roche 2003). Using HRV parameters, they obtained a sensitivity of 83% and specificity of 96% on an independent test set of 52 subjects, using a criterion of AHI>10 as defining OSAS. Using inter-beat intervals, they obtained sensitivity of 87% and specificity of 52% on 124 subjects. Finally, their wavelet based analysis applied to 147 subjects yielded a sensitivity and specificity of 92% and 90%. Their techniques do not provide any temporal information about the occurrence of the apneic events, nor do they attempt to map their output variables to AHI. They do not report on differences between subjects with primarily central versus obstructive apnoea. Stein et al. reported on a technique to identify subjects with OSAS by visual inspection of RR tachograms (Stein 2003). A human scorer was trained to recognise characteristic Cyclical Variations in Heart Rate (CVHR) associated with obstructive events. The magnitude and frequency of occurrence of these CVHRs were then used to classify 11 control subjects and 46 clinical subjects in terms of OSAS. Of the 46 clinical subjects, 33 had significant OSAS (AHI>15). The positive predictive accuracy was 86%, and negative predictive accuracy was 94% (which correspond to a sensitivity of 97% and specificity of 77%) in distinguishing subjects with AHI>15 from those with lower AHIs. However, in this study the decision criteria was arrived at retrospectively. Moreover the need for human scoring may reduce its potential clinical utility.

2) Oximetry

The oximetry-only system did not analyse three subjects (3%) due to poor signal quality. Since these studies were carried out in an attended setting, we expect that the true data loss rate may be even higher (this is consistent with previous field trials of unattended nocturnal oximetry). Of the remaining subjects the system made four false positive and two false negative classification errors. The sensitivity was 94% with a specificity of 93%. The expert determined AHIs of the false negatives range between 15 and 25 and the vast majority of events for these records were either hypopneas or central apneas. These results suggest (not surprisingly) that the oximetry system performs best at identifying obstructive apneas and less well at hypopneas. This is consistent with the comments of Levy et al. (Levy 1996) who also noted that hypoponeas often lead to minimal desaturations. It is also worth noting that the system successfully classified all the severe apnea cases (AHI>30). These results compare well with other published results. Levy et al. reported on a study of 301 adults and found a sensitivity of 98% with a specificity of 46% in distinguishing AHI>15. Yamashiro et al. used nocturnal oximetry as a screener for OSA and achieved a sensitivity of 90% and specificity of 75%; however, they had to reject 10% of subjects from the study due to poor SpO2 signal quality. In an adult population, Zamarron et al. used spectral analysis of the oximetry signal to screen for OSA and found a sensitivity of 78% and a specificity of 89% (Zamarron 1999).

3) Combined ECG-Oximetry System

Given the previously published evidence on the use of ECG and oximetry as screening tools, a hypothesis of this research work was that the combination of ECG and oximetry would provide the highest performing and most robust system. Our classifier system was designed so that if one of the channels had acquisition problems (e.g., leads falling off, poor perfusion) then the system would automatically switch to the other channel and continue its analysis. So for example if the oximeter probe detached during the night the system still continued to use the ECG to identify the apnoeas. When both channels were available then the system used simultaneous ECG and oximetry measurements to obtain the most accurate analysis.

The results show that the ECG-oximetry system successfully analysed all subjects thus demonstrating the robustness of the system to individual signal acquisition problems. Both the specificity and sensitivity were 94% representing a notable improvement on the ECG only system and slightly exceeding the performance of the oximetry-only system. The principal benefit of the ECG-oximetry system is its reliability, and the possibility of linking periods of apnea to any associated

arrhythmia events. There are no other published ECG-oximetry systems we are aware but our results compare well to the oximetry systems described in the previous system and with the added benefit of immunity to poor oximetry signal quality.

C. Clinical Utility

Sleep apnoea impacts significantly on the cardiovascular system (Newman 2001, Shamsuzzaman 2003, Phillips 2002, Sjostrom 2002,Neito 2000). Moreover, many sufferers of cardiovascular disease have associated sleep apnea. For example 50% of subjects with atrial fibrillation and 50% of subjects with congestive heart failure have sleep apnea (Wolf 2003a). Thirty percent of patients with coronary arterial disease also have sleep apnea (Wolk 2003b). Given the strong association between sleep apnoea and cardiovascular health, it makes sense from a clinical perspective to perform cardiac monitoring and sleep-apnea screening analysis simultaneously. Since Holter monitoring is so widely carried out (nearly one million Holter tests are carried out per annum in the USA), and it has been shown by this work (and others) to be a reasonably sensitive test for obstructive sleep apnea, it makes sense to promote the integration of apnea-screening methods into standard Holter software packages. In this way, cardiologists could routinely identify candidates for full evaluation by sleep laboratories, who have a high likelihood of having sleep apnea. Secondly, we have shown that a more clinically accurate and physiologically meaningful screening tool can be provided by combining a standard Holter ECG recorder with a simultaneous oximetry recording. A benefit of an ECG-oximeter system over existing ECG- or oximetry-only systems is its robustness to signal acquisition problems. It also allows the investigation of nocturnal arrhythmias in relation to whether their underlying cause is apnea-related.

V. CONCLUSION

To conclude, we have presented three automated screening algorithms based on ECG, oximetry, or combinations of ECG and oximetry. These algorithms have been trained and validated on a clinically significant cohort of 125 subjects. The performance of the reported screening systems either matches or surpasses other systems that have previously been reported. Moreover, it provides both an estimated AHI, and the temporal sequence of apnea events during the night, which can assist a clinician in forming a diagnosis. The systems' performance has been best validated in non-apneic subjects and subjects with obstructive and mixed sleep apnea; there were insufficient numbers of subjects in this database with primarily central sleep apnea to provide meaningful performance figures. Thus, the system may be of particular clinical value as a screening tool among snoring subjects to evaluate the presence or absence of obstructive sleep apnea. In particular, the profile of sleep apnea has risen in recent years among cardiologists due to increased recognition of the disorder as an important contributing factor to cardiovascular morbidity. The proposed system may facilitate a higher involvement of cardiologists in the clinical management of sleep disordered breathing since the evaluation of possible OSA by this technique can be performed as part of routine Holter monitoring.

VI. REFERENCES

[1] American Academy of Sleep Medicine Task Force. (1999). Sleep-related breathing disorders in adults: Recommendations for syndrome definition and measurement techniques in clinical research. *Sleep, 22(5)*, 667-689.

[2] Bishop, C. M. (1995). *Neural Networks for Pattern Recognition*. New York: Oxford University Press.

[3] Bloch, I. (1996). Information combination operators for data fusion: a comparative review with classification. *IEEE Transactions on Systems, Man and Cybernetics, part A, 26(1)*, 52 – 67.

[4] Brouillette, R. T., Morielli, A., Leimanis, A., Waters, K. A., Luciano, R., Ducharme, F. M. (2000). Nocturnal pulse oximetry as an abbreviated testing modality for pediatric obstructive sleep apnoea. *Pediatrics, 105(2)*, 405-12.

[5] Chesson, A.L. Jr, Berry, R.B., Pack, A. (2003). Practice parameters for the use of portable monitoring devices in the investigation of suspected obstructive sleep apnoea in adults. *Sleep, 26(7)*, 907-13.

[6] Chiner, E., Signes-Costa, J., Arriero, J. M., Marco, J., Fuentes, I., Sergado, A. (1999). Nocturnal oximetry for the diagnosis of the sleep apnoea hypopnoea syndrome: a method to reduce the number of polysomnographies? *Thorax, 54(11)*, 968-71.

[7] Davila, D.G., Richards, K.C., Marshall, B.L., O'Sullivan, P.S., Gregory, T.G., Hernandez, V.J., Rice, S.I. (2002). Oximeter performance: the influence of acquisition parameters. *Chest, 122(5)*, 1654-60.

[8] Dingli, K., Assimakopoulos, T., Wraith, P. K., Fietze, I., Witt, C., Douglas, N. J. (2003). Spectral oscillations of RR intervals in sleep apnoea/hypopnoea syndrome patients. *Eur. Respir. J., 22*, 943-950.

[9] DeBoer, R. W., Karemaker, J. M., Strackee, J. (1984). Comparing spectra of a series of point events particularly for heart rate variability data. *IEEE Trans. Biomed. Eng., BME-31*, 384–387.

[10] de Chazal, P., Heneghan, C., Sheridan, E., Reilly, R., Nolan, P., O'Malley, M. (2003). Automated processing of the single lead electrocardiogram for the detection of obstructive sleep apnoea. *IEEE Trans. Biomed. Eng., 50*, 686-696.

[11] Duda, R. O., Hart P. E., Stork, D. G. (2001). *Pattern Classification*. New York: John Wiley and Sons.

[12] Gislason, T. & Benediktsdottir, B. (1995). Snoring, apneic episodes, and nocturnal hypoxemia among children 6 months to 6 years old. An epidemiologic study of lower limit of prevalence. *Chest, 107*, 963-6.

[13] Golpe, R., Jimenez, A., Carpizo, R., Cifrian, J. M. (1999). Utility of home oximetry as a screening test for patients with moderate to severe symptoms of obstructive sleep apnoea. *Sleep, 22(7)*, 932-7.

[14] Guilleminault, C., Connolly, S. J., Winkle, R., Melvin, K., Tilkian, A. (1984). Cyclical variation of the heart rate in sleep apnoea syndrome. Mechanisms and usefulness of 24h electrocardiography as a screening technique. *Lancet, 1*, 126-131.

[15] Gyulay, S., Olson, L. G., Hensley, M. J., King, M. T., Allen, K. M., Saunders, N. A. (1993). A comparison of clinical assessment and home oximetry in the diagnosis of obstructive sleep apnoea. *Am. Rev. Respir. Dis., 147(1)*, 50-3.

[16] Hilton, M. F., Bates, R. A., Godfrey, K. R., Chappell, M. J., Cayton, R. M. (1999). Evaluation of frequency and time-frequency spectral analysis of heart rate variability as a diagnostic marker of the sleep apnoea syndrome. *Med. Biol. Eng. Comput., 37(6)*, 760-769.

[17] Johns M.W. (1991). A new method for measuring data and sleepiness: Epworth Sleepiness Scale. *Sleep, 14*, 540-545.

[18] Kohavi, R. (1995). A study of cross validation and bootstrap for accuracy estimation and model selection. In: Proc. of 14th Int. Joint Conference on Artificial Intelligence, 1137-1143.

[19] Levy, P., Pepin, J. L., Deschaux-Blanc, C., Paramelle, B., Brambilla, C. (1996). Accuracy of oximetry for detection of respiratory disturbances in sleep apnoea syndrome. *Chest, 109(2)*, 395-9.
[20] Magalang, U. J., Dmochowski, J., Veeramachaneni, S., Draw, A., Mador, M. J., El-Solh, A., Grant, B. J. (2003). Prediction of the apnoea-hypopnoea index from overnight pulse oximetry. *Chest, 124(5)*, 1694-701.
[21] Mark, R. & Moody, G. (1997). MIT-BIH Arrhythmia Database. http://ecg.mit.edu/dbinfo.html
[22] Moody, G. B., Mark, R. G., Zoccola, A., Mantero, S. (1985). Derivation of respiratory signals from multi-lead ECGs. In: *Computers in Cardiology, 12*, 113-116, Washington, DC: IEEE Computer Society Press.
[23] Moody, G. B., Mark, R. G., Zoccola, A., Mantero, S. (1986). Clinical validation of the ECG-derived respiration (EDR) technique. In: *Computers in Cardiology, 13*, 507-510, Washington, DC: IEEE Computer Society Press.
[24] Moody, G. B., Mark, R. G., Goldberger, A. L., Penzel, T. (2000). Stimulating rapid research advances via focused competition: the Computers in Cardiology Challenge 2000. In: *Computers in Cardiology, 27*, 207-210, Piscataway, NJ: IEEE Press.
[25] Nieto, F. J., Young, T. B., Lind, B. K., Shahar, E., Samet, J. M., Redline, S., D'Agostino, R. B., Newman, A. B., Lebowitz, M. D., Pickering, T. G., for the Sleep Heart Health Study. (2000). Association of sleep-disordered breathing, sleep apnoea, and hypertension in a large community-based study. *JAMA, 283*, 1829-1836.
[26] Newman, A. B., Nieto, F. J., Guidry, U., Lind, B. K., Redline, S., Shahar, E., Pickering, T.G., Quan, S.F., for the Sleep Heart Health Study Research Group. (2001). Relation of sleep-disordered breathing to cardiovascular disease risk factors - The Sleep Heart Health Study. *Am. J. Epidemiol., 154*, 50-59.
[27] Oeverland, B., Skatvedt, O.. Kvaerner, K. J., Akre, H. (2002). Pulse oximetry: sufficient to diagnose severe sleep apnoea. *Sleep Med. 3(2)*, 133-8.
[28] Penzel, T. (2000). The Apnoea-ECG Database. In: *Computers in Cardiology, 27*, 255-258, Piscataway, NJ: IEEE Press.
[29] Penzel, T., McNames, J., de Chazal, P., Raymond, B., Murray, A., Moody, G. (2002). Systematic comparison of different algorithms for apnoea detection based on ECG recordings. *Med. Biol. Eng. Comp., 40*, 402-407.
[30] Phillips, B. G. & Somers, V. K. (2002). Sleep disordered breathing and risk factors for cardiovascular disease. *Curr. Opin. Pulm. Med., 8*, 516-520.
[31] Ripley, B. D. (1996). *Pattern Recognition and Neural Networks*. Cambridge: Cambridge University Press.
[32] Roche, F., Gaspoz, J. M., Court-Fortune, I., Minini, P., Pichot, V., Duverney, D., Costes, F., Lacour, J.-R., Barthélémy, J.-C. (1999). Screening of obstructive sleep apnoea syndrome by heart rate variability analysis. *Circulation, 100*, 1411-1415.
[33] Roche, F., Duverney, D., Court-Fortune, I., Pichot, V., Costes, F., Lacour, J.-R., Antoniadis, A., Gaspoz, J.-M., Barthelemy, J.-C. (2002). Cardiac interbeat interval increment for the identification of obstructive sleep apnoea. *Pacing Clin. Electrophysiol., 25*, 1192-9.
[34] Roche, F., Pichot, V., Sforza, E., Court-Fortune, I., Duverney, D., Costes, F., Garet, M., Barthelemy, J.-C. (2003). Predicting sleep apnoea from heart period: a time-frequency wavelet analysis. *Eur Respir J., 22*, 937-942.
[35] Series, F. (2002). Interpretation of home oximetry tracings. *Chest, 121(3)*, 1006-7.
[36] Series, F., Kimoff, R. J., Morrison, D., Leblanc, M. H., Smilovitch, M., Howlett, J., Logan, A. G., Floras, J. S., Bradley, T.D. (2005). Prospective evaluation of nocturnal oximetry for detection of sleep-related breathing disturbances in patients with chronic heart failure. Chest. 127(5), 1507-14.
[37] Shamsuzzaman, A, S., Gersh, B. J., Somers, V. K. (2003). Obstructive sleep apnoea: implications for cardiac and vascular disease. *JAMA, 290*, 1906-14.
[38] Sjostrom, C., Lindberg, E., Elmasry, A., Hagg, A., Svardsudd, K., Janson, C. (2002). Prevalence of sleep apnoea and snoring in hypertensive men: a population based study. *Thorax, 57*, 602-607.
[39] Stein, P. K., Duntley, S. P., Domitrovich, P. P., Nishith, P., Carney, R. M. (2003). A simple method to identify sleep apnoea using Holter recordings. *J. Cardiovasc. Electrophyiol., 14*, 467-473.
[40] Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology. (1996). Heart rate variability – standards of measurement, physiological interpretation and clinical use. Eur. Heart J., 17, 354-382.
[41] Teich, M. C., Lowen, S. B., Jost, B. M., Vibe-Rheymer, K., Heneghan, C. (2000). Heart rate variability: measures and models. In M. Akay (Ed.), Nonlinear biomedical signal processing, vol. II. Piscataway, NJ: IEEE Press.
[42] Travaglini, A., Lamberti, C., DeBie, J., Ferri, M. (1998). Respiratory signal derived from eight-lead ECG. In: *Computers in Cardiology, 25*, 65-68, Piscataway, NJ: IEEE Press.
[43] Wolk, R., Shamsuzzaman, A. S., Somers, V. K. (2003a). Obesity, sleep apnea, and hypertension. *Hypertension, 42(6)*, 1067-74.
[44] Wolk, R., Kara, T., Somers, V. K. (2003b). Sleep disordered breathing and cardiovascular disease. *Circulation, 108(1)*, 9-12.
[45] Yamashiro, Y. & Kryger, M. H. (1995). Nocturnal oximetry: is it a screening tool for sleep disorders? *Sleep, 18(3)*, 167-71.
[46] Young, T., Palta, M,, Dempsey, J., Skatrud, J., Weber, S., Badr, S. (1993). The occurrence of sleep-disordered breathing among middle-aged adults. *N. Engl. J. Med., 328*, 1230-1235.
[47] Young, T., Evans, L., Finn, L., Palta, M. (1997). Estimation of the clinically diagnosed proportion of sleep apnoea syndrome in middle-aged men and women. *Sleep, 20*, 705-706.
[48] Zamarron, C., Romero, P. V., Rodriguez, J. R., Gude, F. (1999). Oximetry spectral analysis in the diagnosis of obstructive sleep apnoea. *Clin. Sci. (Lond), 97(4)*, 467-73.