Tsimenidis C, Murray A.

[False alarms during patient monitoring in clinical intensive care units are highly related to poor quality of the monitored electrocardiogram signals.](#)

Physiological Measurement 2016, 37, 1383-91.

Copyright:

This is an author-created, un-copyedited version of an article accepted for publication in *Physiological Measurement*. IOP Publishing Ltd is not responsible for any errors or omissions in this version of the manuscript or any version derived from it. The Version of Record is available online at

<http://dx.doi.org/10.1088/0967-3334/37/8/1383>

DOI link to article:

<http://dx.doi.org/10.1088/0967-3334/37/8/1383>

Date deposited:

03/08/2016

Embargo release date:

25 July 2017

This work is licensed under a [Creative Commons Attribution-NonCommercial 3.0 Unported License](#)

False alarms during patient monitoring in clinical intensive care units are highly related to poor quality of the monitored electrocardiogram signals

Charalampos Tsimenidis¹, Alan Murray^{1,2}

¹Electrical & Electronic Engineering, Newcastle University, UK

²Faculty of Medical Sciences, Newcastle University, UK

Addresses for correspondence:

Dr Charalampos Tsimenidis
Senior Lecturer
School of Electrical & Electronic Engineering
Newcastle University, UK
charalampos.tsimenidis@newcastle.ac.uk

Prof Alan Murray
Professor, School of Electrical & Electronic Engineering, and
Professor of Cardiovascular Physics, Faculty of Medical Sciences
Newcastle University, UK
alan.murray@newcastle.ac.uk

Abstract

Electrocardiograms (ECGs) recorded from patients in intensive care were investigated to quantify any relationship between ECG signal quality and false monitoring alarms. False alarms are a considerable problem for nursing and medical staff as they distract from clinical care, and are also a problem for patients as they disturb rest, which is important for clinical recovery. ECG and alarm data were obtained for 750 patient alarms from the PhysioNet database. The final 8 s period before the alarm was triggered was investigated. All but one ECG channel in 38 ECG recordings with out-of-range data were associated with false positive alarms ($p < 0.0001$). The frequency contributions for baseline (BL) instability, electromyogram (EMG) muscle noise, and high frequency (HF) noise were calculated. For all three frequency bands, the contributions associated with false positive alarms were very significantly greater than for true positive alarms ($p < 0.0001$). The greatest difference was for BL with a mean level for false positive alarms 4.0 times greater than for true positive alarms, followed by EMG and HF at 1.6 times and 1.4 times respectively. These results confirm that attention needs to be taken to improve ECG signal quality to reduce the frequency of clinical false alarms, and hence improve conditions for clinical staff and patients.

1. Introduction

Patient monitoring was introduced several decades ago to enable clinical staff to be aware of the clinical condition of patients under their care (Murray 1982). Initially this was for only ECG and blood pressure monitoring, but has evolved to include a vast array of medical devices monitoring different physiological functions. This has enabled medical and nursing staff to attend quickly to changes in patient conditions, and medical monitoring devices have been developed to detect unwanted clinical conditions and hence initiate alarms. Although these alarms often work well, even a small percentage of false alarms can lead to significant frustration and fatigue with increased workload in already busy medical units, such as surgical theatres or intensive care units (Chambrin 2001, Donchin & Seagull 2002, Imhoff & Kuhls 2006, Cvach 2012).

Although the quality of physiological data has been studied in ambulatory monitoring (Di Marco et al 2012) and in Intensive Care (Allen and Murray 1996), there is as yet no adequate compromise between reliable detection of genuine alarms and the elimination of false alarms. False alarms are still a very significant problem. Aboukhalil et al (2008) obtained multiple expert reviews of 5386 critical ECG arrhythmia alarms from a total of 447 adult patient records. They found that on average 43% of the critical ECG arrhythmia alarms were false, with rates for different arrhythmia types varying between 23% and 91%.

Published studies have looked at ECG quality, which is known to be associated with false monitoring alarms. Allen and Murray (1996) studied all patients monitored in a single coronary care unit bed continuously for 10 weeks, and found that night-time noise levels, low frequency content and out-of-range events, were lower than day-time levels. These are the factors that are associated with patient movement and false alarms. Langley et al (2011) developed techniques for identifying 'unacceptable' ECG quality as categorised by expert annotators. These techniques included baseline instability, amplitude saturation and fast high frequency changes, and the research was extended by Di Marco et al (2012). However, these studies were not able to relate quality to false alarms.

Some useful work has been reported on false alarms. Graham and Cvach (2010) recorded alarms in a medical progressive care unit during an 18-day period, and found that there were 16,953 alarms, equating to 942 alarms per day or 1 every 92 seconds. After implementing a procedure for setting the monitoring alarm conditions for each patient the total equivalent number of alarms reduced to 9647, which was a 43% reduction, but still a very significant number. Behar et al (2013) studied three ECG databases, using one to train an ECG quality assessment model with human annotation of quality. The other databases were then used to test and develop a classification tool. They were able to conclude that their result “suggests that the quality assessed by the classifier is associated with the alarm status”.

However, there is little specific data on the association between ECG quality and false alarms. To help reduce the problems of false alarms, a PhysioNet / Computing in Cardiology Challenge was introduced to aid the development of improved alarm monitoring (Clifford et al 2015). Our initial study of these alarm data showed the difficulty in reliably discriminating between noisy ECG signals and genuine clinical conditions requiring alarm activation (Tsimenidis & Murray 2015).

This research has now been extended in the study reported here to investigate ECG signal quality with the aim of quantifying its relationship to true and false alarms, and hence being able to demonstrate that paying attention to improving signal quality could have a significant effect on reducing false monitoring alarms.

2. Methods

2.1. Alarm data

The ECG data analysed in this study were obtained from the PhysioNet website (<http://www.physionet.org/challenge/2015/>). They were made available for a PhysioNet / Computing in Cardiology Challenge. The recordings were obtained from three of the largest intensive care monitor manufacturers' bedside monitoring units in four hospitals in the USA and Europe. These were standard commercial monitors. The true alarms covered five groups of arrhythmias: asystole, bradycardia, tachycardia, ventricular tachycardia and ventricular flutter. In total there were 750 ECG recordings for which the alarm condition was given; 456 were false positive (FP) alarms and 294 were true positive (TP) alarms. These alarm conditions had been checked and confirmed by at least two experienced clinical staff. Each ECG recording was of 2 channels, and contained 5 min of ECG immediately preceding the alarm indicated by the monitor. Pulse data were also provided, but not analysed in our study.

2.2. Analysis

All alarm conditions could be satisfied by 8 s of data, and so this study analysed only the last 8 s of data. The data were downloaded, and analysed in Matlab. Our previous study (Tsimenidis & Murray 2015) had shown that many of the false alarms were associated with unwanted noise, including baseline instability induced by electrode movement, high frequency muscle disturbance, and very high frequency disturbances associated with poor and intermittent electrode contact.

To analyse the level of these noise conditions, both ECG channels were analysed by Fast Fourier Transform in Matlab. The mean signal level of each 8 s data file was subtracted so that the mean of each recording was zero, and the data files were then zero padded to 8192 sample points, with 2000 ECG samples (8 s at 125 Hz) and 6192 zeros. This gave a frequency resolution of 0.015 Hz.

The frequency bands selected were designed to include simple broad ranges. The bands selected were low frequency baseline (BL) instability (0.1-1 Hz), electromyogram (EMG) muscle noise (10-20 Hz) and high frequency (HF) disturbances (20-40 Hz), leaving 1-10 Hz for the ECG (Figure 1). It was accepted that ECG harmonics could spread beyond 10 Hz and that noise could lie in our ECG band, but these bands were deemed acceptable for an initial study. For each frequency band, on every alarm recording, the mean signal content across the frequency band was calculated.

Figure 1. Example frequency analysis showing the FFT spectral density and its contribution to the three frequency bands analysed: BL baseline instability 0.1-1 Hz, EMG muscle noise 10-20 Hz, HF high frequency disturbances 20-40 Hz.

2.3. Statistical comparison

Analysis of variance was used to investigate differences between false and true positive alarms for each frequency band. Chi square analysis was used for categorical data. A probability $p < 0.05$ was taken as significant.

4. Results

4.1 Out of range ECG signals

Some of the ECG recordings contained non-numeric samples, and these were mostly associated with poor recordings that went out of the range. With only one exception all such recordings were associated with false positive alarms. For ECG channel 1, there were 20 (20/456 FP v 0/294 TP, chi square 13, $p = 0.0003$), and for channel 2 there were 35 (34/456 FP v 1/294 TP, chi square 20, $p < 0.0001$). These were from 38 alarm recordings as most with problems had them on both channels. The single true positive with this problem on channel 2 had a satisfactory channel 1, enabling the true alarm to be detected.

4.2 Signal content of the noise frequency bands

There were highly significant differences in frequency content between ECG FP and TP alarms for all frequency bands (BL, EMG and HF), and for both channels (all $p < 0.0001$) (Table 1).

Frequencies analysed	ECG channel	Alarm: True positive (TP) or false positive (FP)	Number of alarm recordings	Mean	95% Confidence interval of mean	Probability of difference
Low frequency baseline (BL)	1	TP	294	1.03	0.40-1.65	<0.0001
		FP	436	4.01	3.50-4.53	
	2	TP	293	1.03	0.42-1.65	<0.0001
		FP	422	4.21	3.70-4.72	
Muscle frequencies (EMG)	1	TP	294	4.04	3.51-4.56	<0.0001
		FP	436	5.52	5.09-5.95	
	2	TP	293	3.80	3.27-4.33	<0.0001
		FP	422	5.52	5.08-5.96	
High frequencies (HF)	1	TP	294	2.80	2.18-3.42	<0.0001
		FP	436	4.44	3.94-4.95	
	2	TP	293	2.33	1.95-2.70	<0.0001
		FP	422	3.81	3.49-4.12	

Table 1. Comparison of ECG signal frequency content between true positive and false positive alarms.

4.3 Which frequency bands give better separation of false and true alarms?

To enable the effectiveness of each frequency band analysis to be compared, the data for the mean of both channels is shown in Figure 2. The greatest separation of TP and FP, was for BL, with a mean value for FPs 4.0 times greater than for TPs. The corresponding values for EMG and HF were 1.6 and 1.4 respectively.

Figure 2. Comparison of the separation between true positive (TP, ○) alarms and false positive (FP, □) alarms for each of the three frequency bands (BL, EMG and HF). The mean values for both channels are indicated, with 95% confidence intervals of the means. The best separation was for BL with the contribution to the BL band for FP alarms 4.0 times greater than for TP alarms.

4.4. Comparison of ECG channels

There was no significant difference between the two channels for BL ($p=0.74$) and EMG ($p=0.60$). There was however a difference for HF ($p<0.001$). This is perhaps not surprising as any patient movement is likely to result in all electrodes moving, inducing baseline (BL) instability and muscle (EMG) noise seen in both channels, but the HF content can relate to poor electrode contact with one electrode and not the other.

4.5. Correlations between frequency bands

Correlation between the frequency band results were all significant ($p<0.001$). However, there was less association between BL and the other two bands, than between EMG and HF (r^2 BL v EMG 31%, BL v HF 29%, EMG v HF 74%). The correlation between EMG and HF is shown in Figure 3.

Figure 3. Correlation between EMG and HF content for both false positive and true positive alarms ($r^2=0.74$, $p<0.001$).

4.6 Results for different arrhythmia alarms

The contribution to BL baseline instability was assessed separately for the five different arrhythmia alarms for channel 1; there being no overall difference between channels. Results are shown in Figure 4. There were very significant differences between the five clinical arrhythmia alarm groups ($p<0.0001$). The separation between the mean frequency content for TPs and FPs was significant for asystole and ventricular tachycardia ($p<0.001$). Unsurprisingly, with long gaps in the signal, the presence of these alarms was predominantly associated with false asystole alarms, with TP alarms having particularly small low frequency BL baseline content.

Figure 4. Comparison of the separation in mean frequency content between true positive (TP, $\circ$) alarms and false positive (FP, $\square$) alarms for each of the five arrhythmia alarms. The mean contributions over the BL baseline band is indicated, with 95% confidence intervals of the means. The arrhythmia groups are Asys (asystole), Brady (bradycardia), VFlut (ventricular flutter), Tachy (tachycardia), VTachy (ventricular tachycardia). The separation was significant for Asys and VTachy ($p < 0.001$).

4.7. Causes of poor ECG signal quality

Typical ECG examples for combinations of good and poor quality inducing false positive alarms are shown for BL, EMG and HF in Figure 5. Quality was based in these examples as being below or above the lower and upper 95% confidence intervals for good quality and poor quality respectively. The upper example shows poor baseline stability, the next example shows significant higher frequencies (both EMG and HF) noise, followed by an example with both low frequency and high frequency noise. The final example is for very high noise levels in all frequency bands.

Figure 5. Examples of ECGs from false positive alarms. Good quality is chosen with a frequency band contribution below the lower limit of the 95% confidence interval of the mean, and poor quality with a frequency band contribution above the upper limit of the 95% confidence interval of the mean, and very poor quality well beyond the poor level. ECGs are shown for (starting at top): poor BL, good EMG and HF; good BL, poor EMG and HF; poor BL, EMG and HF; very poor BL, EMG and HF.

Figure 6 gives two example for true positive alarms. One example is for a good quality signal and one for poor quality, with these conditions applying to all three frequency bands. It can be noted that in the latter case the monitor was still able to identify the correct ventricular tachycardia condition.

Figure 6. Example of a true positive alarm with asystole for a good quality ECG in all three frequency bands; and one with tachycardia for a poor quality ECG in all three bands, indicating that it is possible for device algorithms to identify alarm conditions correctly, even in the presence of noise. Definitions of good and poor quality are as in Figure 5.

5. Discussion

It is universally accepted that the level of false alarms in clinical intensive care areas and other high dependency wards is far too high (Aboukhalil 2008, Clifford et al 2015). This has the negative effect of distracting clinical staff from the care they hope to give, and can result in alarms being ignored, even when there is the potential that they might be genuine. Manufacturers continually strive to improve the accuracy of alarm recognition, and the aim of the PhysioNet challenge was to provide data and give an impetus for researchers to team up to improve alarm detection algorithms. Clifford et al (2015) have provided a summary paper of results so far, and some encouraging results have been obtained. Nevertheless even a high accuracy in detecting true alarms could still leave significant numbers of false alarms.

In this initial study only frequency contributions were assessed, and we acknowledge that additional time domain techniques could improve the accuracy of alarm detection. In addition, we note that the limits of the frequency bands were set to simple values, and that results might be improved with adjustment of these values.

With the knowledge that many false alarms are the result of poor electrode contact, inappropriate siting of electrodes over muscle, and electrode cables poorly positioned resulting in electrode movement artifact, we planned this study to investigate how poor ECG signal quality relates to false alarms. The results showed that there was a very significant relationship between poor quality and false alarms. This was particularly so for low frequency baseline disturbances, which were four times higher with false alarms than with true alarms.

6. Conclusion

Poor quality of ECG signals is strongly related to false positive alarms. Detection of these poor quality conditions should provide an opportunity for improving ECG signal quality and hence reducing false alarms.

References

- Aboukhalil A, Nielsen L, Saeed M, Mark RG and Clifford GD 2008 Reducing false alarm rates for critical arrhythmias using the arterial blood pressure waveform. *J. Biomedical Informatics* 41 442-51
- Allen J and Murray A 1996 Assessing ECG signal quality on a coronary care unit *Physiological Measurement* 17 249-58
- Cvach M 2012 Monitor alarm fatigue: an integrative review *Biomedical Instrumentation and Technology* 268-77
- Chambrin MC 2001 Alarms in the intensive care unit: how can the number of false alarms be reduced? *Critical Care* 5 184-8
- Di Marco LY, Duan W, Bojarnejad M, Zheng D, King S, Murray A and Langley P 2012 Evaluation of an algorithm based on single-condition decision rules for binary classification of 12-lead ambulatory ECG recording quality *Physiological Measurement* 33 1435-48
- Donchin Y and Seagull FJ 2002 The hostile environment of the intensive care unit *Current Opinion in Critical Care* 8 316-20
- Clifford GD, Silva I, Moody B, Li Q, Kella D, Shahin A, Kooistra T, Perry D and Mark RG 2015 The PhysioNet/Computing in Cardiology Challenge 2015: reducing false arrhythmia alarms in the ICU *Computing in Cardiology* 42 273-6
- Graham KC and Cvach M 2010 Monitor alarm fatigue: standardizing use of physiological monitors and decreasing nuisance alarms *American J. of Critical Care* 19 28-34
- Imhoff M and Kuhls S 2006 Alarm algorithms in critical care monitoring. *Anesth Analg* 102 1525-37
- Langley P, Di Marco LY, King S, Duncan D, Di Maria C 2011 An algorithm for assessment of quality of ECGs acquired via mobile telephones *Computing in Cardiology* 38 281-4
- Murray A 1982 Coronary care unit ECG monitoring *Journal of Medical Engineering & Technology* 6 53-61
- Tsimenidis C and Murray A 2015 Reliability of clinical alarm detection in intensive care units 2015 *Computing in Cardiology* 42 1185-8