Patient-Specific Biomedical Condition Monitoring in Post-operative Cancer Patients

David A. Clifton\textsuperscript{1}, Lei A. Clifton\textsuperscript{1,2}, Lionel Tarassenko\textsuperscript{1}, Peter J. Watkinson\textsuperscript{2}, Vicki S. Barber\textsuperscript{2}, and Jonathan Salmon\textsuperscript{2}

\textsuperscript{1}Institute of Biomedical Engineering
Department of Engineering Science, University of Oxford
Old Road Campus, Roosevelt Drive, Oxford, OX3 7DQ, UK
davidc@robots.ox.ac.uk

\textsuperscript{2}Nuffield Department of Anaesthetics, University of Oxford,
John Radcliffe Hospital, Oxford, OX3 9DU, UK

Abstract

Large numbers of preventable deaths occur in hospitals each year, due to adverse events such as cardiac arrest and unplanned admission into Intensive Care Units (ICUs) from other hospital wards. The majority of these patients exhibit physiological deterioration in their vital signs prior to onset of the adverse event, which can be detected by condition monitoring. This paper describes a multivariate, multimodal approach to condition monitoring that may be performed in real-time, which has been previously shown to provide early warning of patient deterioration, while generating a small number of false alarms. We describe a clinical trial currently being undertaken in which post-operative cancer patients are monitored in bed for the first day of their recovery period, and then monitored using telemetry for the remainder of their stay in hospital, during which they may be ambulatory. The use of such telemetry requires monitoring techniques that are robust in the presence of signal artefact introduced by patient movement. We also motivate the use of a patient-specific approach to condition monitoring for improved identification of physiological deterioration.

1 Introduction

1.1 Avoidable Deaths in Hospitals

Every year, large numbers of preventable deaths occur in hospitals. It has been estimated that 23,000 cardiac arrests that occur in UK hospitals could be avoided with improved patient care \cite{1}. Over 20,000 unforeseen admissions to ICUs occur in the UK each year \cite{2}, a patient group for which the mortality rate is approximately 50\% \cite{3}. The majority of these adverse events are preceded by physiological abnormalities evident in vital-sign data \cite{4}, and it has been estimated that such events could be avoided by early identification of patient deterioration \cite{5}. 
1.2 Conventional Patient Monitoring

Patients in UK NHS hospitals are categorised as follows:

- **Level 3:** ICU patients, who may have suffered failure of at least two organ systems, and who may require constant respiratory support;

- **Level 2:** Patients in “step-down” wards, including the Emergency Department, who require frequent observation, and who may have suffered failure of a single organ system;

- **Level 1:** Patients in “acute” wards, who may have been “stepped down” from higher levels of care, and who can be managed with occasional intervention from critical care teams; and

- **Level 0:** Patients in general wards, for whom “normal ward care” is deemed sufficient.

Conventional methods for monitoring the condition of patients in levels 2 and 3 involve continuous acquisition of vital-sign data, which are then used to trigger alarms if any single parameter exceeds a fixed threshold. Such univariate approaches typically suffer from large numbers of false alarms, and a clinical study has estimated that 86% of alarms from conventional patient monitors are false-positive [6]. Typically, this results in clinical staff being trained to ignore alarms generated by patient monitoring equipment.

Patient vital signs are observed by clinical staff at periodic intervals. In level 1 wards, the typical nurse:patient ratio is 1:1, but this declines to 1:4 in level 2 wards, and to 1:10 in level 0 wards. Outside the ICU, observations of patient vital-signs are typically made every 4 to 6 hours [7]. These lower nurse:patient ratios, and infrequent patient observations, can lead to increased numbers of adverse patient outcomes.

Furthermore, conventional patient monitoring systems are used only for those patients confined to beds, in levels 2 and 3. Patients in levels 0 and 1 can be ambulatory, and so no continuous monitoring of physiological data is available.

Thus, there is a great need for patient monitoring systems that are (i) continuous, (ii) applicable to both patients confined to bed and those who are ambulatory, (iii) capable of providing early warning of patient deterioration, and (iv) sufficiently robust such that low numbers of false alarms are generated.

In section 2, we describe an approach to patient monitoring that aims to satisfy these needs, and section 3 describes an ongoing clinical trial for the evaluation of that technology. Section 4 discusses extensions to our approach such that the needs identified above may be met, and conclusions are drawn in section 5.

2 Condition Monitoring with Novelty Detection

We have previously proposed [8, 9, 10] an approach to patient monitoring that is based on novelty detection, in which a multivariate, multimodal model of the

...
distribution of vital-sign data from “normal” high-risk patients is constructed. Multivariate test data $\mathbf{x}$ are then compared with this model to give a novelty score $z(\mathbf{x})$, and an alert is generated when the novelty score exceeds some threshold $z(\mathbf{x}) > \kappa$.

### 2.1 Data Acquisition and Pre-processing

Previously, this system has been applied to the monitoring of patients confined to beds in level 2 wards [10, 11], in which vital-sign data were acquired from existing bed-side patient monitors. These multivariate data were comprised of heart rate (HR), breathing rate (BR), peripheral oxygen saturation (SpO$_2$), skin temperature (Temp), and the systolic-diastolic average blood pressure (SDA), which is the mean average of systolic and diastolic blood pressure.

A training set of “normal” data $\mathbf{X} \in \mathbb{R}^D$, with dimensionality $D = 5$, was obtained from a large number of high-risk patients who had stable physiological conditions. Each dimension $d = 1 \ldots D$ of $\mathbf{X}_d$ was normalised using a zero-mean, unit-variance transformation, $(\mathbf{X}_d - \mu_d) \sigma_d^{-1}$, where $\mu_d, \sigma_d$ were the mean and standard deviation of $\mathbf{X}_d$, respectively. Thus, each dimension of $\mathbf{X}$ varied over approximately similar ranges.

### 2.2 Constructing a Model of Normality

In existing work, the training set $\mathbf{X}$ consisted of approximately $2.3 \times 10^6$ data, obtained from over 3,000 hours of patient recording [10]. $k$-means clustering [12] was used to reduce this large dataset to $k = 500$ prototype vectors, $\mathbf{X}'$, where the 100 prototype vectors furthest from the centroid of $\mathbf{X}'$ were discarded in order to ensure that only “normal” prototype vectors were used for construction of a model of normality.

The Parzen window method [12] was used to estimate the probability density function $p(\mathbf{x})$ of the remaining $N = 400$ prototype vectors, $\mathbf{x}_1, \ldots, \mathbf{x}_N$,

$$p(\mathbf{x}) = \frac{1}{N(2\pi)^{D/2}\sigma^D} \sum_{i=1}^{N} \exp\left(-\frac{|\mathbf{x}-\mathbf{x}_i|^2}{2\sigma^2}\right)$$  \hspace{1cm} (1)

which is a weighted sum of Gaussian kernels centred at the 400 prototype vectors, $\mathbf{x}_i$, and where each kernel has variance $\sigma^2$. This variance was estimated using the nearest-neighbour method proposed in [13].

A novelty score $z(\mathbf{x})$ was defined,

$$z(\mathbf{x}) = -\ln p(\mathbf{x})$$  \hspace{1cm} (2)

such that “normal” data, which have higher probability density values $p(\mathbf{x})$, take low novelty scores, and “abnormal” data, which have lower probability density values, take high novelty scores. An example of this is shown in figure 1, which shows the distribution $p(\mathbf{x})$ for $D = 2$ dimensions, $\{\text{HR,BR}\}$. Several modes in the estimate of $p(\mathbf{x})$ may be seen, and a positive correlation can be seen between HR and BR.

A novelty threshold $z(\mathbf{x}) = \kappa$ was set such that test data $\mathbf{x}$ were classified “abnormal” if $z(\mathbf{x}) > \kappa$, and an alert was raised when $z(\mathbf{x}) > \kappa$ for at least 3
Figure 1: Probability density function $p(x)$ for $D = 2$ dimensions of example training data, showing HR and BR, measured in beats per minute (BPM) and respirations per minute (RPM), respectively. A novelty threshold $z(x) = \kappa$ is shown by the red line.

minutes in any 5-minute window of data. The value of $\kappa$ was selected such that the maximum number of data corresponding to “abnormal” events in a large dataset of 1,000 patients were correctly classified, while false-alerts were minimised. The figure shows a novelty threshold set in the tails of the $p(x)$ distribution. This would, for example, cause an observation of HR = 120, BR = 10 to be classified “abnormal” with respect to the model of normality, because the point (120, 10) lies outside the novelty threshold in the figure.

This system was used for monitoring patients in a clinical trial at a 24-bed level 2 ward at the University of Pittsburgh Medical Centre in the USA [11]. During Phase I of the trial, in which over 18,000 hours of data were recorded from 336 patients, the system provided an average of 6.33 hours of early warning of adverse physiological events, with a false-alarm rate of 1 every 2.6 patient-days.

3 Post-operative Cancer Patients

The investigation described by this paper seeks to apply a similarly principled monitoring approach to patients who are recovering from surgery for the removal of upper gastrointestinal cancer. These patients are confined to a hospital bed for the first day of their recovery, equivalent to the level of care in a level 2 ward. For the remainder of their recovery period (expected to be approximately 4 to 5 days), patients are ambulatory, equivalent to those in a level 1 ward.

Figure 2 shows the monitoring infrastructure of the clinical trial, in which bed-confined and ambulatory patients are monitored in parallel. While the former are monitored, as described in section 2, using vital-sign data acquired by conventional
Figure 2: Infrastructure for a post-operative clinical trial, in which patients confined to beds are connected to a conventional vital-sign monitor (denoted “Mn”), which in turn are wired to a patient monitor on which the system described in section 2 is implemented. A population of ambulatory patients (denoted “An”) have vital signs measured using mobile sensors, which are connected wirelessly to the patient monitors. Solid lines denote wired connections and dashed lines denote wireless connections.

Figure 3: Novelty scores for all bed-confined and ambulatory patients are reported on a centrally-located display during phase II of the trial.
bed-side monitors, the latter are monitored using mobile devices capable of measuring the subset of parameters \{HR, BR, SpO\textsubscript{2}\}. These mobile devices connect wirelessly to the patient monitors on which the system described in section 2 is implemented.

In order to examine the ability of the novelty detection system to identify patient deterioration, a two-phase methodology is adopted:

- **Phase I**: patients are monitored using the infrastructure shown in figure 2, but the displays of the systems are blanked such that clinical staff cannot see the novelty scores \(z(x)\) associated with each patient. The standard level of care is provided to patients.

- **Phase II**: patients are monitored as in phase I, but novelty scores \(z(x)\) are displayed to the clinical staff, as shown in figure 3. Staff act on the novelty scores and any alarms that are generated by the system, rather than using standard procedures.

This allows a “baseline” period of data to be acquired during phase I, against which the outcome from phase II may be compared in order to determine the effect of introducing the novelty detection system.

However, the use of ambulatory data in this way requires several extensions to the approach, as does the application of the technique to a patient group for whom “normal” vital signs may differ to those used for construction of the “normal” model in the existing system.

### 4 Extensions to Existing Methods

#### 4.1 Patient Heterogeneity

The patient population used to construct the model of normality described in section 2 consisted of both medical and surgical “high-risk” patients. However, the patient population in the trial described by this paper will all have had similar invasive upper gastrointestinal surgery, and will be in the process of recovery from that surgery. This may lead to the distribution \(p(x)\) of vital signs for the target patient group being significantly different to that of the training data used in the development of the existing system. In order to achieve successful identification of physiological deterioration for the target patient group, models of normality may need to be constructed using the new data acquired in phase I of the trial.

Furthermore, the use of ambulatory data, which is 3-dimensional \{HR, BR, SpO\textsubscript{2}\} as described in section 3, requires modification of a model trained using 5-dimensional data \{HR, BR, SpO\textsubscript{2}, Temp, SDA\}.

Previous methods for changing the dimensionality \(D\) of models of normality have been considered in [10], in which model dimensionality was reduced during periods in which channels of data ceased to be acquired. For example, in the absence of the acquisition of SpO\textsubscript{2}, perhaps due to a measurement probe becoming disconnected from the patient, the original 5-dimensional model was reduced to 4 dimensions, omitting SpO\textsubscript{2}. This was performed by marginalising out the dimension corresponding to the
missing data in equation (1), which was shown to be equivalent to using \( D = 4 \) in (1), instead of \( D = 5 \), in the case of our example.

In order for novelty detection to take place using the 3-dimensional data acquired from ambulatory patients in the trial described by this paper, it may be necessary to marginalise out \{Temp, SDA\} from a 5-dimensional model constructed using data from phase I using techniques similar to those described in [10].

### 4.2 Patient-Specific Models of Normality

The approach taken in existing work, and that described in the previous subsection, uses a population-generic model of normality, constructed using “normal” data from a training population of patients. However, previous work in novelty detection of other safety-critical applications [14, 15, 16, 17] has shown that early warning of deterioration can be improved, and the number of false-positive alarms reduced, using models of normality that are specific to the system-under-test. In the case of the clinical trial described by this paper, this would be a patient-specific approach, in which models of normality are constructed using data acquired from the patient being monitored. Furthermore, it is an on-line learning problem, in which a model of normality must be constructed during the monitoring of a patient.

Previous work involving novelty detection in jet engines [16] has taken a Bayesian approach, in which data from the engine-under-test are initially compared against a generic model, constructed using prior information obtained from a representative training population of engines from the same class. As data are acquired from the engine-under-test, the model of normality is updated using those data, and novelty detection thus tends towards a system-specific approach as the model of normality tends towards the distribution \( p(x) \) of data from the monitored engine. It is likely that such an approach can be similarly applied to the condition monitoring of patients in the clinical trial described by this paper.

### 4.3 Artefactual Data

The use of ambulatory data introduces the potential for significant signal artefact caused by patient movement. Existing work in the bed-side monitoring of patients in level 2 [10] and level 3 [18] has used Kalman filtering [12] to estimate the variance of measurement and process noise in vital-sign data, such that artefact may be identified and replaced with estimated values. The approach of [18, 19] extends this technique to perform fusion of estimates of derived parameters, such as HR (which is derived from chest-mounted electrocardiogram electrodes and finger-mounted pulse oximeters) and BR (which is derived from frequency and amplitude modulation of the electrocardiogram, and from the photoplethysmograph acquired by pulse oximeters). In this work, the use of multiple signal sources allows signal artefact to be identified, and the final parameter estimate produced using a weighted combination of the trusted, non-artefactual signals.
5 Discussion

This paper has described methods for novelty detection in patient vital-sign data such that patient deterioration can be identified in order to avoid adverse events. We have described how such techniques will be applied to monitoring patients in a clinical trial where those patients are initially confined to beds. However, patients will then be monitored using telemetry during the remaining ambulatory phase of their recovery period. We have described the infrastructure required for this trial, which spans levels 1 and 2 of the conventional patient-criticality categorisation. The use of mobile acquisition technology introduces further problems that must be overcome, such as signal artefact and varying model dimensionality, and we have introduced methods by which such problems may be overcome.

Future work will concentrate on the refinement of these techniques for the target population group, and in the improvement of model construction, allowing patient-specific monitoring to take place.

Acknowledgements

This research project is supported by the NIHR Biomedical Research Centre, Oxford. The authors wish to thank Oxford BioSignals Ltd., and Marc Smith in particular, for technical support of this project, and to Sarah Vollam and Deborah Evans for clinical support.

References


