

Modelling physiological deterioration in post-operative patient vital-sign data

Marco A. F. Pimentel · David A. Clifton ·
Lei Clifton · Peter J. Watkinson · Lionel Tarassenko

Received: 20 March 2012 / Accepted: 5 March 2013 / Published online: 21 March 2013
© The Author(s) 2013. This article is published with open access at Springerlink.com

Abstract Patients who undergo upper-gastrointestinal surgery have a high incidence of post-operative complications, often requiring admission to the intensive care unit several days after surgery. A dataset comprising observational vital-sign data from 171 post-operative patients taking part in a two-phase clinical trial at the Oxford Cancer Centre, was used to explore the trajectory of patients' vital-sign changes during their stay in the post-operative ward using both univariate and multivariate analyses. A model of normality based vital-sign data from patients who had a "normal" recovery was constructed using a kernel density estimate, and tested with "abnormal" data from patients who deteriorated sufficiently to be re-admitted to the intensive care unit. The vital-sign distributions from "normal" patients were found to vary over time from admission to the post-operative ward to their discharge home, but no significant changes in their distributions were observed from halfway through their stay on the ward to the time of discharge. The model of normality identified patient deterioration when tested with unseen "abnormal" data, suggesting that such techniques may be used to provide early warning of adverse physiological events.

Keywords Patient monitoring · Early warning scores · Novelty detection

1 Introduction

Delayed detection of clinical deterioration has been repeatedly associated with high rates of avoidable in-hospital death and intensive care unit (ICU) readmissions (which are associated with a substantially increased mortality rate) [8, 11, 15]. According to large national surgical audits such as the UK National Confidential Enquiry into post-operative deaths, current systems of post-operative care fail to detect or respond appropriately to early signs of critical illness [17]. Such failures have been explained by lack of experienced senior nursing staff, inexperienced trainee medical staff [17], poor quality of care offered to critically ill patients [6, 15], and, more importantly, the inability of current systems to recognise clinical deterioration early. All of these factors can lead to deterioration in a patient's condition and admission to the ICU, or death.

The UK National Institute for Health and Clinical Excellence (NICE) [16] has recommended that physiological track and trigger (T&T) systems should be used to monitor all adult patients in acute hospital units, to promote the recognition of patient deterioration early enough to allow proper intervention by medical staff. These systems are based on early warning scores (EWS) calculated from the values of physiological variables observed periodically. Univariate scoring criteria are applied to each physiological variable (vital sign) in turn, and then care is escalated to a higher level if any of the scores assigned to individual vital signs, or the sum of all such scores, exceed some threshold. There is widespread interest and clinical utilisation of these scores in countries across Europe and

M. A. F. Pimentel (✉) · D. A. Clifton · L. Clifton ·
L. Tarassenko
Institute of Biomedical Engineering, Department of Engineering
Science, University of Oxford, Old Road Campus Research
Building (Off Roosevelt Drive), Oxford OX3 7DQ, UK
e-mail: marco.pimentel@eng.ox.ac.uk

P. J. Watkinson
Nuffield Department of Clinical Neurosciences, Oxford
University Hospitals NHS Trust, Oxford OX3 9DU, UK

Australasia, and increasingly in North America [5]. However, the quality of evidence supporting the use of T&T systems is poor [5], and they have a number of disadvantages. The thresholds and ranges of these EWS systems are mostly determined heuristically (although evidence-based methods have recently been proposed [18, 23]). Furthermore, each vital sign is treated independently and correlations between them are not taken into account. Also, the clinical setting from which data are acquired for either validating or designing the EWS system is an important consideration. Many studies have been conducted in medical assessment units [5, 18], and it is questionable whether the scores can be extrapolated to other medical units; for example, post-operative wards, general wards, or other settings.

An alternative approach to detecting patient deterioration from changes in vital signs is that of novelty detection [2, 20], or one-class classification, which involves the construction of a multivariate, multimodal model of normality using examples of “normal” vital signs. This then allows the classification of test data as either “normal” or “abnormal” with respect to that model. Several approaches to novelty detection have been proposed, and an extensive review of these techniques is presented in [13, 14]. We have shown how novelty detection can be combined with continuous vital-sign monitoring of acutely ill in-hospital patients [7, 9, 10, 21].

In this paper, we investigate models of normality tuned to a specific post-operative patient population, recovering from gastro-intestinal surgery. Following surgery, patients start in their most acute state and gradually stabilise. We hypothesise that models of the distribution of vital-sign data from “normal” patients may be used to describe the physiological trajectory associated with “normal” recovery of these patients. These models may then be used to identify “abnormal” trajectories in patients who experience major deterioration and have to be re-admitted to the ICU.

2 Methods

2.1 Dataset

Vital-sign data (heart rate, HR, measured in beats per minute; respiratory rate, RR, measured in breaths per minute; arterial blood oxygen saturation, SpO₂, measured as a percentage; systolic blood pressure, SysBP, measured in mmHg; core temperature measured with a tympanic thermometer in °C; and a level of consciousness assessed typically with the Glasgow Coma Scale,¹ GCS) were recorded by nursing staff during their regular observations of post-operative patients in the upper gastrointestinal (GI)

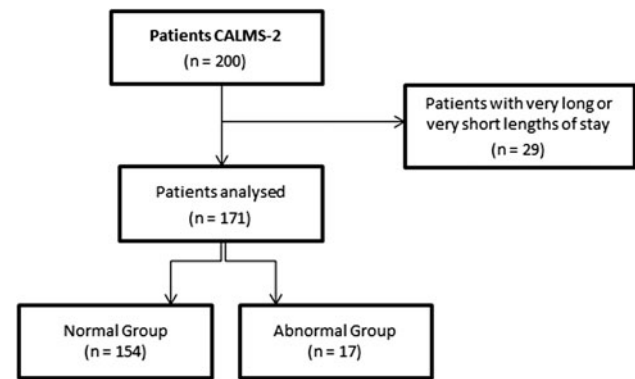


Fig. 1 Flow diagram showing the steps involved in creating the “normal” and “abnormal” patient datasets, including the number of patients included in the study

ward at the Oxford Cancer Centre, Oxford University Hospitals NHS Trust, Oxford, UK. The dataset used for the work described by this paper comprises measurements of HR, RR, SpO₂, SysBP and temperature (dimensionality of the input space, $D = 5$) acquired by ward staff every hour or every 2 h in the days immediately following the patient admission to the ward (depending on patient’s condition), and every 4 h in the last few days of the patient’s stay on the ward. These measurements were then transcribed by two independent research nurses into an electronic database.

200 patients were recruited during Phase I of the CALMS2 clinical trial in the upper GI ward (approved by the local research ethics committee, REC reference: 08/H0607/79). This dataset was firstly refined to include only observations with no missing physiological variables (for example, if an observation from a patient does not include HR, it was removed from the dataset). The median length of stay on the ward was 9 days, and we selected those patients who stayed on the ward for a minimum of 4 days (which corresponds to the 10th percentile) and a maximum of 29 days (90th percentile), to construct our model of normality, for the purposes of novelty detection. This reduces the number of patients from 200 to 171 (see Fig. 1).

From the original dataset, a set of 12,797 observations $\mathbf{X} \in \mathbb{R}^5$ obtained from the 171 patients was then analysed. From the patients analysed, those who were either admitted to the ICU or died on the ward were labelled as belonging to the “abnormal group” of patients (17 patients), while the remainder were labelled as being part of the “normal group” (154 patients). The patient population characteristics in each group are shown in Table 1. We note that the mortality rate in the “abnormal” set of patients was

¹ The Glasgow Coma Scale is a scale from 3 to 15, in which 15 indicates “alert” consciousness and 3 indicates complete unresponsiveness.

Table 1 Patient demographics for the “normal” and “abnormal” groups of patients

	Normal group	Abnormal group
Number of patients	154	17
Number of observations	10,299	2,498
Age (mean \pm SD)	61 \pm 12	67 \pm 10
Sex (male)	90 (58.4 %)	10 (58.8 %)
Length of stay (median \pm IQR)	9 \pm 5	5 \pm 4 ¹
Length of stay (25th/75th percentile)	7/12	4/8 ¹
Mortality	0	6 (35.3 %)

¹ As described in Sect. 3.3, the “length of stay” for the abnormal group is the time to the first event (re-admission to ICU or death)

35.3 %, which shows the severity of the risk associated with ICU re-admission.

2.2 Vital-sign distributions

The changes in vital-sign distributions between admission to the upper GI ward and subsequent discharge, when the patient was deemed sufficiently stable to go home, were evaluated for the patients in the “normal group”. Normalised histograms (unit area under the curve) and cumulative distribution functions (cdfs) were plotted for each physiological variable (HR, RR, SpO₂, SysBP and temperature), using the average value for each variable on the admission and discharge days.

The trajectory of each vital sign throughout the patient’s stay on the ward was evaluated by examining the following subgroups of observations:

- G₁: the set of averages of all observations performed on the first day of the patient’s stay on the ward (admission day);
- G₂: the set of averages of all observations performed on the day that corresponds to a quarter (25 %) of the length of the patient’s stay on the ward;
- G₃: the set of averages of all observations performed on the day that corresponds to half (50 %) of the length of the patient’s stay on the ward;
- G₄: the set of averages of all observations performed on the day that corresponds to 75 % of the length of the patient’s stay on the ward;
- G₅: the set of averages of all observations performed on the last day of the patient’s stay on the ward (discharge day).

These subgroups were defined in this way because of the different lengths of patient stay on the ward (which varied between 4 and 28 days). Three different metrics were used to compare the resulting vital-sign distributions: the

Kolmogorov–Smirnov (KS) metric [4], the symmetrical Kullback–Leibler (KL) distance [12, 25] and the Bhattacharyya (*Bhat*) distance [1].

The KS distance is a non-parametric metric that quantifies the distance between the empirical distribution functions of two sample sets [4]. Considering two probability densities, p and q , if P and Q are the respective cdfs, the KS distance (ΔKS) between them is defined by

$$\Delta KS(p, q) = \sup(|P(x) - Q(x)|) \quad (1)$$

where $\sup(y)$ is the supremum of the set of distances y .

The KL divergence [12] compares the entropy of two distributions over the same random variable. It measures the number of additional bits required when encoding a random variable with a distribution $p(x)$ using the alternative distribution $q(x)$. This measure is asymmetrical, but it can be modified to be the symmetrical KL distance (ΔKL_S) [25], defined as

$$\Delta KL_S(p, q) = \sum_{x \in X} (p(x) - q(x)) \log \frac{p(x)}{q(x)} \quad (2)$$

The *Bhat* distance ($\Delta Bhat$) [1] measures the amount of overlap between two distributions, and is defined by

$$\Delta Bhat(p, q) = -\log[BC(p, q)] \quad (3)$$

where $BC(p, q) = \sum_{x \in X} \sqrt{p(x)q(x)}$.

In order to study the physiological trajectory of the “normal” patients, the distributions of each vital sign, for each of the first 4 subgroups described (G₁, G₂, G₃, G₄) were compared with G₅ (which contains the average of the vital signs from the most physiologically stable period of the patient stay), using the three metrics defined by (1), (2) and (3).

2.3 Data visualisation

The first stage in constructing a model of normality for novelty detection usually consists of obtaining more insight into the structure of the data [22]. Procedures for visualisation of the data in their original high-dimensional space are therefore required.

Data in high-dimensional space ($D > 3$) can be visualised through a non-linear projection from \mathbb{R}^D to \mathbb{R}^2 . Sammon’s method [19] seeks to create a mapping such that the distances between pairs of image points in a projection plane (y_i, y_j) are as close as possible to the distances between the corresponding pair of points in data space (x_i, x_j). The following error function, known as the Sammon stress metric, is defined as

$$E = \frac{1}{\sum_{i < j} d_{ij}^*} \sum_{i < j}^N \frac{(d_{ij}^* - d_{ij})^2}{d_{ij}^*} \quad (4)$$

with $d_{ij} = \|x_i - x_j\|$ and $d_{ij}^* = \|y_i - y_j\|$, where $\|\bullet\|$ is the Euclidean norm. The Sammon mapping aims to minimise the error metric (4), which can be achieved by initialising the image points y to have random locations in a 2-D map and by iteratively adjusting these locations in the direction which gives the maximum change in E using a gradient descent method.

It is assumed a priori that each of the five vital signs has equal importance in the model of normality. Each variable was therefore scaled to have approximately the same dynamic range to ensure that variables with large changes (e.g., blood pressure in mmHg) do not dominate parameters with smaller changes (e.g., temperature in °C). Every vital-sign measurement, x , was normalised using a zero-mean, unit-variance transformation, $x_n = (x - \mu)/\sigma$, where x_n is the normalised value and μ and σ are the mean and standard deviation of the vital sign, respectively, in the overall dataset (171 patients).

The Sammon mapping algorithm was then applied to the 770 normalised vectors contained in the 5 subgroups (G_1 , G_2 , G_3 , G_4 and G_5) from the 154 patients included in the “normal” subset of patients.

2.4 Model of normality

We now consider the construction of a model of normality, based on all observations made on the last day on the ward (discharge day) of each patient from the “normal group”. This dataset contains the vital signs from the most physiologically stable period of the patient stay, because these data were acquired immediately prior to discharge from the ward, when the patient is at their most “normal” after recovering from surgery. This set of “normal” pre-discharge data contains 1,100 vital-sign vectors, $\mathbf{X} \in \mathbb{R}^5$, which were subsequently used for the construction of our model of normality.

A kernel density estimate [3] is a technique that allows the underlying 5-dimensional vital-sign pdf to be estimated from training data. A kernel density estimate was chosen because it is a non-parametric method, so no a priori assumptions are made about the form of the probability distribution.

The pdf of the set of $N = 1,100$ “normal” vectors, $\mathbf{x}_1, \dots, \mathbf{x}_N$, was estimated using the following equation:

$$p(\mathbf{x}|\mathbf{x}_i, \sigma^2) = \frac{1}{N(2\pi)^{D/2}\sigma^D} \sum_{i=1}^N e^{-\frac{\|\mathbf{x}-\mathbf{x}_i\|^2}{2\sigma^2}} \quad (5)$$

which is a weighted sum of Gaussian kernels centred on the 1,100 vectors, \mathbf{x}_i , and where each kernel is isotropic with variance σ^2 . The variance was determined using the nearest-neighbour method proposed by Bishop [2], in which the average of the squared Euclidean distance to the

set of 10 nearest neighbours $\{NNs\}$ is determined for each point $\mathbf{x}_1, \dots, \mathbf{x}_N$ in \mathbf{X} ,

$$\Delta_i = \frac{1}{10} \sum_{j \in \{NNs\}} \|\mathbf{x}_i - \mathbf{x}_j\| \quad (6)$$

and σ^2 is estimated by calculating the average over all points:

$$\sigma = \frac{1}{N} \sum_{i=1}^N \Delta_i \quad (7)$$

The likelihood for all data from the “normal” group of patients was then calculated using (5). The likelihood of all data from the “abnormal” group of patients, prior to the occurrence of an adverse event (either death or ICU admission) was also evaluated using the same model of normality.

In order to estimate the “abnormality” of a data point \mathbf{x} , the departure from normality is usually quantified using a novelty score defined as follows

$$z(\mathbf{x}) = -\log p(\mathbf{x}|\theta) \quad (8)$$

where $z(\mathbf{x})$ is the novelty score and $\theta = \{\mathbf{x}_i, \sigma\}$. “Normal” data, which have higher likelihoods $p(\mathbf{x}|\theta)$, therefore generate low novelty scores $z(\mathbf{x})$; conversely, “abnormal” data, which have lower likelihoods, generate high novelty scores $z(\mathbf{x})$.

3 Results

3.1 Vital-sign distributions

Empirical pdfs (histograms) and cdfs for each physiological variable for each of the two subgroups G_1 (average of observations at admission) and G_5 (average of observations at discharge day) are shown in Fig. 2. Table 2 gives the corresponding means and standard deviations.

Figure 3 shows the KS, KL and *Bhat* distance-maps between each of the distributions, for the 4 subgroups (G_1 , G_2 , G_3 , and G_4) and the distribution for the G_5 subgroup. In each distance-map, the subgroups involved (G_5 — G_i with $i = \{1, 2, 3, 4\}$) are represented on the x -axis, and physiological variables are represented on the y -axis. The colour code is associated with the values of the calculated distances (blue indicates small distances, and red indicates large distances).

It may be seen that, apart from the HR, the distributions represented for each of the other 4 vital signs vary from admission to discharge, as the patient recovers from surgery.

The results obtained by the three different metrics are very similar, in the sense that the patterns in the distances

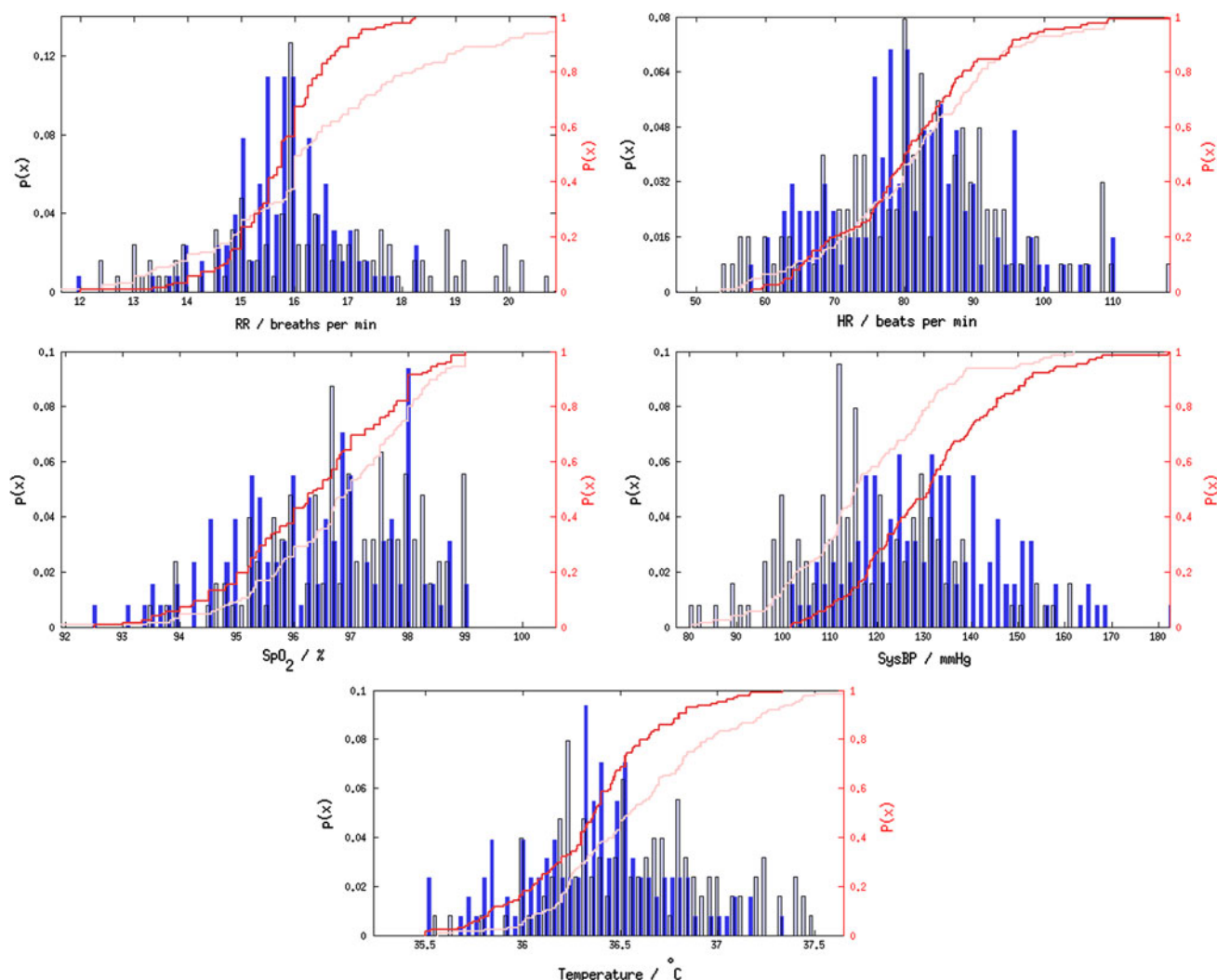


Fig. 2 Histograms for respiratory rate, heart rate, blood oxygen saturation and temperature, computed from the average of vital-sign data acquired from patients at admission to the post-operative ward (light blue) and time near discharge (dark blue). Cumulative

distribution functions $P(x)$ for each vital sign from patients at admission (light red) and time near discharge (dark red), are also represented (refer to the right vertical axis) (colour figure online)

Table 2 Vital-sign means (SD) for admission and discharge days for “normal” patients

	RR (breaths per minute)	HR (beats per minute)	SpO ₂ (%)	SysBP (mmHg)	Temp. (°C)
Admission	16.7 (2.5)	80.6 (12.8)	97.0 (1.3)	115.7 (16.0)	36.6 (0.5)
Discharge	15.7 (1.0)	81.2 (11.7)	96.3 (1.5)	132.1 (16.6)	36.4 (0.4)

for each physiological variable are identical. The distances between the G_1 and G_5 distributions are greater than the distances between the G_3 and G_5 distributions.

3.2 Data visualisation

The resulting Sammon maps obtained are shown in Fig. 4. Represented in each map are the projected data points from

G_1 , G_2 , G_3 , G_4 (red crosses) superimposed on the projected data points from G_5 (blue points).

3.3 Model of normality

The novelty scores are computed each day for each patient, by averaging the scores for each set of vital-sign observations that day. The group mean novelty scores $z(\mathbf{x})$ for

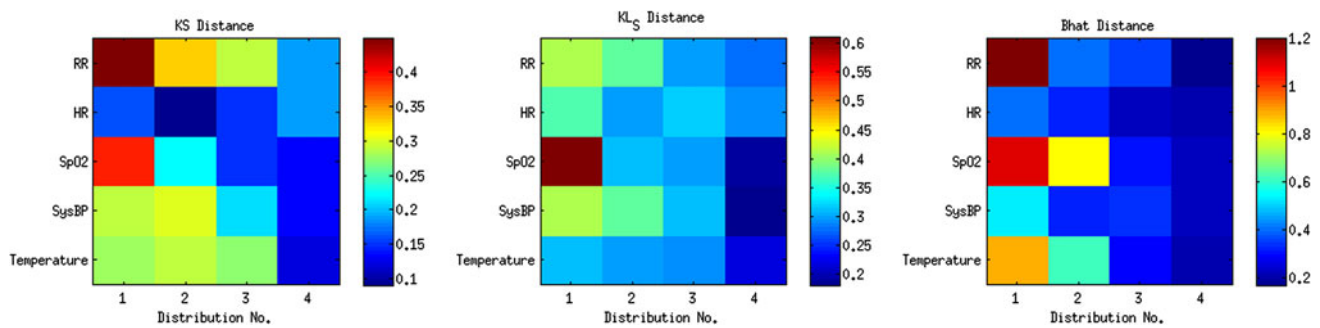
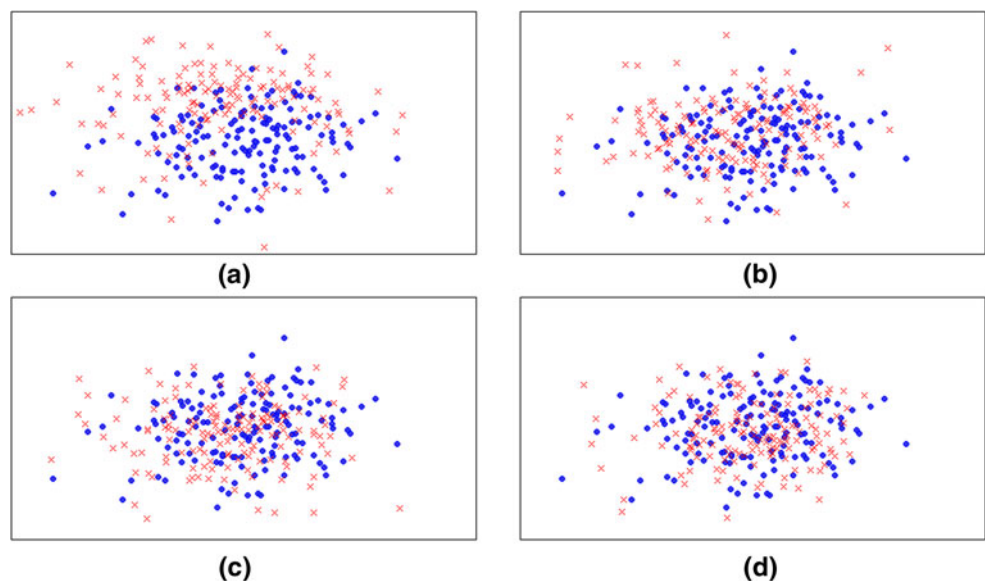


Fig. 3 Representation of the distances between each of the 4 groups (G_1 , G_2 , G_3 and G_4 represented in the horizontal axis as 1, 2, 3 and 4) and the G_5 group. The Kolmogorov–Smirnov (KS), the symmetrical Kullback–Leibler (KL_S) and the Bhattacharyya metrics for each

distribution are shown in the form of matrices, in which the colour is associated with the values of the calculated distances (colour figure online)

Fig. 4 Sammon maps obtained for the groups: **a** G_1 and G_5 , **b** G_2 and G_5 , **c** G_3 and G_5 , and **d** G_4 and G_5 . Projected data from G_5 are shown by blue •; projected data from all other groups are shown by red x (colour figure online)



each day are shown in Fig. 5 for “normal” and “abnormal” patients. The median length of stay on the ward for the “normal” group of patients is 9 (see Table 1). For the “abnormal” group, we considered the length of stay on the ward prior to the event (either admission to the ICU or death). The median time to event for this group is 5 days. The novelty scores are displayed in Fig. 5 for the length of stay (or time to event in the case of the “abnormal” group) up to the 75th percentile (12 and 8 days, respectively) for each group of patients.

Figure 6 shows the change in novelty score over time for two example patients from the “abnormal” group who deteriorate sufficiently after surgery to be re-admitted to ICU. A threshold $z(\mathbf{x}) = k$ was determined using $k = \mu + 3s.d.$, where μ is the average of the density estimates $z(\mathbf{x})$ for the 154 “normal” patients in the model of normality, and where $s.d.$ is one standard deviation of $z(\mathbf{x})$ for these “normal” patients.

The first example (Fig. 6a) shows a patient who deteriorated 2 days after admission to the upper GI ward and was then admitted to the ICU. The patient was sent back to the upper GI ward after 2 days in the ICU. The patient stayed a further 4 days before being discharged. During the first 2 days after surgery, the patient exhibits physiological instability (which is more significant at the end of the second day) showing indications of tachycardia (HR reaching 150 beats per min) and tachypnea (RR reaching almost 40 breaths per min). It can be seen that $z(\mathbf{x})$ increases in value at approximately 12 h before ICU admission, indicating physiological deterioration. After the stabilisation of the patient in the ICU, $z(\mathbf{x})$ remains close to the “normal” trajectory.

The second example (Fig. 6b) shows a patient who had some periods of instability after being admitted to the upper GI ward, following surgery. After 7 days, the patient was re-admitted to the ICU, and then died 1 month later.

In this case, large variations are observed in $z(\mathbf{x})$ during the post-surgical period of abnormality, which are caused by episodes of elevated blood pressure (SysBP at around 190 mmHg) and bradycardia (HR decreasing to 45 beats

per minute). These exceed the threshold defined by $z(\mathbf{x}) = k$.

4 Discussion

The histograms and cdfs shown in Fig. 2 indicate that the HR distributions are similar and approximately symmetrical. The distributions for SpO₂ are one-sided because they are limited to the maximum value SpO₂ = 100 %. For the distribution of SpO₂ values at admission, a mode occurs at SpO₂ = 97 %. Patients are likely to achieve 100 % oxygen saturation only if they are receiving additional oxygen through an oxygen mask. Therefore, the distributions shown in Fig. 2 exclude values of SpO₂ >99 %. RR distributions are similar between admission and discharge. Tympanic temperature and SysBP distributions show that patients are, in general, mildly pyrexial (high temperature) and hypotensive (low systolic blood pressure) when admitted to the ward following surgery. They subsequently show decreasing temperature (back to “normal” values) and increasing blood pressure (back to “normal” values) by the last day of their stay on the ward.

From the distances between the distributions calculated with the three different metrics (Fig. 3), we can easily see the pattern of recovery with time: the distance between the G₁ and G₅ distributions is greater than, for example, the distance between the G₃ and G₅ distributions. If we consider the SysBP, for example, the KS, KL and *Bhat* distances between the G₁ and G₅ distributions are 0.29, 0.41

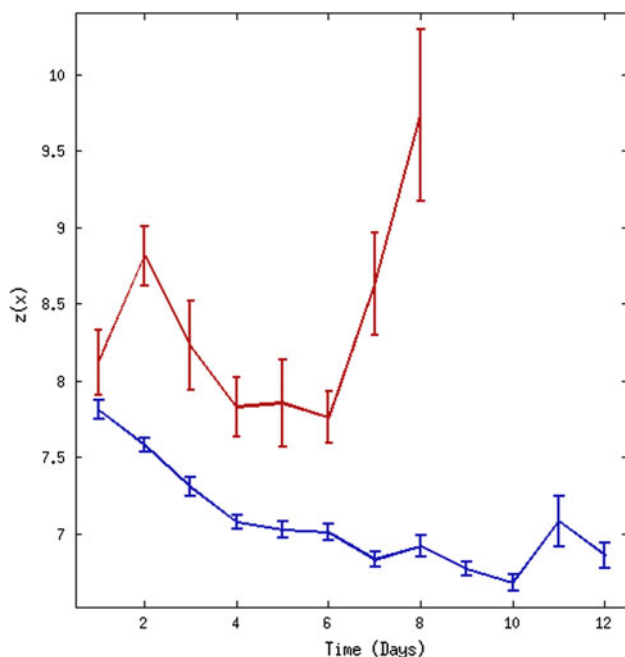


Fig. 5 Representation of average (per day) of novelty scores $z(\mathbf{x})$ against time for the “normal” group of patients shown in dark blue and the “abnormal” group of patients shown in dark red. Error bars denote one SE of the group mean (colour figure online)

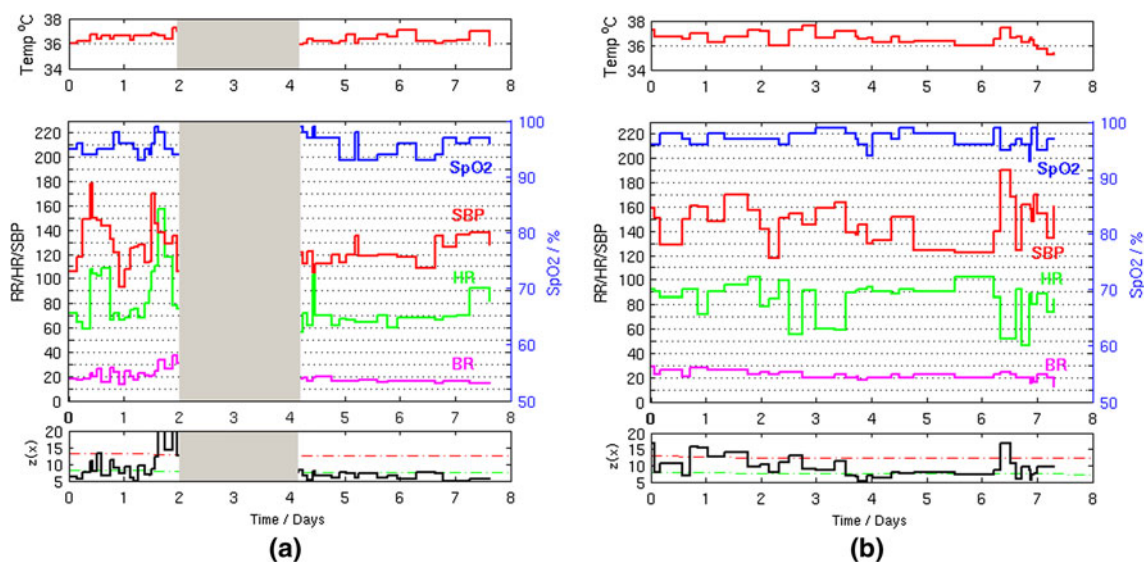


Fig. 6 Two example patients are shown in (a) and (b). The upper plots show the observations of vital signs with time: temperature is shown in the upper plot in each column; RR, HR, SysBP and SpO₂ are shown in the middle plot of each column (refer to the right vertical axis for SpO₂). The lower plot in each column shows the novelty

score $z(\mathbf{x})$ determined using the model of normality: the decision threshold (dashed red line) was determined from the average of the density estimates for “normal” patients (dashed green line). The grey box in (a) indicates the patient’s admission to the ICU, during which period no vital-sign data are available (colour figure online)

and 0.54, respectively, whereas the distances between the G_3 and G_5 distributions are 0.21, 0.31 and 0.34.

The Sammon maps represented in Fig. 4 show that the projected data from the five groups form clusters with some overlap between them, but that there are groups with visually separable distributions. The G_1 cluster is the most diffuse (shown in red, in the upper-left plot in Fig. 4), while the projected data from G_3 , G_4 and G_5 are more concentrated, and similar to each other in their locus in the projection plane. This suggests that there are no large changes in the vital-sign distributions from halfway through a patient's stay to the time of their discharge from the ward. That is, "normal" patients appear to have stabilised at around halfway through their stay on the ward. These results suggest that patients included in the "normal group" could have been considered for earlier discharge, or provided with a lower level of care from halfway through their stay.

From the trajectory of $z(\mathbf{x})$ for the "normal" group of patients (Fig. 5a) we can see a significant decrease in $z(\mathbf{x})$ in the first 4 days, after which $z(\mathbf{x})$ is approximately constant for $t \geq 4$ days. The first 4 days correspond to patient recovery immediately following surgery [24]. After day 4, the majority of patients included in the "normal" group appear to have fully physiologically recovered from surgery and are physiologically stable. It could be argued that these patients are sufficiently stable for early discharge to be considered, or for them to be provided with a lower level of care should they need to remain in hospital for reasons not related to physiological instability. Conversely, $z(\mathbf{x})$ for the "abnormal" group of patients, suggests that the physiological trajectory for these patients is significantly different to that of "normal" patients with a sudden increase in novelty in the last 48 h, following the gradual decrease prior to this. These results suggest that patients' criticality could be assessed by evaluating the distribution of their vital signs using the novelty score of Eq. (8) after their admission to the post-operative upper GI ward, following major surgery.

In summary, this study indicates that multivariate models of normality may be used to assess post-operative patients' criticality. A multivariate model of the distribution of vital-sign data from "normal" patients was constructed using a kernel density estimate, and tested using "abnormal" data from patients who deteriorate sufficiently after surgery to be re-admitted to the ICU. Significant differences were found between the physiological trajectories for "normal" patients and those for "abnormal" patients.

Acknowledgments The work described in this paper was funded by the NIHR Biomedical Research Centre Programme, Oxford. MAFP was supported by the RCUK Digital Economy Programme grant

number EP/G036861/1 (Oxford Centre for Doctoral Training in Healthcare Innovation) and FCT—*Fundação para a Ciência e Tecnologia* under grant SFRH/BD/79799/2011. DAC was supported by the Centre of Excellence in Personalised Healthcare funded by the Wellcome Trust and EPSRC under grant number WT 088877/Z/09/Z, and LC was supported by the NIHR Biomedical Research Centre Programme, Oxford. The authors wish to thank all the clinical staff involved in the collection of the clinical data analysed in this paper.

Open Access This article is distributed under the terms of the Creative Commons Attribution License which permits any use, distribution, and reproduction in any medium, provided the original author(s) and the source are credited.

References

1. Bhattacharyya A (1946) on a measure of divergence between two multinomial populations. *Sankhyā Indian J Stat* 7(4):401–406
2. Bishop CM (1994) Novelty detection and neural network validation. *IEE Proc Vis Image Signal Process* 141(4):217–222
3. Bishop CM (2007) Pattern recognition and machine learning. Springer, Berlin
4. Chakravarti IM, Indra M, Roy J, Laha RG (1967) Handbook of methods of applied statistics. Wiley, New York
5. Gao H, McDonnell A, Harrison DA, Moore T, Adam S, Daly K, Esmonde L, Goldhill DR, Parry GJ, Rashidian A, Subbe CP, Harvey S (2007) Systematic review and evaluation of physiological track and trigger warning systems for identifying at-risk patients on the ward. *Intensive Care Med* 33(4):667–679
6. Garrard C, Young D (1998) Suboptimal care of patients before admission to intensive care. *BMJ* 316(7148):1841–1842
7. Hann A (2008) Multi-parameter monitoring for early warning of patient deterioration. PhD Thesis, University of Oxford, Oxford
8. Hodgetts TJ, Kenward G, Vlackonikolis I, Payne S, Castle N, Crouch R, Ineson N, Shaikh L (2002) Incidence, location and reasons for avoidable in-hospital cardiac arrest in a district general hospital. *Resuscitation* 54(2):115–123
9. Hravnak M, Edwards L, Clontz A, Valenta C, DeVita MA, Pinsky MR (2008) Defining the incidence of cardiorespiratory instability in patients in step-down units using an electronic integrated monitoring system. *Arch Intern Med* 168(12):1300–1308
10. Hravnak M, DeVita MA, Clontz A, Edwards L, Valenta C, Pinsky MR (2011) Cardiorespiratory instability before and after implementing an integrated monitoring system. *Crit Care Med* 39(1):65–72
11. Kause J, Smith G, Prytherch D, Parr M, Flabouris A, Hillman K (2004) A comparison of antecedents to Cardiac arrests, deaths and emergency intensive care admissions in Australia and New Zealand, and the United Kingdom—the ACADEMIA study. *Resuscitation* 62(3):275–282
12. Kullback S, Leibler RA (1951) On information and sufficiency. *Ann Math Stat* 22(1):79–86
13. Markou M (2003) Novelty detection: a review—part 2: neural network based approaches. *Signal Process* 83(12):2499–2521
14. Markou M, Singh S (2003) Novelty detection: a review—part 1: statistical approaches. *Signal Process* 83(12):2481–2497
15. McQuillan P, Pilkington S, Allan A, Taylor B, Short A, Morgan G, Nielsen M, Barrett D, Smith G, Collins CH (1998) Confidential inquiry into quality of care before admission to intensive care. *BMJ* 316(7148):1853–1858
16. National Institute for Clinical Excellence (2007) Acutely ill patients in hospital: recognition of and response to acute illness in adults in hospital. NICE, Guidance/Clinical Guidelines CG50

17. NCEPOD (2005) An acute problem? A report of the National Confidential Enquiry into Patient Outcome and Death (NCEPOD). NCEPOD, London
18. Prytherch DR, Smith GB, Schmidt PE, Featherstone PI (2010) ViEWS—Towards a national early warning score for detecting adult inpatient deterioration. *Resuscitation* 81(8):932–937
19. Sammon JW (1969) A nonlinear mapping for data structure analysis. *IEEE transactions on computers* C-18(5):401–409
20. Tarassenko L, Hayton P, Cerneaz N, Brady M (1995) Novelty detection for the identification of masses in mammograms. Fourth international conference on artificial neural networks, IET 1995 p. 442–447
21. Tarassenko L, Hann A, Young D (2006) Integrated monitoring and analysis for early warning of patient deterioration. *Br J Anaesth* 97(1):64–68
22. Tarassenko L, Clifton DA, Bannister PR, King S, King D (2009) Novelty detection. *Encyclopedia of structural health monitoring*. Wiley, London
23. Tarassenko L, Clifton DA, Pinsky MR, Hravnak MT, Woods JR, Watkinson PJ (2011) Centile-based early warning scores derived from statistical distributions of vital signs. *Resuscitation* 82(8):1013–1018
24. Thompson JS, Baxter BT, Allison JG, Johnson FE, Lee KK, Park WY (2003) Temporal patterns of postoperative complications. *Arch Surg* 138(6):596–603
25. Veldhuis R (2002) The centroid of the symmetrical Kullback–Leibler distance. *IEEE Signal Process Lett* 9(3):96–99