PAPER

Feature clustering of intracranial pressure time series for alarm function estimation in traumatic brain injury

IOP Publishing | Institute of Physics and Engineering in Medicine Physiological Measurement

Physiol. Meas. **38** (2017) 2015–2043 https://doi.org/10.1088/1361-6579/aa8a51

Feature clustering of intracranial pressure time series for alarm function estimation in traumatic brain injury

M Teplan[1], I Bajla[1], R Rosipal[1] and M Rusnak[2]

[1] Department of Theoretical Methods, Institute of Measurement Science, Slovak Academy of Sciences, Bratislava, Slovakia
[2] Department of Public Health, Trnava University, Trnava, Slovakia

E-mail: michal.teplan@savba.sk

Received 15 July 2016, revised 17 August 2017
Accepted for publication 5 September 2017
Published 31 October 2017

Abstract

Objective: The conventional application of intracranial pressure (ICP) monitoring of traumatic brain injury (TBI) patients consists merely in the acquisition of ICP values in discrete time and their comparison to the established ICP threshold. An exceeding of this threshold triggers a special emergency treatment protocol. This paper addresses the possibility of making use of the rich information latent in the ICP records of known vital and fatal outcomes gathered during real clinical practice of treating TBI patients. Our assumption was that the proposed algorithmic procedure derived from this information could, in addition to ICP monitoring itself, provide a complementary added value. This might help clinicians to make better decisions during a patient's treatment. *Approach*: We concentrated on studying specific clustering schemes for subsequences of ICP time series. The clusterization problem was formulated for feature vectors which are introduced to represent ICP time subsequences. The ICP transformation to a feature space uses global and local definitions of time subsequences. For clusterization itself, we adopted hierarchical Gaussian mixture models (hGMMs). By using posterior probabilities of the clusters, we introduced three novel alarm functions. We explored two alternative methods of searching for optimum alarm function thresholds (ROC analysis and a novel efficiency measure). *Main results*: We performed extensive cross-validation experiments on a clinical retrospective data set. The results of the optimization over several hGMMs, various feature space dimensionality and all the types of the novel alarm functions show the potential of the novel alarm functions for supplementing conventional ICP monitoring. *Significance*: In conclusion, the paper provides a prospective extended ICP monitoring technique for real TBI patients, based on the

proposed methodology of ICP subsequence clustering and thresholding of the optimum novel alarm function.

Keywords: intracranial pressure (ICP) measurement, ICP time-series, ICP feature clustering, Gaussian mixture model, alarm functions

S Supplementary material for this article is available online

(Some figures may appear in colour only in the online journal)

1. Introduction

1.1. Overview of some ICP monitoring methods

According to the findings of the United States Centers for Disease Control and Prevention, each year about 1.5 million people sustain traumatic brain injury (TBI). Approximately 800 000 receive early outpatient care in the USA. Of that number, 50 000 patients die and another 85 000 people suffer various long-term disabilities including cognitive deficits. Unfortunately, brain injuries do not heal like injuries to any other part of the body. Before medical knowledge and technology advanced to control breathing using automatic respirators and decrease intracranial pressure (ICP), the death rate from TBI was very high. To reduce the mortality and morbidity rate of patients with severe TBI, much effort has been devoted by physicians and medical organizations to search for a remedy to this situation. To help medical practitioners, several different systems of input diagnosis of the symptoms of TBI have been introduced (Teasdale and Jennett 1964). The most popular and practical application at the admission of patients with TBI at hospitals is achieved by using the so-called Glasgow coma scale (GCS) (Teasdale and Jennett 1964) or the Glasgow outcome scale (GOS) (Marmarou *et al* 1991).

In critical care medicine (mainly intensive care units—ICUs), the severe TBI associated with the GCS score of 3–8 (within the scale of 1–15 points) has become a major challenging problem. Based on progress achieved in this area, the Brain Trauma Foundation (BTF) published the first official guidelines on the management of severe TBI in 1996 (Bullock *et al* 1996). The latest fourth edition was published in 2016 by Carney *et al* (2017). The European Brain Injury Consortium has published similar guidelines (Maas *et al* 1997). A number of studies and papers have reported the significant impact of implementation of the guideline-based protocols for severe TBI patients' treatment and outcome (e.g. Vukic *et al* (1999), Palmer *et al* (2001), Hesdorffer *et al* (2002), Fakhry *et al* (2004) and Arabi *et al* (2010)). In addition, in Eastern European countries the positive influence of the practical application of the Guidelines on reduction in mortality and morbidity of TBI patients was confirmed in Vukic *et al* (1999).

TBI is characterized by two discrete periods: *primary* and *secondary* brain injury. The primary brain injury occurs during a traumatic event at an accident scene resulting in the physical damage of brain tissues and vessels. Neurological damage invoked by these dangerous conditions may not occur immediately at the moment of primary injury, but evolves over time as a secondary injury. This is the result of a complex process, following the primary brain injury in the ensuing hours or days. Secondary intracranial brain insults include cerebral edema, hematomas, cerebral venous outflow obstruction and disturbances in cerebrospinal fluid (CSF) circulation, and it is the leading cause of in-hospital deaths after brain trauma. The secondary brain injury is directly associated with the increase of ICP and subsequent decrease

in cerebral perfusion that brings about brain tissue ischaemia. Patients with severe TBI have a significant risk of hypotension, hypoxaemia and brain swelling. The incidence of raised ICP in these patients has been reported at over 50% (Marshall *et al* 1979, Narayan *et al* 1982) and therefore such patients are suitable for routine ICP monitoring.

ICP is a robust predictor of outcome from TBI and threshold values for treatment are recommended based on this evidence (Marmarou *et al* 1991). ICP is measured in millimetres of mercury (mmHg) and, at rest, normally varies between 7–15 mmHg for adults. The range of 20–40 mmHg represents the upper limit of the ICP normality, namely, the ICP value of 22 mmHg is generally considered to be the critical (life-threatening) ICP level threshold *Thr* (Mathew 2005, Carney *et al* 2016).

Marmarou *et al* (1991) demonstrated the strongest association between the outcome, measured by the GCS and ICP equal to 20 mmHg that was generally accepted threshold at which interventions should started. Although ICP represents a global diagnostic measure that cannot identify the specific mechanisms of pressure elevation, subsequent neuromonitoring and assessment of cerebral autoregulation may help to individualize treatment of severe TBI patients. The guidelines for the management of severe TBI (Bullock and Povlishock 2007, Carney *et al* 2016) recommend that ICP should be monitored in all salvageable patients with a severe TBI with abnormal computed tomography (CT), or normal CT scan, if the patient is over 40 years old, has motor posturing or systolic blood pressure (BP) <90 mmHg. Regarding the ICP monitoring, there has been no novel additional information on ICP waveforms structures or statistical characteristics up to now. It should be noted, that the conventional application of ICP monitoring (in accordance with the mentioned Guidelines) consists merely in the utilization of the discrete mean ICP values calculated over agreed periods of time (hours, as a rule) and decision on triggering of the necessary intervention procedures is made only by comparing the actual mean value to the ICP threshold Thr; sometimes the duration of the state when ICP is over the threshold Thr is also considered. In Ghajar (2000) the first experiences with applications of the abovementioned guidelines on treating TBI patients in practice was summarized.

Particular attention has been devoted to regular ICP monitoring and to the succeeding treatment that had increased the likelihood of a favourable outcome. Mathew (2005) systemized important evidences of indication of the benefit of ICP monitoring to the patient. He remarked that in addition to the actual number of ICP alone, the utility of the ICP waveforms should be a subject of further research activities. His conclusion was as follows: 'Admittedly the process of ICP measuring is not as easy or risk-free as the measurement of mean arterial pressure (MAP), but with indications and techniques for ICP measurement becoming increasingly refined, exposure of patients to the risk of undertreatment or unnecessary empirical treatment of raised ICP is not an acceptable solution'. In the review article of Chesnut *et al* (2014), a well-systemized review is provided that focuses on the influence of the ICP monitor-based management on outcome in TBI. The authors assert that observational data supports the opinion that the ICP management has the potential to influence outcome, particularly when care is targeted and individualized and supplemented with data from other clinical examinations and imaging.

In their retrospective study (Balestreri *et al* 2004), the authors focused attention on severe TBI patients with ICP continuously above 25 mmHg for more than 4 hours and defined specific patterns for physiological variables. Although systemic hypotension is considered as one of the major factors resulting in secondary ischaemic injury, the authors could not find a significant difference in the mean values and time trends of ICP between the two outcome categories, namely, for 39 patients with favourable outcome and 57 patients with fatal outcome. Recently, a topical comparative review was published Zhang *et al* (2017) on measuring

ICP. They performed a comprehensive literature review on how to measure ICP invasively and noninvasively. In the review, the extreme importance of ICP measuring for patients with severe TBI was again demonstrated.

We encountered two papers which dealt with clusterization of the ICP signal. The authors of Lee *et al* (2016) addressed extraction of morphological landmarks, as peaks, troughs, and flats, and proposed a specific peak clustering. The second paper of Novak *et al* (2004) presents a clustering algorithm based on continuous hidden Markov models (HMMs) to automatically classify ICP beats based on their morphology. Both papers use the signal domain of ICP and did not use feature representation of the ICP signal in a vector space with vectors constituted by the whole retrospective ICP signals (records), neither did they use any Gaussian mixed model approach for clustering the vectors in the vector space. No suggestions of the construction of an additional alarm function based on such clustering was presented in the papers. Thus the motivation of our research was to propose novel alarm functions based on an original feature vector space clusterization derived from the ICP retrospective records.

The main goal of the research for our retrospective study of ICP records—provided kindly from the archive of the Lorenz Böhler Unfallkrankenhaus Wien, Vienna, Austria—was as follows:

- to extend the existing emergency indicator for severe TBI cases, that is given by the detection of the instant ICP value greater than the allowed ICP threshold, by a *novel alarm function*; this function should reflect a deeper information inherent in already available (retrospective) clinical ICP data, in particular, this information is hidden in ICP subsequences of time series (TS) by which vital and fatal ICP records can be represented.

This goal comprises the following particular subgoals, namely, to explore possibilities of transformation of the ICP time series to a feature vector space in which various clusters can be utilized for the definition of novel alarm functions; applicability of a hierarchical Gaussian mixture models (hGMMs) to the feature vector space with the aim to develop a specific ICP signals clustering scheme; alternative alarm functions, based on values of the individual posterior probability functions for feature vectors, calculated for instant ICP values; and possible thresholds of alarm functions and their optimization on the basis of the receiver operating characteristic (ROC), or using a specific efficiency measure.

In order to provide the interested reader by additional useful, but not essential information, some complementary considerations related to motivation, data description and details on methodology of sliding window principle applied to clusterization of time series, disputed by Keogh and Lin (2005), have been moved to the supplementary material (stacks.iop.org/PM/38/2015/mmedia).

We would like to emphasize that our paper reports the results of our basic research in the field of emergency medicine oriented towards monitoring and treating patients with severe brain injuries. As such, it represents an attempt to extend the methodological and mathematical aspects of contemporary approaches and proposes several novel solutions. However, at this point of research it cannot serve for direct application to clinical practice.

1.2. Methodological basis of our research and its main elements

Since in the case of ICP recorded during hospitalization of patients with severe TBI, any planning of experiments is fundamentally excluded, the only possibility of exploring the internal regularities of the ICP records is to perform a so-called retrospective study with all the data available. Then, each ICP record (for vital or fatal patients) can be regarded as a time series (TS) and we can propose an appropriate clustering scheme based on a Gaussian mixture model of a set of the ICP TS data.

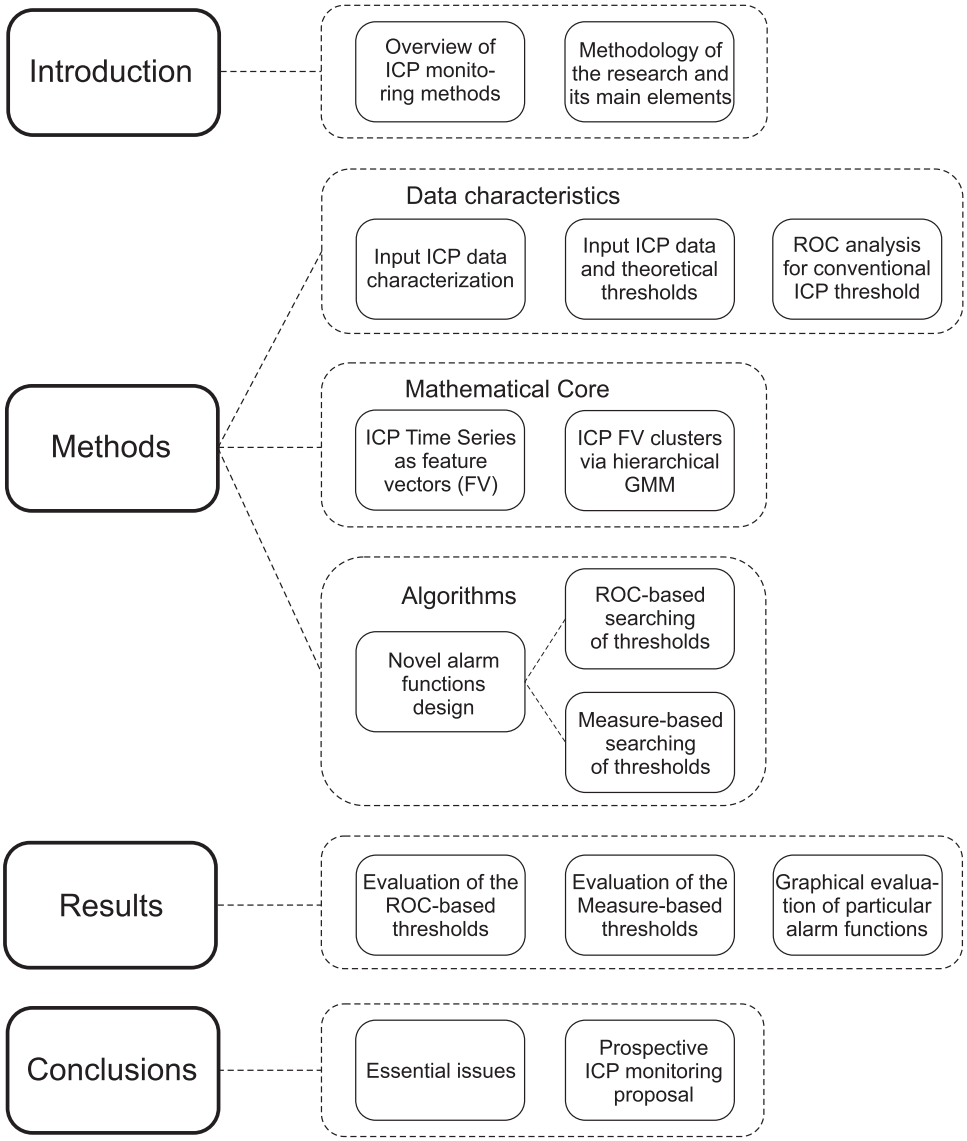

Figure 1. The scheme of the individual methodological steps used in our research.

Although we could define a conventional task of supervised classification of ICP TS, our task differs essentially in the purpose of final utilization of alarm functions constructed on the basis of the ICP TS clustering. Mainly, we are interested in two questions: whether we could make use of *a priori* existing complex information about the entire set of the ICP records of vital and fatal cases, and whether it provides useful auxiliary information on the danger state of a patient at any time moment of his/her ICP monitoring. The idea is that such a type of information could extend the common protocol of identification of the ICP threshold overrun by relevant decision that precedes such a moment.

In figure 1 the overall structure of our research activities is depicted. Some details can be explained more precisely:

- generation of the feature vectors for all subsequences defined for each ICP record that results in a feature vector subset of the multi-dimensional vector space, whereby any feature vector is accompanied by a label of the subsequence origin (a vital or fatal patient) (section 2.4);
- for clusterization of feature vector data we propose to use the hierarchical Gaussian mixture model (hGMM) (MingLiu *et al* 2002), i.e. clusters are searched as Gaussian clusters—separately for data originated from vital and fatal records (section 2.5);
- the obtained parameters of the hGMM model are afterwards used for calculation of the corresponding posterior probabilities: (i) first, within the cross-validation framework we select a testing ICP record, (ii) then, for any time moment of this record, we calculate the feature vectors of the subsequence associated with this moment, (iii) finally, we calculate the posterior probabilities of the given feature vector with regard to the clusters found within the obtained hGMM model (section 2.5);
- the ultimate element of our research methodology consists of searching for an appropriate 'alarm' time function that would be derived from the mentioned posterior probabilities (section 2.6);
- besides this basic methodology, an additional problem of searching for an optimum value of the threshold to which a novel alarm function should be compared at every time moment had to be solved. For this task, we used the whole set of the ICP data again, and developed two alternative approaches: (1) a special ROC analysis for all possible alarm function thresholds, (2) a measure of the alarm function efficiency with regards to the conventional ICP (zero) alarm function has been proposed (section 2.7).

It should be pointed out that all the above depicted steps within section 2 are applied in parallel to subsequences of an ICP record which are defined for time support of two kinds: global and local.

2. Methods

2.1. Input ICP data characterization

ICP data were recorded for patients with severe TBI admitted to the ICU of the Lorenz Böhler Unfallkrankenhaus Wien hospital. For our study, a set of clinical records of ICP for 45 severe TBI patients in the period of maximally 10 d were available. Within these data, 30 records belonged to surviving patients who were discharged from the ICU in a stabilized state and 15 records related to deceased patients. From all the patients, data from three fatal and two vital subjects were excluded, because of their extreme shortness (less than 70 h) for our analysis to be advisable. For each patient, there is a label of his/her state (fatal or vital) assessed at the time he/she left the ICU. Time resolution of the data was hours, similarly as in research studies published in Marmarou *et al* (1991), and Jun-Yu-Fan *et al* (2010). Altogether, 5575 h instances for the ICP records of vital patients (survived outcomes), as well as 2006 h instances for the ICP records of fatal patients (deceased outcomes) were accepted. Due to various clinical limitations, the ICP recording was performed as non-overlapping 1 h intervals characterized by maximum values within these intervals. The maxima represented the discretized version of an ICP record (sequence of time series). The starting time of the ICP recording, in relation to the moment of the TBI episode, varied for individual patients. For the surviving patients, the maximum duration of the ICP measurement was set to 10 d (represented by 240 h instants). The clinical protocol has led to occasional discontinuities in the ICP recording caused by the necessity to disconnect the patient for the period of the examination at another clinical

Table 1. The demographics and accompanying characteristics of all the available patients included in the retrospective study. In the table, there are six rows in which the data of various diagnostic trauma score are contained. The listed numbers represent mean values evaluated for all patients within the given category, while the numbers listed in the parentheses represent the standard deviation of the corresponding score.

	Survivors	Non-survivors
Characteristic	No. of patients (%) (total = 28)	No. of patients (%) (total = 12)
Women (%)	6 (21%)	2 (17%)
Age in years (range)	39.4 (17–93)	58.1 (20–82)
Trauma type	22×SBT/spinal trauma 5×polytrauma 1×skelet trauma	9×SBT/spinal trauma 3×polytrauma
GCS 3–4	8 (29%)	8 (67%)
GCS 5–6	8 (29%)	1 (8%)
GCS 7–9	12 (43%)	3 (25%)
APACHE II score	20.8 (4.5)	15.7 (4.8)
APACHE II mortality	27.9 (16.4)	17 (10.2)
SAPS II score	54.3 (9.6)	39.9 (10.7)
SAPS II mortality	54.8 (19.6)	28.3 (17.9)
ISS	32.8 (13.3)	31.3 (10.3)
TRISS survived	40.8 (25)	71.7 (24.9)
% of missing time points[a]	5 %	7.3 %
% of monitoring time above 20 mmHg (range)	14 % (0 - 67)	41 % (0–93)

[a] the precise specification is given at the beginning of section 2.1.

department. The reparation of all missing ICP values has been made by linear interpolation of data. Most of the gaps with missing points represent only isolated points. The gaps with three and more adjacent missing points occupy 5% of all time instants (5575) for survived (vital) and 7.3% of all time instants (2006) for non-survived (fatal) patients. In the following in table 1 the data on demographics and several diagnostic characteristics of the patients studied are summarized. The patient data are reported separately for survivors and non-survivors.

2.2. Statistic of input ICP data with regard to theoretical thresholds

Whereas it can be assumed that in all cases included in the retrospective data set the ICP threshold Thr= 20 mmHg has been used for triggering the special medical intervention procedure, an image of the relations between the acquired ICP values and the various critical (life-threatening) ICP values, was of our interest. Namely, for each individual ICP record from

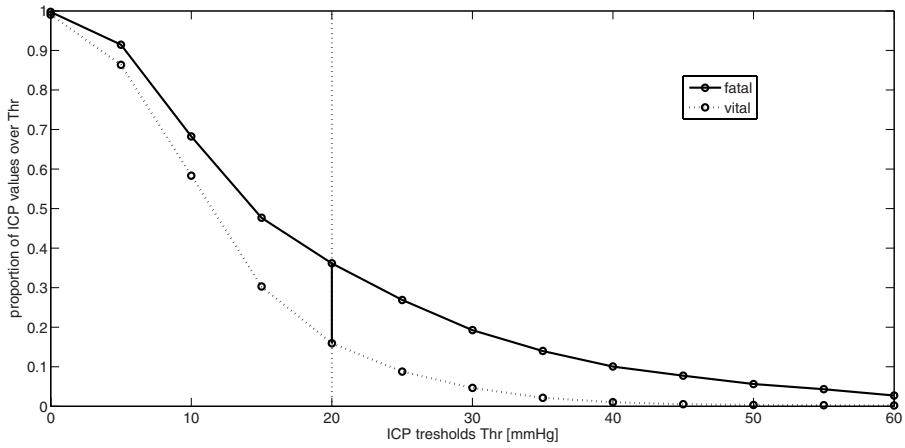

Figure 2. The proportions of the ICP values in individual ICP records which are over the fixed critical ICP thresholds Thr (5 mmHg steps were used for the abscissa). The averaged values over the fatal cases and vital cases have been calculated separately.

the retrospective study, we were interested in a proportion of the ICP values, acquired for the given record, which exceeded the variable ICP critical value Thr$\in \{0, 5, 10, 15, 20, \ldots, 60\}$, to other ICP values. Two plots of such proportions averaged over the whole set of fatal or vital records, respectively, are depicted in figure 2.

The overall trend of decreasing proportions seen in figure 2 for both types of records is in accordance with our intuition. However, the number of ICP values above critical values, beginning at the indicator threshold Thr$= 20$ up to the value 50 mmHg, confirm that in reality many TBI patients survived with significantly increased ICP values, though this proportion continually decreases. In vital cases, no values above the 50 mmHg occur. On the other hand, we can see that in spite of triggering the intervention procedure by the same threshold value Thr$= 20$, the proportion of the ICP values over a critical level is high in all fatal records. In more detail, e.g. for the critical value Thr$= 25$ mmHg, the proportion of the ICP values greater than this threshold, which are present in fatal cases, equals 0.28, while the proportion calculated for the vital cases equals only to 0.09. This fact serves as further motivation to search for a more complex indicator derived from the ICP record ensemble that could trigger the intervention procedure earlier than the ICP threshold itself. Considering the differences of the ICP proportions over Thr between vital and fatal cases, we see that the maximum of this quantity is reached exactly for the used threshold Thr$= 20$ mmHg.

2.3. Input ICP data characterization based on ROC analysis for the conventional ICP threshold Thr value

There are a number of papers dealing with the ICP monitoring of TBI patients using ROC analysis to solve various tasks. Chambers *et al* (2001) applied this standard statistical tool to explore whether there are significant threshold levels in the determination of outcome and whether the quantity of ICP is a better predictive outcome than the quantity of CPP. Using ROC curves, the authors found out, that the threshold value of $\mathrm{ICP}_{\max} = 35$ mmHg and the threshold value for $\mathrm{CPP}_{\min} = 55$ mmHg appear to be the best predictors in adults. In the paper of Steyerberg *et al* (2008), ROC was used to explore a prognostic model that combined age, motor score and pupillary reactivity. An area under the ROC curve (AUC) between 0.66 and 0.84 was achieved at cross-validation. Guiza *et al* (2015) investigated the relationship

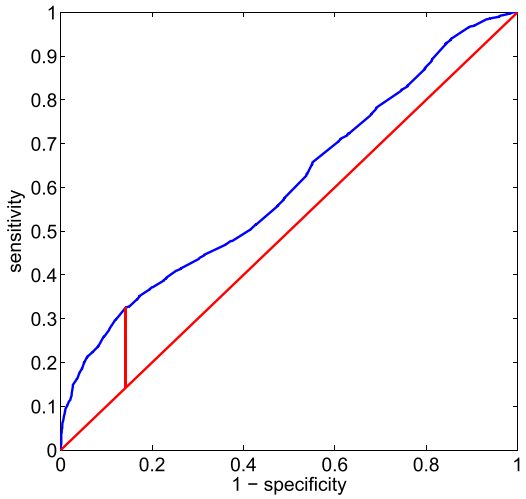

Figure 3. The ROC plot calculated for the retrospective data and a set of the ICP thresholds Thr. The Youden index (YI= 0.19) is displayed as the red vertical line segment, the corresponding value of *sensitivity* is 0.33 and of *specificity* is 0.86. The optimal ICP threshold associated with the Youden index is Thr= 19.9.

of episodes of elevated ICP above a certain threshold during a certain time with a 6 month Glasgow outcome scale by means of visualization in a colour-coded plot. However, to the best of our knowledge, no relevant information is available in the literature that is related exactly to the ROC analysis of the conventional criterion of $\mathrm{ICP}(k) > \mathrm{Thr}$. Therefore, we decided, to carry out the ROC analysis of the zero-alarm function, i.e. all ICP records from the retrospective data set were considered. For the ICP threshold values Thr_v for $v = 1, 2, \ldots, 10$, we selected the following set of ICP integer values (mmHg): $\{5, 10, 15, 20, 25, 30, 35, 40, 45, 50\}$ which, for completeness, span the extremum ICP values, though they are out of the clinically permitted range. In this situation, the vital and fatal cases, as well as prediction outcomes are associated with the values of $\mathrm{ICP}(k)$ for all the individual indices k encountered in the set of ICP records.

In figure 3 the ROC curve, obtained for all the ICP values and the range of the selected thresholds, is plotted. It should be pointed out that the threshold Thr= 19.9 was found as the optimum via the ROC analysis (with YI= 0.19) and it coincides with the lower boundary of the interval $\langle 20, 40\rangle$ mmHg considered in the guideline-based protocols of Fakhry *et al* (2004), and Arabi *et al* (2010). According to Mathew (2005), the most-accepted ICP threshold is 20 mmHg. In the clinical study of Honda *et al* (2017), published quite recently, in which ICP monitoring was accompanied by parallel-acquired CT neuroimaging and perfusion study, the same value of ICP threshold is recommended.

Let us recall the definition of the basic notions of the ROC analysis adjusted for our situation. The *sensitivity* is defined as the proportion: $\sum$*true positive predictions*/$\sum$*condition positives*. Here we interpret the denominator as the overall sum of the cases of the $\mathrm{ICP}(k)$ (k-time instant) values associated with all the available fatal records. Although the fatal records are known only in the retrospective view, we can consider all evolving patient states during his/her hospitalization, specified by the $\mathrm{ICP}(k)$ values, as 'life threatening' that represents for us the notion of the variable *condition positives*. The nominator represents, on the other hand, the number of all cases of the $\mathrm{ICP}(k)$ selected from the fatal records for which the relation $\mathrm{ICP}(k) > \mathrm{Thr}$ is valid. Then, the sensitivity (= 0.33) value found by ROC

(retrospective) analysis for the available ICP data can be interpreted such that in 33% of cases, the relation ICP(k) $>$ Thr indicated a life-threatening patient state and the medical intervention was triggered or has been continued. On the other hand, it can be claimed that, in 67% of the cases selected from the fatal records, the state of the patient in the given moment was life-threatening, but the relation ICP(k) $\leqslant$ Thr did not indicate it and therefore no medical intervention could be triggered on an ICP-only basis. The introduction of a novel alarm function that could help to indicate critical patient states also in these cases is a basic motivation of our research. The position of such a novel alarm function to the conventional ICP versus Thr indicator should not therefore be seen as its opponent, but on the contrary, it should be a suitable complementary emergency indicator.

2.4. Representation of the ICP original data by means of vector features

Based on state-of-the-art ICP monitoring of TBI patients and taking into account the results of the detailed statistical analysis of the available ICP data included in the previous section, we concluded that a further advanced technique of ICP monitoring analysis should provide relevant information for any moment of ICP measurement of an actual patient. This could serve as an impulse (alarm) for triggering additional medical intervention applicable to the patient. To reflect the complex shape of the ICP records, we propose to represent ICP time segments by globally and locally computed vector features. Namely, under the global characterization of the evolving ICP record we mean the transformation of the ICP values into vector features defined on a discrete time support N.1 given by all time moments from the beginning up to its instant time. Thereby we want to utilize the entire history of the ICP TS up to the given moment. By using the local information we understand a shorter time segment of ICP, ending again at the instant ICP value, for which the same type of the feature vector is calculated (the only methodological difference consists of a different time support). It should be noted, that in view of the critical claim of Keogh and Lin (2005) related to the useless application of the sliding window principle to time series clusterization, by transition from the ICP time series to a vector feature characterization we avoid the situation described and analysed by authors. Then,

- the first type of the support (N.1) is derived from the time series $T_n = (x_1, x_2, \ldots, x_n)$ with components x_i equal to ICP values measured at the subsequent discrete time moments with the identical beginning and variable length k; i.e. we write the corresponding ICP subsequence as $S_k = (x_1, x_2, \ldots, x_k)$, where $k = 12$, 13, ..., $n \leqslant 240$; the value 12 was empirically set as a minimum subsequence length, for which the individual (*global*) features acquire yet reasonable values and which could also be clinically relevant; while the value 240 was found to be a convenient fixed length of the ICP records available in our study;
- the second type of support (denoted N.2) is given as the latest s time samples for which the ICP subsequence is written as $L_{k-s} = (x_{k-s+1}, \ldots x_{k-1}, x_k)$, where we select $s = 12$, as an appropriate value; the features calculated for these cases are called *local*;
- similarly as in Kuncheva and Rodriguez (2013), we define $2m$ appropriate feature vectors for each ICP record subsequence given by individual supports N.1 and N.2.

For each subsequence S_k and L_{k-s} of the whole ICP sequence T_n, we define the global and local sets of the following seven features: f_k^j, $(j = 1, \ldots, 7)$:

1. Maximum value:
$$f_k^1 = \max_{i=1,\ldots,k} \{x_i\},$$
2. Mean value:
$$f_k^2 = 1/k \sum_{i=1}^{k} x_i = \bar{x}_k,$$
3. Standard deviation:
$$f_k^3 = \sqrt{1/(k-1) \sum_{i=1}^{k} (x_i - \bar{x}_k)^2},$$
4. Covariance coefficient (Kuncheva and Rodriguez 2013):
$$f_k^4 = 1/k \sum_{i=1}^{k} (i.x - \bar{x}_k.(k/2)),$$
5. Coefficient derived from the ICP absolute differences along the segment:
$$f_k^5 = 1/(k-1) \sum_{i=1}^{k-1} |x_{i+1} - x_i|,$$
6. Autocorrelation coefficient of t-order ($t = [k/2]$):
$$f_k^6 = \sum_{i=1}^{k-t} (x_i - \bar{x}_k)(x_{i+t} - \bar{x}_k) / \sum_{i=1}^{k-t} (x_i - \bar{x}_k).$$
7. Instant ICP value:
$$f_k^7 = x_k.$$

The latter feature, as can be seen, is nothing more than a simple ICP value for the time instant k that does not depend on any time segment, i.e. it is a pure point characteristic. Thus, our experiments can be related to a maximum seven-dimensional feature space. As will be described later, some combinations of the defined vector features of lesser dimension m will be selected.

2.5. ICP feature vector clusterization using the hierarchical Gaussian mixture models (hGMM)

A number of clustering schemes have been developed (e.g. Jain (2010), Fu (2011) and Gama (2012)) in the domain of machine learning and data mining. One prominent branch of methods is known as distribution (density)-based clustering, within which, the GMMs (Bishop 2006) have been used extensively in various application fields. These semi-parametric estimation techniques represent a very general class of functions for the density model in which it is possible to increase a number of adaptive parameters in a systematic way. Thus the GMM can be made arbitrarily flexible and it fits clustering tasks like ours very well. Let us recall the basic formulation of the GMM. It is defined as a weighted sum of M components multivariate Gaussian densities given by the equation:
$$p(\boldsymbol{x} \mid \boldsymbol{\lambda}) = \sum_{i=1}^{M} w_i g(\boldsymbol{x} \mid \boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i),$$

where $\boldsymbol{x}$ is a D-dimensional random variable (feature vector), w_i, $i = 1, \ldots, M$, are the mixture weights and $g(\boldsymbol{x} \mid \boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i)$,$i = 1, \ldots, M$, represents Gaussian distribution densities with means $\boldsymbol{\mu}_i$ and covariance matrices $\boldsymbol{\Sigma}_i$. The clustering of the data in D-dimensional feature vector space can then be viewed as searching for an optimum GMM with fixed number of components M (corresponding to the fixed number of data clusters). The solution of such an optimization problem is carried out by finding the unknown parameters $\boldsymbol{\lambda} = \{w_i, \boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i\}$, for $i = 1, \ldots, M$. As will be shown later, this parametrization of the clustering task will serve as an appropriate tool for the construction of an alarm function by means of posterior probability functions for a new input feature vector data of the ICP subsequence.

In the previous section we proposed representing the ICP TS by feature vectors which are defined for global support and for local support. Thus our basic clusterization task needs to be split into two independent GMMs. As the applications of these models are identical, we will describe the formalism only for one such model. Next, to combine these two GMMs, we propose to apply the so-called *hierarchical* Gaussian mixed model (hGMM) (e.g. in MingLiu *et al* (2002)) with various numbers M of components. As this clustering task is of a supervised type, the hGMM enables straightforward labelling of clusters, or posterior probabilities of an input feature vector. The basic scheme of our implementation of the supervised hGMM to the D-dimensional feature vector space can be described as follows. First, we consider a class decomposition of the set of all feature vectors calculated for the entire ICP data set: let C^V denote the class of all feature vectors $\boldsymbol{x}^V$ associated with the ICP records for vital patients, and let C^F denote the class of all feature vectors $\boldsymbol{x}^F$ associated with the ICP records for fatal patients; let M represent the same number of GMM components for modelling the vital as well as fatal data. Then within the overall hGMM scheme, we formulate two parallel GMM tasks. The formal notation for the vital case can be written as

$$p^V(\boldsymbol{x}^V \mid \boldsymbol{\lambda}^V) = \sum_{i=1}^{M} w_i^V g(\boldsymbol{x}^V \mid \boldsymbol{\mu}_i^V, \boldsymbol{\Sigma}_i^V)\,,$$

and similarly we define the GMM for the fatal case

$$p^F(\boldsymbol{x}^F \mid \boldsymbol{\lambda}^F) = \sum_{i=1}^{M} w_i^F g(\boldsymbol{x}^F \mid \boldsymbol{\mu}_i^F, \boldsymbol{\Sigma}_i^F)\,.$$

Having found two separate GMMs we can construct a final mixture considering the second level of the hierarchy that can be expressed as a weighted mixture of the two first level hierarchy GMMs:

$$p(\boldsymbol{x} \mid \boldsymbol{\lambda}) = q^V p^V(\boldsymbol{x}^V|\boldsymbol{\lambda}^V) + q^F p^F(\boldsymbol{x}^F|\boldsymbol{\lambda}^F)\,.$$

In our approach we estimate the weighting coefficients (priors) q^V and q^F directly from the given data. The ratio of the number of feature vectors calculated for fatal and vital patient records in the retrospective data set is $r = N^F : N^V = 1 : 3.2$, (where N^F and N^V denote the number of fatal and vital feature vectors, respectively), from which we obtain the weights $q^F = 1/4.2 = 0.24$, and $q^V = 3.2/4.2 = 0.76$, preserving the condition $q^V + q^F = 1$. Secondly, we have to ensure that the values of $2M$ posterior probabilities for an arbitrary input vector $\boldsymbol{x}$ sum up to the value one. We achieve this by using the normalizing coefficient $\mathrm{NORM}(\boldsymbol{x})$ for our hGMM, comprising $2M$ Gaussian densities from the first level of hierarchy, and calculating $2M$ posterior probabilities for any vector $\boldsymbol{x}$ for $i = 1, 2, \ldots M$:

$$p_i^V(w_i^V g_i^V|\boldsymbol{x}) = \tfrac{q^V}{\mathrm{NORM}(\boldsymbol{x})}\, w_i^V g_i^V(\boldsymbol{x}), \qquad p_i^F(w_i^F g_i^F|\boldsymbol{x}) = \tfrac{q^F}{\mathrm{NORM}(\boldsymbol{x})}\, w_i^F g_i^F(\boldsymbol{x}),$$

where the coefficient $\mathrm{NORM}(\boldsymbol{x})$ is defined as

$$\mathrm{NORM}(\boldsymbol{x}) = q^V \sum_{i=1}^{M} w_i^V g_i^V(\boldsymbol{x}) + q^F \sum_{i=1}^{M} w_i^F g_i^F(\boldsymbol{x}) .$$

Next, recall that two sets of features are considered, representing local and global ICP properties. Each set is modelled by its own hGMM. In this study we consider two ways of combining hGMMs calculated for the global support and local support. (i) First, up to the normalization step for posterior probabilities we estimate parameters of each hGMM in parallel; then to accomplish the last step of mixing two hGMMs we use the following formula:

$$\mathrm{NORM}_{\mathrm{G}}^{L}(\boldsymbol{x}) = q^V[v_{\mathrm{G}}(\sum_{i=1}^{M} w_i^V g_i^V(\boldsymbol{x}))_{\mathrm{G}} + v_{\mathrm{L}}(\sum_{i=1}^{M} w_i^V g_i^V(\boldsymbol{x}))_{\mathrm{L}}]$$

$$+q^F[v_{\mathrm{G}}(\sum_{i=1}^{M} w_i^F g_i^F(x))_{\mathrm{G}} + v_{\mathrm{L}}(\sum_{i=1}^{M} w_i^F g_i^F(x))_{\mathrm{L}}],$$

where the symbol G stands for Global, and the symbol L for Local support; the all global and local posterior probabilities (vital and fatal) can be in this case weighted by additional weights v_{G} and v_{L}. (ii) Second, according to the above described normalization procedure of hGMM we can calculate posterior probabilities for global and local support separately. Then we propose to combine pairs of the corresponding posterior probabilities from global and local branches of the hGMM procedure as a linear combination. After evaluating pilot experiments with both approaches, we preferred the latter approach.

2.6. Introduction of novel alternative alarm functions

2.6.1. Research methodology. For any of the selected number $2M$ of clusters, the solution of the hGMM problem provides the full description (parameters values) of the Gaussian clusters. These parameters, together with the selection of M posterior probabilities for fatal clusters only, constitute initial characteristics (an imprint) of the whole data set at hand. All this information can then be used for the calculation of auxiliary functions $F^*(k)$, $V^*(k)$ which can be used afterwards to calculate the value of the selected alarm function $Al_j(k)$ for the actual ICP(k) value acquired in the instant k for the given patient. To carry out a systematic exploration of the behaviour of the proposed alternative alarm functions and accompanying operations, it was necessary to propose a cross-validation scheme. For $N = 40$ valid ICP records (of those, 12 were fatal cases and 28 were vital cases) we repeatedly detached one record that was stored for testing purposes. The remaining 39 records are used as input data for solving the individual hGMM task for the chosen number of clusters. Having obtained a full description of the clusters it is possible to look at each testing ICP record as a fictively evolving ICP record and for each index k calculate all quantities defined for the ICP segments (ICP time subsequences up to the index k) which are necessary for obtaining the value of the alarm function $Al_j(k)$. In the following section we will propose three alternative alarm functions, each coupled with its own threshold TA_j. The optimum values of these thresholds will be searched (i) by means of the corresponding ROC analysis, that includes the values of the given alarm function for each time index k, and, alternatively, and (ii) by means of the optimization of certain efficiency measures of the proposed alarm functions.

2.6.2. Design of alternative alarm functions. Let us recall that the coupling of the information represented by a time-dependent ICP record with the information on an *a priori* established

threshold Thr, and evaluating their mutual relation with the evolving ICP discrete time record, serves in contemporary medicine as a basic tool for triggering a necessary medical intervention. Thus it is natural to call the ICP record itself a *zero-alarm* function and denote it as $Al_0(k) = ICP(k)$. In particular, the most important moment (ALARM) for treating patients with severe TBI is represented by the time instant when the value ICP(k) crosses the threshold value Thr for the first time in any ICP time subsequence of ascending values. At this very moment a sequence of specific medical measures (according to the guideline-based protocols of Fakhry *et al* (2004), and Arabi *et al* (2010)) starts. The ICP monitoring continues with increased attention. For this online case, it is natural to interpret all the subsequent ICP values over the threshold as an indication of persisting state of emergency for the patient. Of course, the dynamics of the ICP time curve, as well as medical supervision of other medical parameters of the patient, can serve as feedback and to control the intervention process alone. All these aspects represent relevant issues of specific medical research, however, these were beyond the scope of our research project related exclusively to the ICP retrospective study. At the time moment, when the ICP(k) drops below the threshold Thr, the emergency level can be weakened. Monitoring of the ICP(k) curve continues under stabilized conditions (and probably aborted intervention actions) up to the possible next ALARM.

All preceding considerations are concerned equally with the concept of the application of novel alarm functions to the ICP monitoring process. However, in our research project, a new problem of finding the optimum threshold level of each individual alarm function arises. In our situation, the ICP records (and consequently the values of the alarm functions) are defined as posterior curves for the whole definition range. Then for searching for optimum thresholds, we can consider any individual value of an alarm function separately. For the corresponding optimization procedure we make a theoretical assumption that each alarm function value exceeding the variable threshold can be a possible first indication of an ALARM in a future application in the actual case of patient ICP monitoring. In other words, within this optimization problem, we deem an arbitrary alarm value exceeding the threshold equally important.

For all the ICP records included in the retrospective study, we will introduce and explore three alarm functions $Al_j(k)$ for $j = 1, 2, 3$ which will be based on the proposed methodology of feature space clusterization using hGMMs. These functions will differ from each other by their local or cumulative nature and by the general or selective application to an ICP record. The prerequisite of the definition of all these functions is the construction of auxiliary functions using the corresponding posterior probabilities calculated from the obtained hGMM for the given feature values of the evolving ICP record in each time instant. Let us suppose we have found Gaussian clusters in the feature vector space calculated for the known ICP records of the given data set. Then, for any given subsequence $\boldsymbol{x}_k$ (that represents the instant state of an evolving ICP record in time, indexed by the symbol $k = 12, 13, ..., n \leqslant 240$), the posterior probability of the given ICP feature vector related to the fatal cluster C_i^F is denoted as $p_i^F(k \mid \boldsymbol{x}_k)$, and the posterior probability of the given ICP feature vector related to the vital cluster C_i^V is denoted as $p_i^V(k \mid \boldsymbol{x}_k)$, where $i = 1, \ldots, M$.

Finally, the values of the auxiliary function $F^*(k)$ for the given subsequence $\boldsymbol{x}_k$ are defined using the following formula:

$$F^*(k) = \sum_{i \text{ for fatal}} p_i^F(k),$$

and similarly, the auxiliary function $V^*(k)$ is defined as

$$V^*(k) = \sum_{i \text{ for vital}} p_i^V(k),$$

where, for the sake of simplicity, we omitted the symbol of the feature vector $\boldsymbol{x}_k$. It should be noted that within the framework of hGMM the equality $F^*(k) + V^*(k) = 1$ is always true.

As formulated in our basic methodology, we wish to evaluate the value of the alarm function at any discrete time of the evolving ICP record of the monitored patient, based on all previous values of the auxiliary function from this single ICP record, up to the index k. The obtained value $Al_j(k)$ can be interpreted as a measure of danger in which the monitored patient finds him or herself at the moment k. It is compared to the threshold alarm value (a constant TA_j for the whole alarm function range) aimed at triggering a special medical intervention procedure. Thus the relation $Al_j(k) > TA_j$ should be tested for the subsequent values of the index k. Intuitively, the most important cases (indices k) in practice are when this relation is valid for the first time in each subsequence of the ascending values of any alarm function. However, for theoretical purposes we need to consider also the states of lasting danger, i.e. all subsequent time instants when the relation $Al_j(k) > TA_j$ is still valid. For the reflecting such patient states, we define the following mutually alternative alarm functions for all indices k available in the given ICP record, where the attributes are related to the time moments:

1. The *local* alarm function $Al_1(k)$:

$$Al_1(k) = F^*(k).$$

2. The *selective cumulative index* alarm function $Al_2(k)$:

$$Al_2(k) = 1/k \sum_{F^*(i)>V^*(i)}^{i=1,\ldots,k} i.$$

3. The *selective cumulative* alarm function $Al_3(k)$:

$$Al_3(k) = 1/k \sum_{F^*(i)>V^*(i)}^{i=1,\ldots,k} F^*(i).$$

2.7. Optimization of the thresholds for novel alarm functions

In this section the main goal is to propose a suitable characterization of various alarm states in monitoring ICP by means of the retrospective information on ultimate outcomes of medical treatment of the patients included, and by utilization of the three novel alarm functions together with their optimum thresholds. It appeared that for such a characterization, the methodology analogous to measures of medical diagnosis accuracy based on ROC can be used efficiently. Originally, the ROC curve shows the characteristics of a diagnostic test by graphing the false-positive rate (1-specificity) on the horizontal axis and the true-positive rate (sensitivity) on the vertical axis for various cut-off values. A useful diagnostic test should have a cut-off value at which the true-positive rate is high and the false-positive rate is low. In medical practice, searching for an optimal threshold point from the ROC curve (*the best cut-off point*) (Fawcett 2006, Indrayan 2013) represents the important diagnostic task.

In our study, the evidence of vital or fatal patient outcomes is seen as analogous with actual values (of binary classification into healthy or sick person class), whereas the states (for each time instant k) when the particular alarm function exceeds the given threshold or it does not, serves as analogous with the prediction outcomes (positive or negative) of medical diagnostic tests. For use in the monitoring procedure of TBI patients, finding the best cut-off point on the ROC curve means that we have found the threshold value of the particular alarm function

that is optimal in the described sense. As we indicated in section 2.6.1, each individual alarm function needs its own optimum threshold. For searching for the optimum threshold values, we proposed two approaches: first, the direct application of the ROC methodology has been developed and evaluated; second, we proposed and explored novel suitable measures for alarm function efficiency which make use of specific ROC-like characteristics of the ICP data. Details of both methods are described in the following sections.

2.7.1. Determination of the optimum thresholds TA_j^{opt} for the proposed alarm functions via the corresponding ROC analysis. The critical value of any of the novel alarm functions $Al_j(k)$, which have different and more complex methodological nature than the quantity of ICP, could be derived from the values of the alarm functions calculated for all data using the cross-validation scheme. Since in our case *a priori* information on vital/fatal type of the patient ICP record is available, it is possible to calculate a contingency table of actual values and test (prediction) outcomes. We proposed to perform the ROC analysis by including discrete values of the threshold TA_j, originating from the whole ranges of the corresponding alarm functions $Al_j(k)$ calculated for all available patient ICP records and for all k time arguments. We will describe this procedure in detail.

For the given type of hGMM, and for hypothetical threshold values $TA_j(i), i = 1, 2, \ldots, N_{TA_j}$, $TA_j(i) \in < \min Al_j, \max Al_j >$ of the selected alarm function $Al_j(k)$, one step of the cross-validation scheme is performed for each patient p. Then, we calculate the values of the alarm function $Al_j^p(k)$ for the given patient and for all available values of the time instant k. After completing the cross-validation scheme for all patients (fatal as well as vital) $p = 1, 2, \ldots, 40$, we obtain values of all alarm functions $Al_j^p(k)$. Afterwards the corresponding quantities of the ROC analysis are calculated separately for the fatal and vital patient ICP records and the corresponding alarm functions according to the formulae:

Fatal records: for all p and k:

- True positive (TP) $= \#\,(Al_j^p(k) > TA_j)$
- False negative (FN) $= \#\,(Al_j^p(k) \leqslant TA_j)$.

Vital records: for all p and k:

- False positive (FP) $= \#\,(Al_j^p(k) > TA_j)$
- True negative (TN) $= \#\,(Al_j^p(k) \leqslant TA_j)$.

Based on a preliminary exploration we found out that the selection of all seven basic features defined in section 2.4 do not represent the best decision for our task. Therefore we decided to experiment with a reduced set of these features. Namely, we considered the following three feature vectors $H_{3D} = (f_k^2, f_k^3, f_k^5)$, $H_{4D} = (f_k^2, f_k^3, f_k^4, f_k^5)$, and $H_{5D} = (f_k^2, f_k^3, f_k^4, f_k^5, f_k^7)$. The ROC analysis for the whole available data set can be described in the following way.

For each given threshold value $TA_j(i)$, and for all values of the alarm functions $Al_j^p(k)$, and for all patients p and the corresponding time indices k we calculated the values TP and FN (from fatal records), and the values FP and TN (from vital records). We obtained unique values of sensitivity (Se) and specificity (Sp) characterizing the whole dataset. Consequently, for N_{TA_j} pairs of these quantities, we obtained the N_{TA_j} values of the ROC curve cut-off points with the coordinates $[x, y] = [(1 - Sp), Se]$. According to the standard ROC technique, the goal is to find the best cut-off from the ROC curve. We applied the frequently used Youden index (YI) that maximizes the vertical distance from the line of equality (diagonal line in ROC) to the points $[x, y]$ of the curve plot. The Youden index maximizes the difference between the TP rate (Se) and the FP rate $(1 - Sp)$.

2.7.2. Determination of the optimum thresholds TA_j for the proposed alarm functions using a suitable efficiency measure. The efficiency of a novel alarm function, in relation to the existing ICP alarm, could be characterized by measures which are also suitable for determining an optimum thresholds TA_j^{opt} of the proposed alarm functions. For such an alternative to the basic ROC approach we need to introduce the following terms. Having selected an alarm function $Al_j^p(k)$, calculated for the pth patient, and the variable threshold value TA_j for the pth patient, given the variable threshold Thr_v for the whole ICP record of this patient, we can consider the time instant k when the condition $\mathrm{ICP}^p(k) > Thr_v$ becomes true for the first time in any ascending ICP value subsequence as the conventionally justified moment for triggering the initial alarm. Similarly, the moment of the satisfaction of the condition $Al_j^p(k) > TA_j$ can be designated to be a suitable moment for triggering a novel alarm. The periods of time including the very first moments and the following time points with lasting valid conditions can be called 'conventional alarm active' in the former case, and 'novel alarm active' in the latter. For the index subsequences when the abovementioned conditions are not valid, we introduce the terms 'conventional alarm passive' and 'novel alarm passive'.

Let us consider a hypothetic situation that for clinicians, besides the information on the conventional alarm active, also the information on the novel alarm active is available for each time instant k in the given ICP record. Then, using the above introduced terms and the retrospective view we get four specific sets of indices k:

(i) $E_{Alj}^p = \{k:\ (Al_j^p(k) > TA_j^{\mathrm{opt}}) \& (\mathrm{ICP}^p(k) \leqslant \mathrm{Thr}_v)\}$—union of all exclusive novel alarm active subsequences ($N_{Alj}^p = \#\ E_{Alj}^p$),

(ii) $E_{\mathrm{ICP}}^p = \{k:\ (\mathrm{ICP}^p(k) > \mathrm{Thr}_v) \& (Al_j^p(k) \leqslant TA_j^{\mathrm{opt}})\}$—union of all exclusive conventional alarm active subsequences ($N_{\mathrm{ICP}}^p = \#\ E_{\mathrm{ICP}}^p$),

(iii) $A_{Alj\&\mathrm{ICP}}^p = \{k:\ (Al_j^p(k) > TA_j^{\mathrm{opt}}) \& (\mathrm{ICP}^p(k) > \mathrm{Thr}_v)\}$—subsequences when both alarms are simultaneously active ($M_{Alj}^p = \#\ A_{Alj\&\mathrm{ICP}}^p$),

(iv) $P_{Alj\&\mathrm{ICP}}^p = \{k:\ (Al_j^p(k) \leqslant TA_j^{\mathrm{opt}}) \& (\mathrm{ICP}^p(k) \leqslant \mathrm{Thr}_v)\}$—subsequences when both alarms are simultaneously passive ($M_{Pj}^p = \#\ P_{Alj\&\mathrm{ICP}}^p$).

Also in this case it is reasonable to search for optimum thresholds of the alarm functions by using the information contained in the whole dataset. It means that the definition of the efficiency measure requires extension of the above-introduced quantities to all patients. Once the cross-validation process is finished and the values of the jth alarm function are available for all time arguments k, the following sets can be defined:

(i) $E_{Alj} = \bigcup_p E_{Alj}^p$,

(ii) $E_{ICP} = \bigcup_p E_{\mathrm{ICP}}^p$,

(iii) $A_{Alj\&\mathrm{ICP}} = \bigcup_p A_{Alj\&\mathrm{ICP}}^p$,

(iv) $P_{Alj\&\mathrm{ICP}} = \bigcup_p P_{Alj\&\mathrm{ICP}}^p$.

Consequently, we obtain the total number of the subsequence positions related to the corresponding sets defined above: $N_{Alj} = \#\ E_{Alj} = \sum_p N_{Alj}^p$, $N_{ICP} = \#\ E_{\mathrm{ICP}} = \sum_p N_{\mathrm{ICP}}^p$, $M_{Alj} = \#\ A_{Alj\&\mathrm{ICP}} = \sum_p M_{Alj}^p$, $M_{Pj} = \#\ P_{Alj\&\mathrm{ICP}} = \sum_p M_{Pj}^p$.

Let us consider now only the subset of the *fatal* patient records. We can say that in spite of the treatment activities following the moments of conventional alarm activation, the individual

patient ultimately deceased. This means that all time indices with ICP *alarm active* can be considered technically as deficient in fatal outcomes. Therefore we can hypothesize that the presence of an ICP subsequence with the *exclusive* novel alarm active (we denote the number of all these incidences as N_{Alj}^{fat}) could have helped clinicians to make earlier decision on triggering medical intervention, which could have possibly increased the patients survival chances. On the other side, in cases when both alarms are simultaneously active, related to retrospective fatal records, there is clearly no benefit of the novel alarm.

Now, let us consider the subset of the *vital* patient ICP records. We can affirm that during the ICP monitoring, no additional successful ICP-based alarms were applied (this does not exclude application of the other types of medical alarms, based, e.g. on checking biochemical parameters). Further, there are certainly some vital patients for whom the conventional alarms were triggered, but it is impossible to identify which of them actually had life-saving effects. For research purposes, we can employ a more rigorous approach and consider all index subsequences for the exclusive novel alarms active to be purely redundant ('false') alarms for the vital patients. The number of all the incidences in this case is denoted as N_{Alj}^{vit}.

Having these two different views on the sets of fatal and vital ICP records, we can ask how to construct an alternative (to the ROC) optimization procedure that can yield optimum thresholds of the individual alarm functions for (ad hoc) indication of the alarm situation of the patients included in the retrospective study. We would like to trade-off these two opposite indications and therefore we propose a measure μ_r as the following proportion:

$$\mu_r(\text{Thr}, TA_j) = \frac{N_{Alj}^{\text{fat}}}{N_{Alj}^{\text{vit}}}.$$

For the already fixed threshold Thr of the ICP values we have a freedom to set the threshold values for each alarm function within its range. The optimum threshold TA_j can be found as such a value for which the measure $\mu_r(\,)$ reached the maximum for all variable parameters, similar to the case of the ROC optimization approach.

3. Results

We ran computer experiments organized as a set of cross-validation computations in which the 'one-leave-out' scheme was applied sequentially to the ICP record of each patient. As variable parameters and method branches, we selected the global (G) and local (L) approach and their combinations, which resulted in characteristics of posterior probabilities (PP) calculated as the pointwise (with regards to time k) arithmetic mean of the G and L values of the PP, and the pointwise maximum of the G and L values of the PP for the given patient. Further, we selected 3D, 4D and 5D feature spaces of the ICP records and used the following numbers of clusters $M = 1, 2, 3, 4, 5, 6$. As mentioned above, evaluation of the preliminary experiments yielded three recommended types of feature vectors: $H_{3\text{D}} = (f_k^2, f_k^3, f_k^5)$, $H_{4\text{D}} = (f_k^2, f_k^3, f_k^4, f_k^5)$, and $H_{5\text{D}} = (f_k^2, f_k^3, f_k^4, f_k^5, f_k^7)$ (extracted from the formulae defined in section 2.4). To ensure numerical consistency, the values of each feature included in the feature vector of the individual set of experiments have to be normalized into the interval $< 0, 1 >$. This normalization is accomplished for all patients and all time instances k. The goal of these computer experiments was to calculate the individual hGMMs of the input ICP data for all variable parameters which had been used afterwards for the calculation of several novel alarm functions. The resulted alarm functions had to be included into the ROC analysis and into the optimization

calculations of efficiency measure. The ultimate goal was to select the optimal variable parameters of the method, as well as the optimum alarm functions together with the corresponding optimal thresholds.

For our experimental setup we used the Matlab-based (Mathworks, version R2012b) NETLAB toolbox (Nabney 2002). In this implementation, an expectation maximization (EM) iterative algorithm is implemented that minimizes the negative log-likelihood function of the data set. Next, we constructed novel alarm functions of time, given through the indices $k = 12$, 13,..., $n \leqslant 240$ of the discrete time instants, using auxiliary functions $F^*(k)$ or $V^*(k)$ computed from the posterior hGMM probabilities (see section 2.6.2). In association with computational aspects of the implementation of the EM method, which includes the minimization of the negative log-likelihood function, it should be emphasized that this operation is not a well-posed problem (e.g. Bishop (2006)). To manage this problem when calculating the corresponding hGMM by the EM (before the final combination of the hierarchical outputs), we decided to repeat this critical step s times. Our experimentation (with $s = 100$) showed a certain variability of the obtained results (different cluster centres). In most cases several groups of identical centres were observed.

We ordered the results according to the terminal-likelihood values of the individual attempts (in our case of negative log-likelihood in increasing order). The first $t = 10$ values served then for the selection of the respective 10 hGMM runs and for the calculations of the corresponding posterior probabilities for the ICP record of the given patient. These probabilities also manifested certain variability, so a representative curve needed to be determined. We proposed to define this curve as a posterior probability that had a minimum sum of L_2 deviations with respect to the all of the remaining nine curves. As a final choice, we selected hGMM parameters which corresponded to the found representative posterior probability.

3.1. Evaluation of the ROC technique for calculation of the alarm function thresholds

According to the standard approach, the optimum value TA_j^{opt} for the given set of parameters dim, M and all considered alarm functions (and their variations) is found as a value TA_j, which corresponds to the Youden index (YI) of the previously calculated ROC curve (figure 4). The optimum thresholds TA_j^{opt} for the individual cases of ROC associated with the mentioned choice of the parameters are given in table 2. The alarm function Al_3GL with the highest YI $= 0.37$ (sensitivity $Se = 0.71$ and specificity $Sp = 0.67$) is located in the table cell with dim $= 3$ and $M + M = 3 + 3$. However, we see that the corresponding optimum threshold TA_3^{opt} is very close to the boundary value zero that makes its application in practice unacceptable. Therefore, we also considered different cases for which the YI is close to the maximum value and the corresponding optimum thresholds for the alarm functions are more suitable for practice. The following alarm functions and characteristic values were obtained: for dim $= 3$, $M + M = 2 + 2$, alarm function Al_1G, YI$= 0.26$ (sensitivity $Se = 0.79$, specificity $Sp = 0.48$), and the threshold $TA_1^{\text{opt}} = 0.12$; and for dim $= 5$, $M + M = 4 + 4$, alarm function Al_2L, Youden index $YI = 0.33$ (sensitivity $Se = 0.48$, specificity $Sp = 0.85$), and the threshold $TA_2^{\text{opt}} = 0.21$. The names of four possible types of posterior probabilities, used in the alarm function definition, are substituted in the calculated tables by acronyms with the following meanings: G—global, L—local, GL—arithmetic mean of the global and local PP, MGL—maximum of the global and local PP value for the given moment.

For a case when sensitivity is preferred to the Youden index, we created table 3 in which the maximum sensitivity value is reached for the parameters dim $= 3, M + M = 2 + 2$, Al_1G

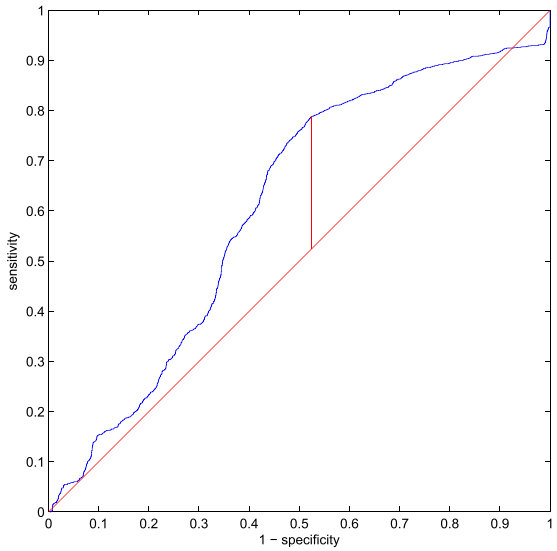

Figure 4. The plot of the ROC curve for the case with preferred maximum of sensitivity (see table 3). The corresponding parameters were: dim $= 3, M + M = 2 + 2$, Al_1G with the optimum threshold $TA_1^{\text{opt}} = 0.12$. The Youden index achieved was YI $= 0.26$ which is represented by the vertical line. The corresponding value of *sensitivity* is 0.79 and of *specificity* is 0.48.

Table 2. The results of computer experiments with hGMM models generated for the input data and for four choices of the numbers of vital and fatal clusters. They are arranged in three separate blocks, each for the optimum combinations of the alarm function and feature vector space dimension. The individual columns contain the corresponding Youden index (YI) values together with the ROC *optimum threshold* TA_j^{opt} values of the corresponding alarm function. The individual rows are associated with one of four possible combinations of the local and global approaches used within the hGMM calculations.

		hGMM (cluster numbers: $M + M$)							
		$2 + 2$		$3 + 3$		$4 + 4$		$5 + 5$	
Al_j & dim	Al-type	YI	TA_j^{opt}	YI	TA_j^{opt}	YI	TA_j^{opt}	YI	TA_j^{opt}
	G	**0.26**	**0.12**	0.12	0.64	0.05	0.60	0.16	0.74
Al_1	L	0.11	0.48	0.18	0.30	0.16	0.31	0.17	0.38
& $3D$	GL	0.11	0.15	0.17	0.43	0.09	0.42	0.17	0.42
	MGL	0.11	0.61	0.18	0.62	0.13	0.60	0.19	0.64
	G	0.22	0.03	0.19	0.01	0.22	0.35	0.08	0.65
Al_2	L	0.17	0.28	0.21	0.21	**0.33**	**0.21**	0.23	0.21
& $5D$	GL	0.19	0.03	0.23	0.15	0.24	0.29	0.19	0.32
	MGL	0.19	0.05	0.21	0.22	0.22	0.35	0.14	0.51
	G	0.15	0.01	0.33	0.02	0.02	0.41	0.18	0.61
Al_3	L	0.16	0.14	0.29	0.01	0.37	0.01	0.33	0.02
& $3D$	GL	0.14	0.24	**0.37**	**0.01**	0.15	0.20	0.19	0.38
	MGL	0.12	0.47	0.37	0.01	0.13	0.01	0.18	0.61

Table 3. The maximum ROC sensitivity (*Se*) and specificity (*Sp*) values and the corresponding optimum threshold TA_j^{opt}. For each of the six types of hGMM according to cluster number (*M* vital and *M* fatal clusters in the first column) and each of the feature space dimensionalities, 3D, 4D or 5D, there are 12 possible combinations of alarm function types (three basic forms and four combination types) which are expressed via the symbol Al_jY in the table (*Y* stands for one of the acronyms *G, L, GL*, or *MGL* mentioned in the text). All these variations were included in the ROC analysis, and the number listed in the table represents the case for which the maximum Youden index was achieved. The alarm functions are accompanied in the table rows by the corresponding values of the ROC sensitivity and specificity with the corresponding optimum threshold. The maxima of the sensitivity are enhanced in each row (corresponding to the individual choice of the cluster numbers).

hGMM	3D				4D				5D			
$M+M$	Al_jY	*Se*	*Sp*	TA_j^{opt}	Al_jY	*Se*	*Sp*	TA_j^{opt}	Al_jY	*Se*	*Sp*	TA_j^{opt}
$1+1$	Al_1MGL	0.44	0.79	0,20	Al_3GL	0.42	0.82	0.06	Al_3GL	**0.77**	0.47	0.01
$2+2$	Al_1G	**0.79**	0.48	0.12	Al_3MGL	0.78	0.39	0.01	Al_2GL	0.59	0.60	0.03
$3+3$	Al_3GL	**0.71**	0.67	0.01	Al_3L	0.41	0.86	0.06	Al_3G	0.64	0.55	0.01
$4+4$	Al_3MGL	0.54	0.58	0.01	Al_1L	0.53	0.60	0.24	Al_1L	**0.55**	0.64	0.23
$5+5$	Al_3L	**0.57**	0.76	0.02	Al_1L	0.52	0.66	0.22	Al_1GL	0.34	0.80	0.38
$6+6$	Al_2L	0.62	0.71	0.03	Al_3G	**0.69**	0.49	0.01	Al_3L	0.61	0.66	0.05

with the optimum threshold $TA_1^{\text{opt}} = 0.12$. Furthermore, in table 3 the maxima of sensitivity achieved for the individual choices of cluster numbers (table rows) are in bold. This suboptimal, but fully acceptable, case is illustrated in the following figures: figure 5(B2) for vital patient no. 9, and figure 6(B2) for the fatal patient no. 11.

3.2. Evaluation of the calculation of the alarm function thresholds based on the efficiency measurement

The aim of our computer experiments with efficiency measurement was to determine the necessary thresholds for individual novel alarm functions. We recall that the essential element of such a measurement is the relation between the real ICP alarm indication (derived by comparing the ICP value with the threshold Thr) and the novel alarm indication (derived by comparing the actual alarm value with the specific threshold TA_j). According to the recommendations (Mathew 2005), for the ICP threshold, we used the value Thr= 20 mmHg. For the experiments we established a parametric set consisting of $3 \times 6 \times (3 \times 4) = 216$ triplets of the quantities: (dim, *M*, *j*), where dimensions of the feature vector spaces are dim = 3,4,5, the number of hGMM clusters are $M = 1, 2, \ldots, 6$, and, finally, we considered three alternative novel alarm functions Al_j, $j = 1, 2, 3$, each having four versions (global, local, average, maximum). The optimization procedure consists of

- taking a triplet from the parametric set (216 possibilities);
- calculating the corresponding hGMM for features derived from ICP data;
- calculating the values of the chosen alarm functions (for each time instant);
- varying the values of the thresholds TA_j;
- calculating the corresponding values of the efficiency measure $\mu_r(20, TA_j)$ for all possible values of the thresholds TA_j; and
- finding the maximum of $\mu_r(20, TA_j)$.

Table 4. The threshold values TA_j^{opt} of the alarm functions corresponding to the maximum values of the efficiency measure $\mu_r^{\max}$ calculated for various parameters ($\dim = 3D, 4D, 5D$, $M + M = 1 + 1, \ldots, 6 + 6$), the reference ICP threshold $\text{Thr} = 20$, and for the range of all thresholds $TA_j \in (0, 1)$ and each alarm function $Al_j(k)$. The maxima of $\mu_r^{\max}$ obtained for each dimension are in bold.

hGMM	3D			4D			5D		
$M + M$	Al_jY	$\mu_r^{\max}$	TA_j^{opt}	Al_jY	$\mu_r^{\max}$	TA_j^{opt}	Al_jY	$\mu_r^{\max}$	TA_j^{opt}
$1 + 1$	Al_3L	1.58	0.54	Al_3L	1.79	0.54	Al_3L	1.85	0.58
$2 + 2$	Al_3L	2.33	0.48	Al_2L	1.81	0.59	Al_2L	2.09	0.56
$3 + 3$	Al_2L	1.61	0.21	Al_1L	2.45	0.64	Al_2L	1.5	0.27
$4 + 4$	Al_3L	1.84	0.43	Al_1G	3.64	0.99	Al_2GL	2.06	0.51
$5 + 5$	Al_2L	1.75	0.21	Al_2L	1.78	0.36	Al_2L	1.75	0.56
$6 + 6$	Al_2L	2.07	0.56	Al_1GL	2.28	0.75	Al_2L	2.05	0.59

For finding the local maxima of the measure $\mu_r(\,)$ we used the available Matlab function *peakdet* and the following heuristics:

- the function *peakdet* uses a control parameter delta for setting an appropriate scale of the local maxima; on the basis of computer experiments with various values of the delta we found the optimum value $\delta = 0.001$;
- next, we removed the local maxima of $\mu_r(\,)$ which were encountered on the borders of the range of the threshold TA_j because we regarded them as methodologically unacceptable – they would either cause a situation (at very low threshold TA_j values) with an almost permanent active novel alarm, or a situation (at very high threshold TA_j values) with almost no novel alarm possible;
- thus, for each triplet of parameters we obtained several peaks of the local maxima of the measure μ_r (on average 4.2 such peaks have been observed); the highest local maximum of the obtained peaks was taken as the final solution.

The results of the maximization procedure performed for all 216 triplets of the parameters, described above are listed in table 4. The range of the measure $\mu_r(\,)$ is $(0.3, 3.64)$ and the threshold TA_j is within the interval $(0, 0.99)$. From the table we can draw the following conclusions:

- the global maximum of the measure $\mu_r^{\max} = 3.64$ is achieved for the parameters: $\dim = 4, M = 4$, Al_1G with the optimum threshold $TA_1^{\text{opt}} = 0.99$;
- similarly to the case of the ROC optimization, we obtained a threshold value corresponding to the global maximum that is very close to the boundary value that can cause tedious and ambiguous implementation in clinical practice; therefore we analysed the table with the aim to find such values of the measure $\mu_r^{\max}$ which differ from the general maximum minimally, but which yield more acceptable thresholds TA_j^{opt};
- the following values are the best candidates for consideration: 2.45 (Al_1L ,4D, 3+3), 2.33 (Al_3L ,3D, 2+2) and 2.28 (Al_1GL ,4D, 6+6).

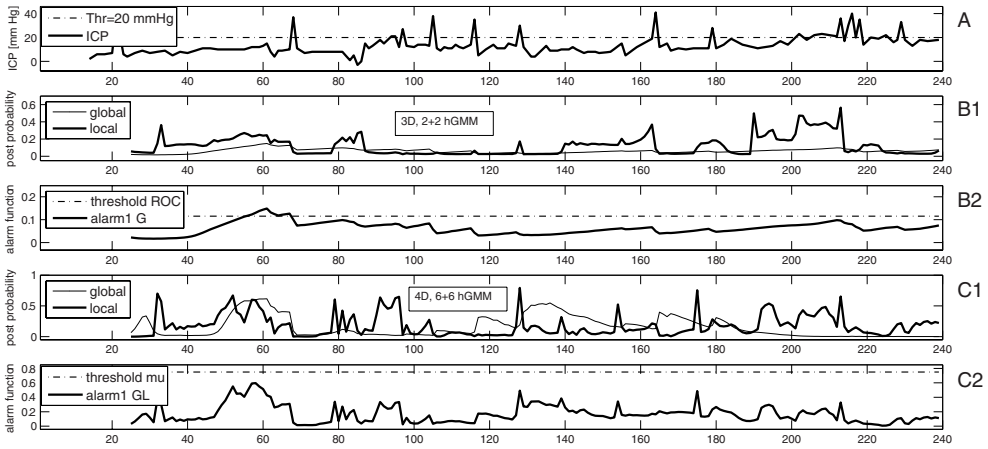

Figure 5. An example of the set of summary plots for the vital patient (no. 9).

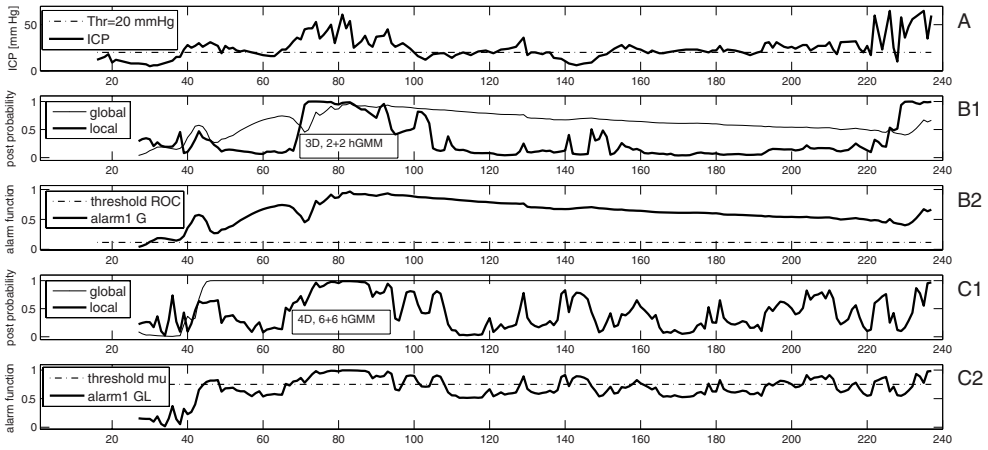

Figure 6. An example of the set of summary plots for the fatal patient (no. 11).

3.3. Graphical evaluation of particular alarm functions and thresholds

Before we characterize selected graphical outputs from our computer experiments, two important methodological aspects should be emphasized:

(i) All theoretical considerations and calculations needed for the proposed method resulted in the development of three novel alarm functions that could serve as a complementary tool for monitoring of the ICP records of real TBI patients with the aim of extending the possibilities of reasonable triggering of additional medical intervention.
(ii) Since the concept of the novel alarm functions is similar to the conventional detection of the ICP values exceeding a predefined threshold, the search for a reasonable threshold for each individual alarm function represents an inherent part of the proposed alarm function methodology.

In figures 5 and 6 we illustrate two examples in which the results achieved for the optimization approach based on the ROC analysis are grouped with the results achieved by

optimization via the efficiency measure, μ_r, for one representative case of a vital patient (no. 9) and one representative case of a fatal patient (no. 11). The organization of the plots in these pictures is identical, and it can be described jointly as follows:

(i) The first box, A, plots the basic ICP record with the commonly established threshold Thr= 20 mmHg.
(ii) The second box, B1, plots for the posterior probabilities calculated for the 3D result, and the hGMM clusterization with $2+2$ clusters are depicted: one curve for the global approach, and one for the local approach.
(iii) The third box, B2, shows the plot of the optimum alarm function Al_1G derived from the posterior probabilities given in B1, together with its threshold $TA_1^{\text{opt}} = 0.12$, found for the abovementioned parameters.
(iv) In the fourth box, C1, the plots of the posterior probabilities calculated for dim $= 4$ and the hGMM clusterization with $6+6$ clusters are depicted using the global and local approach; again, one curve depicts the global approach, and the other the local approach.
(v) The fifth box, C2, represents the plot of the optimum alarm function Al_1GL derived from the posterior probabilities given in C1, together with its threshold $TA_1^{\text{opt}} = 0.75$, calculated via the efficiency measure $\mu_r(\,)$ ('mu').

The illustrated plots of the novel alarm functions, as well as their thresholds, reveal the following particular findings:

- In the case of vital patient no. 9 (figure 5), despite the fact that the plot of the ICP record often exceeds the conventional threshold, the patient survived; though it is very likely that some of the triggered alarms contributed to the improvement of the patient's health condition, a number of alarms could still be false, thereby leading to an unnecessary high clinical burden on the patient; on the other hand, both novel alarms show the alarm function values under the critical threshold for most of the time instants.
- In the case of fatal patient no. 11 (figure 6), similarly to the preceding vital case, a number of alarms should have resulted in triggering of the medical intervention procedure, but this time, unfortunately, without the favourable effect on the patient's survival. Therefore a question arises as to whether there are some additional alarms indicated by novel alarm functions; in the first case (B2), obtained by application of the ROC analysis, we see that for many time instants when ICP values are below the 20 mmHg threshold, the novel alarm function Al_1G yields values over the threshold $TA_1^{\text{opt}} = 0.12$ indicating an alarm state, in the second case (C2), obtained through the efficient measure$\mu_r(\,)$ maximization, there are only time instants round the 140 (± 5) h for which we get values of the novel alarm function Al_1GL over the threshold $TA_1^{\text{opt}} = 0.75$.

3.4. Summary of the results

The optimum values of the ROC method applied to the entire data set are achieved for lower dimensionality (3D) and the lower clustering model ($2+2$ clusters of the hGMM). Moreover, its advantage is in model simplicity, lower computational costs and higher stability in comparison with the maximization of the efficiency measure $\mu_r(\,)$ (4D, $6+6$ hGMM clusters).

The ROC analysis of the entire retrospective ICP data set considering possible values of the conventional ICP threshold (section 2.3) yielded the optimum threshold Thr= 19.9 mm Hg for Youden index YI= 0.19, with sensitivity Se = 0.33 , and specificity Sp = 0.86. This result is in good agreement with the clinically established threshold Thr= 20 mm Hg. However, by

the application of our alternative alarm functions to the same ICP data set based on the proposed novel method (section 3.1) we reached better results of the ROC analysis, i.e. Youden index YI= 0.33 (sensitivity Se = 0.48 and specificity Sp = 0.85) and Youden index YI= 0.26 (sensitivity Se = 0.79 and specificity Sp = 0.48).

4. Conclusions

(i) The measured ICP data have been thoroughly analysed together with the accompanying information on medical intervention. No significant values of correlation between ICP data and 12 clinical parameters were observed. Despite the high overlap between ICP data sets for vital and fatal patients, the results of the nonparametric Kruskal–Wallis test proved that these two sets are significantly different ($p < 0.001$).

(ii) We studied various clustering schemes defined on the basis of different types of segmentation, grouping and representing the individual ICP records considered as time series. Keogh and Lin (2005) carried out a comprehensive survey of 'sliding window' methods applied to TS resulting in the conclusion that 'clustering of time series subsequences is meaningless'. We have tested and analysed Keogh's study. We proposed a novel ICP subsequence representation using vector features in the four-, five- and six-dimensional vector spaces that make the objections of Keogh *et al* not applicable to our clustering scheme.

(iii) We formulated a specific clusterization problem for feature vector spaces. In section 2.4 we proposed six numerical features defined for each time segment of ICP.

(iv) We proposed the use of a hierarchical Gaussian mixture model (hGMM) for clustering the extracted features which represent ICP time subsequences. The model defines independent clusters for vital and fatal patients. This is the advantage of the hGMM model, since the basic GMM model (albeit they both are supervised) does not provide such a possibility and an extra procedure has to be developed for labelling output clusters into vital and fatal classes.

(v) We proposed combining a global approach with a local approach of feature generations in the following way: (1) in the global approach, a feature is defined on a discrete support of time instants (indices) with the fixed beginning where the ICP record start; (2) in the local approach, a feature is defined on a discrete support of time instants with a width of 12 indices, ending at the moment where the last ICP value is recorded. Although the shifting of the individual supports resembles the sliding window approach, it transforms the ICP values into a low-dimensional feature space instead of dealing with the raw ICP values.

(vi) We proposed the following procedure for processing the real ICP values recorded during the TBI treatment of a patient: for any time moment of the ICP record, the feature vectors for all the ICP subsequences preceding and associated with the given time moment are calculated. Next, posterior probabilities of hGMM are calculated over the time.

(vii) In section 2.6 three alarm functions were proposed which were later explored by extensive computer experiments.

(viii) In section 2.7 we developed two novel techniques: (i) the first one is related to the calculation of the optimum thresholds TA_j^{opt} for the proposed alarm functions via the ROC analysis (that is not standard in using specific relations for definition of true-positive and false-positive rates), and using the Youden index; (ii) the second technique

consists of determination of the thresholds TA_j using a suitable efficiency measure. The results obtained by each of these two alternative techniques were evaluated in sections 3.1 and 3.2 and grouped in the corresponding tables 2–4.

(ix) The research study reported in this paper has a limitation relating to the extent of the input retrospective ICP data. This was caused by the limitation of the clinical archive available to us. However, it should be emphasized that the essential points of the proposed and explored methodology are independent of the magnitude of the processed data. To achieve a stronger generalization of the obtained results, the developed algorithms can be applied to any much greater ICP data cohort that was beyond the feasibility of this study.

(x) In relation to the cost of the necessary computational operations, there are two aspects that should be mentioned: (i) first, it is obvious that the most demanding calculations are related to the preparation of the hGMM for the whole set of the retrospective ICP data, however, these are needed just once (in offline mode) and therefore do not influence the ICP monitoring itself; (ii) second, having all the posterior probabilities ready-to-use (according to the number of the chosen clusters within the already found hGMM), the calculation of the instant value of the alarm function (in real time) for the ICP record being evolved represents a very simple operation.

(xi) It is not possible to interpret the proposed novel alarm functions directly in the view of the characteristics of the clinical state of the patients with severe TBI. In the following, we describe our attitude towards understanding of the possible use of the proposed alarm functions in the context of future clinical practice.

4.1. Proposal of a prospective ICP monitoring technique.

On the basis of the research results, we propose to introduce a monitoring technique that could extend conventional ICP measuring itself. Such an extended technique might be helpful to clinicians in making a more sophisticated and efficient decision on necessary consecutive intervention steps at any moment of TBI patient treatment. First of all, we assume the availability of a new extensive validation of the proposed methodology on an ICP data cohort of a greater size. This activity can be carried out in offline mode (A). Then a suitable protocol (B) has to be developed, *including the conventional ICP alarm*, completed by *novel alarm functions*. The protocol could be applied in the required online mode when monitoring actual TBI patients. We describe these two steps in detail.

A

(i) Our ICP data set was of limited size especially in regard to fatal cases that necessitated using the principle of cross-validation. This principle could be avoided by acquisition of a larger dataset. In such a case it is necessary, first, to set apart data that will be used exclusively for the calculation of the chosen hGMM of ICP data, while the remaining part of ICP data (testing data) will be used merely for the calculation of the selected alarm functions and subsequently for searching for the optimum values of the alarm function thresholds. Similar experimentation with the parameter set of triplets, as was proposed in the paper, would be needed.

(ii) For determination of the optimum thresholds of the alarm function(s), which would serve as characteristics of a new ICP data cohort, two possibilities are available: the ROC approach using the Youden index, or the approach using efficiency measure $\mu_r(\,)$. As the former approach is less computationally demanding, we recommend using calculations needed for finding the Youden index of the ROC for each of the

three alarm functions (section 2.7.1). The final decision on selecting clinically acceptable values of the optimum set of parameters found in the experimental offline part of processing will depend on the interdisciplinary expertise of clinicians and experts in machine learning.

B

(i) The conventional alarm function, consisting of comparing the actually evolving ICP record with the established ICP threshold, could be extended by a novel alarm function and its threshold. Provided that all needed parameters from the **A** step are available, for the ICP record, existing in the time instant *k,* the calculations of the corresponding posterior probabilities for the fixed hGMM can be carried out. Then the value of the selected novel alarm function could be computed and compared to its threshold known from the **A** step. All these calculations can be performed in real time and simultaneously with the ICP measurement. Each additional alarming situation should be evaluated by a clinician before triggering additional urgent medical intervention.

(ii) All necessary calculations could be automatized and implemented as an extension of the standard software used in the ICP monitoring process.

(iii) All interventions, evoked either by the exceeding of the conventional ICP threshold, or the novel alarm(s) threshold(s) should be strictly recorded—not only due to legal reasons, but also with the aim of accumulating useful data for further improvement of TBI patient treatment.

Future research

As the method based on the Youden index is the most frequently used one in biomedical statistical applications, we adopted it in our explorations. Nevertheless, for future research it would be helpful to also consider complementary approaches (Leeflang *et al* 2008), because the applied method does not consider any costs of wrong classifications or benefits of correct classifications. For some purposes, higher sensitivity may be more important than a higher specificity (or vice versa). A pre-specified property of a cut-off point that is relevant to the context of the task, in which the test will be applied, could be preferred. For example, a cut-off could be based on whether optimizing sensitivity or specificity has greater practical value related to the ICP monitoring of the TBI patient (Xinhua 2012). Obviously, this problem represents an open issue of the prospective research in the given field, for which the close collaboration of clinicians with experts in machine learning is inevitable.

Acknowledgments

This work has been supported by the Slovak Grant Agency for Science (projects VEGA 2/0138/16, VEGA 2/0011/16) and by the MZ 2013/56-SAV-6, APVV-0668-12, and APVV-14-0875 grants. We appreciate the helpful assistance of Prof Dr Med. Walter Mauritz, PhD, primarius of the Lorenz Böhler Unfallkrankenhaus Wien, Vienna, Austria, who provided us with an archive data set of ICP records. We appreciate the valuable comments and recommendations of the anonymous reviewers which made it possible to considerably increase the technical quality and readability of the paper. We are grateful to the postdoctoral fellow Dr Radoslav Skoviera who carried out computer experiments aimed at testing Keogh's (Keogh and Lin 2005) claim that 'clustering of time series subsequences is meaningless'.

References

Arabi Y, Haddad S, Tamim H, Al-Dawood A, Al-Qahtani S and Ferayan A 2010 Mortality reduction after implementing a clinical practice guidelines-based management protocol for severe traumatic brain injury *J. Crit. Care* **25** 190–5

Balestreri M *et al* 2004 Intracranial hypertension: what additional information can be derived from ICP waveform after head injury? *Acta Neurochir. (Wien)* **146** 131–41

Bishop C 2006 *Pattern Recognition and Machine Learning* (New York: Springer)

Bullock M R and Povlishock J T 2007 Guidelines for the management of severe traumatic brain injury *J. Neurotrauma* **24** S1–106

Bullock M R, Chestnut R M, Clifton G, Ghajar J, Marion D W and Narayan R K 1996 Guidelines for the management of severe head injury *J. Neurotrauma* **13** 643–734

Carney N *et al* 2017 Guidelines for the management of severe traumatic brain injury, fourth edition *Neurosurgery* **80** 6–15

Chambers I R, Treadwell L and Mendelow A D 2001 Determination of threshold levels of cerebral perfusion pressure and intracranial pressure in severe head injury by using receiver-operating characteristic curves: an observational study in 291 patients *J. Neurosurg* **94** 412–6

Chesnut R *et al* 2014 Intracranial pressure monitoring: fundamental considerations and rationale for monitoring *Neurocrit. Care* **21** S64–84

Fakhry S M *et al* 2004 IRTC neurotrauma task force: management of brain injured patients by an evidence-based medicine protocol improves outcomes and decreases hospital charges *J. Trauma* **56** 492–9

Fawcett T 2006 An introduction to ROC analysis *Pattern Recognit. Lett.* **27** 861–74

Fu T 2011 A review on time series data mining *Eng. Appl. Artif. Intell.* **24** 164–81

Gama J A 2012 A survey on learning from data streams: current and future trends *Prog. Artif. Intell.* **1** 45–55

Ghajar J 2000 Traumatic brain injury *Lancet* **356** 923–9

Guiza F *et al* 2015 Visualizing the pressure and time burden of intracranial hypertension in adult and paediatric traumatic brain injury *Intensive Care Med.* **41** 1067–76

Hesdorffer D *et al* 2002 Predictors of compliance with the evidence-based guidelines for traumatic brain injury care: a survey of United States trauma centers *J. Trauma* **52** 1202–9

Honda M *et al* 2017 Consideration of the intracranial pressure threshold value for the initiation of traumatic brain injury treatment: a xenon CT and perfusion CT study *Neurocrit. Care* accepted (http://doi.org/10.1007/s12028-017-0432-5)

Indrayan A 2013 *Medical Biostatistics* 3rd edn (London: Chapman and Hall)

Jain A K 2010 Data clustering: 50 years beyond K-means *Pattern Recognit. Lett.* **31** 651–66

Jun-Yu-Fan R N, Kirkness C, Vicini P, Burr R and Mitchell P 2010 An approach to determining intracranial pressure variability capable of predicting decreased intracranial adaptive capacity in patient with traumatic brain injury *Biol. Res. Nurs.* **11** 317–24

Keogh E and Lin J 2005 Clustering of time-series subsequences is meaningless: implications for previous and future research *Knowl. Inf. Syst.* **8** 154–77

Kuncheva L I and Rodriguez J J 2013 Interval feature extraction for classification of event-related potentials (ERP) in EEG data analysis *Prog. Artif. Intell.* **2** 65–72

Lee H J *et al* 2016 Morphological feature extraction from a continuous intracranial pressure pulse via a peak clustering algorithm *IEEE Trans. Biomed. Eng.* **63** 2169–76

Leeflang M M G, Moons K G M, Reitsma J B and Zwinderman A H 2008 Bias in sensitivity and specificity caused by data-driven selection of optimal cutoff values: mechanisms, magnitude, and solutions *Clin. Chem.* **54** 729–37

Maas A I *et al* 1997 EBIC-guidelines for management of severe head injury in adults *Acta Neurochir. (Wien)* **139** 286–94

Marmarou A *et al* 1991 Impact of ICP instability and hypotension on outcome in patients with severe head trauma *J. Neurosurg.* **69** 15–23

Marshall L F *et al* 1979 The outcome with aggressive treatment in severe head injuries. Part I: The significance of intracranial pressure monitoring *J. Neurosurg.* **50** 20–5

Mathew J 2005 Intracranial pressure monitoring: vital information ignored *Indian J. Crit. Care Med.* **9** 35–41

MingLiu C E and Beiqian D 2002 Hierarchical Gaussian mixture model for speaker verification *Proc. 7th Int.Conf on Spoken Language Processing* pp 1353–6

Nabney I 2002 *NETLAB: Algorithms for Pattern Recognition* (Berlin: Springer)
Narayan R K *et al* 1982 Intracranial pressure: to monitor or not to monitor? A review of our experience with severe head injury *J. Neurosurg.* **56** 650–9
Novak D *et al* 2004 Clustering of intracranial pressure using hidden Markov models *17th European Meeting on Cybernetics and Systems Research (Vienna, 13–16 April 2004)* pp 1–10
Palmer S *et al* 2001 The impact of outcomes in a community hospital settging using the AANS Traumatic Brain Injury guidelines *J. Trauma* **50** 657–64
Steyerberg E W *et al* 2008 Predicting outcome after traumatic brain injury: development and international validation of prognostic scores based on admission characteristics *PLOS Med.* **5** 1251–61
Teasdale G M and Jennett B 1964 Assesment of coma and impaired consciousness: a practical scale *Lancet* **2** 81–4
Vukic M, Negovetic L, Kovac D, Ghajar J, Glavic Z and Gopcevic A 1999 The effect of implementation of guidelines for the management of severe head injury on patient treatment and outcome *Acta Neurochir (Wien)* **141** 1203–8
Xinhua L 2012 Classification accuracy and cut point selection *Stat. Med.* **31** 2676–86
Zhang X *et al* 2017 Invasive and noninvasive means of measuring intracranial pressure: a review *Phys. Meas.* **38** R143–82

Feature Clustering of Intracranial Pressure Time-Series for an Alarm Function Estimation in Traumatic Brain Injury

Part 2 : Electronic supplement

M Teplan[1], I Bajla[1], R Rosipal[1] and M Rusnak[2]

[1]Institute of Measurement Science, Department of Theoretical Methods, Slovak Academy of Sciences, Bratislava, Slovakia,

[2]Department of Public Health, Trnava University, Trnava, Slovakia

E-mail: michal.teplan@savba.sk

Abstract. This is the supplementary part to the article "Feature Clustering of Intracranial Pressure Time-Series for an Alarm Function Estimation in Traumatic Brain Injury". This part represents an extension of selected items and ideas presented in the main text, based on reviewers recommendations. It comprises necessary references to the corresponding positions in the main text, and vice versa, necessary references to supplement text are given in the main text.

1. Introductory information for the electronic supplement

Additionally to the subgoals of our research study (mentioned in Section 1.1 of the main text) we set two complementary subgoals:

- to avoid problems caused by using sliding window principle for subsequence generator, as addressed in the paper of Keogh & Lin (2005), we set the next additional goals: i) testing this principle in the initial ICP domain, as well as using various simulated signals, ii) transforming the original ICP-time domain to a *multi-dimensional* feature vector space, and testing the principle in the new setting,
- based on the reviewer's recommendation, we amend the paper text by the figure examples, where the alarms do not behave ideally as expected; in particular, we provide the plots of posterior probabilities and alarm functions for a surviving patient with a relatively long period of elevated ICP over the established threshold, as well as the plots for a fatal patient whose ICP record manifested none of time instants with the elevated ICP value above the threshold.

2. Methods - ICP data statistics

2.1. Statistics of various characteristics of retrospective ICP data

The plots of typical ICP data for fatal and vital cases are presented in Fig. 1 and Fig. 2. Along with the ICP measurement, four different types of intervention have been recorded - administration of the insulin (Ins), norepinephrine (NorEp), Manitol, and Hypersalinity. In our calcultaions, for single patient correlation, only data with at least 20 data points were taken into acount (available from 22 up to all 40 patients, depending on the selected parameter). Thus, for Manitol and Hypersalinity no correlation could be calculated, as these two types of intervention data were too sparse. Moreover, six diagnostic parameters have been recorded as well: heart rate (HR), body temperature (BT), PaO2, PaCO2, % of glucose (Gl) and lactate (Lac). The interventions have been characterized by a total inserted amount of the intervening substance administered during each hour.

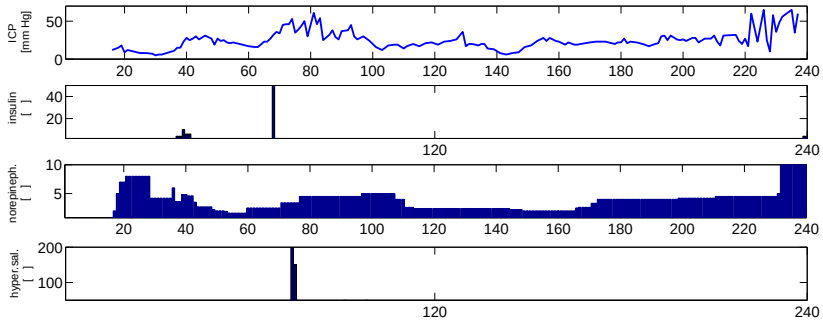

Figure 1. A sample of ICP time record of a fatal patient together with the time course (in hours) of the interventions.

We have computed correlation of ICP data with each of 8 above mentioned parameters (Tab.1). No significant values of correlation have been observed.

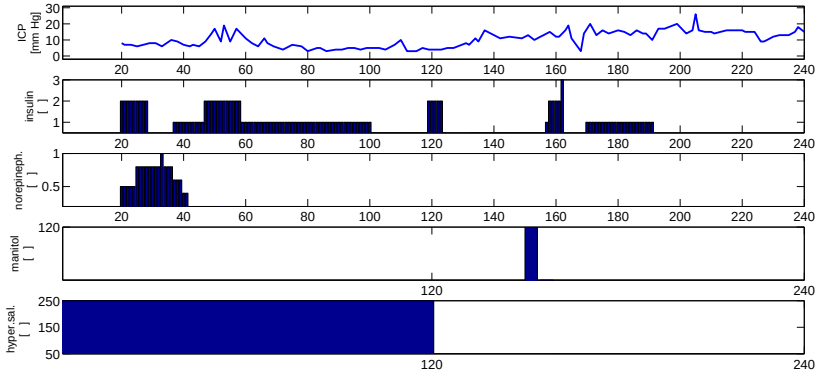

Figure 2. A sample of ICP time record of a vital patient together with the time course (in hours) of the interventions.

variables	intervention type		physiological parameters					
	Insulin	NorEp	HR	BT	PaO2	PaCO2	%Gl	%Lac
correlation	0.00	-0.06	0.07	0.09	0.01	-0.09	-0.03	-0.13
stdv of corr.	0.22	0.30	0.26	0.25	0.24	0.25	0.20	0.26

Table 1. The values of individual correlation coefficients between ICP and variables (the first two types of intervention and six physiological parameters).

To support a clinical decision on the subsequent treatment of TBI patients, several TBI scores of various types are used as auxiliary information to the measured ICP values. For our retrospective study the following scores were available: APACHE II Score, APACHE II mortality, GCS b. A.: total, SAPS II Score, SAPS II mortality, ISS, TRISS survive. We have analysed all ICP records for fatal and vital patients in the view of relations between individual scores and the basic record attributes (fatal or vital outcome). We have found no unique indication of the fatal case by the critical score value, i.e., the subset of the fatal ICP records could not be separated from the set of the vital ICP records based on any score. Therefore we hypothesize, that clinicians undertook intervention action whenever they felt, that the development of the patient state had satisfied conditions fixed in a conventional protocol (including ICP threshold and practically independently on TBI score values). Thus our analysis has been accomplished solely on ICP data, without taking into account any additional information mentioned above.

The plots of normalized histograms in Fig.3 demonstrate several facts: i) taking into account the standard information that the critical ICP value threatening the survival of TBI patients $Thr = 25$ mmHg, we can see that in our retrospective cohort of the ICP data the values from the interval $< 25{,}45 >$ occur several times in the set of vital patients, and there is negligible number of the values greater than 45 present in the same set. It means, the ICP overun of the *Thr* cannot be in any case declared as an absolute indicator of approaching terminal decease of the patient.

Note that this imbalance is inherent in the whole set of the available ICP data that reflects approximately the overall priors of these two categories of treated cases. The proportion of the ICP samples: fatal/vital patients is approximately 1/3.2, therefore, first, we had multiplied the absolute frequencies of the fatal ICP samples by this coefficient. Then we calculated two histograms of

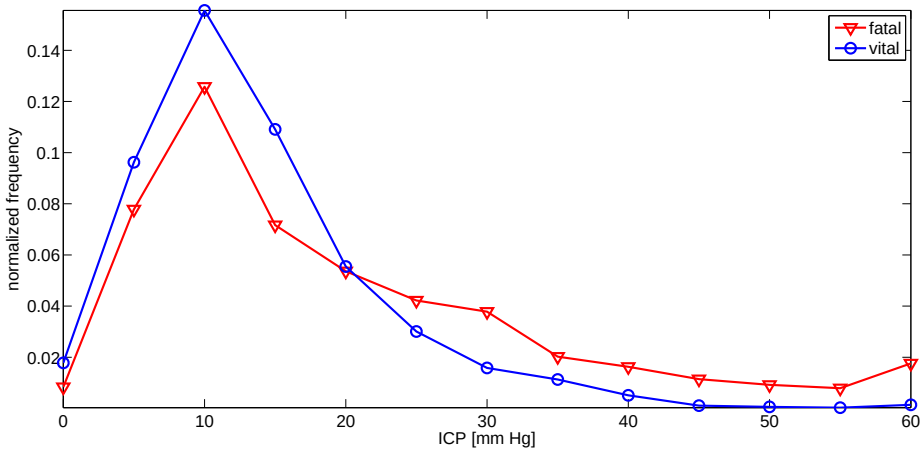

Figure 3. The normalized histograms of the ICP values incident in the set of ICP records included in the retrospective study. Red: for fatal cases, Blue: for vital cases.

relative frequencies for both cases by their normalization to the overall number of the ICP samples in the available input set. In Fig.3, these histograms are depicted separately for vital and fatal patients. From the histograms, it can be seen that for the smaller ICP values - up to the value of 20 mmHg - more frequent values originate from vital patients. On the other hand, the higher ICP values correspond to the fatal patients. This fact is in full accordance with the situations in clinical practice with the patients sufferred from severe TBI.

Finally, in spite of a high overlap between ICP data records for vital and fatal patients, it can be stated that according to the results of the nonparametric Kruskal-Wallis test, these two sets are significantly different ($p < 0.001$). This test has been used, because neither of the data sets of the two categories obeyed normal distribution.

2.2. ROC for the conventional criterion of comparing the ICP values to threshold values Thr - Add on

To our best knowledge, we did not find any specific example in the available literature that would provide ROC analysis of the **conventional criterion** of $ICP(k) > Thr$ for particular ICP data of a clinical (retrospective) study. Therefore, we decided, to carry out the ROC analysis of all ICP records from our retrospective data set. For the ICP threshold values Thr_v for $v = 1, 2, \ldots, 10$, we selected the following set of ICP integer values (mmHg): {5,10,15,20,25,30,35,40,45,50} which, for completeness, span the extremum ICP values, though there are out of clinically permitted range. In this situation, the vital and fatal cases, as well as prediction outcomes are associated with the values of $ICP(k)$ for all the individual indices k encountered in the set of ICP records. The details on the results of this ROC analysis are provided in the main paper body, however, for the sake of terminology unambiguity we add to the supplement the following basic ROC terms definitions.

(i) If Se and Sp denote sensitivity and specificity, respectively, the distance between the point [0,1] and any point on the ROC curve (on the plot of real-valued pairs $[x, y] = [1 - specificity, sensitivity]$) is $d = \sqrt{[(1 - Se)^2 + (1 - Sp)^2]}$. To obtain optimal cut-off point that

defines the optimum threshold value of the given alarm function, it is necessary to calculate the distance for each observed cut-off point of the ROC, and locate the point where the distance d is minimum.

(ii) The second method is the so-called *Youden index* J that maximizes the vertical distance from the line of equality (diagonal line) to the point $[x, y]$. The main aim of Youden index is to maximize the difference between True Positive Rate (Se) and False Positive Rate ($1 - Sp$). The little algebra yields $J = max(Se + Sp)$.

2.3. Representation of the ICP original data by means of vector features

In Section 2.4 we introduced a set of seven vector features based on the ICP values. We introduced these quantities by reasons of avoidance of the problems related to the sliding window principle in time-series clusterization. After having finished preliminary experiments with the testing of the Keogh'claim (Keogh & Lin (2005)) (see the next section), we eventually confined ourselves to only five most informative features $\{f_k^2, f_k^3, f_k^4, f_k^5, f_k^7\}$.

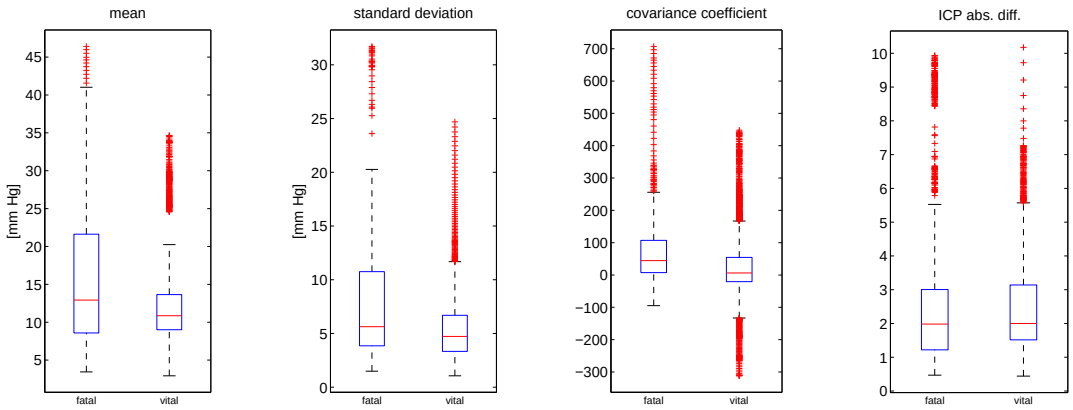

Figure 4. The boxplots calculated for four main features: *mean, standard deviation, covariance cofficient,* and *ICP absolute difference* calculated for the support S_k. All subsequences of the available ICP records were used. In each boxplot the separate characteristics of the fatal and vital records are depicted.

To illustrate the behavior of the most relevant real-valued features, we selected four selected features, namely: *mean value, standard deviation, covariance coefficient,* and *ICP absolute difference*. Then we calculated the feature values for the subsequences S_k of vital and fatal categories of the ICP records and displayed the ranges by means of boxplots in Fig.4. For the first three features the means calculated for the subsequences of vital and fatal categories differed significantly (Kruskal-Wallis test, $p < 0.001$). Whereby, the feature mean values for the fatal category were higher than feature mean values for the vital category. In the displayed boxplots, three features - mean, standard deviation, and covariance coefficient showed partial data separation ($p < 0.001$). Having defined two m-dimensional feature vectors for the subsequences $\{S_k\}$and $\{L_{k-s}\}$, generated for every ICP record, we can proceed to the question of selecting an appropriate clustering scheme in this feature space.

2.4. Examination of the "sliding window"principle applied to TS according to Keogh's results

As outlined in Keogh & Lin (2005), clustering of subsequences of individual or streamed TS, extracted via a "*sliding window*" technique, has received much attention. Based on a comprehensive survey of methods in the domain of data-mining and machine learning, the authors exposed this technique to a scrupulous analysis with several unexpected conclusions confirming their basic claim, that: "*clustering of time series subsequences is meaningless*". Naturally, before starting our research into clustering the specific type of TS, namely ICP records, we were interested in three points: i)validating the conclusions of Keogh et al in the case of simulated synthetic data they used, ii)analyzing our specific situation with respect to their claims, iii)examination of our ICP clustering scheme, that will differ from sliding window method used in their research.

2.4.1. Testing the eponymous pattern data of cylinder, bell, an funnel

After adopting three basic definitions of *time series* (*TS*), *subsequence, and sliding window* (Keogh & Lin (2005)), we concentrated on a simulated data set with three categories of input TS, i.e., data synthetically generated as real-valued vectors (with the dimension 128) having the patterns of cylinder, bell, and funnel. Besides the calculations discussed by Keogh et al, we decided to accomplish also a cross-validation scheme for a classification task defined for segmented (sTS) data into three classes. For each class (consisting of 30 samples of the given pattern), 2/3 of the generated TS data were used for training, and the rest for testing. For a streamed TS (individually generated training vectors have been concatenated into a final (128×60) dimensional streamed vector. First, we applied the sliding window procedure to this long vector (for the window length values $w = 16, 32, 64, 128$) and generated sets of sTS subsequences. Afterwards, the standard k-means ($k = 3$) clustering was performed on every set of (training) subsequences corresponding to the selected window length. We observed the sinusoidal shape of centroids of the three obtained clusters. Consequently, the claim of Keogh related to the independence of obtained sinusoidal centroids of input sTS data, was confirmed with minor deviations for all the sliding window lengths. However, the claim has been confirmed without any exception, if we enlarged the simulated data sets up to 2000 training samples for each class. Since in the paper Keogh & Lin (2005) no information was given related to any real task in which a sTS method had been used, we were also interested in a question, how the sinusoidal shape of cluster centroids can affect results of the actual TS segments classification task. Although the claim of Keogh on meaningless clustering of sTS applied to streamed TS is relevant, we argue that cannot be taken in an absolute sense, as the values we obtained for classification accuracy (CA) were greater than simple 1/3 for a random classification. Additionally, we went through the streamed TS scheme to the application of the sliding window separately to the individual simulated TS vectors. For the larger set of training samples (2000), the independence of sinusoidal centroids of all three clusters has been completely confirmed for all sliding window lengths w. We note that the obtained values of CA, averaged over three clusters, were as follows: $CA = 0.17$ (for $w = 16$),$CA = 0.46$ (for $w = 32$), $CA = 0.62$ (for $w = 64$), and $CA = 0.76$ (for $w = 128$). This results justify us to affirm that the classification realized for the given STS sets with sinusoidal centroids is meaningless only for the very small window length in relation to the dimension of the initial input TS.

2.4.2. Analysis of the STS of ICP data of the retrospective study

Keogh et al (Keogh & Lin (2005)) state, if we run sTS k-means on any dataset T (with the length m) with an overall trends of 0, with $k = 1$, we will always end up with a horizontal line as the

cluster centre (centroid). Therefore, we would like to test this claim on our specific data, i.e., on a set of ICP records. More specifically on a set of the sTS derived by a sliding window method with $w \ll m$. First of all, we should emphasize that in the case of ICP TS, the application of the sliding window technique to a streamed TS, generated from all the available ICP records, does not come on force. The reason is that each ICP record uniquely originates from one individual patient and any approach to clusterization of any time segments of the ICP has to preserve this uniqueness of the origin. First, we calculated zero-mean (centered) ICP records. Based on preliminary experiments with clusterization of the dataset using various values of w, we selected the value $w = 18$ on the empirical basis. Further, we performed sliding window procedure for each ICP record separately.

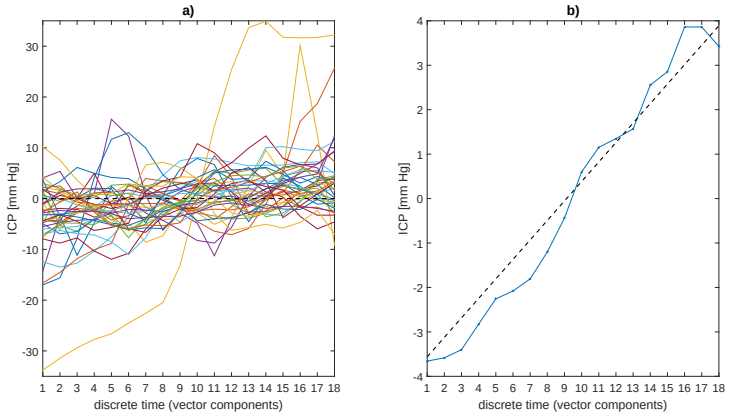

Figure 5. a) The plots of centred (zero-mean) ICP values generated by the sliding window width w=18 as subsequences and averaged over the whole ICP record. **b)** The plot of the global mean of these subsequences calculated over the all ICP records.

In Fig.5a, the obtained values of the centred subsequences for all ICP records (18-D real-valued ICP vectors) are displayed as individual broken lines. Each line represents the data derived from the ICP record of one patient. Each ICP component $(1 \div 18)$ has been calculated as an average value of the values of the same component considered over the whole set of the TS subsequences for the given ICP record. The global mean vector of the ICP subsequences for the given window, illustrated in Fig.5b, is not a tilted straight line. Finally, we can observe that after performing k-means clustering of the obtained sTS data (using the windov length $w = 18 \ll 240$) into two categories of the assumed fatal and vital patients, the nonlinear shapes of the obtained cluster centroids (18-D vectors) do not satisfy the claim of Keogh, i.e. they did not end up as a horizontal line.

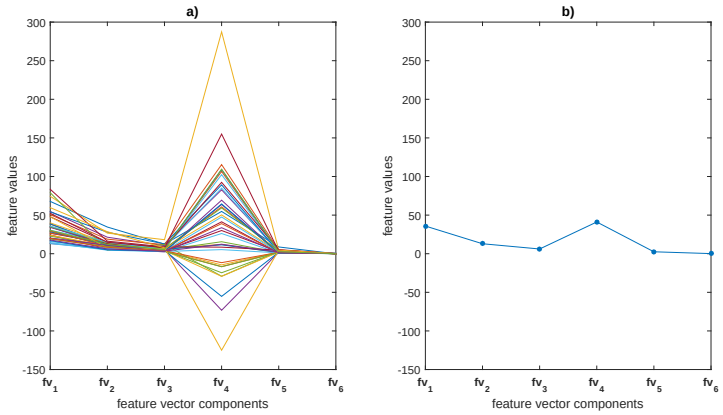

Figure 6. **a)** The plots of the 6-D feature vectors components derived from the initial ICP values based on a system of time segments with fixed beginning and various length, **b)** the plot of the global mean vector of the feature vectors calculated by averaging of each feature vector component over all ICP subsequences.

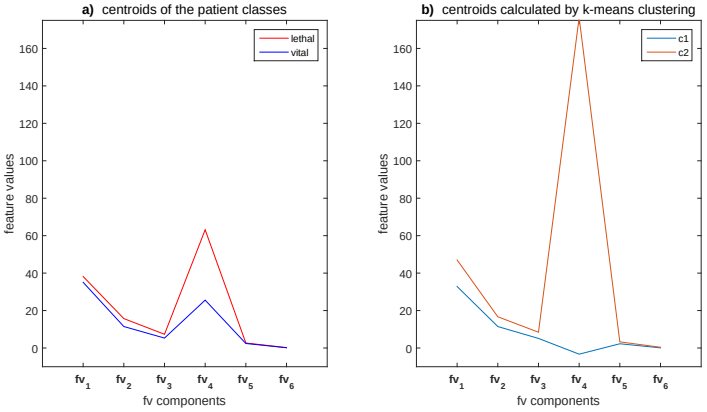

Figure 7. The plots of the centroids of the clusters of the feature spaces.

Now we can recall the Keogh's claim presented in the publication Keogh & Lin (2005) on segmented TS we mentioned above. The claim states the meaningless of clustering the sTS generated by the sliding window approach. In comparison to the scheme described in the paper, we have already showed one difference to our approach. Namely we do not use streamed TS, i.e., the sliding window has to be applied always to an individual ICP record of the patient separately. We proposed a different method that served a basis for our research into subsequent ICP time segment clustering. Instead of the original ICP values we represent each ICP segment by its m-D feature vector (as defined above). As the seventh feature represents just point value of the ICP in its time record and has nothing to to with the sliding window principle, for our experiments we used merely first six features. In the next two figures, the averaged values of the individual feature components are displayed for all ICP records. In Fig.6a each broken line represents one patient and the particular

numbers for the indices (1÷6) represent mean values of the corresponding component calculated for all segments pertinent to the given patient. The global mean vector of the feature vectors obtained for all ICP subsequences is illustrated in Fig.6b.

Finally, we calculated centroids of two clusters: - for the clusters apriori given according to the vital and fatal outcomes of ICP monitoring - and for the clusters found by k-mean clustering method applied to the set of features values. The cluster centroids of both clusters are depicted in Fig.7, the results of sTS clusterization, accomplished according to our TS segmentation method without sliding window technique, are significantly different from those declared in the paper Keogh & Lin (2005).

2.5. Graphical evaluation of particular alarm functions and thresholds

In addition to Figures 8 and 9, presented as samples of simulated monitoring with "appropriate" or advantageous performance of several new alarm functions, we illustrate in the graphical form two interesting cases of posterior probabilities and alarm functions. These are, the plots for the survived patient (N.45), whose ICP record has longer periods of elevated ICP over the threshold, and the plots for the deceased patient (N.7) having the ICP record with zero or minimum number of episodes overrunning the ICP threshold.

The organization of the plots in the following pictures is identical as in the paper text, and it can be described jointly as follows:

(i) The first box - A, involves the plot of the basic ICP record with the commonly established threshold $Thr = 20$ mmHg.

(ii) In the second box - B1, the plots of the posterior probabilities calculated for the dimension 3D, and the hGMM clusterization with 2+2 clusters are depicted: one curve for the global approach, and one for the local approach.

(iii) The third box - B2, represents the plot of the optimum alarm function Al_1G derived from the posterior probabilities given in B1, together with its threshold.

(iv) In the fourth box - C1, the plots of the posterior probabilities calculated for the dimension 4D, and the hGMM clusterization with 6+6 clusters are depicted; again, one curve belongs to the global approach, and another to the local approach.

(v) The fifth box - C2, represents the plot of the optimum alarm function Al_1GL derived from the posterior probabilities given in C1, together with its threshold.

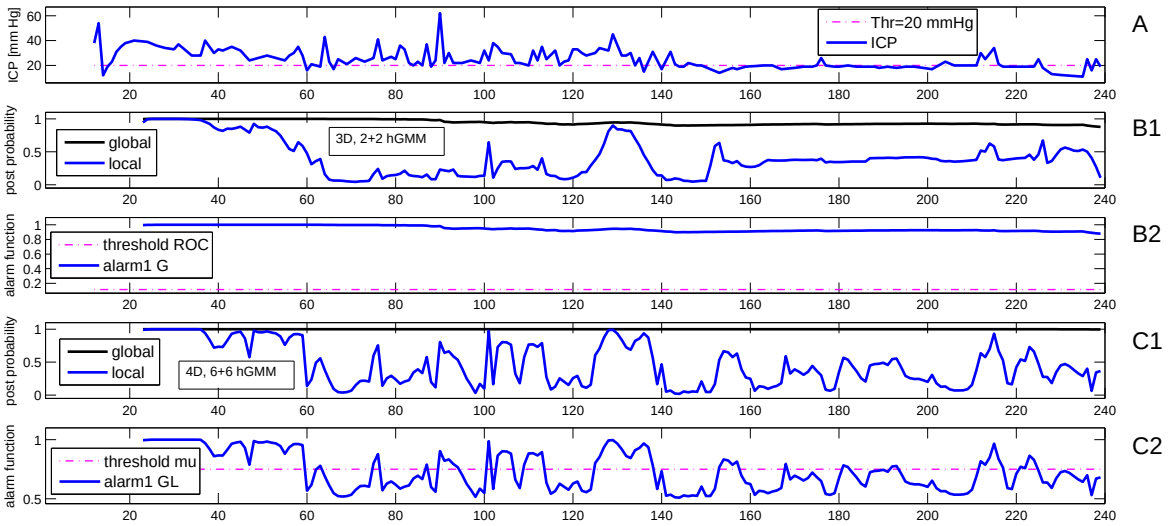

Figure 8. An example of the set of summary plots for the VITAL patient (N.45).

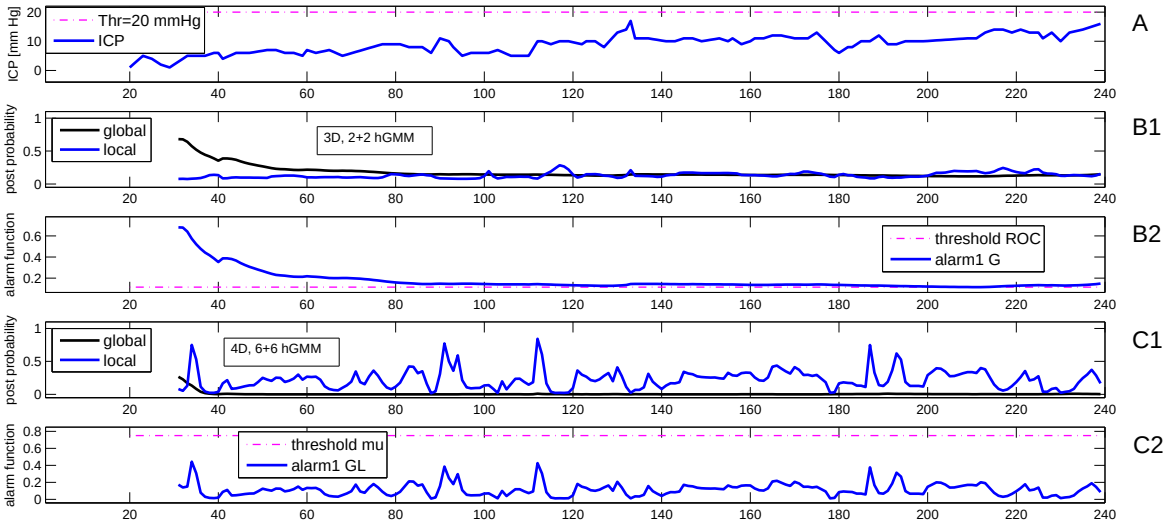

Figure 9. An example of the set of summary plots for the FATAL patient (N.7).

The illustrated plots of the novel alarm functions, as well as their thresholds, reveal the following particular findings. The plots depicted in Fig. 8 represent the "worst" case of the survived patients set; we can see that the traditional alarm, i.e. ICP (part A) holds above the threshold of 20 mmHg for maximum proportion of the recording time (67 %) among all the vital patients. In Fig. 9, on the contrary, the "worst" case for non-survived patient is presented. Here, the traditional ICP alarm (part A) holds below the conventional ICP threshold of 20 mmHg for maximum proportion of the recording time (100 %). It is obvious that the novel alarms (found as optimal for both methods of the alarm function threshold optimization) do not behave in favour of correct indication of an emergency state of the patient being monitored.

In the former case, the first alarm (alarm1 G) (Fig. 8, part B2) is remarkably elevated above its threshold as it only reflects ICP curve that is continually increased over the 20 mmHg threshold during the major part of the recording time. The second alarm (alarm1 GL) oscillates around its threshold (part C2) and rests much more inside the vital zone (below the threshold) than the conventional ICP alarm (part A).

In the latter case, while the first alarm (alarm1 G) Fig. 9, part B2) wanders most of the time slightly above its threshold, the second alarm (alarm1 GL, part C2) holds all the range uniquely under its threshold. Due to the extreme vital-like ICP record, though it belongs to the non-survived patient, the additional alarm behaves similarly to the conventional ICP and it is difficult to find proper clues for the intervention procedure just on the basis of the ICP monitoring.

References

Keogh, E, & Lin, J. 2005. Clustering of time-series subsequences is meaningless: implications for previous and future research. *Knowledge and Information Systems*, **8**, 154–177.