

Review

Medical device alarms*

Matthias Borowski^{1,a}, Matthias Görge^{2,a}, Roland Fried¹, Olaf Such³, Christian Wrede⁴ and Michael Imhoff^{5,**}

¹ Technische Universität Dortmund, Dortmund, Germany

² University of Utah, Salt Lake City, UT, USA

³ Philips Medical, DA Best, The Netherlands

⁴ Helios Klinikum Berlin-Buch, Berlin, Germany

⁵ Ruhr-Universität Bochum, Bochum, Germany

Abstract

The high number of false positive alarms has long been known to be a serious problem in critical care medicine – yet it remains unresolved. At the same time, threats to patient safety due to missing or suppressed alarms are being reported. The purpose of this paper is to present results from a workshop titled “Too many alarms? Too few alarms?” organized by the Section Patient Monitoring and the Workgroup Alarms of the German Association of Biomedical Engineering of the Association for Electrical, Electronic and Information Technologies. The current situation regarding alarms and their problems in intensive care, such as lack of clinical relevance, alarm fatigue, workload increases due to clinically irrelevant alarms, usability problems in alarm systems, problems with manuals and training, and missing alarms due to operator error are outlined, followed by a discussion of solutions and strategies to improve the current situation. Finally, the need for more research and development, focusing on signal quality considerations, networking of medical devices at the bedside, diagnostic alarms and predictive warnings, usability of alarm systems, education of healthcare providers, creation of annotated clinical databases for testing, standardization efforts, and patient monitoring in the regular ward, are called for.

Keywords: alarm systems; equipment design; intensive care unit; noise prevention and control; physiologic monitoring; safety management.

*This article is based on a position paper titled “Alarmgebung medizintechnischer Geräte”, published by the VDE Association for Electrical, Electronic and Information Technologies.

^a Matthias Borowski and Matthias Görge contributed equally to this manuscript.

**Corresponding author: Prof. Dr. med. Michael Imhoff, Deutsche Gesellschaft für Biomedizinische Technik im VDE, Fachausschuss Methodik der Patientenüberwachung, Abteilung für Medizinische Informatik, Biometrie und Epidemiologie, Ruhr-Universität Bochum, Universitätsstraße 150, 44801 Bochum, Germany
Phone: +49-231-9730220
E-mail: mike@imhoff.de

Introduction

The high number of false positive alarms has long been known to be a serious problem in critical care medicine – yet it remains unresolved. At the same time, threats to patient safety due to missing or suppressed alarms are being reported. This position paper focuses on the conflict of having “too few” and “too many” alarms, the question of how to generate the “right” alarms, and how to effectively communicate these alarms.

A workshop titled “Too many alarms? Too few alarms?” organized by the Section Patient Monitoring (Fachausschuss Methodik der Patientenüberwachung) and the Workgroup Alarms of the German Association of Biomedical Engineering (DGBMT) of the Association for Electrical, Electronic and Information Technologies (VDE) was held on November 25th, 2009. In this workshop, experts from academia, research and development, industry, clinical medicine, notified bodies, and competent authorities discussed questions concerning medical device alarms. The workshop, along with additional material provided by the attendees, results from clinical trials and other research studies, as well as additional scientific literature provide the foundation of this paper.

Quickly detecting the critical condition of a patient can be lifesaving. Therefore, patient monitors, which monitor not only the state of the patient but also its own device function, are frequently used in intensive care units (ICUs), operating rooms (ORs), emergency departments (EDs) and other areas of critical care. These systems inform healthcare providers about deviations from a patient’s “normal” state¹ using acoustic and visual signals. However, problems with these alarm systems are well documented in the literature; a review is given in the following section. The reported problems are mainly caused by a high number of alarms and/or a high number of clinically irrelevant alarms [9–11].

The sheer number of alarms creates acoustic overload, which not only causes permanent stress for patients and staff, but can also overwhelm caregivers [17, 31, 33, 41]. This is expected to worsen in the future, as the staffing ratio will decrease and the patient population is getting older which increases demand for patient care. Finally, the introduction of patient monitoring to regular wards with hopes to improve patient care and safety will cause additional problems: patient monitors need to be mobile and especially robust

¹It is important to note that for critically ill patients the “normal” state can lie outside the range specified in textbooks. This is due to the patient’s disease state defining a new “normal” for this individual.

against motion artifacts to work well in this setting with at least partially ambulatory patients. Atzema et al. [3] ask the question whether the disadvantages caused by false positive alarms justify the advantages gained by monitoring these low-risk patients.

The high number of false positive (clinically irrelevant) alarms leads to a desensitization of caregivers against alarms, also known as the crying wolf phenomenon [30]. This poses a serious well-known but still unresolved problem. Additionally, there are consistent reports of incidents, where missing or suppressed alarms lead to endangerment of patients. This field of tension between “too many irrelevant” and “too few relevant” alarms, and the question of how to “correctly” generate and communicate these alarms were the topics of the DGBMT workshop. This paper summarizes the main findings.

Problems with patient monitoring and resulting consequences

Patient monitoring systems are used to continuously monitor patients in order to quickly detect critical and life-threatening conditions. Acoustic alarms are being used to communicate a multitude of alarm conditions, e.g., if a vital sign parameter exceeds a given alarm threshold or if a technical problem occurs. A high sensitivity² in detecting clinically relevant situations is desired, as this is important to guarantee patient safety. However, although this requirement is generally fulfilled in modern alarm systems, this leads to problems in the real-world clinical application.

Several studies have demonstrated that the majority of alarms created by patient monitoring systems have no clinical relevance ([9–11]; see also Table 1). Siebig et al. [38] recorded patient monitoring data of 68 patients in a medical ICU: 5934 alarms were recorded over a period of 982 h, which is equivalent to an average alarm rate of 6 alarms per patient per hour [37]. However, only 15% of all alarms were found to be clinically relevant. Among others, Koski et al. [27], Lawless [30], Tsien and Fackler [42], and Chambrin et al. [10] found high percentages of clinically irrelevant alarms in a multitude of ICUs. A 2008 questionnaire in the US found that only 1% of all alarms resulted in a therapeutic intervention [26], which was found to be even lower (0.6%) in ED patients with angina and low coronary risk [3]. Although most of these studies have been performed in adult settings, some work with pediatric patients [30, 32] exists and demonstrates similar problems in both settings.

However, as most latching and many non-latching alarms have to be manually acknowledged and silenced, this increases the workload of healthcare team members. With a staffing ratio of two or more patients for each nurse, not all alarms can be attended to immediately. Additionally, the chaining of alarms and alarm showers, the collection of multiple alarms which need to be acknowledged, lead to delays

in responding to critical patient states. This observation is consistent with results of a representative survey of German ICUs [39], where the majority of monitoring users found the quality of current patient monitoring alarms to be inadequate; 88% of all users found less than 50% of all alarms to be clinically relevant. Similar results were reported in the US [26].

Owing to the high number of clinically irrelevant alarms (false positive alarms or false alarms) healthcare providers are becoming desensitized. Additionally, constant readiness causes a reduction of the attention threshold. This desensitization results in inadequate responses to alarms or even the complete lack of response. In a US based study, Reslan [34] found alarm response times of up to 40 min. Furthermore, Chambrin et al. [9] reported that only 10% of all alarms were attended to by caregivers, and Cropp et al. [11] found that 50% of all relevant alarms were not correctly identified by caregivers. Therefore, the sensitivity of the alarm system “patient monitor and healthcare provider”, which is the sensitivity experienced by the patient, is by no means close to 100%.

Secondary to the high rate of false positive alarms, there is the absence of alarms in alarm-warranting conditions. During 2007–2009, 75 critical situations, where no alarm occurred, were reported to the Federal Institute for Drugs and Medical Devices (BfArM) in Germany. In total, 18 deaths and 6 delayed cases of cardiopulmonary resuscitation, 2 resulting in permanent brain damage, were attributed to the lack of alarms. In 14 cases, reasons for the lack of alarms could not be found. For the remaining 61 cases the following causes were identified: 19 cases of the alarm being disabled, 23 cases of incorrect configuration or lack of knowledge concerning the alarm function by the users, and 10 cases of technical device error, none of which lead to harm for the patient. The remaining 9 cases had miscellaneous sources of error, including incorrect configuration of pacemaker detection. Therefore, the majority of cases were attributed to operator error.

The US Food and Drug Administration’s (FDA) Manufacturer and User Facility Device Experience (MAUDE) database shows similar numbers of events reported by hospitals in the US. Between 2002 and 2004, 237 deaths related to device alarms, the majority concerning patient monitor alarms, were reported. Again, the majority was attributed to operator error and incorrect configuration of the alarm system. However, a high number of unreported and unnoticed situations are likely because only incidents with severe consequences have to be reported to the respective government authorities.

With more than 40 different alarm and information signals occurring in ICUs [11], potential danger is not only based on desensitization of caregiver and operator errors. Acoustic alarm and information signals are frequently misinterpreted, as even very experienced physicians and nurses have problems attributing alarms to their respective sources [11]. Additionally, alarms are frequently attributed to the incorrect patient’s room or device. Product specific alarms, as used in some mechanical ventilators, could be beneficial here.

²Sensitivity notes the conditional probability of correctly detecting a condition warranting and alarm (alarm triggering) when such a condition is present.

Table 1 Studies regarding alarm quality of medical alarm systems in ICUs.

Source	Setting	Number of patients, duration, number of alarms	Annotation method used	Definitions of true positive (TP) and false positive (FP) alarms	Results
O'Carroll, 1986 [32]	General purpose ICU (mixed adult and pediatric)	26 patients, 3 weeks, 1455 alarms	Nursing staff recorded: alarm source, cause, and the silencing of alarms	Not provided	8/1455 alarms indicated potentially life-threatening problems
Koski et al., 1990 [27]	Cardiac surgery postoperative car unit (PACU)	10 patients, 400 h, 1307 alarms	Clinician records: alarm source, time, and classification of the alarm	TP: diagnosis made or action performed FP: alarm caused by artifact or no action performed	TP: 139 (10.6%)
Lawless, 1994 [30]	Pediatric ICU	928 h, 2176 alarms	ICU staff recorded: time, alarm source and classified the alarm	TP: resulted in a change in therapy FP: technical error or no change in therapy	TP: 5.5% FP: 68% Sens.: 31% Spec.: 82% PPV: 5% NPV: 98%
Tsien and Fackler, 1997 [42]	Multidisciplinary, pediatric ICU	298 h, 2942 alarms	Trained observers recorded: alarm source, reason and validity of alarm. Additional verification by nurse at the bedside	TP-R: measured correctly and clinically relevant TP-I: measured correctly, but clinically irrelevant FP: measurement incorrect	TP-R: 8% TP-I: 6% FP: 86%
Chambrin et al., 1999 [10]	Five adult ICUs	131 patients, 1971 h, 3188 alarms	Trained nurse recorded: type, source, and consequence of the alarm	TP: intervention, including solving technical problems FP: no intervention	TP: 880 FP: 2308 FN: 24 TN: 3196 Sens.: 97% Spec.: 58% PPV: 27% NPV: 99%
Biot et al., 2000 [6]	Surgical ICU	25 patients with ARDS, 15013 min, 3665 alarms	Bedside clinical observer recorded detection of absence of clinical events	TP: alarm based on correct measurement FP: alarm based on technical problem	TP: 54% TP with direct medial consequence: 5% FP: 46%
Atzema et al., 2006 [3]	ED	72 patients, 371 h, 1762 alarms	Low-risk patients with chest pain. Trained observers counted alarms, changes in management, and monitor detected adverse events	TP: detection of adverse event with or without change in management FP: no change in management	TP: 0.7% FP: 99.3%
Görges et al., 2009 [20]	Medical ICU	21 patients, 200 h, 1271 alarms	Trained observer recorded: alarm source, alarm reason, and validity of the alarm	TP: "effective" = treatment performed FP: "ineffective" = no treatment owing to absence of provider "ignored" = no treatment while provider present	TP: 23% FP: 77%

(Table 1 continued)

Source	Setting	Number of patients, duration, number of alarms	Annotation method used	Definitions of true positive (TP) and false positive (FP) alarms	Results
Siebig et al., 2010 [37]	Medical ICU	68 patients, 982 h, 5934 alarms	Experienced physician annotated video recordings, and data from patient monitoring system	TP: diagnosis or treatment FP: no clinical intervention	TP: 15% FP: 85%

TP, true positive alarm; FP, false positive alarm; FN, false negative alarm; TN, true negative alarm; PPV, positive predictive value; NPV, negative predictive value; Sens., sensitivity; Spec., specificity.

The alarm volume, with up to 80 dB(A) is an additional problem [4]. This noise³ can cause significant stress, ranging from burn-out in healthcare providers to delayed recovery, sleep deprivation, and even psychosis in patients [17, 31, 33, 41]. Here, Kahn et al. [24] report that the majority of noise in an ICU is due to alarms. Frequent and loud alarms pose a large psychological burden for patients and healthcare providers. Busch-Vishniac et al. [8] found that the average noise level in large ICUs frequently exceeds the threshold recommended by the World Health Organization sometimes by more than 30 dB(A). The largest contributor of medical device related noise were alarms by patient monitors. Among healthcare providers, the number of burn-out cases has been steadily increasing over the past couple of years [33], most likely to be attributed to problems in the workplace. The reduction of noise-related stress is one of the most important measures to improve the situation at the workplace in most industries [43].

Noise poses equally significant problems for ICU patients. Convalescence was found to be delayed due to noise interrupting the patient's sleep cycle [5, 17]. Here, acoustic alarms of medical devices are again a major contributor to this problem. Finally, noise was found to have a negative influence on patient satisfaction [34].

Although the conditions described are the rule for the majority of ICUs, there are some ICUs where very few alarms are triggered. This is most probably attributed to the fact that no standardized alarm settings are available. Such standardization might not even be possible owing to large variability in patient populations and disease states. Therefore, the individual "philosophy" of every ICU with regard to alarm settings has a huge influence on the number of clinically irrelevant alarms. However, it is still unclear if a reduction of (false) alarms will lead to a reduced or increased

sensitivity of the combined "patient monitor and healthcare provider" alarm system. There is certainly a need for future work in this area.

Causes of problems with (clinical) patient monitoring

Although a multitude of alarms are being generated in the ICU, some critical and life-threatening situations remain undetected and do not result in alarms. However, this does not present a contradiction: critical situations can remain undetected if the alarming function of the device is disabled or compromised due to alarm thresholds set too wide. Alarms are often deactivated in order to reduce the number of false alarms, or a lack of knowledge concerning the alarm system settings. There are multiple reasons for the high number of false alarms, which will be discussed next.

Comprehensive monitoring of the patient's condition is performed not only to improve bedside monitoring of the patient but also for liability reasons. For instance, pulse oximetry, a monitoring modality known to be associated with many false alarms, is recommended by guidelines. Additionally, there has been an increase in the number of vital signs monitored in a single patient, resulting in an increase of the number of alarms, and thereby the number of false alarms. However, most medical devices are not interconnected, which frequently results in the same situation triggering multiple alarms from different devices. Finally, the devices used in the ICU are normally not standardized; even products of the same manufacturer vary within one product line. This leads to predictable operator errors, which are discussed in the usability standard DIN EN IEC 60601-1-6.

A sensitivity as high as possible is a desired feature of alarm systems, as clinically relevant situations should never remain undetected. By contrast, the specificity⁴ or false alarm rate is of lesser importance. This results in numerous false

³According to DIN EN IEC 60601-1-8, concerning alarm systems in medical equipment, and DIN EN 60601-2-49, targeting specifically multi-parametric monitors, the operating institution has the ability to set upper and lower limits on the range of alarm volumes, which can be selected by the end-user. This has significant influence whether the base noise level of each ICU is high or low.

⁴Specificity is the conditional probability of correctly not detecting an alarm relevant condition (no alarm is triggered) when actually no condition warranting an alarm is present.

alarms, alarms that are irrelevant or have no clinical consequences, especially in the monitoring of patients who no longer require close monitoring or would not benefit from it [23].

Incorrect operation and/or configuration of an alarm system, along with low specificity, are further reasons leading to incidents. Setting and/or changing alarm thresholds requires navigating through multiple levels of menus, which emphasizes the need to use usability and other human factor approaches to solve this problem effectively. User-friendliness and usability are clear and basic requirements of the Medical Device Directive of the European Union (93/42/EEC). New alarm systems entering the market must comply with this directive by having a user interface which is designed to minimize the errors by fulfilling ergonomic requirements. This is to be achieved through a usability process during the development of new products. In practice, however, devices from different manufacturers are used at the bedside and their user interfaces are not matched/harmonized. This leads to problems in the application of alarm systems, which are intensified owing to the increasing complexity of these systems.

Owing to the high complexity of the devices, manuals for patient monitors are lengthy. This leads to them seldom being read, even though the German Ordinance on the Installation, Operation and Use of Medical Devices (MPBetreibV) mandates that the devices be used according to the instructions provided in the manual. Consequently, the nursing staff is not always knowledgeable of the functions and adjustments of the alarm device. This results in the alarm system not being adapted to the individual patient, resulting in an increased number of false alarms if the alarm thresholds are too narrow, and potentially missing alarms, if the thresholds are too wide.

Training and briefing of staff regarding medical devices, including the respective alarm systems, is mandated by the German Medical Device Act (MPG) and the German Ordinance on the Installation, Operation and Use of Medical Devices (MPBetreibV) and offered by the manufacturers according to their legal obligation; this is similarly done in other EU countries and the US. However, these training courses are frequently not attended by the nursing staff owing to the time commitment required. In addition, there is an increasing fluctuation of nursing staff in intensive care. Consequently, the experience of the healthcare providers in dealing with medical equipment is constantly decreasing [33].

The high false alarm rate of patient monitoring systems in critical care is of concern to notified bodies, e.g., the TÜV NORD. However, it is clear that the regulatory approval of a medical device (CE, FDA) focuses only on the device itself, not its integration in clinical practice. Until recently, this led to the practicality of the alarm system being only of minor concern of development departments. To counteract this problem, DIN EN 62366:2008-09 defines that the usability of the product (fitness for use) has to be included in the functional specification for the developers. Similar requirements were previously covered by the risk management guideline DIN EN ISO 14971:2007-07. The goal is to min-

imize risks, e.g., distractions from unnecessary noise sources. Because alarm systems used in clinical practice do not always follow the latest technical standard, it remains to be seen whether the above standards can achieve the desired effect.

Approaches to improving alarm systems

Ergonomics and human factors engineering

Alarm systems should be intuitive, easy to use and user-friendly. Current alarm systems show potential for improvements not only in the areas of alarm quality and clinical effectiveness but also in the area of usability.

Instead of currently used single-piercing alarm sounds, one can improve recall of alarms by using softer alarm melodies. Newer generations of medical devices already implement this approach. DIN EN 60601-1-8 alarm systems define alarm tones and alarm melodies to be used. However, this approach has only been mandatory since December 2007. One should also note that the correct identification of alarm events and patterns was found to be <50% in a study setting [36]. Here, the unification of alarm signals for devices of the same type, such as patient monitors, mechanical ventilators, and infusion pumps, of all manufacturers would be highly welcomed. However, such standardization would only be possible by introducing a new product standard, which needs to be in sync with actual practice. Here the involvement of users of the different products would be very helpful.

One step further than using melodic alarms is the approach of acoustically representing measurements using melodies or music [36]: changes in tonal patterns, pitch or rhythm are used to communicate changes in measured values. However, this approach creates a constant exposure to noise, which makes this approach less feasible for critical care monitoring. For anesthesiologists in the operating room, however, there is high potential for this application. Watson and Sanderson [44] demonstrated that a respiratory sonification allowed anesthesiologists to maintain high levels of awareness of a simulated patient's state while performing other tasks more effectively than when relying upon visual monitoring.

Although interesting, the approach of using ‘softer’ acoustic alarms targets only the consequences but not the sources of false alarms with its resulting noise pollution and desensitization of the user. It would be better to restrict acoustic alarms only to relevant situations. By networking alarm devices, one could reduce the large number of alarms in alarm chains, where one underlying condition triggers multiple alarms, into one single and more meaningful alarm. However, this approach poses high regulatory and medico-legal challenges: standardized and certified interfaces would be necessary owing to the high reliability required for this approach, but also because of potential liability issues. Although each manufacturer is responsible for the proper function of its device in the framework of appropriate use according to the German Medical Devices Act (MPG), this changes when combinations of multiple devices or systems

in a network, which was not manufactured or certified by its manufacturer, are used. Here the operator of the network/system, e.g., the hospital, becomes the manufacturer of the system and is responsible for its function as a new medical product under the MPG. The operator is now responsible for product conformity assessment and risk management, and possibly liable. Future work concerning IEC 80001-1, which standardizes risk management of information networks with medical devices, is of great importance here. Although being of great significance for the future creation and use of medical networks, discussing the standard is beyond the scope of this paper.

Another approach, which minimizes the noise pollution, is the use of centralized alarms. This approach is currently being used in Japan and the US. Here, alarms are no longer acoustically announced at the bedside but instead redirected to a central monitoring station. A scope-watcher, typically a specially trained healthcare provider, evaluates the situations and if necessary notifies/pages nursing staff for evaluation and treatment. Although this approach certainly reduces noise for the patient and relieves nurses at the bedside, it is more of an abatement of symptoms than a fix of the problem.

Fineman [15] suggests a system where healthcare providers wear armbands on their wrists to receive patient-specific information. In the case of a patient in a critical state, the armband notifies the responsible nurse using a vibrotactile alarm. The armband shows the patient involved, where he is located, and what type of alarm is being triggered. Implementing this system is technically feasible, however, with no small effort. Additionally, it must be ensured that alarms are being received when the responsible caregiver is busy elsewhere. Moreover, compliance with the hospital hygiene standards must be ensured. However, this example demonstrates that the current requirements of acoustical alarms are not necessarily the only viable and reasonable modality to communicate alarms. In the area of mobile communication, vibrotactile alarms have long been customary and are widely accepted. In the medical field, Ford et al. [16] demonstrated that vibrotactile communication in the OR can reduce response time to critical incidents, such as the detection of anaphylaxis by anesthesiologists.

Alarm algorithms

In addition to the improvements in usability, interoperability, and optimization of alarm modalities, there are technically feasible methods to generating better, as in more relevant, alarms today: combining existing data sources in a smart and algorithmic way.

The following subsections discuss improvements of alarm algorithm. Working online, which means in real time, they extract relevant information from measured values or manipulate alarm thresholds automatically.

A simple method to reduce the number of false alarm is triggering an alarm only when a threshold has been crossed for a defined period of time: these alarm delays are already implemented in currently available alarm systems. However, current standards define relative short limits, such as no more than 10 s of delays for invasive arterial blood pressure

alarms. G6rges et al. [20] demonstrated that increasing the alarm delay to 19 s can result in a reduction of the false alarm rate of up to 80%. However, given the current standards such an increase in alarm delays has to be explicitly declared and justified. Increasing the maximum length of alarm delays in the standard specifications is worth considering. With the exception of life-threatening alarms, such as asystole or atrial fibrillation which need to be triggered immediately, the average heart rate alarm could be delayed by up to 30 s. This would pose no risk to the patient, and most clinically irrelevant alarms would not occur in the first place. The extent of how long a user can set his alarm delays should be configurable on a system-wide setting by the responsible organization.

Another highly effective approach also uses the idea of alarm delays. Here, a combination of alarm delay and alarm severity, as the distance of the measured value from the alarm threshold, is used to ensure patient safety. This results in extremely high or low values triggering an alarm much faster than a moderate alarm threshold violation.

The concept of introducing an alarm delay poses a simple but highly effective way of reducing false alarms. Next, two promising approaches for improvements of alarm systems are presented; both use online transformations of measured values to present the user only with relevant information.

One approach is using an early warning system, which predicts the current trend from a series of measured monitoring values. An overview of such systems can be found in the review of Imhoff and Kuhls [22]. If the trend predicts that a threshold violation will occur in the near future, a warning is issued by the system. The advantage of such a system is that a patient's treatment can be initiated before his condition reaches a critical condition. However, to ensure that such a system does not generate more alarms than current patient monitors, wider fixed alarm thresholds must be chosen. Yang et al. [45] expanded the idea of trend monitoring and developed a procedure to detect clinically relevant changes in measured monitoring time series online. Consecutive work by the same group [2, 13, 46] showed excellent performance when comparing the trend detection of their system using previously recorded changes during pediatric anesthesia [2], and in real time during a clinical evaluation in a pediatric OR [13]. High agreement in trend decision with opinions of the anesthesiologist was reported [46]. Finally, Tappan et al. [40] demonstrated that adding a visual cue (triangle) to vital sign parameters on a patient monitor significantly reduced the detection time to a change.

Another approach to reduce false alarms is using robust statistical signal extraction [7, 12, 18]. This approach assumes that the measured values consist of a clinically relevant but not directly observable signal, which is superimposed by noise and outliers. Outliers are extremely short deflections in measurements, which are not caused by physiologic changes, but rather caused by movement artifacts or other technical reasons, such as flushing an arterial line. Using robust regression methods, with sliding windows, signal extraction is performed online and compared with alarm thresholds. Here the signal is separated from noise and out-

liers and the resulting cleaned dataset is used for the alarming decision. As outliers are a main source of false alarms, this online signal extraction method causes a reduction in the false alarm rate. Additionally, the data are smoothed, i.e., a signal extraction method removes irrelevant fluctuations (the noise).

To gain acceptance, it is important that the implemented algorithms are transparent and can be comprehended by the user. Additionally, it is important that nurses and physicians are able to access the raw measurement data. It is important to note that for all algorithms the reduction in false alarm rates is accompanied with a loss in sensitivity. However, there is hope that owing to the decreased false alarm rate the desensitization of healthcare providers to alarms can be reduced or reversed, leading to an increased sensitivity of the “patient monitor and healthcare provider” system.

Most of this research has been done in adult patient populations. It can be expected that in pediatric or neonatal settings requirements for sensitivity, time delays, individual thresholds, and other algorithm parameters are different. For instance, desaturation in pediatric patients is quicker as a result of different pulmonary properties in infants and children compared to adults. By contrast, the underlying algorithmic concepts, e.g., robust signal extraction, are expected to be the same between different patient populations.

Finally, there is much need for additional research. First, the different algorithms need to be compared for their specificity and rate of reduction in false alarms, while considering the decrease in their sensitivity. For some algorithms these data are available and an overview can be found in Imhoff and Kuhls [22]. Second, the question of how a reduction in the false alarm rate influences the sensitivity of the “patient monitor and healthcare provider” system, a non-trivial task, needs to be addressed.

Intelligent alarm systems

“Intelligent” alarm systems use a special type of signal processing that uses algorithms. These algorithms interpret the raw data, thus providing a higher level of abstraction.

Many alarms can be suppressed using the concept of alarm validation⁵. Implementations of validation algorithms are commercially available. For example, a system which validates asystole alarms by comparing the heart rate from ECG signal extraction with the pulse rate from the pulse oximeter or the pulse rate from the arterial blood pressure waveform exist. This system uses the fact that an asystole with concurrent stable blood pressure pulsation is highly unlikely. In this case the asystole alarm will be suppressed.

A relevant number of false alarms are caused by manipulation, and the clinical significance of an alarm depends heavily on the current clinical context. Therefore, validation

of alarms can be based on available context information. For example, pulse oximetry alarms, obtained from a finger, can be suppressed if non-invasive blood pressure is measured on the same arm. However, using this context information is only useful when its detection happens automatically and does not require interaction or manual setting by a healthcare provider. Automatic detection of non-invasive blood pressure measurements is available in a commercially available system.

Root cause analysis is another method, which can be used to interpret alarms intelligently. This method is commercially available in areas other than medicine, such as for nuclear power plants or chemical plants. In this method, processes, conditions, and dependencies are modeled to identify (trace) the source of an alarm. For example, the increase of the arterial blood pressure above a set threshold can lead directly to the diagnosis “hypertension”. Root cause analysis is particularly useful for interpreting critical conditions eliciting alarm chains. Instead of signaling a multitude of separate alarms, an alarm system using root cause analysis would notify only one single alarm and also indicate the underlying reason for this problem. As early as 1987–1993, a group of computer scientists and physicians from Stanford developed a system for the cardiac surgical ICU, which correctly detected and diagnosed approximately 30% of the typical ICU complications [21, 28]. Although this system was never commercially available, it shows potential for using root cause analysis in medicine.

“Diagnostic” alarms are an even higher level of abstraction than root cause analysis. Such systems detect pathophysiologic or technical solutions for alarms and imply a relevant solution or treatment. For example, instead of presenting the diagnosis “hypotension” the diagnosis “hypotension due to hypovolemia” is presented, from which the user can imply that volume replacement therapy is a relevant treatment option. However, diagnoses cannot be made with 100% certainty. Therefore, statistical methods such as Bayesian networks [29] are used to associate potential diagnoses with a probability. Knowledge bases, which are used to infer classes from currently available data and information, are needed for this method to work. Here, Dunsmuir et al. [14] created a knowledge authoring engine, a software tool which allows easy and practical collection of expert knowledge using simple if...then rules.

The major challenge in using knowledge based alarm systems is the high patient to patient variability. It is desirable that alarm systems work even for “untypical” patients, even if they deviate from the general populations. The use of patient-specific learning using artificial intelligence is an interesting approach to this problem [47].

Technical validation (see Footnote 5) of complex intelligent alarm systems is challenging as a proper gold standard does not exist. Clinically annotated databases, where physicians decided on the correct outcome, have to be used as a reference. Additionally, there are large international differences in clinical culture and in the willingness of physicians to engage in and trust interpreting systems. This situation will certainly pose interesting challenges to the introduction

⁵Validating an alarm is determining whether this alarm actually describes a clinically alarm-requiring situation. To distinguish the term of validating in the sense of product development and product management, a validation which compares a product with its specification will be referred to as a “technical validation”.

of such systems by medical device manufacturers operating worldwide.

The question of safety of intelligent alarm systems in clinical practice comes to mind when these systems are technically validated. The fact that an increase in specificity always goes along with a decrease in sensitivity is still true for intelligent alarm systems. Additionally, the influence of decision support on healthcare provider training needs to be investigated, and one has to pose the question: What will happen if the system fails but the user cannot do without the decision support system? Also, it is still unknown how and if the user can decide on the quality of the decision presented.

The current situation, where many healthcare providers are overloaded with information, could be mitigated by the introduction of decision supporting systems. However, it is important to note that the system provides a justification of each diagnosis presented or alarm triggered and that the option of users to visualize the raw data is not removed. Many users still look at these to quickly check whether they should trust the decisions made or whether artifacts might be present. By providing access to raw data one could also address legal issues that could occur under the circumstances where no clear and unambiguous diagnosis is given.

Finally, it is important to acknowledge that nurses are the best monitors. Providing them with the right tools, such as mobile decision support systems or personalized alarms, has high potential to improve their situational awareness and efficacy, thereby improving patient safety. Special care should be taken to avoid replacing experienced nurses with a combination of less experienced healthcare providers and additional patient monitoring equipment. This is important, as Aiken et al. [1] found that improving minimum nurse-to-patient ratios is likely to save thousands of lives each year in the US alone.

Healthcare providers' knowledge about alarm systems

Inadequate configuration and use of the alarm systems lead to unnecessary alarms on the one hand and also result in critical situations not detected on the other hand. Therefore, a higher general awareness, and increased knowledge, of healthcare providers regarding the function of the alarm system is of interest.

In addition to the comprehensive and complete manual, manufacturers should provide an abridged (quick glance) version. Alternative media, such as videos or PC-based interactive training, simplify and facilitate learning about the function of the device, and are oftentimes already available. A self-explanatory alarm system, where context sensitive help of frequently asked questions and problems is provided, is also helpful.

Additionally, targeted training sessions can be used to improve awareness of the patient monitoring systems. On the one hand, device manufacturers are required by law to provide these; and on the other hand, these training sessions have to be accepted by healthcare providers who need to actively engage in them. To facilitate this, the training sessions should be reasonable and targeted towards the right

audience, and should be considered additional work hours (overtime) to encourage participation by interested healthcare professionals. However, this approach will work only by making training mandatory and requiring frequent participation. Finally, a supervisory body and a certifiable level of training are needed. This approach is problematic, as training sessions, continued education sessions, and refresher courses are associated with high and regular costs, which hospitals often try to avoid.

Although training, with known problems in its current implementation, certainly can be improved, improving alarm systems with regard to alarm quality, networking of alarms, and improvements in usability needs to be of highest priority. By improving the inherent safety of these alarm systems, an unnecessary burden on users and patients can be avoided. Finally, the challenge for inherent safety is already an essential requirement of the medical devices directive 93/42/EEC.

Need for more research and development

There is no doubt that there is a need for major research and development efforts for medical device alarm systems – particularly those involving complex human-machine interactions. The entire chain, starting with the selection of appropriate alarm settings for a patient, continuing with the signal acquisition, and ending with the communication of the alarm message, needs to be carefully examined.

The quality of measured systems with regard to false alarms remains a problem in patient monitoring. This also concerns the ever spreading non-invasive measurement techniques. Additional research and development in the areas of signal acquisition, signal validation, and artifact detection and removal is needed.

The increased networking of medical devices at the bedside poses additional problems. Questions concerning integration, analysis, and interpretation of highly complex data streams need to be investigated. Additionally, questions regarding networking and interoperability of the devices with regard to patient and equipment safety need to be addressed. However, the integration of data and alarm messages from different devices could lead to a reduction in the number of alarms.

A new and largely unexplored area is diagnostic alarms and methods of early detection and warning (predictive warnings). There is a significant need for research in the areas of algorithm development, knowledge discovery, and technical validation. Additionally, knowledge concerning the areas where patient monitoring had little or no use in the past is sparse. Here, complex and structured data acquisition in the real clinical context would be recommended.

Furthermore, considerable progress in the area of usability is needed. Research endeavors, targeting specifically the usability of medical alarm systems, is needed, as the “bedside” environment poses special and very complex demands. This research should focus on nurses as the primary user of these devices, and a user-centered development cycle. Experience with research methods, technical approaches, and results

from cognitive psychology obtained in other areas should be incorporated.

Another area for research is the training of healthcare providers. Structural investigations on how the user's knowledge influences the function of the system "patient monitor and healthcare provider" are needed.

New alarm algorithms, alarm systems, usability improvements to existing systems, as well as new approaches to healthcare provider education have to be clinically validated. Therefore, the creation of databases with clinically annotated and validated data, which can be used for research and development purposes, should be an additional goal of clinical trials. Although some databases, such as IMPROVE [25], MICIC II [35], or PhysioBank [19] exist, these new databases could be used as a reference for researchers, developers, and industry.

Research, on topics such as alarm detection, acoustic, optic or tactile alarm notification, or device interoperability, has direct implications for the development of related standards. This is important as current standards, however carefully crafted, have not been able to improve the problems related to alarms in clinical practice. This leads to the conclusion that current research, knowledge about clinical necessities and problems, and standardization have to be more closely interlinked. With regard to standardization, it is also important to differentiate between results, which yield into the primary standard for all medical products (DIN EN IEC 60601-1-8), and results for which relevance applies to only one specific standard, such as the standard for patient monitoring (DIN EN IEC 60601-2-49).

Although improvements in alarm algorithms and more intelligent systems will reduce the number of false alarms, this is always accompanied with a certain amount of loss of sensitivity. Therefore, it is of importance to investigate if a loss in sensitivity is justified by the gain through the reduction of false alarms. This comparison has to be performed on multiple levels: the legal, ethical, and occupation psychological perspectives. The question needs to be answered whether a reduction in the false alarm rate influences the sensitivity of the system "patient monitor and healthcare provider".

Finally, it should be noted that research to improve alarm systems for medical applications has implications for areas far exceeding the medical domain. These findings are also very relevant for other complex high-risk applications.

Summary and conclusion

Alarm systems are employed to detect critical states in ICU patients and therefore have to meet the highest safety requirements. However, the alarm systems used to monitor ICU patients are burdened with problems. The sheer number of alarms causes a decrease in the attention threshold, and acoustic alarm systems are responsible for a majority of the noise in an ICU, posing a high burden on patients and healthcare providers. Additionally, the high rate of false alarms

causes the medical staff to become desensitized, leading to a delayed or missing response to an alarm. This can lead to missing critical or even life-threatening conditions, so that the real sensitivity will certainly be below 100%.

One reason for the flood of alarms is the increasing use of monitoring, which often is unnecessary and poses no additional benefit to the patient. A large number of false alarms are caused by the low specificity of the alarm system, especially in mobile patients. Additionally, the complexity of alarm systems is increasing while no adequate education of healthcare providers is performed. Inadequate knowledge about the alarm system as well as resulting misconfiguration and improper use of alarms can lead to both missed alarms and an increased false alarm rate. Finally, the usability of alarm systems is less than optimal and should be improved.

Different approaches to improve currently used alarm systems have been identified and discussed, among others:

- "softer" alarm melodies instead of single piercing alarms,
- vibrotactile alarms allowing silent notification of a responsible provider,
- networking of alarm devices,
- algorithms which reduce the number of false alarms (alarm delays, online signal extraction, ...),
- intelligent alarm systems (context aware alarms, alarms based on root cause analysis, diagnostic alarms, ...),
- an improvement of the knowledge of healthcare providers regarding the function of the alarm system (user-friendly manuals, self-explanatory systems, and providing adequate training).

There is still a lot of research needed to address the proposed approaches to improve current alarm systems. New alarm systems are strongly needed to provide adequate patient monitoring, as human and financial resources in healthcare are in shorter supply. However, positive effects and safety of these applications has to be evaluated carefully, always keeping in mind the best interest for the patient.

Finally, the role of standards for medical devices and medical information systems needs to be taken into account as standards have a direct influence on medical devices functions for everyday clinical life. It is imperative that findings and insights presented here, as well as future research, are incorporated directly into the work of relevant standardization bodies.

Acknowledgments

The authors would like to acknowledge contributions by Dr.-Ing. Thomas Becks, VDE; Stephan Corr, Nihon Kohden Europe GmbH; Thomas Falck, Philips Research Europe; Dipl.-Ing. Frank Franz, Drägerwerk AG & Co. KGaA; Dipl.-Ing. Jörg Hidding, mdc medica device certification GmbH; Dipl.-Ing. Klaus Jung, TÜV NORD CERT GmbH; Dr.-Ing. Jens Mühlsteff, Philips Technologie GmbH; Dr. med. Thomas Rölleke, Bundesinstitut für Arzneimittel und Medizinprodukte (BfArM); Andreas Schäfer, DGF Asklepios Kliniken Schwalm-Eder; and PD Dr. med. Sylvia Pemmerl, Universitätsklinikum Regensburg.

References

- [1] Aiken LH, Sloane DM, Cimiotti JP, et al. Implications of the California nurse staffing mandate for other states. *Health Serv Res* 2010; 45: 904–921.
- [2] Ansermino J, Daniels J, Hewgill R, et al. An evaluation of a novel software tool for detecting changes in physiological monitoring. *Anesth Analg* 2009; 108: 873–880.
- [3] Atzema C, Schull M, Borgundvaag B, Slaughter G, Lee C. ALARMED: adverse events in low-risk patients with chest pain receiving continuous electrocardiographic monitoring in the emergency department. A pilot study. *Am J Emerg Med* 2006; 24: 62–67.
- [4] Balogh D, Kittinger E, Benzer A, Hackl J. Noise in the ICU. *Intensive Care Med* 1993; 19: 343–346.
- [5] Barach P, Potter Forbes M, Forbes I. Designing safe intensive care units of the future. In: Gullo A, Besso J, Lumb PD, Williams G, editors. *Intensive and critical care medicine*. Milan, Italy: Springer 2009: 525–542.
- [6] Biot L, Carry P, Perdrix J, Eberhard A, Baconnier P. [Clinical evaluation of alarm efficiency in intensive care]. *Ann Fr Anesth Reanim* 2000; 19: 459–466.
- [7] Borowski M, Schettlinger K, Gather U. Multivariate real-time signal extraction by a robust adaptive regression filter. *Communicat Stat Simulat Comput* 2009; 38: 426–440.
- [8] Busch-Vishniac I, West J, Barnhill C, Hunter T, Orellana D, Chivukula R. Noise levels in Johns Hopkins Hospital. *J Acoust Soc Am* 2005; 118: 3629–3645.
- [9] Chambrin M. Alarms in the intensive care unit: how can the number of false alarms be reduced? *Crit Care* 2001; 5: 184–188.
- [10] Chambrin M, Ravaux P, Calvelo-Aros D, Jaborska A, Chopin C, Boniface B. Multicentric study of monitoring alarms in the adult intensive care unit (ICU): a descriptive analysis. *Intensive Care Med* 1999; 25: 1360–1366.
- [11] Cropp A, Woods L, Raney D, Bredle D. Name that tone. The proliferation of alarms in the intensive care unit. *Chest* 1994; 105: 1217–1220.
- [12] Davies P, Fried R, Gather U. Robust signal extraction for online monitoring data. *J Stat Plann Inference* 2004; 122: 65–78.
- [13] Dosani M, Lim J, Yang P, et al. Clinical evaluation of algorithms for context-sensitive physiological monitoring in children. *Br J Anaesth* 2009; 102: 686–691.
- [14] Dunsmuir D, Daniels J, Brouse C, Ford S, Ansermino J. A knowledge authoring tool for clinical decision support. *J Clin Monit Comput* 2008; 22: 189–198.
- [15] Fineman B. MetaMonitor: a system for patient monitoring in ICUs. Pittsburgh, PA: Carnegie Mellon University 2004.
- [16] Ford S, Daniels J, Lim J, et al. A novel vibrotactile display to improve the performance of anesthesiologists in a simulated critical incident. *Anesth Analg* 2008; 106: 1182–1188.
- [17] Gabor J, Cooper A, Crombach S, et al. Contribution of the intensive care unit environment to sleep disruption in mechanically ventilated patients and healthy subjects. *Am J Respir Crit Care Med* 2003; 167: 708–715.
- [18] Gather U, Schettlinger K, Fried R. Online signal extraction by robust linear regression. *Computat Stat* 2006; 21: 33–51.
- [19] Goldberger A, Amaral L, Glass L, et al. PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals. *Circulation* 2000; 101: E215–E220.
- [20] Görges M, Markewitz B, Westenskow D. Improving alarm performance in the medical intensive care unit using delays and clinical context. *Anesth Analg* 2009; 108: 1546–1552.
- [21] Hayes-Roth B, Washington R, Ash R, et al. Guardian: a prototype intelligent agent for intensive-care monitoring. *Artif Intell Med* 1992; 4: 165–185.
- [22] Imhoff M, Kuhls S. Alarm algorithms in critical care monitoring. *Anesth Analg* 2006; 102: 1525–1537.
- [23] Institute E. The hazards of alarm overload. Keeping excessive physiologic monitoring alarms from impeding care. *Health Dev* 2007; 36: 73–83.
- [24] Kahn D, Cook T, Carlisle C, Nelson D, Kramer N, Millman R. Identification and modification of environmental noise in an ICU setting. *Chest* 1998; 114: 535–540.
- [25] Korhonen I, Ojaniemi J, Nieminen K, van Gils M, Heikelä A, Kari A. Building the IMPROVE data library. *IEEE Eng Med Biol Mag* 1997; 16: 25–32.
- [26] Korniewicz D, Clark T, David Y. A national online survey on the effectiveness of clinical alarms. *Am J Crit Care* 2008; 17: 36–41.
- [27] Koski E, Mäkivirta A, Sukuvaara T, Kari A. Frequency and reliability of alarms in the monitoring of cardiac postoperative patients. *Int J Clin Monit Comput* 1990; 7: 129–133.
- [28] Larsson JE, Hayes-Roth B, Gabac D, Smith B. Evaluation of a medical diagnosis system using simulator test scenarios. *Artif Intell Med* 1997; 11: 119–140.
- [29] Laursen P. Event detection on patient monitoring data using causal probabilistic networks. *Methods Inf Med* 1994; 33: 111–115.
- [30] Lawless S. Crying wolf: false alarms in a pediatric intensive care unit. *Crit Care Med* 1994; 22: 981–985.
- [31] Novaes M, Knobel E, Bork A, Pavão O, Nogueira-Martins L, Ferraz M. Stressors in ICU: perception of the patient, relatives and health care team. *Intensive Care Med* 1999; 25: 1421–1426.
- [32] O’Carroll T. Survey of alarms in an intensive therapy unit. *Anaesthesia* 1986; 41: 742–744.
- [33] Organisation for Economic Co-operation and Development. The looming crisis in the health workforce: how can OECD countries respond? Paris: OECD 2008.
- [34] Reslan ZA. *Clinical alarm management and noise reduction in hospitals*. Storrs, CT: University of Connecticut 2007.
- [35] Saeed M, Lieu C, Raber G, Mark R. MIMIC II: a massive temporal ICU patient database to support research in intelligent patient monitoring. *Comput Cardiol* 2002; 29: 641–644.
- [36] Sanderson P, Wee A, Lacherez P. Learnability and discriminability of melodic medical equipment alarms. *Anaesthesia* 2006; 61: 142–147.
- [37] Siebig S, Kuhls S, Imhoff M, Gather U, Schölmerich J, Wrede C. Intensive care unit alarms – how many do we need? *Crit Care Med* 2010; 38: 451–456.
- [38] Siebig S, Kuhls S, Imhoff M, et al. Collection of annotated data in a clinical validation study for alarm algorithms in intensive care – a methodologic framework. *J Crit Care* 2010; 25: 128–135.
- [39] Siebig S, Sieben W, Kollmann F, et al. Users’ opinions on intensive care unit alarms – a survey of German intensive care units. *Anaesth Intensive Care* 2009; 37: 112–116.
- [40] Tappan J, Daniels J, Slavin B, Lim J, Brant R, Ansermino J. Visual cueing with context relevant information for reducing change blindness. *J Clin Monit Comput* 2009; 23: 223–232.
- [41] Topf M. Hospital noise pollution: an environmental stress model to guide research and clinical interventions. *J Adv Nurs* 2000; 31: 520–528.
- [42] Tsien C, Fackler J. Poor prognosis for existing monitors in the intensive care unit. *Crit Care Med* 1997; 25: 614–619.

-
- [43] Ulrich R, Quan X, Zimring C, Joseph A, Choudhary R. The role of the physical environment in the hospital of the 21st century: a once-in-a-lifetime opportunity. Concord, CA: The Center for Health Design 2004.
- [44] Watson M, Sanderson P. Sonification supports eyes-free respiratory monitoring and task time-sharing. *Hum Factors* 2004; 46: 497–517.
- [45] Yang P, Dumont G, Ansermino J. Adaptive change detection in heart rate trend monitoring in anesthetized children. *IEEE Trans Biomed Eng* 2006; 53: 2211–2219.
- [46] Yang P, Dumont G, Ansermino J. A Cusum-based multilevel alerting method for physiological monitoring. *IEEE Trans Inf Technol Biomed* 2010; 14: 1046–1052.
- [47] Zhang Y, Szolovits P. Patient-specific learning in real time for adaptive monitoring in critical care. *J Biomed Inform* 2008; 41: 452–460.

Received November 1, 2010; accepted February 16, 2011; online first March 3, 2011