Intensive Care Med (1999) 25: 1360–1366

ORIGINAL

M.-C. Chambrin
P. Ravaux
D. Calvelo-Aros
A. Jaborska
C. Chopin
B. Boniface

Multicentric study of monitoring alarms in the adult intensive care unit (ICU): a descriptive analysis

Received: 19 March 1999
Accepted: 16 August 1999

This study was supported by grants from the University Hospital of Lille and PHRC no. 9501

M.-C. Chambrin (✉) · D. Calvelo-Aros · A. Jaborska · C. Chopin
UPRES 2689, IFR 22, University of Lille 2, F-59037 Lille, France

P. Ravaux · B. Boniface
Laboratoire de Biomathématiques, University of Lille 2, F-59037 Lille, France

Address for correspondence:
M.-C. Chambrin, UPRES 2689, ITM, Bâtiment Vancostenobel, CH&U de Lille, F-59037 Lille, France,
e-mail: chambrin@lille.inserm.fr,
Tel: + 33-3-20446733,
Fax: + 33-3-20446745

Abstract *Objectives*: To assess the relevance of current monitoring alarms as a warning system in the adult ICU.
Design: Prospective, observational study.
Settings: Two university hospital, and three general hospital, ICUs.
Patients: Hundred thirty-one patients, ventilated at admission, from different shifts (morning, evening, night) combined with different stages of stay, early (0–3 days), intermediate (4–6 days) and late (> 6 days).
Interventions: Experienced nurses were asked to record the patient's characteristics and, for each alarm event, the reason, type and consequence.
Measurements and main results: The mean age of the patients included was 59.8 ± 16.4 and SAPS1 was 15.9 ± 7.4. We recorded 1971 h of care. The shift distribution was 78 mornings, 85 evenings and 83 nights; the stage distribution was 88 early, 78 intermediate and 80 late. There were 3188 alarms, an average of one alarm every 37 min: 23.7 % were due to staff manipulation, 17.5 % to technical problems and 58.8 % to the patients. Alarms originated from ventilators (37.8 %), cardiovascular monitors (32.7 %), pulse oximeters (14.9 %) and capnography (13.5 %). Of the alarms, 25.8 % had a consequence such as sensor repositioning, suction, modification of the therapy (drug or ventilation). Only 5.9 % of the alarms led to a physician's being called. The positive predictive value of an alarm was 27 % and its negative predictive value was 99 %. The sensitivity was 97 % and the specificity 58 %.
Conclusions: The study confirms that the level of monitoring in ICUs generates a great number of false-positive alarms.

Key words Monitoring, alarm · ICU · Multicentric study

Introduction

The development of new technologies has dramatically increased the number of audio and/or visual alarms available in intensive care units (ICUs) to alert staff of either changes in a patient's condition or equipment malfunction.

The alarms are expected to monitor vital cardiocirculatory and respiratory functions and to increase the patients' safety and quality of care by allowing early detection of any abnormality. About 40 alarms, including ventilatory alarms, electrocardiogram, arterial pressure and pulse oximetry (Table 1) can be activated in routine practice for a patient undergoing mechanical ventilation. Alarms generated by perfusion pump, nutrition pump, automatic syringe and dialysis system, among others, must be added to this list.

In a pediatric study by Lawless, 68 % of the alarms were not justified [1]. They produce noise louder than 80 dB, leading to sleep deprivation [2, 3, 4], and continuous stress for both nurses and patients [5, 6]. As a result of such constant demand, nurses may delay their inter-

Table 1 Description of the available alarms in monitoring systems within an intensive care environment

cardiocirculatory function			
ECG	heart rate (HR)	high and low limits	2
		anormal QRS complex	1
		successive anormal complex	1
	ST segment		1
		noisy signal	1
Arterial pressure	systolic	high and low limits	2
	diastolic	high and low limits	2
	mean	high and low limits	2
Pulmonary arterial pressure	systolic	high and low limits	2
	diastolic	high and low limits	2
	mean	high and low limits	2
respiratory function			
Pulse oximetry	heart rate (HR)	high and low limits	2
	SpO2	low limit	1
	noisy signal		1
Capnography	respiratory rate (RR)	high and low limits	2
	etCO2	high and low limits	2
	Zero		1
Ventilator	RR	high and low limits	2
	VT	high and low limits	2
	VE	high and low limits	2
	Paw	high and low limits	2
	Peep	low limit	1
	FiO2	high limit	1
	apnea		1
Temperature		high and low limits	2
Total			n = 40

vention, trying to recognize life-threatening alarms by sound only. A recent study demonstrated that experienced nurses are able to recognize only 38 % of vital alarms [7]. Therefore, this practice could have severe consequences when the patient's condition is deteriorating.

The microprocessor-based technology used in current monitors and ventilators makes it possible to design an alarm hierarchy according to the part of the monitoring that is activated. Usually the central monitor is also a microcomputer acting as a repeater. In commercial systems, all data and trends are available but no treatment has yet been implemented to produce "intelligent" alarms. The aims of this study were (1) to determine the number and type of alarms triggered per hour by the monitoring system in adult ICUs, (2) to determine alarm relevance by computing their predictive values, and (3) to emit suggestions to reduce the number of false alarms.

Materials and methods

This study was initiated in the context of a hospital program for clinical research. Five medical ICUs located in two university hospitals and three general hospitals in Northern France were asked to participate. A schedule for protocol experiments was established to record observations at different stages of the patient's stay in the ICU and during the different nurses' shifts. The stages were defined as: early (0–3 days), intermediate (4–6 days) and late (more than 6 days). The shifts were: morning (7:00 a.m.–2:00 p.m.), evening (2:00 p.m.–9:00 p.m.) and night (9:00 p.m.–7:00 a.m.).

Voluntary experienced nurses, paid to record the observations outside their normal job, acted as observers and were asked to record the following data:

1. The characteristics (sex, age, diagnosis and Simplified Acute Physiologic Score (SAPS1) of each patient at admission [8]);
2. The stage, shift, date and starting time of observation, SAPS1, drug therapy for each observation;
3. The type, reason, and consequence for each alarm.

At patient's inclusion, the monitoring should include at least: electrocardiogram (HR), invasive or non-invasive systemic arterial pressure (SAP), pulse oximetry (SpO_2), expired minute ventilation (V_E), maximum airway pressure (Paw, max), and respiratory rate (RR). The monitoring of pulmonary arterial pressure (PAP) as well as capnography ($etCO_2$), and RR measured on capnogram (RR_{etCO2}) was not mandatory, but recorded when available. The monitoring integrated into the ventilator was used, whatever the mode of ventilation. The make of the monitoring equipment was irrelevant, but was to be noted. At the start of the protocol, alarm limits were set by the nurse in charge of the protocol, at a fixed percentage of the initial value observed during a stable period (Ta-

Table 2 Description of the alarm limits initially used in the protocol. Values are given in percentage of the initial value, except for SpO2 and Paw

cardiocirculatory function			
ECG	heart rate (HR)	high and low limit	± 30 %
Arterial pressure	systolic	high and low limits	± 30 %
Pulmonary arterial pressure (facultative)	mean	high and low limits	± 30 %
respiratory function			
Pulse oximetry	(HR) (facultative)	high and low limits	± 30 %
	SpO2	low limit	(x – 5) %
Capnography (facultative)	rhythm (RR)	high and low limits	± 30 %
	etCO2	high and low limits	± 20 %
Ventilator	RR	high limit	± 30 %
	VE	high and low limits	± 30 %
	Paw	high limit	x + 10 cmH20

Table 3 Value of limits characterizing vital alarms

60	≤	HR	≤	140 b/min
80	≤	systolic SAP	≤	160 mmHg
85	≤	SpO2		
3.5	≤	etCO2	≤	9.8 %
5	≤	VE	≤	20 l/min
8	≤	RR	≤	30 c/min
8	≤	PAPmean	≤	30 mmHg

ble 2). When this computation led to a value beyond the vital threshold (Table 3), the alarm limit was set to the corresponding vital threshold value. The medical staff in charge of the patient's care could, after this initial setting, modify the alarm limits according to their medical practice. For each alarm a form had to be filled out. Alarms were classified as: due to staff manipulation, due to technical problems and due to the patient himself (including cough, agitation and complication). The consequence of the alarm was recorded as "nothing done" when appropriate, or the acts performed following the alarm were noted. For data analysis, an alarm was classified according to its degree of emergency: an alarm was considered as vital when triggered by a parameter value outside the limits reported in Table 3. Otherwise it was classified as an alert. On discharge from the ICU, the patient protocol book was completed with duration of stay and outcome.

The alarm was considered as:

1. True-positive (TP) when followed by an action whatever the reason, including technical problems solved by staff (e. g. sensor repositioning);
2. False-positive (FP) when leading to "no action";
3. False-negative (FN) when, apart from the planned therapy, a drug or a ventilatory setting modification induced by clinical status or laboratory results or a modification of alarm limits was performed without prior audible alarm;
4. True-negative (TN) was assumed for this study to correspond to one potential alarm sounding per time interval calculated as the mean interval between alarm occurrences as observed during the whole study.

Any patient undergoing mechanical ventilation on admission could be included in the study. When weaned during the protocol, the patient remained in the study. For a given patient, observations were made in the early stage during one shift, then in the intermediate stage during another shift and finally in the late stage during the third shift. As some patients could have left the ICU before the end of the protocol, some could be included at the intermediate, or even late, stage. The study was conducted in the five ICUs over a 16-month period, between November 1996 and February 1998. One hundred thirty-one patients were included (76 male (58 %) and 55 female (42 %)) corresponding to 246 observations: one patient was present for four shifts, 36 for three, 40 patients for two shifts and 54 for one shift. The distribution by shifts was 78 mornings, 85 evenings, 83 nights; the distribution by stages was 88 early, 78 middle, 80 late, totaling 1971 h of care.

The data issued from the manual records were codified for statistical analysis performed with the SAS software (SAS Institute, Carry, N. C.). Descriptive and hypothesis testing methods such as the chi-squared test were used. Predictive values were computed as follows: positive predictive value PPV = TP/(TP + FP), negative predictive value NPV = TN/(TN + FN), sensitivity Se = TP/(TP + FN) and specificity Sp = TN/(TN + FP).

Results

The patients

The patients' mean age was 59.8 ± 16.4 (median = 64) years, stay duration was 22.1 ± 19.5 (median = 16) days and SAPS1 at admission was 15.9 ± 7.4 (median = 15). The main etiologies were: acute cardiovascular failure (19.1 %), acute respiratory distress syndrome (20.6 %), acute respiratory failure of chronic pulmonary diseases (32 %), severe neurologic diseases (25.2 %), polytrauma (3.8 %), peripartum distress (1.5 %), self-poisoning (2.3 %), miscellaneous (3.1 %). Of the 131 patients, 87 survived (66.4 %) and 44 died (33.6 %).

The alarms

There were 3,188 alarms. The mean interval between alarm occurrences was 37 min per patient. In terms of drug therapy, 69 % of alarms occurred in non-sedated

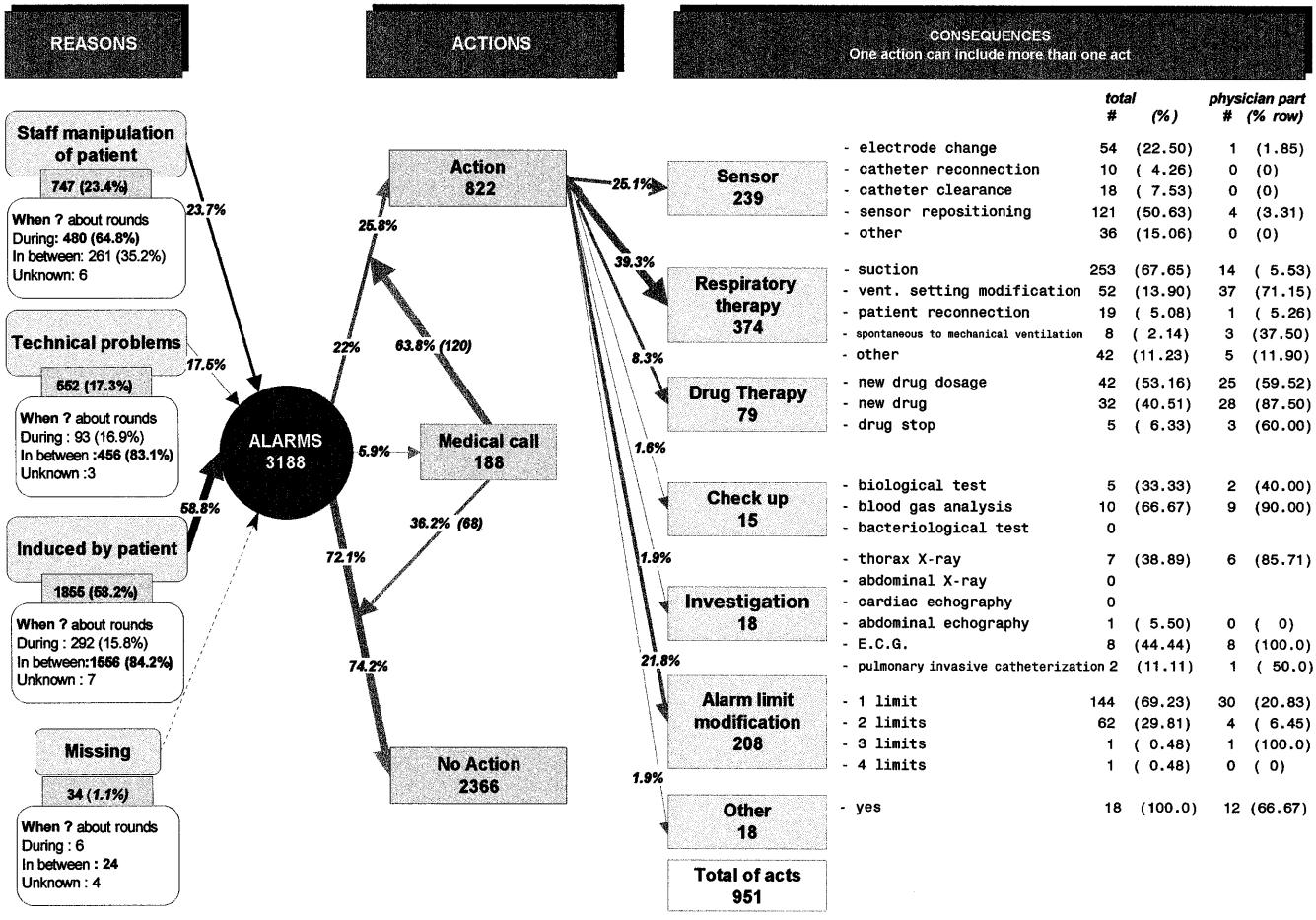

Fig. 1 Summary of the reasons and consequences of the alarms

patients. The distribution of alarms according to nurses' shifts was significantly different (morning: 37.81 %, evening: 34.21 %, night: 26.95 %, chi-squared test $p = 0.001$ computed with a distribution equivalent to 7/24, 7/24 and 10/24, respectively, i.e. the duration of each shift). The distribution of alarms according to the stages was also significantly different (early: 31.93 %, intermediate: 28.26 %, late: 39.81 %, chi-squared test $p = 0.001$ computed with a distribution equivalent to 88/246, 78/246 and 80/246, respectively, i.e. the number of shifts for each stage).

An overview of the alarm generation and its consequences is given in Fig. 1. The reason for the alarm was known in 99 % of cases: 24 % were induced by staff manipulation, 17 % by technical problems and 59 % by the patient. Ventilators were the source of the alarm in 37.8 % (26.8 % of which led to action), cardiovascular monitors in 32.6 % (25.6 % of which led to action), pulse oximeters in 14.9 % (39 % of which led to action) and capnography in 13.5 % (8.8 % of which led to action). Only 5.9 % of the alarms led to calls for a physician (among which 51.3 % were from cardiovascular monitors, 23.0 % from pulse oximeters, 20.9 % from ventilators and 4.8 % from capnography), who modified the therapy in about 2/3 of the cases. Consequences such as sensor repositioning, suction or modification of the therapy (drug or ventilation) arose from 25.8 % of the alarms.

The value of the parameter triggering the alarm was known for only 1982 (62.2 %) alarms: 68.8 % were classified as alerts and 31.8 % as vital alarms. Among the vital alarms, 28.5 % were induced by staff manipulation, 12.3 % by technical problems and 59.2 % by the patients (Table 4). According to the number of TP (880), FN (2308), FN (24) and TN (3196) alarms, the PPV was 27 %, NPV was 99 %, sensitivity was 97 % and specificity was 58 %.

Discussion

The characteristics of our patients are similar to those of the population of patients admitted into French ICUs [9]. The death rate corresponds to that predicted by the SAPS1. For a 10-bed ICU, on average, 390 alarms are

Table 4 Distribution of alerts and vital alarms according to the alarm reasons

type of alarm	alerts					vital alarms				
	staff	techn	patient	missing	total	staff	techn	patient	missing	total
HR (ECG)	42	19	76	0	137	28	52	61	0	141
SAP	22	10	119	2	153	54	15	70	1	140
SPO2	50	35	133	0	218	9	7	22	0	38
HR (SPO2)	0	0	1	0	1	0	0	0	0	0
etCO2	29	1	91	1	122	8	0	14	0	22
RR (etCO2)	8	0	77	0	85	10	0	84	0	94
VE	82	8	216	1	307	48	0	36	0	84
Paw, max	53	1	223	1	278	2	0	21	0	23
RR	9	6	39	0	54	8	1	39	0	48
PAP	4	1	4	0	9	9	1	18	0	28
total	299	81	979	5	1364	176	76	365	1	618

activated a day. Only one quarter of these are followed by a nurse's or physician's action. A few (6 %) lead to a call for the physician. In our study, the nurses considered only 28 % of the alarms as requiring any action, mainly concerning sensors or alarm limits (47 %). Bronchial aspiration represents 26.6 % of the interventions. Physicians called in decided to perform medical action in 64 % of the cases, mainly concerning drug or ventilatory therapy planning.

During the protocol, no major event related to worsening of patients' status occurred without previous alarm, suggesting that the current monitoring is effective in detecting vital problems and that the work was well carried out by the medical staff. This is corroborated by the very high NPV. However, too many alarms were activated in the ICU. Considering the huge number of false-positive alarms reported in this study, one could suppose that many of them were induced by the design of our protocol. Indeed the level of initial mandatory monitoring was perhaps higher than required by the patient's status and the alarm limits were arbitrarily set at ± 30 % of the monitored value as long as they fell into the vital thresholds. However, the minimal level of monitoring was defined according to practice and was established based on the physicians' opinions on a correct setting, and our results are quite similar to those obtained in a study conducted in a pediatric ICU, where 94 % of the alarms were reported to have no clinical relevance [1]. This last study was performed in a 16-bed ICU over 7 consecutive days, corresponding to 928 h of care. The staff member who silenced the alarm also recorded the sounding and made the classification according to: false (including technical alarms), significant (resulted in change of therapy) and induced (by staff manipulation). On average there was one alarm every 26 min. This average is close to the rate observed in our study.

When computing predictive values we assumed that an alarm was true-positive if it was followed by an action, for whatever reason. A technical problem solved by the staff was therefore considered a true-positive alarm. Other definitions could be proposed: in [1], for example, only alarms resulting in a change of therapy were considered as true-positive, and true-negative was defined as no alarm sounding for 5 min of monitoring time. In our study, we have chosen to compute the true-negative value from mean time interval between two consecutive alarms. Depending on the mode of calculation, therefore, the predictive values are slightly different but the conclusions remain identical.

We tried to identify the main reasons for the alarms by classifying according to the categories staff manipulation, technical problems and complications due to the patient. It can be noticed that staff manipulation was responsible for 23.4 % of alarms, 84 % of which led to no action. This means that 20 % of audible alarms could be avoided by the staff's using the key to suspend the alarm sound for 2 min that is available on all the monitoring systems. Room layout often restricts its use and centralization of this functionality could be helpful without additional risk for the patient as the visual alarm remains active. There were more alarms during morning shifts than during evening and night shifts; this is due to greater staff manipulation during the morning than during the evening and night: 41 %, 29.5 % and 29.5 % in morning, evening and night shifts, respectively; the difference is significant.

Technical problems were also more frequent during morning and evening shifts than during the night (38.8 %, 39.3 % and 21.9 %, respectively, in morning, evening and night shifts; the difference is significant). They were often transient alarms without intervention (in 65 % of the cases), mainly due to SpO_2, HR(ECG) and SAP. These technical problems required an intervention in 45 % of the cases for SpO_2 (sensor repositioning), 53 % for SAP (catheter clearance) and only 24 % for ECG (electrode defects). Such an observation leads to the conclusion that technical improvement of sensors is required.

Patient-induced alarms represented 58.2 % of the alarms, 72 % of which were transient, mainly due to coughing or agitation and not followed by medical staff action. To decrease the number of transient alarms, the development of procedures for trend and multisignal analysis seems necessary. The percentage of sedated patients was 61.4 %, 47.4 % and 35.4 %, respectively, in the early, intermediate and late stages; this significant difference could explain the relatively high rate of alarms observed during the late stage.

Regarding the type of alarms, 27 % of alarms on ventilators led to interventions, half of which were bronchial suction. An alarm on a ventilator is the main reason for modification of the ventilatory therapy. On cardiovascular monitors, the majority of alarms concerned sensor malfunction and 26 % led to interventions. An alarm on SAP is the main reason for the modification of drug therapy. On SpO_2, nearly half of the alarms concerned the sensor itself, so an improvement of sensor fastening is desirable. On capnography, only 9 % of the alarms led to any action and 75 % of these were changes in alarm limits. The reason could be lack of understanding of how to use this parameter.

The classification between alerts and vital alarms was performed to test if the determination of a given range for alarm limits could improve the relevance of the monitoring system. Values of the limits that characterized vital alarms (Table 3) were determined on the strength of clinical practice. There is no standard to determine these thresholds. They vary from one study or one ICU to another and are also dependent on the patient's type. One study conducted on cardiac postoperative patients provided clinically suggestive ranges for alert and alarm limits on hemodynamic parameters [10]. Limit values were given as the median value obtained for 21 clinicians. Except for $etCO_2$, the limits were similar. The difference observed for $etCO_2$ (2.5–6.5 % versus 3.5–9.8 % in our study) is strongly dependent on the patient's pathology. Whatever the threshold values, it could be noticed that quite a lot of alarms related to technical problems were classified as vital alarms (Table 4). Taking into account a larger range obviously induces a decrease in the number of false-positive alarms but, alternatively, an increase in the number of false-negative alarms. So, the classification cannot be done properly using only threshold values and it is necessary to develop more sophisticated criteria to reach a better level of information.

A review of the problems was presented in [11] and some suggestions were given to remedy some of them. They concern attempts to standardize warning by the standard-generating organizations (International Standards Organization (ISO), Comité Européen de Normalisation (CEN) and American National Standards Institute (ANSI), among others). Suggestions concern mainly the design of auditory warnings: "ideally the signal produced by the alarm system should not only indicate that something is wrong but also give some idea of what the problem is. Each auditory warning should have the same meaning wherever it is encountered"[11]. Another way is to produce "intelligent" alarms. A review of this perspective is given in [12]: in classical bedside monitoring systems, the patient's signals are sensed, amplified, filtered and acquired. Afterwards, these signals are processed by algorithms that allow waveform detection. Some data are computed as characteristic parameters. Some of them may trigger alarms, usually for high and low limits. As the number of signals to be handled grows, a large number of alarm thresholds must be set at the start of the monitoring. The quality of such alarms deteriorates quickly as the monitored signals are corrupted by noise. All these problems have led to the proposal of "intelligent" alarm schemes where other factors along with the signal being monitored can be considered, e.g., pathology, clinical interventions and so on. One option should be to compute trends and redundancy procedures and to include some knowledge procedures to produce a higher level of information. Much work remains to be done to improve the alarm specificity and there is a need for research into the optimal use of monitoring equipment.

In conclusion, this descriptive study confirms that the level of monitoring in adult ICUs generates a great number of false-positive alarms. The results suggest taking some steps to improve the relevance of the monitoring: (1) to edit recommendations to define how much monitoring is necessary for patients according to the severity of the disease,(2) to modify nurses' habits and room ergonomics to reduce the number of alarms during nursing,(3) to refine the alarm algorithms in order to reduce markedly the frequency of false-positive alarms. At present we are testing some new procedures based on parameter trends and multiparametric analysis, on data and alarm messages automatically collected from the cardiovascular monitor and the ventilator every 5 s and stored in a database.

Acknowledgements The authors would like to thank the ICU medical staff of the centers who participated in the data collection: Service de Réanimation Polyvalente (Prof. Chopin), CH&U de Lille; Service de Réanimation Médicale (Prof. Beaucaire), CH de Tourcoing; Service de Réanimation Médicale (Dr. Moulront), CH de Dunkerque; Service de Réanimation Médicale (Dr. Thévenin), CH de Lens; Service de Réanimation Médicale (Dr. Socolovsky), CH de Valenciennes. The authors are also grateful to Janette Andre who corrected the English manuscript.

References

1. Lawless ST (1994) Crying wolf: false alarms in a pediatric intensive care unit. Crit Care Med 22: 981–985
2. Balogh D, Kittinger E, Benzer A, Hackl JM (1993) Noise in the ICU. Intensive Care Med 19: 343–346
3. Meyer TJ, Eveloff SE, Bauer MS, Schwartz WA, Hill NS, Millman RP (1994) Adverse environmental conditions in the respiratory and medical ICU settings. Chest 105: 1211–1216
4. McLaughlin A, McLaughlin B, Elliott J, Campalani G (1996) Noise levels in a cardiac surgical intensive care unit: a preliminary study conducted in secret. Intensive Crit Care Nurs 12: 226–230
5. Aaron JN, Carlisle CC, Carskadon MA, Meyer TJ, Hill NS, Millman RP (1996) Environmental noise as a cause of sleep disruption in an intermediate respiratory care unit. Sleep 19: 707–710
6. Kam PCA, Kam AC, Thompson JF (1994) Noise pollution in the anaesthetic and intensive care environment. Anaesthesia 49: 982–986
7. Cropp AJ, Woods LA, Raney D, Bredle DL (1994) Name that tone: the proliferation of alarms in the intensive care unit. Chest 105: 1217–1220
8. Le Gall JR, Loirat P, Alpérovitch A, et al. (1984) A simplified acute physiologic score for ICU patients. Crit Care Med 2: 975–977
9. The French multicenter group of ICU research, INSERM U169 (1989) Factors related to outcome in intensive care: French multicenter study. Crit Care Med 17: 305–308
10. Koski EM, Mäkivirta A, Sukuvaara T, Kari A (1995) Clinician's opinions on alarm limits and urgency of therapeutic responses. Int J Clin Monit Comput 12: 85–88
11. Meredith C, Edworthy J (1995) Are there too many alarms in the intensive care unit? An overview of the problems. J Adv Nurs 21: 15–20
12. Mora FA, Passariello G, Carrault G, Le Pichon JP (1993) Intelligent patient monitoring and management systems: a review. IEEE Engineering in Medicine and Biology December: 23–33