Masking Between Reserved Alarm Sounds of the IEC 60601-1-8 International Medical Alarm Standard: A Systematic, Formal Analysis

Matthew L. Bolton, University at Buffalo, The State University of New York, USA, **Judy R. Edworthy**, University of Plymouth, United Kingdom, and **Andrew D. Boyd**, University of Illinois at Chicago, USA

Objective: In this work, we systematically evaluated the reserved alarm sounds of the IEC 60601-1-8 international medical alarm standard to determine when and how they can be totally and partially masked.

Background: IEC 60601-1-8 gives engineers instruction for creating human-perceivable auditory medical alarms. This includes reserved alarm sounds: common types of alarms where each is a tonal melody. Even when this standard is honored, practitioners still fail to hear alarms, causing practitioner nonresponse and, thus, potential patient harm. Simultaneous masking, a condition where one or more alarms is imperceptible in the presence of other concurrently sounding alarms due to limitations of the human sensory system, is partially responsible for this.

Methods: In this research, we use automated proof techniques to determine if masking can occur in a modeled configuration of medical alarms. This allows us to determine when and how reserved alarm sound can mask other reserved alarms and to explore parameters to address discovered problems.

Results: We report the minimum number of other alarm sounds it takes to both totally and partially mask each of the high-, medium-, and low-priority alarm sounds from the standard.

Conclusions: Significant masking problems were found for both the total and partial masking of high-, medium-, and low-priority reserved alarm sounds.

Application: We show that discovered problems can be mitigated by setting alarm volumes to standard values based on priority level and by randomizing the timing of alarm tones.

Keywords: medical devices and technologies, audition, patient safety, psychophysical methods, computational modeling

INTRODUCTION

Healthcare personnel are failing to react to medical alarms. The Pennsylvania Patient Safety Authority (ECRI Institute & ISMP, 2009) documents 194 incidents of healthcare workers not responding to telemetry monitoring alarms from June 2004 to December 2008 (12 of which produced fatalities). A Sentinel Event Alert (The Joint Commission, 2013a; covering January 2009–June 2012) described 98 alarm nonresponses, where five led to extended hospital stays, 13 resulted in "permanent loss of function," and eight resulted in deaths. Not surprisingly, the ECRI Institute has identified medical alarms as one of the most significant technological hazards to patient safety for over a decade (ECRI Institute, 2014; Stead & Lin, 2009).

There are many reasons why humans may not respond to medical alarms (Edworthy, 2013; Edworthy, McNeer, et al., 2018; Phansalkar et al., 2010). The perceivability of medical alarms in the presence of other medical alarms is a contributor to this problem (ECRI Institute, 2014; The Joint Commission, 2013a, 2013b; Vockley, 2014). In particular, the Joint Commission's National Patient Safety Goal from 2014 to "improve the safety of clinical alarm systems" said that "individual alarm signals are difficult to detect" (The Joint Commission, 2013b).

The IEC 60601-1-8 international medical alarm standard (IEC 60601-1-8:2006+AMD1:2012, 2012) was specifically created to give engineers guidance about how to design and test medical alarms so that they are "readily discernible without being unnecessarily distracting or disturbing." To accomplish this, the standard contains a set of reserved alarm sounds (for common alarm conditions). Unfortunately, the melodic patterns of

Address correspondence to Matthew L. Bolton, Department of Industrial and Systems Engineering, University at Buffalo, the State University of New York, Buffalo, NY 14260, USA; e-mail: mbolton@buffalo.edu

HUMAN FACTORS
2022, Vol. 64(5) 835–851
DOI:10.1177/0018720820967596

tones that are specified within IEC 60601-1-8 make them particularly susceptible to simultaneous masking: a condition where multiple sounds interact in a way that prevents the human sensory system from hearing one of or more of them (Fastl & Zwicker, 2006).

Simultaneous auditory masking has been acknowledged as a hazard of alarms by many experts and researchers (Edworthy & Hellier, 2005, 2006; Edworthy & Meredith, 1994; Konkani et al., 2012, Meredith & Edworthy, 1995; Patterson, 1982; Patterson & Mayfield, 1990) and experimentally detected in clinical settings (Momtahan et al., 1993; Toor et al., 2008). As such, simultaneous masking is a real problem because, if practitioners fail to hear an alarm, they will not be able to respond to it. Thus, it is one of the factors contributing to practitioners failing to respond to medical alarms (The Joint Commission, 2013b) and its associated negative health outcomes (ECRI Institute, 2014; Stead & Lin, 2009).

Simultaneous masking can be very difficult to detect experimentally because it may only occur with very specific interactions of multiple, concurrently sounding medical alarms. For this reason, we developed formal computational methods (Bolton, Edworthy, Boyd, et al., 2018, Bolton et al., 2016, Hasanain et al., 2014, 2016, 2017) that use automated proof techniques and the psychoacoustics of simultaneous masking to determine if masking exists in a modeled configuration of medical alarms. We have validated these methods with human participant experiments (Bolton et al., 2020) and used them to discover serious problems with real alarms designed in conformance with IEC 60601-1-8 (Bolton, Edworthy, Boyd, et al., 2018; Bolton et al., 2016, Hasanain et al., 2017). However, nobody has investigated the masking potential of the reserved alarms (common alarms to be used across devices) that are specified in the standard.

In this work, we use the latest version of our method (Bolton, Edworthy, Boyd, et al., 2018) to systematically evaluate the IEC 60601-1-8 alarm sounds. In what follows, we provide a background for understanding our method and a deeper description of our research objectives. We then describe the methods we employed in our analyses and their results. We then discuss our results and explore future research possibilities.

BACKGROUND

In this section, we discuss the information necessary for understanding our research. This includes background on our method and the IEC 60601-1-8 standard.

Our Method

Our method (Bolton, Edworthy, Boyd, et al., 2018; Bolton et al., 2016; Hasanain et al., 2014, 2016, 2017) detects simultaneous masking by using a unique combination of psychoacoustics and model checking to determine if masking is ever possible in a modeled configuration of medical alarms.

The psychoacoustics of simultaneous masking (Bosi & Goldberg, 2003) mathematically represent how the volume and tone/frequency of a sound cause masking. These are based on how the sensitivity of sensory cells on the inner ear's basilar membrane changes in the presence of other sounds. This threshold shift is represented as a masking curve. For tonal sounds, like those used in IEC 60601-1-8, the masking curve for a given masking sound (the masker) is represented as:

$$\begin{aligned} curve_{masker}(z_{maskee}) = \; & spread_{masker}(z_{maskee} - z_{masker}) \cdot \\ & v_{masker} - 6.025 - 0.275 \cdot z_{masker}, \end{aligned} \tag{1}$$

where z_{maskee} and z_{masker} are the frequency of the potentially masked sound (the maskee) and the masker, respectively, on the Bark scale (Zwicker & Feldtkeller, 1967); and v_{masker} is the volume of the masker in dB. The spreading function spread_{masker} represents how the masking effect changes as the frequency distance between the maskee and masker ($z_{maskee} - z_{masker}$ or δz) changes:

$$\text{spread}_{masker}(\delta z) = \begin{cases} -17 \cdot \delta z + 0.15 \cdot v_{masker} \cdot (\delta z - 1) & \text{for } \delta z \geq 1 \\ -17 \cdot \delta z & \text{for } 0 \leq \delta z < 1 \\ -\left(6 + 0.4 \cdot v_{masker}\right) \cdot |\delta z| & \text{for } -1 \leq \delta z < 0 \\ -6 \cdot |\delta z| - 0.4 \cdot v_{masker} - 11 & \text{for } \delta z < -1. \end{cases} \tag{2}$$

Masking effects represented in masking curves is additive, where masking potential increases with the number of sounds in the environment (Lutfi, 1983). The additive process produces a new absolute threshold of hearing (in dB) of a sound (the potential maskee) in the presence of N masker sounds with the formulation (Bosi & Goldberg, 2003):

$$\begin{aligned}\mathrm{mtresh}_{maskee} = \quad & 10 \cdot \log\Bigg(10^{(\mathrm{abs}_{maskee})/10} \\ & + \Bigg(\sum_{n=1}^{N} \Big(10^{\left(\mathrm{curve}_{maker_n}(z_{maskee})\right)/10} \Big)^{\alpha} \Bigg)^{1/\alpha} \Bigg). \end{aligned} \tag{3}$$

In this, α is a constant that can vary based on the type of sounds. In our work, we use $\alpha = 0.33$ because it is most appropriate for tonal sounds masking other tonal sounds. Furthermore, abs_{maskee} is the original absolute threshold of hearing at the maskee's frequency. This is represented by:

$$\begin{aligned}\mathrm{abs}_{maskee} = \quad & 3.64 \cdot \left(f_{maskee}/1000\right)^{-0.8} \\ & -6.5 \cdot e^{-0.6\left(f_{maskee}/1000-3.3\right)^2} \\ & +10^{-3} \cdot \left(f_{maskee}/1000\right)^4 , \end{aligned} \tag{4}$$

where f_{maskee} is the maskee's frequency in Hz.

With these psychoacoustics, if the potential maskee's volume is less than or equal to the new threshold ($\mathrm{mthresh}_{maskee}$; Equation 3), the maskee will be simultaneously masked.

These psychoacoustics have been validated and used in many applications (Ambikairajah et al., 1997; Brandenburg & Stoll, 1994). This includes being the basis of the MPEG and other "lossy" audio codecs (Bosi & Goldberg, 2003). Specifics about the psychoacoustics used in our method can be found in Bolton, Edworthy, Boyd, et al. (2018) and Hasanain et al. (2017).

Model checking (Clarke et al., 1999) is an automated approach to performing mathematical proofs (called formal verification), which comes from the larger discipline of formal methods. To perform model checking, an analyst must create a formal model that captures a target system's behavior. This is usually represented as a collection of concurrently executing state machines: a set of variables and transitions between variable values. Specification properties are used to assert desirable system conditions using a combination of model variables, Boolean logic operators, and temporal operators (Emerson, 1990). The model check approach to formal verification proves whether or not the model satisfies the specification by exhaustively searching the system model's statespace. The specification is proven true if no violation is found. If one is discovered, the model checker produces a counterexample: a trace through the model that shows how the violation manifested. Model checking is predominately used to verify computer software and hardware. However, research has been showing that it can be used in human factors (Bolton et al., 2013; Bolton, 2017; Weyers et al., 2017) and medical systems engineering (Bolton & Bass, 2009, 2010; Bolton et al., 2012).

When psychoacoustics and model checking are combined together in our method (Figure 1; Bolton, Edworthy, Boyd, et al., 2018; Bolton et al., 2016, Hasanain et al., 2017), analysts use an Excel spreadsheet to model alarms. Software then automatically generates formal models of the represented alarms along with the specification properties for checking whether masking can occur. The method supports the ability to detect both the partial and total masking of each modeled alarm in the configuration. When one of these properties is checked, the model checker will consider all the possible alarm sound interactions to see if it can discover masking. The newest version of the method (Bolton, Edworthy, Boyd, et al., 2018) significantly improves the scalability/efficiency of the approach compared with previous versions (Bolton et al., 2016, Hasanain et al., 2014, 2016, 2017) to enable its use with industrial applications.

It is beyond the scope of this article to describe the details of how the formal model in the method (Figure 1) is constructed. Details about this can be found in Bolton, Edworthy, Boyd, et al. (2018). A conceptual overview is shown in Figure 2. Specifically, the formal model contains variables to represent the state of each of the N alarms as they change over time. Each state where the alarm is producing sound is given a unique identifier (i.e., $Alarm_{1,2}$ is the state when alarm one is sounding its second tone). Whenever any alarm is not producing sound (it is not sounding or is in a pause), it is in state $Alarm_0$. The start time of each alarm is an open parameter that can assume

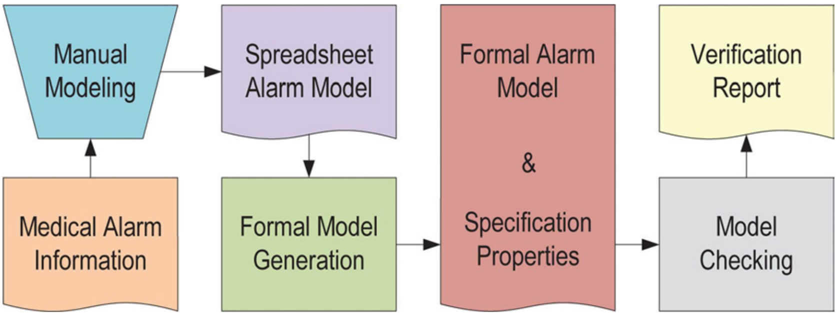

Figure 1. The method for using model checking to discover masking between concurrently sounding medical alarms (Bolton, Edworthy, Boyd, 2018; Bolton, Edworthy, Boyd, et al., 2018).

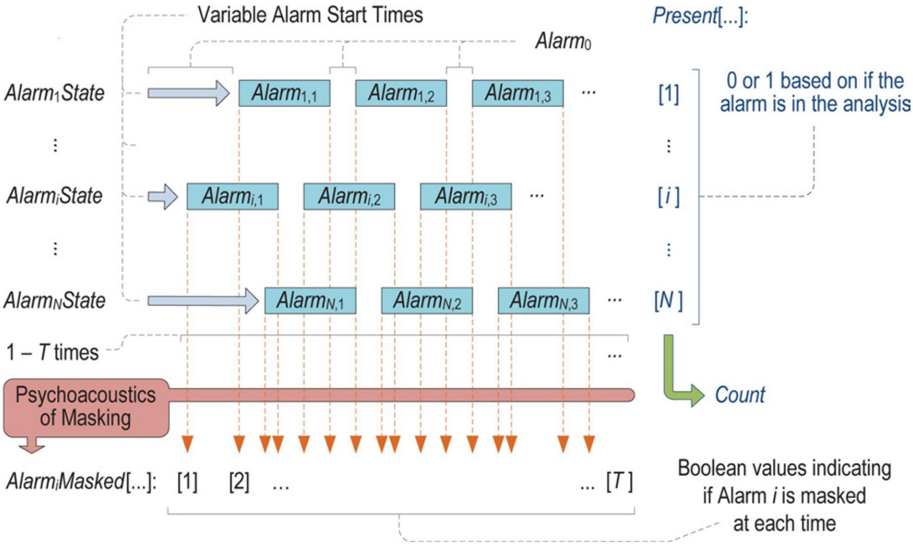

Figure 2. Conceptual overview of how the formal alarm model from Figure 1 works. Note that the *Present* and *Count* variables are introduced as part of the work presented in this paper and are discussed in subsequent sections.

any real-valued time that is ≥ 0 (thus allowing the model checker to consider any potential overlapping of alarms). Based on a given set of alarm start times, across all alarm states, the model computes a set of T discrete times based on the total number of times when change events happen in the alarms. Then, using the psychoacoustics of simultaneous masking and the physical properties (volumes and frequencies) of all the alarms' states at each time, the model computes an array of Boolean variables called $Alarm_i Masked$ (indexed from 1 to T corresponding to each discrete time) to indicate if a given alarm ($Alarm_i$) is masked at each discrete time. Specification properties can then be asserted over this array so that model checking can determine if a given alarm ($Alarm_i$) can ever be partially or totally masked.

We have applied our method to a number of realistic medical alarms (Hasanain et al., 2014, 2016), including alarms that conform to IEC 60601-1-8 (Bolton, Edworthy, Boyd, et al., 2018; Bolton et al., 2016, Hasanain et al., 2017).

TABLE 1: Melodies Used for IEC 60601-1-8 Reserved Alarm Sounds

Reserved Alarm	Alarm Priority												
	Medium			High									
General	C_4	C_4	C_4	-	C_4	C_4	-	C_4	C_4	C_4	-	C_4	C_4
Cardiac	C_4	E_4	G_4	-	G_4	C_5	-	C_4	E_4	G_4	.	G_4	C_5
Perfusion	C_4	$F^{\#}_4$	C_4	-	C_4	$F^{\#}_4$	-	C_4	$F^{\#}_4$	C_4	-	C_4	$F^{\#}_4$
Ventilation	C_4	A_4	F_4	-	A_4	F_4	-	C_4	A_4	F_4	-	A_4	F_4
Oxygen	C_5	B_4	A_4	-	G_4	F_4	.	C_5	B_4	A_4	-	G_4	F_4
Temperature	C_4	D_4	E_4	-	F_4	G_4	-	C_4	D_4	E_4	-	F_4	G_4
Drug delivery	C_5	D_4	G_4	-	C_5	D_4	-	C_5	D_4	G_4	-	C_5	D_4
Failure	C_5	C_4	C_4	-	C_5	C_4	-	C_5	C_4	C_4	-	C_5	C_4
Low	E_4	E_4											
	1	2	3		4	5		6	7	8		9	10

Note. Numbered letters (C_4 and D_4) are musical tones presented using scientific pitch notations. These have the frequencies shown in Table 2. Numbers under the table indicate the order in which a tone is played in an alarm's melody. The first three tones of each high-priority alarm are the same as the corresponding medium-priority alarm. The last five tones of a high-priority alarm (tones 6–10) are a repetition of the melody from the first five tones of the alarm. There is a pause between each tone in an alarm. A – indicates a different pause, while a – represents yet a third pause. The timing of tones and pauses between tones is described in Table 3.

We have also experimentally validated that our method accurately predicts whether medical alarm tones used to compose IEC 60601-1-8 alarm sounds will be audible to human listeners (Bolton et al., 2020). However, it has not been used to analyze the reserved alarm sounds of IEC 60601-1-8.

IEC 60601-1-8 and its Reserved Alarm Sounds

To help alarm designers and improve alarm recognition across the healthcare industry, the IEC 60601-1-8 international standard has a collection of reserved alarm sounds. These represent common types of alarms with specific meanings that can be used across devices. There are eight high-priority alarms, eight medium-priority alarms, and one low-priority alarm (Table 1). The high-priority alarm melodies have 10 sequential tones with pauses between each. The medium-priority alarms have three sequential tones separated by pauses, where these three tones of medium-priority alarms are the first three tones of their high-priority counterparts (though not necessarily with the same timings and volumes). The low-priority alarm has two sequential tones with a pause in between.

The standard does not require specific volumes and timings of these sounds. Rather, there is a range of acceptable values that conform with the general alarm design guidelines of the standard. The relevant parameters are summarized in Table 3.

It is worth noting that the standard does provide requirements for frequencies for tones in designed alarms (something not topical to this discussion). It also requires the inclusion of at least four additional harmonics (additional frequencies) in each tone of any alarm (designed or reserved), with bounds on the acceptable frequencies and volumes of these. These are typically represented as whole number multiples of the tone's primary frequency at lower volumes. These additional harmonics are included to give alarms harmonic complexity and to mitigate the impact of physical interactions (i.e., dissonance, beating, and physical masking, which are explored more in the Discussion) between primary harmonics on alarm perception. We do not consider these additional frequencies in

TABLE 2: Musical Pitches

Note	Frequency (Hz)
C_4	261.63
D_4	293.66
E_4	329.63
F_4	349.23
$F^{\#}_4$	369.99
G_4	392
A_4	440
B_4	493.88
C_5	523.2

the work presented here because our previous experimental results showed that their inclusion did not impact the ability of people to identify the presence of an alarm tone when its primary frequency was totally masked by other alarms (Bolton et al., 2020).

OBJECTIVE

The problems with masking that have been discovered with IEC 60601-1-8 compliant alarms in previous analyses with our method (Bolton, Edworthy, Boyd, et al., 2018; Bolton et al., 2016, Hasanain et al., 2017), the experimental validation of the predictive capabilities of our approach for alarm sounds like those that compose reserved alarm melodies (Bolton et al., 2020), and the internal similarity of the tones used in the reserved sounds (Table 1) suggest that there may be deeper masking problems with the alarms in the standard. In this research, we applied our method to the evaluation of reserved alarm sounds from the standard. Specifically, we sought to characterize the masking potential of the IEC 60601-1-8 alarms by determining the minimum number of alarms required to totally mask each reserved sound (make it inaudible) for different volumes of that alarm. To accomplish this, we extend our method to enable it to only consider

TABLE 3: IEC 60601-1-8 Alarm Design Parameters

	Alarm Priority		
Characteristic	High	Medium	Low
Num. Tones	10	3	1 or 2
Tone Vol. (dB)	v_H	v_M	v_L
Max. Tone Vol. Diff. (dB)	10	10	10
Tone Duration (s)	t_H	t_M	t_M
Tone Spacing (s):			
b/w $1^{s}$and 2nd	x	y	y
b/w 2nd and 3rd	x	y	
b/w 3rd and 4th	$2x + t_H$		
b/w 4th and 5th	x		
b/w 5th and 6th	z		
b/w 6th and 7th	x		
b/w 7th and 8th	x		
b/w 8th and 9th	$2x + t_H$		
b/w 9th and 10th	x		
Time Between Repeat (s)	[2.5, 15]	[2.5, 30]	[15, ∞) or no repeat

where $v_H \geq v_M \geq v_L$ $t_H \in [0.075, 0.2]$ $t_M \in [0.125, 0.25]$ $x \in [0.05, 0.125]$

$y \in [0.125, 0.25]$ $t_M + y \geq t_H + x$ $z \in [0.35, 1.3]$

5% max. x and y variance in an alarm A blank indicates nonapplicability an analyst-specified number of alarms in a given analysis. This allowed us, across multiple model checking verification analyses, to systematically determine the minimum number of alarms that could mask each of the considered reserved alarm sounds at certain volumes. We used this version of the method to determine the minimum number of alarms required to mask each reserved alarm sound in all the following conditions: (1) each high-priority reserved alarm sound was analyzed in the presence of all of the other high-priority reserved alarm sounds; (2) each medium- and low-priority reserved alarm sound was analyzed in the presence of all the other medium- and low-priority reserved alarm sounds; and (3) each high-priority reserved alarm sound was analyzed in the presence of all the other medium- and low-priority reserved alarm sounds. Next, we describe our methods and results. We then discuss them and their implications for the standard. We ultimately outline directions for future research.

METHODS

Method Extension

In the method (Figure 1; Bolton, Edworthy, Boyd, et al., 2018), there are N alarms in a given configuration. In this work, we extended the method to give it the ability to consider all possible subsets of alarms of a particular size $Num \leq N$. This was done to enable us to determine the minimum number of alarms that could totally mask a given alarm in the configuration across multiple model checker verifications.

This extension was accomplished by adding an input variable called *Present* (shown in Figure 2): an array with N elements, one for each alarm. Each entry in this array was an integer from 0 to 1 indicating if the corresponding alarm was (1) or was not (0) allowed in the given model state. The formal model also added a variable $Count$ computed in every model state as the sum of all N integers in *Present* (Figure 2).

The specification properties used for verifying if a given alarm is masked were also reformulated to account for the new modeling constructs. For a given alarm with an array $Alarm_i Masked$ indicating if it is masked at each time in the model and a *Present* index of i, we can assert that the alarm is never totally masked (never completely inaudible) by *Num* or fewer other alarms with

$$\mathbf{G}\left(\begin{array}{l}\left(Count = Num \wedge Present(i) = 1\right) \\ \Rightarrow \neg\left(\forall_{t \in Times}\, Alarm_i Masked(t)\right)\end{array}\right). \quad (5)$$

Explicitly, this asserts that through all paths through the model (G) if $Count = Num$ and alarm i is present ($Present(i) = 1$) then ($\Rightarrow$) it should not be true ($\neg$) that, for all possible times ($\forall_{t \in Times}$), alarm i is masked ($\forall_{t \in Times}$).

We can assert that the alarm is never partially masked (where any part that is not a pause is inaudible) by *Num* other alarms with

$$\mathbf{G}\left(\begin{array}{l}\left(Count = Num \wedge Present(i) = 1\right) \\ \Rightarrow \left(\forall_{t \in Times}\left(\begin{array}{l}Alarm_i Masked(t) \\ \Rightarrow Alarm_i State(t) = Alarm0\end{array}\right)\right)\end{array}\right). \quad (6)$$

Note that $Alarm_0$ represents the alarm state where the alarm is either not sounding or in a pause between tones. Thus, Equation 6 asserts that through all paths througsh the model (G) if $Count = Num$ and alarm i is present ($Present(i) = 1$) then ($\Rightarrow$), for all possible times ($\forall_{t \in Times}$), if alarm i is masked ($Alarm_i Masked(t)$) then ($\Rightarrow$) that means the alarm is not sounding or in a pause ($Alarm_i State(t) = Alarm0$).

In this modified version of our method, *Num* can be specified by an analyst at the time model checking is performed.

Modeling, Experimental Design, and Analysis

Using our modified method, we created three base configurations of alarms that were used for the analyses: one for each of the (1–3) conditions listed in the Objectives section. Each of these configurations contained a set of reserved alarm sounds from Table 1. The first, which was used to evaluate masking between high-priority reserved alarm sounds, Condition (1), contained all the high-priority reserved alarm sounds. The second, which was used to evaluate masking between medium- and low-priority reserved alarm sounds, Condition (2), contained all the low- and medium-priority reserved alarm sounds. The third, which was used to evaluate the masking of high-priority reserved alarm sounds by lower priority sounds, Condition (3), contained all the low- and medium-priority reserved alarm sounds along with a blank entry used for representing a single high-priority alarm sound. In all three base models, frequencies were assigned the value corresponding to the presented notes (Table 2). All sounds in each configuration within an alarm were given a standard volume of 80 dB ($v_H = 80$ and $v_M = 80$ for high- and medium-priority alarms respectively; see Table 3. Further, to maximize the potential of alarms masking other alarms (and to be consistent with standard alarm design practice), timings were kept consistent between configurations for every alarm with tone times of $t_H = t_M = 0.2$ s. Pause times of $x = y = 0.125$ s and $z = 0.35$ s were used.

Using these base configurations, we created scripts that would systematically generate formal alarm models and specification properties, where the volume of each analyzed (masked) alarm would vary while keeping the volume of the other (masking) alarms at the standard level of 80 dB. For the first two configurations, the scripts would also iteratively set the analyzed alarm to each of the alarms contained in the configuration. For the third configuration, the script would iteratively insert a different high-priority reserved alarm sound into the blank entry, where this was always the analyzed high-priority alarm. Using these scripts, we automatically generated formal models for each alarm, where the volume of the analyzed alarm would range between 40 and 80 dB (in increments of 1 dB) between models. This range of volumes was used specifically because 80 dB represents an expected upper bound on the volume of a medical alarm. Further, the considered range accounts for the 10-dB range in variance in alarm volumes within a given designed alarm system. The additional variance further accounts for potential differences that could arise between the volumes of alarms from independently engineered systems. The 35 dB minimum was chosen because it should be a sufficiently low enough volume to ensure the maximum potential for alarm masking. Finally, it should be noted that while the actual volume level of maskers does impact the shape of the masking curve, where higher volumes have more masking than lower volumes, this variation is very minor (Bosi & Goldberg, 2003). Thus, results of analyses presented with this method should provide a good approximation (with a minor bias toward detection) of how the relative volumes of maskers at any given relative volume levels will manifest.

For each of the generated models, we performed multiple verifications to prove whether the analyzed alarm with the varied volume could ever be totally masked (using the specification property pattern from Equation 5) or partially masked (using the specification property pattern from Equation 6). For all models, the value of *Num* was varied between two and the total number of alarms (N) in the model for every analyzed volume. For example, if $Num = 2$, the verifications would consider any possible combination of two sounding alarms from the modeled configuration.

Verification with model checking was performed for the analyzed alarm for each model, for each of the different volumes (40–80 dB) with values of *Num* ranging from two up to the total number of alarms in the associated configuration. In all cases, we recorded the minimum number of alarms ($Num - 1$) it took to both totally and partially mask the analyzed alarm. Verifications were performed in parallel on a desktop workstation computer at the University at Buffalo with a 12-core 3.60 GHz Intel Xeon E5-1650 CPU with 128 Gigabytes of RAM running Linux Mint.

RESULTS

Verifications for the analyses took between 3.52 s and 335,168.36 s ($\approx$ 4 days) each, with a median verification time of 455.63 s ($\approx$ 7.6 min). Next, we present results across all three of the analysis conditions.

Condition (1): High-Priority Sounds Masking High-Priority Sounds

Results showing the minimum number of alarms required to both totally and partially mask each of the high-priority IEC 60601-1-8

(a) Masked Alarm	… 50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	… 73	74	75	76	77	78	79	80
General	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	3	4	5	5	5	5	6	1	2	2	2	2	2	2	2
Cardiac	1	1	1	1	1	1	1	2	2	2	2	2	2	2	2	2	2	3	3	3	3	3	3	4	4	4	4	5	5	5	5	1	2	2	2	2	2	2	2
Perfusion	1	1	1	1	1	1	1	1	2	2	2	2	2	2	2	2	2	2	2	3	3	3	3	3	3	4	4	4	5	5	5	1	2	2	2	2	2	2	2
Ventilation	1	1	1	1	1	1	2	2	2	2	2	2	2	2	2	2	2	3	3	3	3	3	3	4	4	4	5	5	5	5	6	1	2	2	2	2	2	2	2
Oxygen	1	1	2	2	2	2	2	2	2	2	2	2	2	2	3	3	3	3	3	3	3	3	3	4	4	4	4	5	5	6	6	1	2	2	2	2	2	2	2
Temperature	1	1	1	1	1	1	1	2	2	2	2	2	2	2	2	2	2	2	2	3	3	3	3	3	4	4	4	4	4	4	4	1	2	2	2	2	2	2	2
Drug Delivery	1	1	1	1	1	1	1	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	3	3	3	3	4	4	4	5	1	2	2	2	2	2	2	2
Failure	1	1	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	3	3	3	4	4	5	6	1	2	2	2	2	2	2	2

(b) Masked Alarm	… 50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	… 73	74	75	76	77	78	79	80
General	1	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	3	3	3	3	3	4	4	1	2	2	2	2	2	2	2
Cardiac	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	2	2	2	2	2	2	2	2	2	3	3	3	3	3	3	3	1	2	2	2	2	2	2	2
Perfusion	1	1	1	1	1	1	1	1	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	3	3	3	3	3	4	4	1	2	2	2	2	2	2	2
Ventilation	1	1	1	1	1	1	1	1	1	1	1	2	2	2	2	2	2	2	2	2	2	3	3	3	3	3	4	4	4	4	4	1	2	2	2	2	2	2	2
Oxygen	1	1	1	1	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	3	3	3	3	3	4	4	5	5	6	6	1	2	2	2	2	2	2	2
Temperature	1	1	1	1	1	1	1	1	1	1	1	1	1	1	2	2	2	2	2	2	2	2	2	2	3	3	3	3	3	4	4	1	2	2	2	2	2	2	2
Drug Delivery	1	1	1	1	1	1	1	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	3	3	4	4	4	4	1	2	2	2	2	2	2	2
Failure	1	1	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	3	3	3	4	4	4	5	1	2	2	2	2	2	2	2
Low	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	3	3	3	3	3	3	3	3	3	1	2	2	2	2	2	2	2

(c) Alarm	… 50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	… 73	74	75	76	77	78	79	80
General	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	2	2	2	1	2	0	0	0	0	0	0	0	0
Cardiac	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	0	0	1	1	1	1	1	1	2	1	1	1	2	2	2	2	0	0	0	0	0	0	0	0
Perfusion	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	0	1	1	1	2	1	1	0	0	0	0	0	0	0	0
Ventilation	0	0	0	0	0	0	1	1	1	1	1	0	0	0	0	0	0	1	1	1	1	0	0	1	1	1	1	1	1	1	2	0	0	0	0	0	0	0	0
Oxygen	0	0	1	1	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Temperature	0	0	0	0	0	0	0	1	1	1	1	1	1	1	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0
Drug Delivery	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0
Failure	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0

(d) Alarm	… 50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	… 73	74	75	76	77	78	79	80
General	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	1	2	2	2	2	2	2	2
Cardiac	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	1	2	2	2	2	2	2	2
Perfusion	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	1	2	2	2	2	2	2	2
Ventilation	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	1	2	2	2	2	2	2	2
Oxygen	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	1	2	2	2	2	2	2	2
Temperature	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	1	2	2	2	2	2	2	2
Drug Delivery	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	1	2	2	2	2	2	2	2
Failure	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	1	2	2	2	2	2	2	2

Figure 3. Masking analysis results for (a) high-priority alarms masking high-priority alarms, (b) medium- and low-priority alarms masking medium- and low-priority alarms, and (d) medium- and low-priority alarms masking high-priority alarms. All of these show the minimum number (higher is better) of other alarms it takes to both totally and partially mask each of the considered IEC 60601-1-8 reserved alarm sounds (the entries on the y axis) at different volumes (the x axis). A $\emptyset$ indicates that no masking could occur. (c) shows the difference in comparable numbers between the high-priority alarms from (a) and their medium-priority counterparts from (b). In all cases, color is used to indicate entries with the same value. Ellipses … indicate that results reported for the minimum listed volume are identical for lower evaluated volumes.

reserved alarm sounds by other high-priority alarms at each of the considered volumes is shown in Figure 3(a).

In the total masking results, if all the alarms are at the same volume (80 dB), only the Temperature alarm could be totally masked by fewer than five other alarms. With the 10-dB range allowed by the standard (Table 3), three of the alarms (General, Drug delivery, and Failure) can be totally masked by two other alarms. All the other alarm sounds can be totally masked by three other alarm sounds. All the high-priority alarm sounds could be masked by a minimum of two other alarms for volumes of 63 dB or lower (constituting a volume difference of 17 dB). At volumes of 57 dB or lower, all but the General alarm could be totally masked by one other high-priority sound. There were differences in the performance of the alarms. For example, at higher volumes, the Oxygen alarm is the least susceptible to masking.

In the partial masking results, all high-priority alarm sounds could be partially masked by as few as two other high-priority sounds for volumes between 73 and 80 dB. All alarms could be partially masked by at least one other alarm for volumes of 72 dB or lower. Note that this occurs within the 10-dB range in alarm volumes allowed by the standard.

Condition (2): Medium- and Low-Priority Sounds Masking Medium- and Low-Priority Sounds

Results showing the minimum number of alarms required to both totally and partially mask each of the medium- and low-priority IEC 60601-1-8 reserved alarm sounds by other medium- and low-priority alarm sounds at each of the considered volumes is shown in Figure 3(b).

For the total masking analyses, when all the alarms are at the same volume (80 dB), all but the Cardiac and Low alarms could only be totally masked by four or more other alarms. Within the allowable 10-dB range of variance in alarm volume (Table 3), all but one of the alarms (Oxygen) can be totally masked by at least two other alarms. All analyzed alarms below 65 dB can be totally masked by just one other alarm. Finally, there are clear differences in alarm masking susceptibility at different volume levels. At higher volumes, the Oxygen and Failure alarms are the least susceptible and the Cardiac and Low alarms are the most. At lower volumes, the Oxygen, Failure, and Low alarms are the least susceptible, whereas the Cardiac and Temperature alarms are the most susceptible. The Low alarm could never be totally masked by fewer than two other alarms.

The partial masking results match those of the high-priority alarm sounds, where every alarm could be partially masked by at least two other alarms for volumes ranging from 74 to 80 dB and one other alarm for volumes of 73 dB or lower.

It should be noted that the high-priority alarms were either just as susceptible or less susceptible to masking than their medium-priority counterparts. This is illustrated in Figure 3(c). This shows that when the comparable number from Figure 3(b) is subtracted from its counterpart in Figure 3(a), all the numbers are greater than or equal to 0. Thus, the minimum number of alarms required to mask high-priority alarms is always the same or greater than the minimum number required to mask the respective medium-priority sounds.

Condition (3): Medium- and Low-Priority Sounds Masking High-Priority Sounds

In the analyses of the ability for the medium- and low-priority sounds to mask the high-priority alarm sounds, the analyses revealed that the medium- and low-priority alarms could never totally mask the high-priority sounds; see Figure 3(d). This remained true even when the high-priority sounds’ volumes were as low as 35 dB. The partial masking results, shown in Figure 3(d), replicated those observed for the high-priority alarms: every alarm could be partially masked by one other alarm for volumes of 73 dB (or lower) or two others for higher volumes.

DISCUSSION AND CONCLUSIONS

In the work presented here, we used a modified version of our computational formal method to evaluate the medium- and low-priority alarms of the IEC 60601-1-8 international medical alarm standard. Our results show that masking is a concern for the analyzed alarm sounds. It is true that our results are based on a model-based analyses. However, because previous work validated the predictive capabilities of our method for alarm tones that include the notes used is reserved sound melodies (Bolton et al., 2020), we are confident in the accuracy of our findings. Next, we discuss our results, their generalizability, their implications for the standard, and avenues of future research.

Total Masking

The literature does not indicate how many overlapping alarms people are able to differentiate, even if none of them are masked. However, it is clearly better for alarms to have a higher number of simultaneous alarms required for masking to occur. In any case, it seems reasonable that a human would fail to hear one alarm when five or more are sounding. In this respect, if the alarms are kept at the same level (80 dB), the results of the total masking analyses (Figure 3) are encouraging for three reasons. First, for the high-priority alarms, it takes between four and six other alarms (five and seven alarms sounding concurrently) to totally mask any given alarm; see Figure 3(a). For the medium- and low-priority alarms, it takes a minimum of between three and six other alarms (four and seven sounding concurrently) to totally mask any given alarm. Of these, only two (Cardiac and Low) could be masked by a minimum of three others; see Figure 3(b). The low number observed for the Low alarm is not particularly concerning because it is less important than any of the other alarms that may be masking it. However, an examination of the counterexample produced in the Cardiac alarm analysis shows that it is of equal importance to its three other masking alarms. Thus, masking of the medium-priority Cardiac alarm is a higher risk at this volume. Second, the high-priority alarms were consistently as or less susceptible to simultaneous masking as the lower priority alarms, as shown in Figure 3(c). Third, none of the high-priority alarms could be total masked in the presence of the medium- and low-priority alarms; see Figure 3(d).

However, results at the 70-dB level and below are less encouraging. When an alarm is 10 dB below base line (70 dB), three of the high-priority alarms (General, Drug delivery, and Failure), all but one of the medium-priority alarms (Oxygen), and Low, can be totally masked in the presence of only two other alarms at the same priority level. In fact, most of these are masked by two other alarms at higher levels (see the example in Figure 4). This is concerning because the international alarm standard allows for a 10-dB variation in alarm volumes within a given designed configuration (Table 3). Thus, even within the standard alarm sounds of a given device, there are ways for some high-priority alarms and nearly every lower priority alarm to be masked when three alarms sound concurrently. Moreover, alarm resilience to masking decreases with the volume of the masked alarm. For example, it only takes two alarms at the same priority level to mask all the medium- and low-priority alarms at 69 dB (11 dB below the others), the high-priority Perfusion and Delivery alarms at 68 dB (12 dB below the others), and the Cardiac and Ventilation alarms at 66 dB (14 dB below the others). Given that the standard does not provide a reference volume level for alarms, it is completely reasonable to expect such volume differences to manifest between

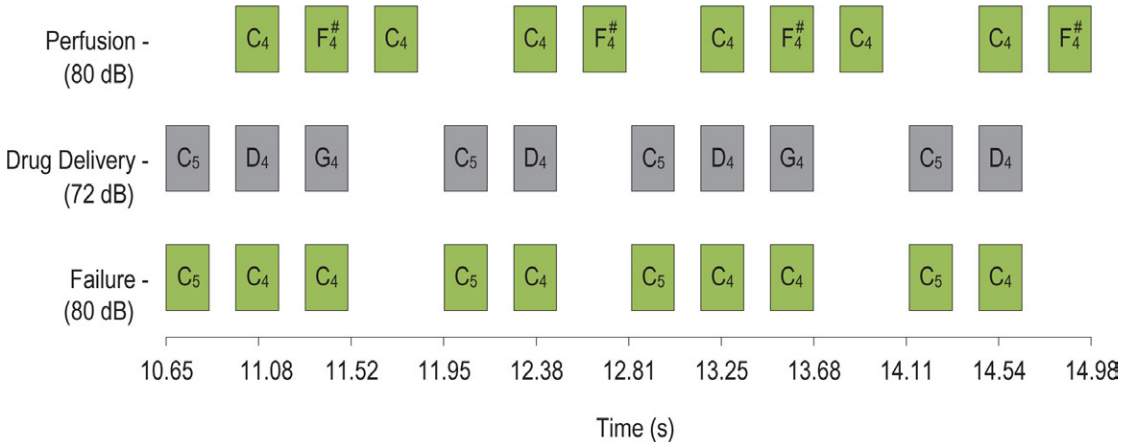

Figure 4. An example visualization of a counterexample that illustrates the sounding pattern for how a high-priority alarm (Drug delivery at 72 dB) can be totally masked by two other high-priority alarms (Perfusion and Failure, both at 80 dB). Note that such a situation could occur when there are three devices, one issuing each of the alarms, one device capable of issuing all three alarms concurrently, or—because all configurations would be consistent with IEC 60601-1-8—any combination of separate device and alarm pairing.

alarms from different devices in a medical environment.

At even lower volumes, many alarms can be totally masked by a minimum of one other alarm at volumes that could be possible between devices. The medium-priority Cardiac alarm (at 64 dB) can be totally masked by a medium-priority Temperature. The medium-priority Temperature (at 63 dB) alarm can be totally masked by medium-priority General. Medium-priority Ventilation (at 60 dB) can be totally masked by medium-priority Perfusion. At lower volumes, all but the low-priority alarm (Low) and the General high-priority alarm can be totally masked by one other alarm at the same level.

Our results also show variation in the ability of alarms to be totally masked. For high-priority alarms, General, Oxygen, and Ventilation were the least susceptible to total masking at higher volumes while Temperature was the most susceptible. However, performance of these alarms varies at lower volumes. For the medium- and low-priority alarms, Oxygen is the alarm that is the most robust to simultaneous masking, with Failure and Ventilation also being fairly robust at higher volumes. Cardiac is clearly the most susceptible medium-priority alarm. In fact, the Cardiac alarm was more susceptible to masking than Low, indicating that a mismatch exists between alarm masking susceptibility and priority.

Partial Masking

Similar to the vagaries associated with total masking, the literature does not provide guidance about how much an alarm needs to be partially masked to be rendered inaudible. However, it is clear that any masking will impair people's ability to perceive an alarm, especially given that there have been noted problems with the alarms in the standard even without the presence of masking (Lacherez et al., 2007; Sanderson et al., 2006). In this regard, the results from partial masking are problematic because they show that it only takes two concurrently sounding alarms (either at the same priority level or between priority levels) to partially mask each of the analyzed alarms. At 73 dB and lower (7 dB below the others, and well within the range allow by the standard), it only took one other alarm to partially mask all the analyzed alarms.

Generalizability

While all of our analyses used 80 dB (a realistic upper bound) as the baseline volume and the volume of a masker does impact masking curve shape (more masking is afforded by higher volumes), these variations are minor (Bosi & Goldberg, 2003). Thus, we would not expect the results to significantly change with decreases in the base level volume with relative decreases in the volume of the analyzed alarm.

It is also worth noting that IEC 60601-1-8 specifies that every alarm sound (including the reserved ones) have at least four additional harmonics (or subfrequencies): tones with frequencies at whole-number multiples of the primary frequency that are at lower volumes. These are intended to improve alarm audibility. However, previous work showed that our method validly predicts simultaneous masking using the primary alarm frequencies irrespective of whether the additional harmonics were included. Thus, the results presented in this paper should be valid even though we only considered the primary frequencies of the reserved alarm sounds.

Given this information, our results collectively indicate that there is potential for masking being a serious problem for devices that are compliant with IEC 60601-1-8. This is because total masking can occur in the presence of three or fewer alarms, and partial masking can occur in the presence of one alarm with variations in volume that are consistent for alarms within and between devices. If practitioners are unable to hear an alarm, they will not be able to respond to them. In a medical environment, where seconds can mean the difference between life and death, this could have profound implications for patient safety and health.

Implications for the Standard

Our results suggest means of interpreting the IEC 60601-1-8 reserved sounds that will minimize the effect of masking. First, for a designed configuration of alarms, alarms at a given priority should be kept as close to the same volume as possible. Second, for alarms between devices, clinical engineers should try to make similar priority alarms as close to the same volumes as possible. They should not increase the volume of a particular alarm to improve its perceivability. These recommendations will reduce the chances that any particular alarm will be masked.

It is important to note that this is not suggesting that all alarms should be kept at the 80-dB base level. This is because such volumes will likely contribute to problems associated with hospital noise (Konkani et al., 2012). Rather, the same or better performance (due to the nature of the psychoacoustics; Bosi & Goldberg, 2003) will be achieved by keeping similarly prioritized alarms at a consistent lower volume level. Third, although not often employed, IEC 60601-1-8 allows designers to use different timings of alarm tones and pauses between them (Table 3). Using different timings across alarms will make it harder for alarms to perfectly overlap and thus reduce the chance of total masking.

To assess this third point, we created new models of the high as well as the medium- and low-priority alarms from analysis conditions (1) and (2), respectively. These models were identical to the original, except that the timing of each alarm's pauses and tones (each alarm's x and t_H or y and t_M; Table 3) were assigned unique random numbers that were consistent with the requirements of the standard (Table 3). The generated values of these were are shown in Table 4. The analyses for conditions (1) and (2) were then re-run. The results of these are shown in Figure 5.

While partial masking results were not affected, using these unique random timings dramatically improved the resilience of the alarms to total masking. In particular, the random timings completely eliminated the ability of high-priority alarms to totally mask other high-priority alarm regardless of the volume of the potentially masked alarm or the number of concurrently sounding alarms; see Figure 5(a). Furthermore, these timings either completely eliminated the total masking of an alarm (for volumes of 55–80 dB for the Oxygen alarm and volumes of 75–80 for the Delivery2 alarm) or increased the number of alarms required for total masking for the medium- and low-priority alarms; see Figure 5(c). These improvements are specifically illustrated in Figure 5(b) and (d), which show how the number of alarms required for masking from Figure 5(a) and (c) improved from the analyses under conditions (1) and (2), respectively. Thus, assigning unique volumes to alarm timing parameters, even within the parameters allowed by the standard, has the potential to significantly improve alarm audibility.

TABLE 4: Random Timing Parameters Used in the Replication Analyses

	Timing Parameters				
Alarm	x	t_H	y	t_M	z
General	0.1158	0.0773	0.1612	0.1627	0.4034
Cardiac	0.1632	0.0763	0.1632	0.1606	1.0788
Perfusion	0.1417	0.0762	0.1417	0.1598	0.7953
Ventilation	0.1474	0.0756	0.1474	0.1664	0.9654
Oxygen	0.1112	0.079	0.1112	0.1602	1.1383
Delivery1	0.1127	0.0781	0.1127	0.1667	0.9097
Delivery2	0.1718	0.0767	0.1718	0.1634	1.0104
Failure	0.0768	0.0782	0.0768	0.161	0.8572
Low			0.1632	0.1605	

Note. x, t_H, y, and t_M are parameters from Table 3.

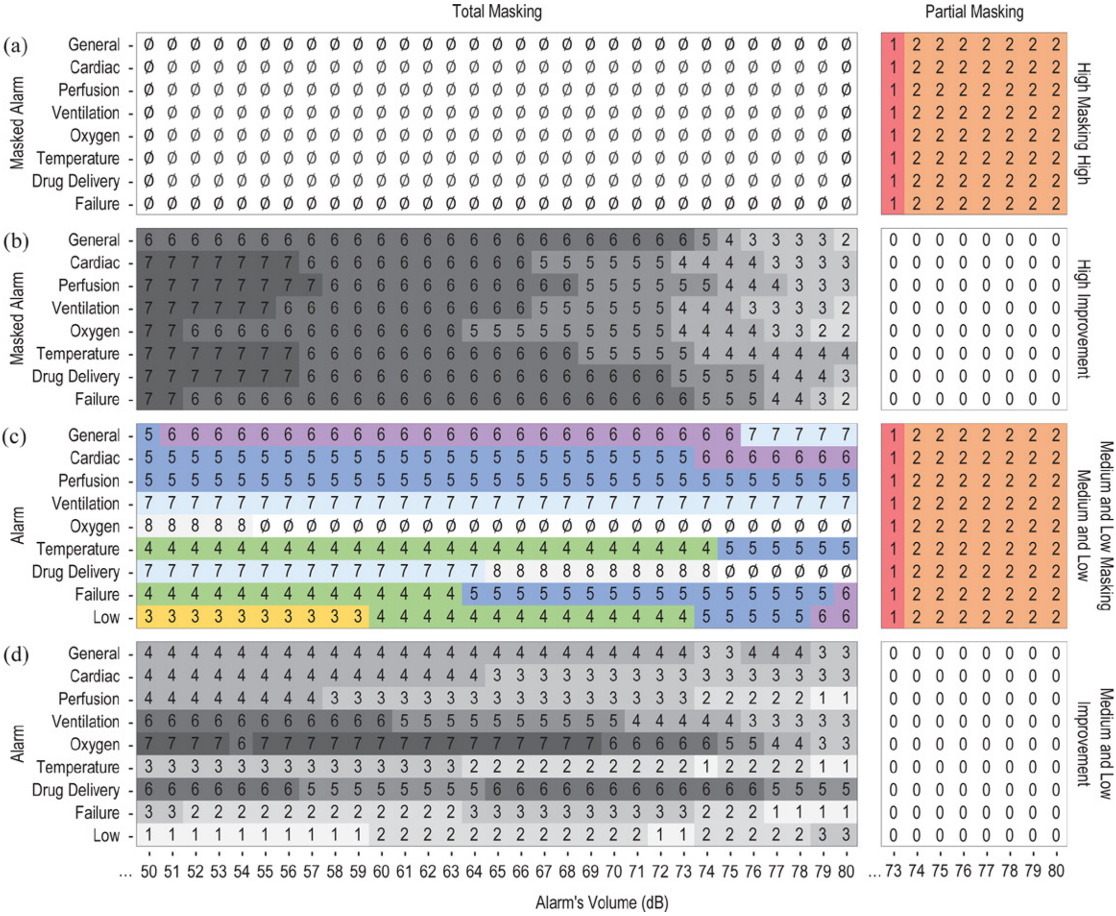

Figure 5. Graphs in (a) and (c) show results from the replication of the analyses from conditions (1) and (2), respectively, but using alarm models with random times assigned to all alarm tone sounding lengths and pauses between tones. A ∅ indicates that no masking could occur. (b) and (d) show the increase/improvement in the minimum number of alarms to produce masking between these new results and the originals. (b) shows the improvement from the result in Figure 3(a) to those in (a). A ∅ was treated as an 8 in the production of this graph. (d) shows the improvement from the result in Figure 3(b) to those in (c). A ∅ was treated as a 9 in the production of this graph.

Additional Audibility Considerations

It is important to note that the work presented here only addresses a very specific problem: when alarms can become unhearable due to simultaneous masking. This is only one way that alarm audibility, perception, and/or identification can be impacted. Alarm signals can be masked by background noise (Konkani et al., 2012). Sound waves can physically interact with each other in ways that create dissonance (a nonharmonic combination of sounds), beating (a perceived periodic variation in volume caused by interference between two sounds with similar frequencies), or physical masking (where sound waves cancel each other out; Fastl & Zwicker, 2006). There are distinct limits on the total number of alarms that people can learn and differentiate (Sanderson et al., 2006). High levels of workload can result in inattentional deafness (Scheer et al., 2018). The number of medical alarms can produce so-called alarm fatigue (Cvach, 2012). The similar nature of IEC 60601-1-8 alarms (designed, reserved, or otherwise), the noisiness of clinical environments, and the number of medical alarms can mean that all of the listed hazards (and potentially many others; Edworthy, 2013) could impact IEC 60601-1-8 alarm detection and response. While these factors are beyond the scope of this paper, research should investigate how to account for them in future iterations of the method. Of the listed considerations, masking caused by background noise and physical interactions between alarm sounds are the most well understood and would likely be the easiest to account for in our method. The other conditions may require additional modeling advances before incorporation can occur.

CONCLUSIONS

It is well established that the audible alarms currently specified in IEC 60601-1-8 have a number of problems. Our efforts are the first demonstration that simultaneous masking is also a concern. The uniformity of these sounds (their rhythm and harmonic structure) is the source of most of these problems. This paper demonstrates that even slight changes in temporal structure significantly reduces the probability of masking.

The current alarms as discussed in this paper will also be allowable in the update of the standard. Considerable effort has been put into the development, testing, and benchmarking of more complex and varied alarm signals. This will accompany an update to the standard that is likely to occur in 2020 (Bennett et al., 2019; Edworthy, McNeer, et al., 2018, Edworthy, Reid, et al., 2018; McNeer et al., 2018). Though the proposed new sounds have been thoroughly evaluated, their masking capabilities have not yet been evaluated as they present new technical and operational challenges for the model. The fact that the proposed new sounds are harmonically more complex and are spectrally and temporally more varied than the current alarm sounds suggests a priori that they will be more resistant to masking. Future work will focus on extending our masking detection approach so that it can be used to evaluate the masking potential of the new alarm sounds.

ACKNOWLEDGMENTS

This research was supported by the Agency for Healthcare Research and Quality under award number R18HS024679. This research complied with the American Psychological Association Code of Ethics. This is an extended version of a conference paper (Bolton, Edworthy, & Boyd, 2018b). The original paper only presented and discussed results of the total masking analyses for the medium- and low-priority alarms. Thus, this journal submission constitutes a significant contribution beyond the original conference paper.

KEY POINTS

- The alarms prescribed by the IEC 60601-1-8 international standard are susceptible to simultaneous masking.
- This work systematically evaluated the masking potential of the standard's reserved alarm sounds.
- The analyses revealed many problems with masking between the alarm sounds for parameters allowed by the standard.
- Discovered problems can be alleviated by setting alarm volumes to standard values based on priority level and by randomizing the timing of alarm tones.

ORCID iD

Matthew L. Bolton https://orcid.org/0000-0002-7943-0497

REFERENCES

Ambikairajah, E., Davis, A. G., & Wong, W. T. K. (1997). Auditory masking and MPEG-1 audio compression. *Electronics & Communication Engineering Journal*, *9*, 165–175. https://doi.org/10.1049/ecej:19970403

Bennett, C., Dudaryk, R., Crenshaw, N., Edworthy, J., & McNeer, R. (2019). Recommendation of new medical alarms based on audibility, identifiability, and detectability in a randomized, simulation-based study. *Critical Care Medicine*, *47*, 1050–1057. https://doi.org/10.1097/CCM.0000000000003802

Bolton, M. L. (2017). Novel developments in formal methods for human factors engineering. In *Proceedings of the human factors and ergonomics Society annual meeting* (Vol. 61, pp. 715–717). Sage. https://doi.org/10.1177/1541931213601664

Bolton, M. L., & Bass, E. J. (2009). A method for the formal verification of human interactive systems. In *Proceedings of the 53rd annual meeting of the human factors and ergonomics Society* (Vol. 53, pp. 764–768). HFES. https://doi.org/10.1177/154193120905301201

Bolton, M. L., & Bass, E. J. (2010). Formally verifying human-automation interaction as part of a system model: Limitations and tradeoffs. *Innovations in Systems and Software Engineering*, *6*, 219–231. https://doi.org/10.1007/s11334-010-0129-9

Bolton, M. L., Bass, E. J., & Siminiceanu, R. I. (2012). Generating phenotypical erroneous human behavior to evaluate human-automation interaction using model checking. *International Journal of Human-Computer Studies*, *70*, 888–906. https://doi.org/10.1016/j.ijhcs.2012.05.010

Bolton, M. L., Bass, E. J., & Siminiceanu, R. I. (2013). Using formal verification to evaluate human–automation interaction in safety critical systems, a review. *IEEE transactions on systems, man, and cybernetics. Systems*, *43*, 488–503.

Bolton, M. L., Edworthy, J., & Boyd, A. D. (2018). A formal analysis of masking between reserved alarm sounds of the IEC 60601-1-8 international medical alarm standard. In *Proceedings of the human factors and ergonomics society annual meeting* (Vol. 62, pp. 523–527). https://doi.org/10.1177/1541931218621119

Bolton, M. L., Edworthy, J., Boyd, A. D., Wei, J., & Zheng, X. (2018). A computationally efficient formal method for discovering simultaneous masking in medical alarms. *Applied Acoustics*, *141*, 403–415. https://doi.org/10.1016/j.apacoust.2018.06.012

Bolton, M. L., Hasanain, B., Boyde, A. D., & Edworthy, J. (2016). Using model checking to detect masking in IEC 60601-1-8-compliant alarm configurations. In *Proceedings of the human factors and ergonomics Society annual meeting* (Vol. 60, pp. 636–640). https://doi.org/10.1177/1541931213601146

Bolton, M. L., Zheng, X., Li, M., Edworthy, J. R., & Boyd, A. D. (2020). An experimental validation of masking in IEC 60601-1-8:2006-compliant alarm sounds. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *62*, 954–972. https://doi.org/10.1177/0018720819862911

Bosi, M., & Goldberg, R. E. (2003). *Introduction to digital audio coding and standards*. Springer.

Brandenburg, K., & Stoll, G. (1994). ISO/MPEG-1 audio: A generic standard for coding of high-quality digital audio. *Journal of the Audio Engineering Society*, *42*, 780–792.

Clarke, E. M., Grumberg, O., & Peled, D. A. (1999). *Model checking*. MIT Press.

Cvach, M. (2012). Monitor alarm fatigue: An integrative review. *Biomedical Instrumentation & Technology*, *46*, 268–277. https://doi.org/10.2345/0899-8205-46.4.268

ECRI Institute. (2014, November). Top 10 health technology hazards for 2015. *Health Devices*. http://www.ecri.org/2015hazards

ECRI Institute & ISMP. (2009). Connecting remote cardiac monitoring issues with care areas. *Pennsylvania Patient Safety Authority*, *6*, 79–83.

Edworthy, J. (2013). Medical audible alarms: A review. *Journal of the American Medical Informatics Association*, *20*, 584–589. https://doi.org/10.1136/amiajnl-2012-001061

Edworthy, J., & Hellier, E. (2005). Fewer but better auditory alarms will improve patient safety. *Quality and Safety in Health Care*, *14*, 212–215. https://doi.org/10.1136/qshc.2004.013052

Edworthy, J., & Hellier, E. (2006). Alarms and human behaviour: Implications for medical alarms. *British Journal of Anaesthesia*, *97*, 12–17. https://doi.org/10.1093/bja/ael114

Edworthy, J., McNeer, R. R., Bennett, C. L., Dudaryk, R., McDougall, S. J. P., Schlesinger, J. J., Bolton, M. L., Edworthy, J. D., Vieira, E. O., Boyd, A. D., Reid, S. K. J., Rayo, M. F., Wright, M. C., & Osborn, D. (2018). Getting alarm sounds into a global standard. *Ergonomics in Design*, *26*, 4–13.

Edworthy, J., & Meredith, C. S. (1994). Cognitive psychology and the design of alarm sounds. *Medical Engineering & Physics*, *16*, 445–449. https://doi.org/10.1016/1350-4533(94)90067-1

Edworthy, J., Reid, S., Peel, K., Lock, S., Williams, J., Newbury, C., Foster, J., & Farrington, M. (2018). The impact of workload on the ability to localize audible alarms. *Applied Ergonomics*, *72*, 88–93. https://doi.org/10.1016/j.apergo.2018.05.006

Emerson, E. A. (1990). Temporal and modal logic. In J. van Leeuwen, A. R. Meyer, M. Nivat, M. Paterson, & D. Perrin (Eds.), *Handbook of theoretical computer science* (pp. 995–1072). MIT Press.

Fastl, H., & Zwicker, E. (2006). *Psychoacoustics: Facts and models* (Vol. 22). Springer.

Hasanain, B., Boyd, A. D., & Bolton, M. L. (2014). An approach to model checking the perceptual interactions of medical alarms. In *Proceedings of the 2014 international annual meeting of the human factors and ergonomics Society* (Vol. 58, pp. 822–826). HFES. https://doi.org/10.1177/1541931214581173

Hasanain, B., Boyd, A. D., & Bolton, M. L. (2016). Using model checking to detect simultaneous masking in medical alarms. *IEEE Transactions on Human-Machine Systems*, *46*, 174–185. https://doi.org/10.1109/THMS.2014.2379661

Hasanain, B., Boyd, A. D., Edworthy, J., & Bolton, M. L. (2017). A formal approach to discovering simultaneous additive masking between auditory medical alarms. *Applied Ergonomics*, *58*, 500–514. https://doi.org/10.1016/j.apergo.2016.07.008

IEC 60601-1-8:2006+AMD1:2012. (2012). *Medical electrical equipment – part 1-8: general requirements for basic safety and essential performance–collateral standard: general requirements, tests and guidance for alarm systems in medical electrical equipment and medical electrical systems*. International Electrotechnical Commission.

Konkani, A., Oakley, B., & Bauld, T. J. (2012). Reducing hospital noise: A review of medical device alarm management. *Biomedical Instrumentation & Technology*, *46*, 478–487. https://doi.org/10.2345/0899-8205-46.6.478

Lacherez, P., Seah, E. L., & Sanderson, P. (2007). Overlapping melodic alarms are almost indiscriminable. *Human Factors*, *49*, 637–645. https://doi.org/10.1518/001872007X215719

Lutfi, R. A. (1983). Additivity of simultaneous masking. *The Journal of the Acoustical Society of America*, *73*, 262–267. https://doi.org/10.1121/1.388859

McNeer, R. R., Horn, D. B., Bennett, C. L., Edworthy, J. R., & Dudaryk, R. (2018). Auditory icon alarms are more accurately and quickly identified than current standard melodic alarms in a simulated clinical setting. *Anesthesiology*, *129*, 58–66. https://doi.org/10.1097/ALN.0000000000002234

Meredith, C., & Edworthy, J. (1995). Are there too many alarms in the intensive care unit? An overview of the problems. *Journal of Advanced Nursing*, *21*, 15–20. https://doi.org/10.1046/j.1365-2648.1995.21010015.x

Momtahan, K., Hétu, R., & Tansley, B. (1993). Audibility and identification of auditory alarms in the operating room and intensive care unit. *Ergonomics*, *36*, 1159–1176. https://doi.org/10.1080/00140139308967986

Patterson, R. D. (1982). *Guidelines for auditory warning systems on civil aircraft. civil aviation authority*. Civil Aviation Authority.

Patterson, R. D., & Mayfield, T. F. (1990). Auditory warning sounds in the work environment. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *327*, 485–492. https://doi.org/10.1098/rstb.1990.0091

Phansalkar, S., Edworthy, J., Hellier, E., Seger, D. L., Schedlbauer, A., Avery, A. J., & Bates, D. W. (2010). A review of human factors principles for the design and implementation of medication safety alerts in clinical information systems. *Journal of the American Medical Informatics Association*, *17*, 493–501. https://doi.org/10.1136/jamia.2010.005264

Sanderson, P. M., Wee, A., & Lacherez, P. (2006). Learnability and discriminability of melodic medical equipment alarms. *Anaesthesia*, *61*, 142–147. https://doi.org/10.1111/j.1365-2044.2005.04502.x

Scheer, M., Bülthoff, H. H., & Chuang, L. L. (2018). Auditory task irrelevance: A basis for inattentional deafness. *Human Factors*, *60*, 428–440. https://doi.org/10.1177/0018720818760919

Stead, W. W., & Lin, H. S. (Eds.). (2009). *Computational technology for effective health care: Immediate steps and strategic directions*. National Academies Press.

The Joint Commission. (2013a). Medical device alarm safety in hospitals. *Sentinel Even Alert*, *50*.

The Joint Commission. (2013b). NPSG.06.01.01: Improve the safety of clinical alarm systems. *Joint Commission Perspectives*, *33*.

Toor, O., Ryan, T., & Richard, M. (2008). Auditory masking potential of common operating room sounds: A psychoacoustic analysis. In *Anesthesiology* (Vol. 109, p. A1207). American Society of Anesthesiologists.

Vockley, M. (2014). *Clinical alarm management compendium*. AAMI Foundation.

Weyers, B., Bowen, J., Dix, A., & Palanque, P. (Eds.). (2017). *The handbook of formal methods in human-computer interaction*. Springer.

Zwicker, E., & Feldtkeller, R. (1967). *Das ohr als nachrichtenempfänger*. Hirzel Verlag.

Matthew L. Bolton is an associate professor of industrial and system engineering at the University at Buffalo, the State University of New York. He received his PhD in systems engineering in 2010 from the University of Virginia, Charlottesville, USA.

Judy R. Edworthy is the director of the Cognition Institute and a professor of applied psychology at the University of Plymouth. She received her PhD in experimental psychology in 1984 from the University of Warwick, UK.

Andrew D. Boyd is an associate professor of biomedical and health information sciences at the University of Illinois Chicago. He received his MD in 2002 from the University of Texas Southwestern Medical School, Dallas, USA.

Date received: December 5, 2019
Date accepted: September 22, 2020