Getting Better Hospital Alarm Sounds Into a Global Standard

Extensive development and testing of alternative hospital alarm sounds shows promise for improving patient safety.

By Judy Reed Edworthy, Richard R. McNeer, Christopher L. Bennett, Roman Dudaryk, Siné J. P. McDougall, Joseph J. Schlesinger, Matthew L. Bolton, Jonathan D. Reed Edworthy, Elif Özcan, Andrew D. Boyd, Scott K. J. Reid, Michael F. Rayo, Melanie C. Wright, & David Osborn

FEATURE AT A GLANCE:
The reserved set of audible alarm signals embodied within the global medical device safety standard, IEC 60601-1-8, is known to be problematic and in need of updating. The current alarm signals are not only suboptimal, but there is also little evidence beyond learnability (which is known to be poor) that demonstrates their performance in realistic and representative clinical environments. In this article, we describe the process of first designing and then testing potential replacement audible alarm signals for IEC 60601-1-8, starting with the design of several sets of candidate sounds and initial tests on learnability and localizability, followed by testing in simulated clinical environments. We demonstrate that in all tests, the alarm signals selected for further development significantly outperform the current alarm signals. We describe the process of collecting considerably more data on the performance of the new sounds than exists for the current sounds, which ultimately will be of use to end users. We also reflect on the process and practice of working with the relevant committees and other practical issues beyond the science, which also need constant attention if the alarms we have developed are to be included successfully in an updated version of the standard.

KEYWORDS:
auditory alarms, technical standards, alarm learnability, clinical alarms, IEC 60601-1-8, medical alarms, patient safety, health care

Audible alarm signals are very important across high-workload industries, and their use in those environments is driven not always by the best science but by other factors, such as customer reaction, budget, lack of expertise in design and application of knowledge, and inflexible and/or conservative approaches to known problems. As a result, one can find many examples of high-workload, safety-critical environments in which the audible alarm signals leave much to be desired in terms of both implementation and design, although increasingly there are many examples of thoughtful and well-designed implementations.

In clinical environments, the problem of bad alarm system implementation has reached colossal proportions, where patient deaths have been attributed to *alarm fatigue* (Drew et al., 2014; Sendelbach & Funk, 2013). Until a national summit in the United States in 2011, little was being done about the general problem of alarm condition overuse. Now, however, there are well-documented and successful attempts to reduce the problem of overalarming in general (Cvach, 2012; Welch 2011; Whalen et al., 2014).

The audible alarm signals that annunciate the hazards traditionally have also left a lot to be desired from the point of view of design, but now that the broader alarm system problems are slowly being resolved, the time is right to improve on the audible alarm signals as well. In this article, we describe a project intended to upgrade and update the audible alarm signals in a global medical device standard.

It is challenging to carry out what is essentially an applied, customer-based problem while maintaining the best scientific approach one can muster. This challenge is highlighted by Morrow and Durso (2011) in their editorial in a special issue of the *Journal of Experimental Psychology: Applied* on cognitive issues in health care. They introduce their paper thus:

> We focus on the need for research that is sufficiently comprehensive to identify threats to patient safety, yet specific enough to explain how provider and patient factors interact with task and health context to engender these threats. Such research should be theory-based, yet also problem-driven; exert experimental control over theoretically relevant variables, yet also involve participants, tasks, and contexts that represent the problems of interest. A tension exists between theory-based, experimentally controlled research on the one hand, and problem-driven research with representative situations on the other. (Morrow & Durso, 2011, p. 191)

The challenge in terms of audible alarm signal design is to bring the scientific evidence to bear on the problem but also to commit, at some point during the process, to a specific set or sets of sounds so that a research database can be built around them.

The evidence base for auditory alarm signal design is considerably more advanced than the typical sorts of alarm signals that are used in practice might suggest. Bridging the "valley of death" between theory and application is always a problem, made more acute in auditory work given the difficulty of talking to nonexperts (often the client) about sound in any abstract way, and given the predisposition that clients have to like or dislike a sound designed for a specific application.

Reactions to alarms can sometimes be colored by the existing, often adverse, alarm

Table 1. IEC 60601-1-8 High-Priority Alarm Signal Characteristics

Function of Alarm	Alarm Signal Characteristics
General	A burst of three regularly spaced pulses, followed by a burst of two regularly spaced pulses in the following pattern: c c c – c c
Power down	A burst of three regularly spaced pulses, followed by a burst of two regularly spaced pulses in the following pattern: C c c – C c
Cardiovascular	A burst of three regularly spaced pulses, followed by a burst of two regularly spaced pulses in the following pattern: c e g – g C
Perfusion	A burst of three regularly spaced pulses, followed by a burst of two regularly spaced pulses in the following pattern: c f# c – c f#
Drug administration	A burst of three regularly spaced pulses, followed by a burst of two regularly spaced pulses in the following pattern: C d g – C d
Oxygen	A burst of three regularly spaced pulses, followed by a burst of two regularly spaced pulses in the following pattern: C b a – g f
Ventilation	A burst of three regularly spaced pulses, followed by a burst of two regularly spaced pulses in the following pattern: c a f – a f
Temperature	A burst of three regularly spaced pulses, followed by a burst of two regularly spaced pulses in the following pattern: c d e – f g

Source. Edworthy, Reid, et al. (2017).
Note. Each regularly spaced pulse ranges between 100 ms and 300 ms.

environment. For example, nurses are typically already overwhelmed with alarm signals (Honan et al., 2015), so anything that looks like an addition to the alarm system environment (such as a new set of audible alarm sounds) needs to be presented within the context of a transition that ultimately will be of benefit to those working with those alarms on a day-in, day-out basis.

THE STANDARD: IEC 60601-1-8

IEC 60601 is a set of standards concerned with the safety of medical electrical equipment, so it covers almost all medical equipment. Part 1-8 specifies the basic safety and essential performance requirements and tests for the alarm systems contained within that equipment. Thus this standard governs almost all medical equipment across the globe. It was published first in 2006, was then updated in 2012, had something of an update in 2015, and is due for another, major update by the end of 2019.

The key feature of the standard in terms of audible alarm signals is that it specifies the acoustic and structural elements of the audible alarm signals that should accompany specific clinical hazards or categories (International Electrotechnical Commission, 2006).

The reserved set of alarm signals was designed with the best of intentions (Block, Rouse, Hakala, & Thompson, 2000), based on some (but not all) aspects of what was known about alarm signal design at the time. The sounds embodied important acoustic features that would increase their resistance to masking (compared at least with single harmonics) and improve their general acceptability over the earlier beeps, buzzers, and bells.

The structure of the alarm signals and their categories is shown in Table 1. Eight categories of risk are specified, each of which has a high- and a medium-priority form. In our studies, only the high-priority version was tested, although generic medium- and low-priority alarm signals were also tested for this update.

A key problem with the design was that the alarm signals, which sound like short, tonal melodies, all possess the same number of pulses and the same rhythm, making them very hard to distinguish one from another (Lacherez, Seah, & Sanderson, 2007; Sanderson, Wee, & Lacherez, 2006; Wee & Sanderson, 2008). The lack of diversity between the sounds is a major contributor to the known problems with learning and recognizing these alarm signals, and the finding is no surprise given that one's ability to distinguish between stimuli depends on the number of dimensions along which they vary (Miller, 1956). A shared rhythm is also a key component of a listener's confusion between sounds (Patterson, 1982). Calls to update and improve the sounds have been numerous, with the designer of the sounds himself issuing an apologia for the current sounds (Block, 2008).

It has become clear that almost anything would be better than the current alarm signals, which present another problem. Atyeo and Sanderson (2015) demonstrated that a similar set of alarm signals designed prior to the 2006 version of IEC 60601-1-8 (designed for an earlier version of the standard; Patterson, Edworthy, Shailer, Lower, & Wheeler, 1986) outperforms the current alarm signals. Other evidence shows

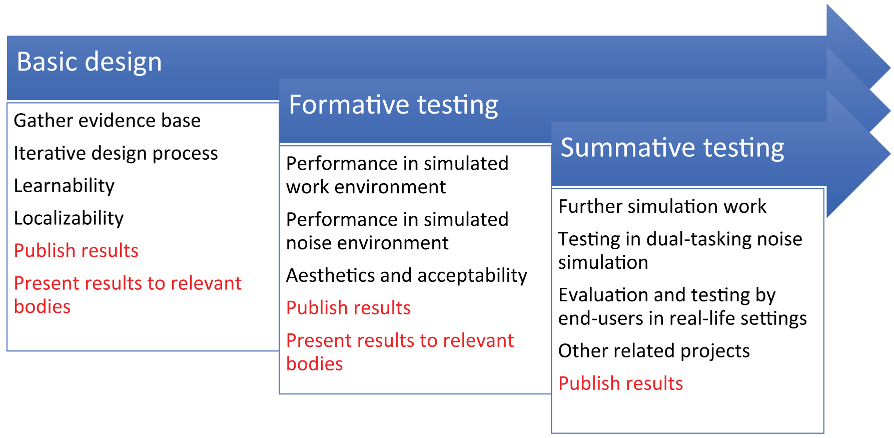


Figure 1. The stages of updating the audible alarm signals for IEC 60601-1-8.

that a random set of audible sounds with no association to the meanings or functions of the alarm conditions was easier to learn than the current alarm signals (Edworthy et al., 2014).

The earlier (1986) set of sounds was rejected on a nonempirical basis, which allowed interested parties to call into a telephone line and listen to the alarm sounds and then to voice an opinion. However, that was the 1980s, and patently replacing the current alarm signals with sounds that simply perform better than the current alarm signals – even those designed in the '80s that turned out to be better than the alarm signals in the standard – is not enough.

Reflection 1. *Despite knowing of the existence of the "IEC 60601-1-8 alarm problem" for years prior to the start of the project, we believed it was important to conduct the project with the endorsement of the body charged with updating the standard, rather than conducting the work in isolation, presenting it to that body, and waiting for a head of steam to build up over any potential replacement. The bodies in this case are the IEC 60601-1-8 and AAMI 60601-1-8 standards committees, which have a common core and some crossover membership through an IEC alarms joint working group. Access to this group was made possible because the Association for the Advancement of Medical Instrumentation (AAMI) has an open policy on membership of its own parallel IEC 60601-1-8 committee, AAMI 60601-1-8. The first author joined and began attending meetings. AAMI later made a grant to the first author to carry out the initial development work.*

Changing and updating standards is akin to the proverbial changing of the course of a ship using a teaspoon. The process of bringing about change in standards is very slow and requires sustained attention. The challenges and demands of achieving global standardization in our increasingly technological world are well documented across several spheres, such as finance and medical devices (Abbot & Snidal, 2001; Cheng, 2003; Mattli & Büthe, 2003). Achieving standardization even of the relatively straightforward and contained issue of medical device alarms inevitably involves stakeholders with many different vested interests, most of which are market and financially driven.

The fate of earlier work heightens our awareness of non–empirically based criticisms and potential scuppering, which are best met with empirically based answers. Thus a key element of our strategy is to create a published and accessible database at every point in the process.

THE PROCESS

Figure 1 shows the process we adopted in developing the alarm sounds. Whereas medical equipment audible alarm signals have traditionally been produced by poor-quality-sounding devices, many medical devices are now equipped with good-quality speakers. Sound storage and reproduction is also much cheaper, all of which means that, potentially, almost any sound can be used as an alarm signal, and the sound reproduction can be of high quality.

This does not make the work of the designer any easier; indeed, it focuses the effort required to demonstrate that any new alarm signals are not only "better" but "the best," or among the best, possible. A key question is what constitutes "best." Here, we have to start with learnability (whether or not it is important, although it probably is), as learnability is the only data we have on the current alarm sounds, and comparisons are a good starting point – indeed, they are essential in making the preliminary arguments for adoption of any new sounds.

BASIC DESIGN

There is ample evidence to show that the concrete–abstract continuum plays a big role in the learnability of sounds. There are many published examples of "auditory icon" alarm signal designs that outperform abstract sounds (Belz, Robinson, & Casali, 1999; Edworthy et al., 2014; Graham, 1999; Keller & Stevens, 2004; Leung, Smith, Parker, & Martin, 1997; Perry, Stevens, Wiggins, & Howell, 2007; Petocz, Keller, & Stevens,

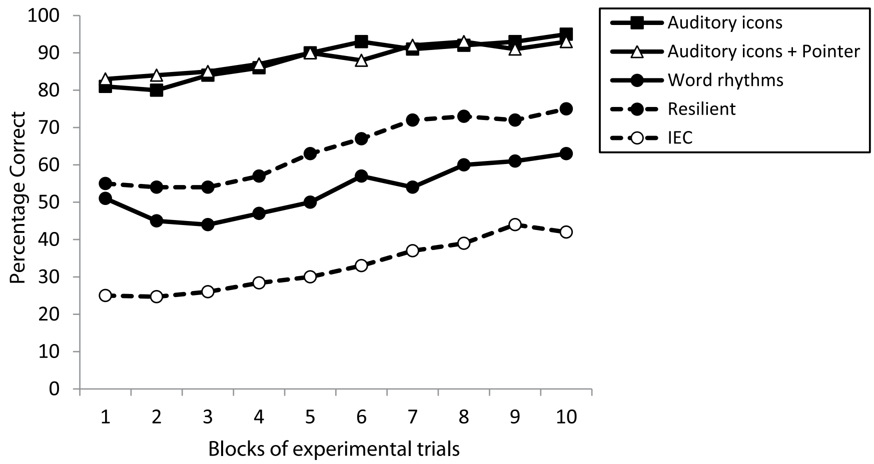


Figure 2. Percentage correct responses for each set of alarm signals, across 10 trials (from Edworthy, Reid, et al., 2017).

2008; Stephan, Smith, Martin, Parker, & McAnally, 2006; Ulfvengren, 2003). Auditory icons, which are typically real-world sounds with direct associations to their meanings, are obviously high in their concreteness, although there are other ways in which metaphors can be achieved.

Actual speech-based sounds (including speech itself) are also readily learnable, as demonstrated in encouraging findings for *spearcons*, speech-based alarm sounds, in a clinical context (Li et al., 2017). For the standard itself, it was felt that speech was not appropriate, but we did include a set of alarms for testing that were based on word rhythms and patterns.

We developed four sets of sounds that used different types of metaphors for the eight alarm conditions and compared them with the current sounds, which have no, or very minimal, metaphors. We tested them in terms of learnability and localizability. In one set ("word rhythms"), the eight words of the functions were imitated and stylized in terms of number of syllables, rhythm, and tonal structure. This set is somewhat closer to concrete on the abstract–concrete continuum than are sounds with no mapping (i.e., the current sounds, which did, however, involve some attempt at mnemonics).

A second set ("resilient") was designed with lower acoustic fidelity, aimed at devices that might have low sound reproduction quality. For these, half again used the word rhythm association, and half used simple metaphors. For example, for temperature, the alarm sound was a tone glide upward, and for power down, it was a tone glide downward. We expected these metaphors to be relatively easy to learn and the word rhythms to be approximately the same as for the word rhythms set. The other two sets were both auditory icons, one set of which contained an abstract "pointer" and one of which did not. The sets were identical otherwise.

For each of the eight alarm categories, a combination of focus groups, questionnaires, and repeated discussions within the research group led to the identification of appropriate metaphors for each of the alarm conditions. For some conditions, the most appropriate metaphor was obvious (for example, a heartbeat sound for cardiovascular), but for others, the most appropriate metaphor was less obvious. Although we refined and tested three metaphors for each function in later testing (see later in this article), it turned out that by and large, we had selected the "best" metaphors at this first attempt. We also took care to ensure that there was acoustic variability across the set of auditory icons, in order to minimize possible confusion.

The learnability data for the sound sets can be seen in Figure 2. All of our designs were more memorable than the existing set (all lines on the graph were significantly different from one another except the two at the top), but there was also variation across our experimental sets, with the auditory icon sets being the most memorable. The performance data suggest that we have covered the range of responses here, in that the performance for the auditory icons was almost at ceiling level from the start, and the current IEC alarm signals were very difficult to learn and retain throughout.

The candidate alarm signals were also varied in their harmonic complexity and denseness, as, by and large, more harmonically dense sounds are easier to localize. Very few tests of alarm signal localizability have been conducted (Alali, 2011; Catchpole, McKeown, & Withington, 2004; Vaillancourt et al., 2013), although localizability is often a pertinent issue in clinical care (for example, in a multibed intensive care unit). Our results confirmed that the more harmonically dense alarm signals were easier to localize and that the least complex – the current alarm signals – were weakest in localizability (Edworthy, Reid, et al., 2017).

Reflection 2. *The findings from the basic design study (Edworthy, Reid, et al., 2017) were presented to the standards alarm systems joint working group in April 2016. They were also presented to the AAMI 60601-1-8 committee in June 2016 and to a meeting of the AAMI alarms coalition in July 2016. The empirical evidence was presented along with the sounds. As a consequence, the alarms joint working group decided that it wished to go ahead with the auditory-icons-plus-pointer design and supplied a list of activities, some formative and some summative, it would like to see undertaken prior to the*

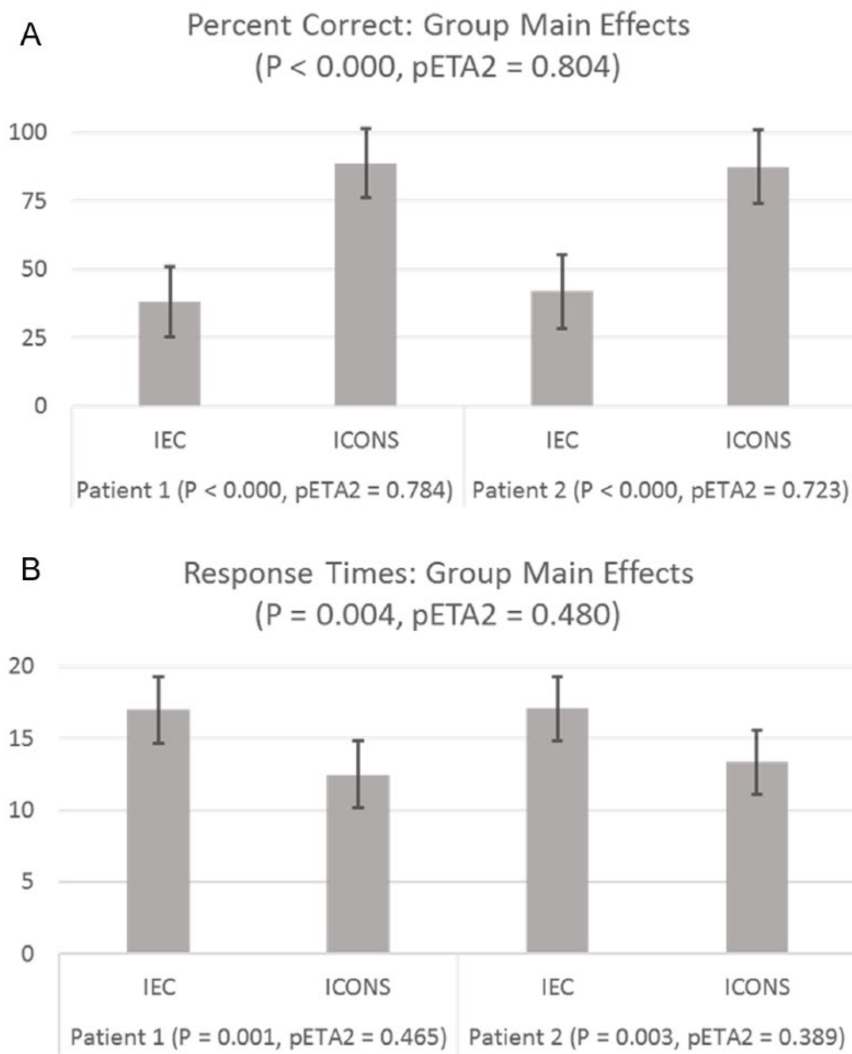


Figure 3. Mean percentage correct identification and mean reaction times to the new alarm signals (ICONS) and the current International Electrotechnical Commission alarm signals (IEC; from McNeer, Bennett, Horn, Dudaryk, & Edworthy, 2017a). *X*-axis = percentage correct; *y*-axis = IEC or ICONS. Patient 1 and Patient 2 refer to the two simulated patients being monitored by participants. Panel A shows percentage correct and Panel B shows reaction time.

committee's recommending the adoption of the alarm signals into the standard. A further grant from AAMI to the first author was negotiated on this basis.

Another unexpected consequence is that there appears to be a substantial amount of dissent over the categories of risk themselves. We have approached this issue by writing a paper to open discussion of the categories themselves (Edworthy, Schlesinger, McNeer, Kristensen, & Bennett, 2017). AAMI has made a grant available to one of the authors (MCW) to carry out research on this issue.

FORMATIVE TESTING

Mindful of Morrow and Durso's (2011) call for the use of contexts, tasks, and participants of relevance, the formative testing involves more realistic tasks using clinically trained participants. Using a range of already developed and published techniques (Bennett & NcNeer, 2012; Bennett et al., 2015); McNeer, Bennett, & Dudaryk, 2016), a paradigm was developed whereby trained anesthesiologists carried out a short clinical simulation task. They were required to monitor two patients and respond to alarm signals by indicating the nature of the alarm condition (its category); their reaction times also were measured. Prior to this task, they were given a brief exposure to either the auditory-icon-plus-pointer alarm signals or the current IEC alarm signals.

Results indicated very early on that the auditory icons produced faster and more accurate responses than the current IEC alarm signals (McNeer, Bennett, Bodzin Horn, Dudaryk, & Edworthy, 2017a, 2017b; McNeer, Bodzin Horn, Bennett, Reed Edworthy, & Dudaryk, 2018). Results of the early trials can be seen in Figure 3.

Secondary workload and fatigue measures were also taken in these studies, and there is some evidence that the auditory icons are less frustrating and impede performance less than

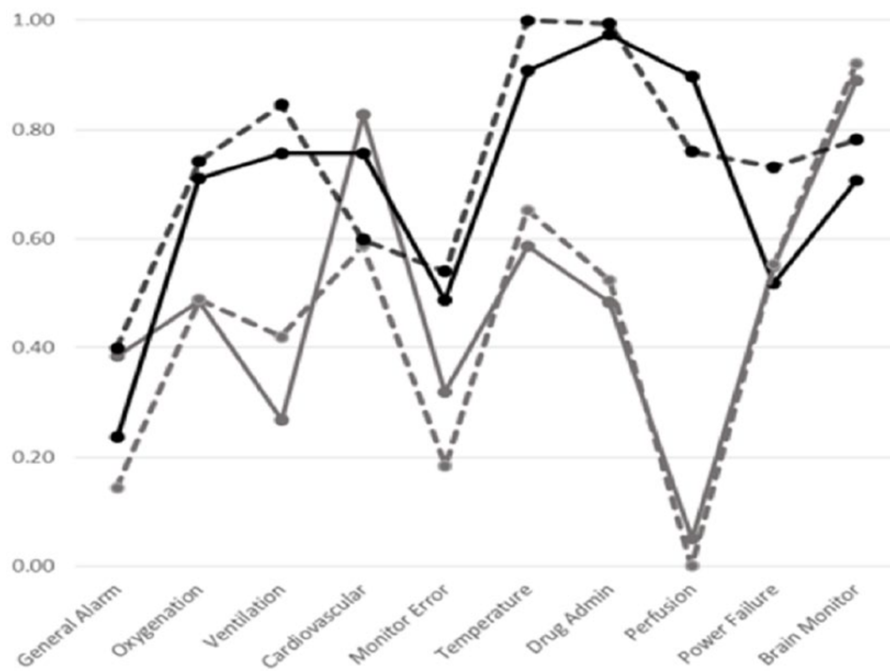


Figure 4. Binary response (a transformed composite of reaction time and accuracy) for "dream" and "nightmare" alarm sets. *X*-axis = accuracy/time index relative to best-performing sound for reaction time (Temperature); *y*-axis = 10 alarm-sound categories.

the current alarm signals. Here, we may be tapping into alarm fatigue. This finding is important because although the concept of alarm fatigue is generally accepted, and there certainly is a clinical alarm problem, the details of its manifestation and dimensions are somewhat sketchy (Deb & Claudio, 2015; Kristensen, Edworthy, & Ozcan Vieira, 2016; Rayo & Moffat-Bruce, 2015).

The final phase of the formative testing in simulation was to test three versions of each auditory icon. Three different auditory icons were generated for each function. (We added two further functions, "brain activity" and "monitor error"; see comments about the categories later.) We tested each of them in the simulation paradigm and derived a "dream" and a "nightmare" set dependent on performance.

The compound results for both reaction time and accuracy in identification are shown in Figure 4, which has undergone a transformation so that for both measures, higher scores are better. Here we see that the dream team outperforms the nightmare team (statistically significantly) and that shorter reaction times are associated with more accurate recognition. Thus, some auditory icons simply work better than others.

Other studies currently being carried out as part of the formative (and, more recently, summative) testing include the audibility of the alarm signals in realistic listening conditions. Findings thus far indicate that the sounds work well in relatively low signal-to-noise ratios (a finding being demonstrated for alarm signals more generally in other studies; Schlesinger et al., 2018; Schlesinger, Stevenson, Shotwell, & Wallace, 2014; Stevenson, Schlesinger, & Wallace, 2013) and that the presence of the pointer enhances audibility. The pointer in particular was found to be audible in noise that was four times louder.

Reflection 3. *Because the alarm signals are intended for the update of the standard, and therefore access to them will be of commercial advantage, the final sounds will be released to medical instrumentation companies via a Web site through AAMI. (The final details of this process are yet to be decided.) Several companies are keen to do their own testing on the sounds once those are released.*

Another aspect of updating the standard is to update and enhance the guidance given to stakeholders, particularly sound designers, human factors engineers working on clinical device safety, medical instrument companies, and test houses, among others.

SUMMATIVE TESTING AND OTHER WORK

Our summative testing follows the broad protocols of the formative testing, with additional researchers testing the sounds in a range of clinical environments using protocols yet to be developed as well as using accepted and published protocols (Schlesinger et al., 2014; Stevenson et al., 2013). There is also other, related work being conducted.

One of the authors (MLB) is leading an Agency for Healthcare Research and Quality–funded project grant looking at the issue of masking of auditory alarm signals with specific reference to IEC 60601-1-8 (Hasanain, Boyd, Edworthy, & Bolton, 2017). This research will fill a large gap in terms of understanding where and when auditory masking will occur, which is somewhat beyond the scope of the immediate project described but is very important in general terms in understanding audible alarms from a human factors perspective.

This model-checking approach uses formal methods involving computing methods used for the specification, verification, and modeling of systems. It works as formal verification by working through all possible configurations of a system to check the propositions of the system. If the properties hold, the model can confirm this result, and if they do not hold (i.e. the model throws up a counterexample), then the specific set of values that gives rise to the counterexample can be checked. Thus it is an efficient way of assessing a system that could otherwise not be achieved. It has often been used to assess automated systems in human factors, but not for auditory masking specifically.

The model uses several submodels, including a clock submodel, an alarms submodel, and a masking computation submodel. Using actual audible alarms as input, the model can predict whether or not alarms will mask one another under specific conditions (for example, the onset of the timing of one alarm relative to one or more others). The model is still in the process of refinement and testing with human participants. Naturally the researchers are aware of both the current alarm signals and the projected new alarm signals, which will help to ensure the model's validity and relevance as a practical instrument, and also pushes the functionality of the model to more complex masking tasks.

We are also carrying out more theoretical studies on the contributions of strength of metaphorical link and auditory diversity in alarm set learning, as these two dimensions are thought to be large contributors to the effectiveness of any set of alarm signals.

Reflection 4*. The work is on track to be completed to the satisfaction of the IEC alarms joint working group well before the updated standard is published in 2019. By that time, many papers will be published documenting the performance of the alarm signals, from basic testing to their performance in simulated environments and their performance in noise and in other, increasingly realistic, tasks. Of course, the project will not have reached a satisfactory conclusion until the alarm signals and the relevant advice are embodied within the standard, and there is still a way to go, with other possible unknown threats along the way.*

We anticipate that our work will improve patient safety and clinical work performance as well as contribute to the science of alarm design and implementation.

ORCID ID

J. R. Edworthy https://orcid.org/0000-0002-2884-5780
M. C. Wright https://orcid.org/0000-0001-7121-504X

REFERENCES

Abbott, K. W., & Snidal, D. (2001). International "standards" and international governance. *Journal of European Public Policy*, *8*, 345–370.

Alali, K. A. (2011). *Azimuthal localization and detection of vehicular backup alarms under electronic and non-electronic hearing protection devices in noisy and quiet environments* (PhD thesis). Virginia Polytechnic Institute and State University, Blacksburg.

Atyeo, J., & Sanderson, P. M. (2015). Comparison of the identification and ease of use of two alarm sound sets by critical and acute care nurses with little or no music training: A laboratory study. *Anaesthesia*, *70*, 818–827.

Belz, S. M., Robinson, G. S., & Casali, J. G. (1999). A new class of auditory warning signals for complex systems: Auditory icons. *Human Factors*, *41*, 608–618.

Bennett, C. L., Dudaryk, R., Ayers, A. L., & McNeer, R. R. (2015). Simulating environmental and psychological acoustic factors of the operating room. *Journal of the Acoustical Society of America*, *138*, 3855–3863.

Bennett, C. L., & McNeer, R. R. (2012). PT-SAFE: A software tool for development and annunciation of medical audible alarms. *Anesthesia & Analgesia*, *114*, 576–583.

Block, F. E., Jr. (2008). "For if the trumpet give an uncertain sound, who shall prepare himself to the battle?"(I Corinthians 14: 8, KJV). *Anesthesia & Analgesia*, *106*, 357–359.

Block, F. E., Rouse, J. D., Hakala, M., & Thompson, C. L. (2000). A proposed new set of alarm sounds which satisfy standards and rationale to encode source information. *Journal of Clinical Monitoring and Computing*, *16*, 541–546.

Catchpole, K. R., McKeown, J. D., & Withington, D. J. (2004). Localizable auditory warning pulses. *Ergonomics*, *47*, 748–771.

Cheng, M. (2003). *Medical device regulations: Global overview and guiding principles*. Geneva, Switzerland: World Health Organization.

Cvach, M. (2012). Monitor alarm fatigue: An integrative review. *Biomedical Instrumentation & Technology*, *46*, 268–277.

Deb, S., & Claudio, D. (2015). Alarm fatigue and its influence on staff performance. *IIE Transactions on Healthcare Systems Engineering*, *5*, 183–196.

Drew, B. J., Harris, P., Zègre-Hemsey, J. K., Mammone, T., Schindler, D., Salas-Boni, R., Bai, Y., Tinoco, A., Ding, Q., & Hu, X. (2014). Insights into the problem of alarm fatigue with physiologic monitor devices: A comprehensive observational study of consecutive intensive care unit patients. *PloS ONE*, *9*(10), e110274.

Edworthy, J., Page, R., Hibbard, A., Kyle, S., Ratnage, P., & Claydon, S. (2014). Learning three sets of alarms for the same medical functions: A perspective on the difficulty of learning alarms specified in an international standard. *Applied Ergonomics*, *45*, 1291–1296.

Edworthy, J., Reid, S., McDougall, S., Edworthy, J., Hall, S., Bennett, D., Khan, J., & Pye, E. (2017). The recognizability and localizability of auditory alarms: Setting global medical device standards. *Human Factors*, *59*, 1108–1127.

Edworthy, J. R., Schlesinger, J. J., McNeer, R. R., Kristensen, M. S., & Bennett, C. L. (2017). Classifying alarms: Seeking durability, credibility, consistency, and simplicity. *Biomedical Instrumentation & Technology*, *51*(s2), 50–57.

Graham, R. (1999). Use of auditory icons as emergency warnings: Evaluation within a vehicle collision avoidance application. *Ergonomics*, *42*, 1233–1248.

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.

Honan, L., Funk, M., Maynard, M., Fahs, D., Clark, J. T., & David, Y. (2015). Nurses' perspectives on clinical alarms. *American Journal of Critical Care*, *24*, 387–395.

International Electrotechnical Commission. (2006). *IEC 60601-1-8: Medical electrical equipment—General requirements, tests and guidance for alarm systems in medical electrical equipment and medical electrical systems*. Geneva, Switzerland: International Electrotechnical Commission.

Keller, P., & Stevens, C. (2004). Meaning from environmental sounds: Types of signal-referent relations and their effect on recognizing auditory icons. *Journal of Experimental Psychology: Applied*, *10*, 3–12

Kristensen, M., Edworthy, J., & Ozcan Vieira, E. (2017). In search of acoustical signs of alarm fatigue. *Sound Effects*, *6*, 88–104.

Lacherez, P., Seah, E. L., & Sanderson, P. (2007). Overlapping melodic alarms are almost indiscriminable. *Human Factors*, *49*, 637–645.

Leung, Y. K., Smith, S., Parker, S., & Martin, R. (1997). Learning and retention of auditory warnings. In *Proceedings of the International Community for Auditory Display*. Retrieved from http://www.icad.org/Proceedings/1997/LeungSmith1997.pdf

Li, S. Y., Tang, T. L., Hickling, A., Yau, S., Brecknell, B., & Sanderson, P. M. (2017). Spearcons for patient monitoring: Laboratory investigation comparing earcons and spearcons. *Human Factors*, *59*, 765–781.

Mattli, W., & Büthe, T. (2003). Setting international standards: Technological rationality or primacy of power? *World Politics*, *56*, 1–42.

McNeer, R. R., Bennett, C. L., & Dudaryk, R. (2016). Intraoperative noise increases perceived task load and fatigue in anesthesiology residents: A simulation-based study. *Anesthesia & Analgesia*, *122*, 2068–2081.

McNeer, R., Bennett, C. L., Horn, D. B., Dudaryk, R., & Edworthy, J. (2017a, May). *The aggregate performance of novel auditory "icon" alarms in a simulated intensive care unit environment*. Poster presented at International Anesthesia Research Society, Washington, DC.

McNeer, R., Bennett, C. L., Horn, D. B., Dudaryk, R., & Edworthy, J. (2017b, May). *Use of ordinal regression analysis for assessing and comparing individual alarm-type performance in a simulated intensive care unit environment*. Poster presented at International Anesthesia Research Society, Washington DC.

McNeer, R. R., Bodzin Horn, D., Bennett, C. L., Reed Edworthy, J., & Dudaryk, R. (2018). Auditory icon alarms are more accurately and quickly identified than current standard medical alarms in a simulated clinical setting. *Anesthesiology*, *129*, 58–66.

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, *63*, 81–97.

Morrow, D. G., & Durso, F. T. (2011). Patient safety research that works: Introduction to the special issue on human performance and health care. *Journal of Experimental Psychology: Applied*, *17*, 191–194.

Patterson, R. D. (1982). *Guidelines for auditory warnings systems on civil aircraft* (Civil Aviation Authority Paper 82017). London, UK: Civil Aviation Authority.

Patterson, R. D., Edworthy, J., Shailer, M. J., Lower, M. C., & Wheeler, P. D. (1986). *Alarm sounds for medical equipment in intensive care areas and operating theatres* (Report No. AC598). Southampton, UK: Institute of Sound and Vibration Research.

Perry, N. C., Stevens, C. J., Wiggins, M. W., & Howell, C. E. (2007). Cough once for danger: Icons versus abstract warnings as informative alerts in civil aviation. *Human Factors*, *49*, 1061–1071.

Petocz, A., Keller, P. E., & Stevens, C. J. (2008). Auditory warnings, signal-referent relations, and natural indicators: Re-thinking theory and application. *Journal of Experimental Psychology: Applied*, *14*, 165–178.

Rayo, M. F., & Moffatt-Bruce, S. D. (2015). Alarm system management: Evidence-based guidance encouraging direct measurement of informativeness to improve alarm response. *BMJ Quality & Safety*, *24*, 282–286.

Sanderson, P. M., Wee, A., & Lacherez, P. (2006). Learnability and discriminability of melodic medical equipment alarms. *Anaesthesia*, *61*, 142–147.

Schlesinger, J. J., Baum Miller, S. H., Nash, K., Bruce, M., Ashmead, D., Shotwell, M. S., Reed Edworthy, J., Wallace, M. T., & Weinger, M. B. (2018) Acoustic features of auditory medical alarms: An experimental study of alarm volume. *Journal of the Acoustical Society of America*, *143*, 3688–3697.

Schlesinger, J. J., Stevenson, R. A., Shotwell, M. S., & Wallace, M. T. (2014). Improving pulse oximetry pitch perception with multisensory perceptual training. *Anesthesia & Analgesia*, *118*, 1249–1253.

Sendelbach, S., & Funk, M. (2013). Alarm fatigue: A patient safety concern. *AACN Advanced Critical Care*, *24*, 378–386.

Stephan, K. L., Smith, S. E., Martin, R. L., Parker, S. P., & McAnally, K. I. (2006). Learning and retention of associations between auditory icons and denotative referents: Implications for the design of auditory warnings. *Human Factors*, *48*, 288–299.

Stevenson, R. A., Schlesinger, J. J., & Wallace, M. T. (2013). Effects of divided attention and operating room noise on perception of pulse oximeter pitch changes: A laboratory study. *Journal of the American Society of Anesthesiologists*, *118*, 376–381.

Ulfvengren, P. (2003). Associability: A comparison of sounds in a cognitive approach to auditory alert design. *Human Factors and Aerospace Safety*, *3*, 313–331.

Vaillancourt, V., Nélisse, H., Laroche, C., Giguère, C., Boutin, J., & Laferrière, P. (2013). Comparison of sound propagation and perception of three types of backup alarms with regards to worker safety. *Noise and Health*, *15*, 420–436.

Wee, A. N., & Sanderson, P. M. (2008). Are melodic medical equipment alarms easily learned? *Anesthesia & Analgesia*, *106*, 501–508.

Welch, J. (2011). An evidence-based approach to reduce nuisance alarms and alarm fatigue. *Biomedical Instrumentation & Technology*, *45*(s1), 46–52.

Whalen, D. A., Covelle, P. M., Piepenbrink, J. C., Villanova, K. L., Cuneo, C. L., & Awtry, E. H. (2014). Novel approach to cardiac alarm management on telemetry units. *Journal of Cardiovascular Nursing*, *29*(5), E13–E22.

Judy Reed Edworthy *is a professor of applied psychology at Plymouth University, United Kindom. She has published many papers on audible alarm design and implementation, as well as papers on visual warnings and music perception. She has worked on many applied projects where audible alarms have been designed, tested, and subsequently implemented. Her recent work has focused on clinical alarms.*

Richard R. McNeer *is a physician scientist and practicing anesthesiologist at a busy Level 1 trauma center. For more than a decade, he has experienced firsthand how noisy and overwhelming critical care settings can be, and his research has focused on improving the interface between caregiver and patient through the auditory medium. He has received funding and investigated topics surrounding clinical environment acoustics, alarm design, and the pressing problem of alarm fatigue.*

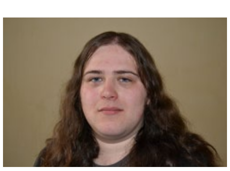
Christopher L. Bennett*, PhD, is an assistant professor in the music engineering technology program at the Frost School of Music at the University of Miami. He teaches classes in psychoacoustics, transducer theory, audio signal processing, and audio programming. His research projects include a collaboration with the Miami VA Hospital, investigating the impact of auditory biofeedback to improve rehabilitation, and an international collaboration investigating audible medical alarms to improve patient outcomes.*

Roman Dudaryk *is an associate professor of anesthesiology in the Department of Anesthesiology and Perioperative Medicine at the University of Miami. He completed his postgraduate training as an anesthesiologist at University of Miami and as an intensivist at Duke University. He has published numerous papers in the areas of noise pollution in the operating room, alarm fatigue, and learning. His research interests are development of new generation of medical alarms, quality improvement, and perioperative crisis management.*

Siné J. P. McDougall *is a professor of psychology at Bournemouth University, United Kingdom. Her research focuses on how icons and signs are learned, understood, and used. This research focus includes how people attend to and search for icons in complex interfaces and, most recently, how visual appeal affects icon search, learning, and comprehensibility.*

Joseph J. Schlesinger *is an assistant professor in the Department of Anesthesiology Division of Critical Care Medicine, at Vanderbilt University School of Medicine, Nashville, Tennessee, where he attends in the operating room, neuro ICU, burn ICU, and VA surgical ICU. His research interests include human factors engineering, auditory perception, temporal precision, and alarm development. His recent work has focused on signal-to-noise ratios in auditory medical alarms.*

Matthew L. Bolton *is an assistant professor of industrial and systems engineering at the University at Buffalo. He obtained his PhD in systems engineering from the University of Virginia, Charlottesville, in 2010. Previously he worked as a senior researcher at the NASA Ames Research Center and as an assistant professor of industrial engineering at the University of Illinois at Chicago. His research focuses on the use of formal methods in human factors engineering.*



Jonathan D. Reed Edworthy *has a BSc in music technology from Plymouth University, United Kingdom. He works as a sound designer, composer, and music producer. He has worked on a number of audible alarm design projects, including the project described in this article.*

Elif Özcan*, PhD, is director of the Critical Alarms Lab at TU Delft. Her research interests include the mental processes underlying product experiences involving sounds and sound-embedded objects. Her recent research focuses on alarm experiences in medical contexts and auditory/multimodal display design for autonomous vehicles. She leads the Silent ICU project at Erasmus Medical Centre, Netherlands.*

Andrew D. Boyd *is an assistant professor in the Department of Biomedical and Health Information Sciences. His research focuses on data simplification to improve clinical outcomes and engages administrators, researchers, and patients. His recent research was in simplifying the transition for International Classification of Disease Version 9 clinical modification (ICD-9-CM) to Version 10 (ICD-10-CM). Another area to improve clinical outcomes is designing audio alarms better to allow clinicians to respond to the alarms.*

Scott K. J. Reid *has a BSc in psychology from Plymouth University, United Kingdom, and is studying for a master's degree at the same institution. He has worked on a number of audible alarms projects, including the clinical alarms work highlighted in this article as well as alarms for the European Space Agency.*

Michael F. Rayo*, PhD, is an assistant professor in the Department of Integrated Systems Engineering at The Ohio State University and a scientific adviser for patient safety at The Ohio State Wexner Medical Center. His research focuses on how the design of cognitive tools influences multiagent teaming architectures and can contribute to overall system resilience. His work focuses on alarm design and management, visual analytics, computerized decision support, and interpersonal communication.*

Melanie C. Wright *has a PhD in industrial engineering from North Carolina State University. She has over 20years of experience in engineering and research in the areas of human performance, usability analysis, and human–machine system design. At Trinity Health and Saint Alphonsus, she leads a research program focused on optimization of human–system performance in health care.*

David Osborn *is the senior manager of global regulations and standards for Philips, where he is responsible for standards development across all of Philips Healthcare. He represents Philips at the national level and the United States at international standards organizations on standards matters affecting the medical device industry. Heavily involved in international standards work, he serves as secretary to the International Standards Organization (ISO) subcommittee TC121/SC3, Anesthesia and Respiratory Equipment, Lung Ventilators, and Related Equipment. He is the co-convener or secretary to several joint working groups and is a U.S. expert to a number of ISO and International Electrotechnical Commission working groups.*

Research reported in this article was supported by the Association for the Advancement of Medical Instrumentation in two grants to Judy Edworthy, "Audible Alarms in Support of IEC 60601-1-8 Phase 1/Phase 2." Research reported in this article was also supported by the Agency for Healthcare Research and Quality (AHRQ) under Award No. R18HS024679 to Matthew Bolton. Michael Rayo's participation in this effort was supported by the Institute for the Design of Environments Aligned for Patient Safety (IDEA4PS) at The Ohio State University, which is sponsored by AHRQ (P30HS024379). The authors' views do not necessarily represent the views of AHRQ.

eid

DOI: 10.1177/1064804618763268