ORIGINAL PAPER

ECG sonification to support the diagnosis and monitoring of myocardial infarction

Andrea Lorena Aldana Blanco[1] · **Steffen Grautoff**[2] · **Thomas Hermann**[1]

Abstract

This paper presents the design and evaluation of four sonification methods to support monitoring and diagnosis in Electrocardiography (ECG). In particular we focus on an ECG abnormality called ST-elevation which is an important indicator of a myocardial infarction. Since myocardial infarction represents a life-threatening condition it is of essential value to detect an ST-elevation as early as possible. As part of the evaluated sound designs, we propose two novel sonifications: (i) *Polarity sonification*, a continuous parameter-mapping sonification using a formant synthesizer and (ii) *Stethoscope sonification*, a combination of the ECG signal and a stethoscope recording. The other two designs, (iii) the *water ambience sonification* and the (iv) *morph sonification*, were presented in our previous work about ECG sonification (Aldana Blanco AL, Steffen G, Thomas H (2016) In: Proceedings of Interactive Sonification Workshop (ISon). Bielefeld, Germany). The study evaluates three components across the proposed sonifications (1) detection performance, meaning if participants are able to detect a transition from healthy to unhealthy states, (2) classification accuracy, that evaluates if participants can accurately classify the severity of the pathology, and (3) aesthetics and usability (pleasantness, informativeness and long-term listening). The study results show that the *polarity* design had the highest accuracy rates in the detection task whereas the *stethoscope sonification* obtained the better score in the classification assignment. Concerning aesthetics, the *water ambience sonification* was regarded as the most pleasant. Furthermore, we found a significant difference between sound/music experts and non-experts in terms of the error rates obtained in the detection task using the *morph sonification* and also in the classification task using the *stethoscope sonification*. Overall, the group of experts obtained lower error rates than the group of non-experts, which means that further training could improve accuracy rates and, particularly for designs that rely mainly on pitch variations, additional training is needed in the non-experts group.

Keywords Electrocardiogram · Sonification · Process monitoring · Myocardial infarction

✉ Andrea Lorena Aldana Blanco
  aaldanablanco@techfak.uni-bielefeld.de

  Steffen Grautoff
  steffen.grautoff@klinikum-herford.de

  Thomas Hermann
  thermann@techfak.uni-bielefeld.de

[1] CITEC, Bielefeld University, Inspiration 1, 33619 Bielefeld, Germany

[2] Klinikum Herford, Schwarzenmoorstrasse 70, 32049 Herford, Germany

1 Introduction

The use of sonification, i.e. the systematic auditory representation of data as sound, for process monitoring has greatly expanded in recent years. At present, there is a wide range of applications where sonification has been presented as a supporting monitoring tool in industrial, domestic or work environments [12,24,25]. One of the advantages of representing data through sound, is that users can focus their attention on a primary task while being aware of changes in the signals of interest. Furthermore, the temporal resolution of the human auditory system provides high performance in detecting changes and patterns in the signals [18], which makes sonification a good asset for time-series monitoring.

In the medical field, sonification is already frequently used as a monitoring tool. One example is the pulse oxime-

ter which is used for measuring oxygen saturation levels of the blood. The device produces a short duration sound of a given frequency [6] that is synchronized with the pulse rate, when the oxygen saturation level decreases, the pitch of the resulting sound decreases as well. Nevertheless, the use of sonification in health does not only provide options for monitoring but it also opens a wide range of applications to support diagnostics and rehabilitation tasks. However, having a larger number of sounding devices in medical scenarios also creates a set of challenges that need to be addressed. For example, constant alarm sounds generate a higher cognitive load in the medical staff leading to delayed responses or underestimation of alarms when treating the patients. Additionally, a noisy environment can cause ear fatigue thus increasing mental distress and irritability to people exposed to the sounds [21].

The current challenges in medical sonification can be addressed from two sides. On the one hand, we need to find methods to convey information reliably so that the resulting sounds can support physicians in their monitoring and diagnostic tasks and, on the other hand, we need to account for aesthetic qualities that take into consideration ear fatigue and other soundscape elements from the medical environment.

In this paper we first provide an overview of ECG signals and explain the ST-elevation pathology. Then, we introduce the four sonification designs: (1) Water ambience, (2) Morph, (3) Polarity pitch and (4) Stethoscope. Subsequently, we describe the study design and present the results obtained. Discussions and conclusions summarize the paper.

2 Related work

Apart from the pulse oximeter, research on biosignals sonification comprises a wider range of applications either for monitoring or as a supporting tool in diagnostic tasks. Methods to sonify electroencephalography (EEG) [26], electromyography (EMG) [19], CT/PET scans [9,20], among others, had been presented.

In terms of ECG sonification, there are two main research lines that can be found in literature. Firstly, there are approaches focused on sonifying temporal features such as the heart rate [22] and the heart rate variability [1] [3,4]. Heart rate sonification is used mainly to track performance in sports as the intensity of training can be estimated if a certain heart rate target is reached whereas heart rate variability is tightly related to diagnostics since it is an important estimator for arrhythmias[2] and other heart-related diseases. The second approach targets the morphology of the ECG signal [2,15,23] since certain pathological states correspond to

---

[1] Measure of the beat-to-beat time variation.

[2] Rhythm disturbances in the heart.

specific changes in the waves that compose the ECG and therefore serve as a supporting tool in diagnosis tasks.

Overall, ECG sonification can be considered to be in early stages since there is still a long way until research prototypes reach people in medical settings. However, results obtained so far show the potential of sonification to monitor and detect cardiac pathologies.

The next section focuses on the main characteristics of ECG signals and explains the characteristics of an ST-elevation in detail.

3 The ECG signal

The electrocardiogram (ECG) is a visual representation of the potential differences of the cardiac muscle cells. These potential differences generate signals which can be measured by placing electrodes on the skin. These signals are usually recorded as a printed set of curves and amplitudes called ECG. The amplitude depends on multiple factors such as the thickness of the muscle, the distance of the recording electrode and the amount of surrounding tissue such as air, fluid or fat [10]. Electrical activity can be assessed from different angles if more than one lead (channel) is used and provide the clinician with important information regarding regional electrical activity [16]. Standard ECGs are performed using 12 leads. The ECG can be divided into several parts which represent different states of the hearts cycle. The periodic rotation from depolarization to repolarization causes a contraction followed by a relaxation of the heart muscle. The first recordable signal is the P wave which results from the depolarization of the left and right atrium. The QRS complex represents the depolarization of the ventricles. The repolarization of the atria is commonly not visible, because they are vanished in the QRS complex. The repolarization of the ventricles is recorded as the T wave.

Physicians use the previously mentioned reference points to detect in which part of the heart abnormalities are located. For example, a healthy ST-segment should be isoelectric. However, when a coronary, the artery of the heart, is blocked, this interval can be suppressed or elevated. The next section will give more detail on ST-elevations and its medical connotations.

3.1 Elevation of the ST-segment

The ST-segment is an isoelectric part of the ECG. Its beginning is located at the J point which marks the end of the QRS complex and ends at the beginning of the T wave. Its isoelectric shape corresponds to the phase in between the ventricular depolarization and the ventricular repolarization (see Fig. 1). In clinical practice the ST-segment of the ECG is most commonly used for the detection of a myocardial ischemia.
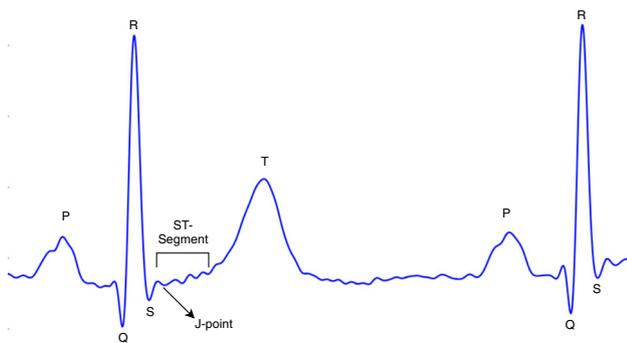
Fig. 1 ECG intervals and reference points during a cardiac cycle (one channel/lead)

Fig. 2 Surrogate signals with several ST-elevation levels

Myocardial ischemia implies that blood supply to the coronaries is occluded and therefore oxygen supply decreases. In this case the ST-segment will show elevation or depression. The reason for ST-segment changes are a transmural[3] conduction slowing and a greater depression of the action potentials of the epicardium [8]. If a myocardial infarction is present, the patient requires urgent revascularization therapy [14].

ST-elevation criteria An ST-elevation suggestive for acute coronary occlusion is defined by the European Society of Cardiology as [14]

- In men < 40 years: at least two contiguous leads with ST-segment elevation ≥ 0.25 mV (2.5 mm)
- In men ≥ 40 years: at least two contiguous leads with ST-segment elevation ≥ 0.2 mV (2 mm)
- In women ≥ 0.15 mV (1.5 mm) in leads V2–V3 and/or ≥ 0.1 mV (1 mm) in the other leads

Stages of ECG changes in myocardial infarction If a coronary is occluded the T wave reacts immediately with an increase of amplitude. In a matter of minutes, the ST segment elevates in leads which represent the affected area of the heart, whereas the leads representing the opposite side of the affected area will show ST segment depression. If the myocardial infarction progresses, T will become negative and the amplitude of Q will increase. On the ECG a life-long deep Q can be seen as a remaining myocardial scar.

4 Preparation of data and ECG features extraction

In this section, we first explain how we created the ECG signals for the study and then go over the feature extraction.
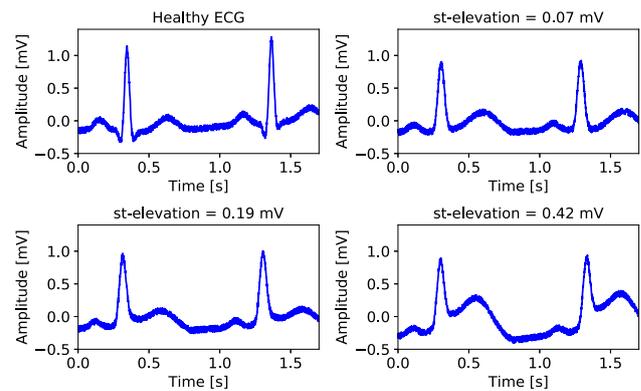
4.1 Creating the surrogate signals

Even though there are large open access ECG databases available, finding real-life signals with specific levels of ST-elevation that range from healthy to severe elevation is not an easy task. For this reason we decided to create ECG surrogate signals for the study in order to have better control of the ST-elevation levels.

To create the surrogate signals we use an ECG waveform generator called *ECGSYN*[4] developed by the Department of Engineering Science in University of Oxford and by the Laboratory for Computational Physiology at the MIT [17]. *ECGSYN* is a versatile tool with a number of parameters to control the location, amplitude and shape of the ECG reference points and intervals, as well as the heart rate of the signal.[5]

Figure 2 depicts four surrogate signals with different levels of ST-elevation (0.0 mV, 0.07 mV, 0.19 mV, 0.43 mV) created using the ECGSYN model. The ST-level is measured at the J-point with respect to isoelectricity. All signals are sampled at 1000 Hz and have a heart rate of 60 beats per minute (bpm).

4.2 ECG signal processing and features extraction

In order to emulate realistic signals the surrogate files also account for signal noise in line with realistic noise levels. Within our data preparation tool-chain as a first step of feature extraction, we remove artifacts and unwanted noise. First we remove the DC component and then apply a low-pass filter with a *cutoff* frequency of 70 Hz. Frequencies above 70 Hz are commonly dismissed because they are outside the range for ECG diagnostic [7]. Then we perform the R-peak detection using the method proposed by Worrall et al. [27] using a time window of 200 ms. We select this window size

---

[3] From the endocardium to epicardium.

[4] https://physionet.org/physiotools/ecgsyn/.
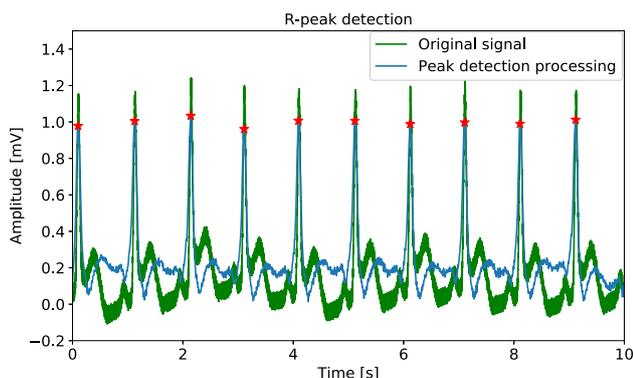
[5] Number of heartbeats per minute.

Fig. 3 Results of R-peak detection: red star symbols mark the detected R-peaks of the original signal (green curve)

Fig. 4 ST-elevation detection

according to the duration of the QRS complex in a 60 bpm signal [6] [7].

Figure 3 shows the ECG signal together with the detected R-peaks. The green color depicts the raw signal and the blue color the transformed signal after applying the method proposed in [27].

Once the R-peaks are detected, the ST-elevation can be calculated as the distance between the J-point and isoelectricity in each heartbeat. The isoelectric reference can be taken either from the PQ segment or the TP segment. Given that the TP segment is longer and the PQ segment might differ from isoelectricity if there is pathologic behavior in the atrium [13], we take the TP segment as the isoelectric reference.

The J-point is determined by applying the method by Al-Kindi and Tafreshi [1] where the first derivative of the ECG signal is calculated, taking the first point after the S-wave where the derivative is zero as J-point. We manually estimate the location and duration of the S-wave—and therefore the J-point search area—using as a reference the regular duration of the QRS complex in a 60 bpm signal [7].

Figure 4 depicts the raw signal and the detected J-points and TP-segments.

When the J-point has been estimated, we calculate the average amplitude in the segment defined between the J-point and a number of samples ahead. We compute

$$\overline{ST} = \frac{1}{t_a - t_{\text{j-point}}} \int_{t_{\text{j-point}}}^{t_a} h(t)\,dt \tag{1}$$

by summing the sampled signal $h(t)$ between the segment borders t_a and $t_{\text{j-point}}$.

In a similar way we calculate the amplitude in the TP segment as follows

---
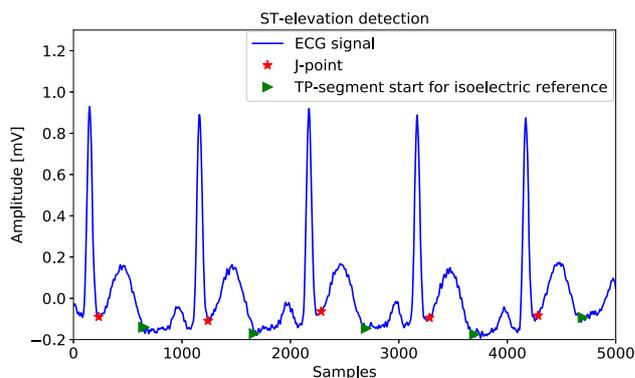
[6] The average QRS complex width is 100 ms

$$\overline{TP} = \frac{1}{t_{\text{TPend}} - t_{\text{TPstart}}} \int_{t_{\text{TPend}}}^{t_{\text{TPstart}}} g(t)\,dt \tag{2}$$

Finally the average ST-segment amplitude is computed by

$$\overline{ST_{amp}} = \overline{ST} - \overline{TP} \tag{3}$$

If ST_{amp} is positive (resp. negative), there is an ST-elevation (resp. ST-suppression). Note that our surrogate signals used for the study only featured zero or positive ST-elevation.

5 Sonification designs

Aesthetics, as well as informativeness are main components that should be taken into account when designing the sonifications [5,11]. We want to create sounds that properly convey information, meaning that they should call the attention from the physician when there are pathological changes in the signal and, at the same time accurately represent the degree of urgency in the signal's change. Furthermore, the sound should be pleasant to listen to, and ideally not contribute to ear fatigue in medical environments.

In order to set a starting point that allows us to better understand which signal features and sounds could adequately represent changes in the ST-segment, we propose four different sonification designs that explore several ECG characteristics and perceptual features.

Polarity sonification The *Polarity sonification* is a basic parameter-mapping approach in which the absolute voltage difference of the ECG signal is mapped to the amplitude and number of harmonics of a Formant oscillator. A Formant oscillator produces a set of harmonics of a fundamental frequency boosting harmonics around a given formant frequency. In our approach, the voltage is mapped to the fundamental, i.e., higher voltages result in higher pitch. Lastly, the direction of the slope is used to control the panning, thus

Table 1 Polarity sonification: parameter-mapping

Data feature	Data range (min, max)	Perceptual parameter	Parameter range (min, max)
Absolute amplitude	(0, data max)	Level (dBFS)	$(-10, 0)$
Amplitude	(data min, data max)	Pitch (MIDI note)	(40, 70)
Amplitude	$(-0.1, 0.1)$	Sharpness (nr. of harmonics)	(1, 10)
Slope	Positive/negative	Location (Stereo panning)	$(-1, 1)$

Table 2 Water ambience sonification: parameter-mapping

Data feature	Data range (min, max)	Perceptual parameter	Parameter range (min, max)
ST-segment amplitude	(0 mV, 0.4 mV)	Number of drops	(1 drop, 5 drops)

Table 3 Morph sonification: parameter-mapping

Data feature	Data range (min, max)	Perceptual parameter	Parameter range (min, max)
ST-segment amplitude	(0 mV, 0.4 mV)	Morph factor	[0.0 (sine), 1.0 (square)]

Table 4 Stethoscope sonification: parameter-mapping

Data feature	Data range (min, max)	Perceptual parameter	Parameter range (min, max)
ST-segment amplitude	(0 mV, 0.4 mV)	Frequency-shift	(100 Hz, 500 Hz)

a positive slope causes the sound to shift to the right audio channel whereas negative slopes shifts the sound to the left channel.

The absolute voltage difference at t_i is given by

$$v = |v_i - \text{ref}| \, , \tag{4}$$

where ref is the signal mean $\bar{f} = \frac{1}{b-a} \int_a^b f(x)\,dx$.

Table 1 explains how the extracted parameters are mapped to the perceptual features.

Sonification example S1[7] corresponds to the *polarity* sonification of a dataset for a healthy condition whereas an ST-elevated signal can be heard in sound example S2.

Water ambience The *Water Ambience* sonification is a parameter-mapping design based on the analogy of blood flowing across the heart. However, in our design, instead of a continuous stream of water sounds we implemented a discrete representation of blood flow using short water drop sounds as the main component. Additionally, in this sonification we take an opposite approach to the analogy of blood flowing and propose that a healthy signal is represented with the least amount of sounds in order to not increase ear fatigue. As a result, in the *water ambience* sonifcation a number of water drops are triggered every heartbeat. The ST-elevation is mapped to the number of drops. For example, a healthy signal results in one water drop sound per cardiac cycle, while an ST-elevated signal triggers more drops. A more detailed explanation of the sonification design can be found in our previous work [2].

The amplitude in the ST-segment is mapped as follows (Table 2):

Sonification examples S3 and S4 illustrate the *water ambience* sonification of a healthy and ST-elevated signal.

Morph The morph sonification produces a short duration synthesized sound for every heartbeat, morphing continuously from a pure sine tone to a square wave signal as function of the ST-elevation. As the square wave is characterized by a richer harmonic series and thus higher spectral spread, this perceptual feature gives the only cue for pathological changes. The amplitude in the ST-segment is mapped as follows (Table 3):

Sonification example S5 and S6 illustrate the *morph* sonification of a healthy and an ST-elevated ECG signal.

Stethoscope The stethoscope design is a combination of the ECG raw signal and a stethoscope recording that triggers a real-heartbeat sound every cardiac cycle measured the ECG. The stethoscope sound is frequency-shifted according to the amplitude in the ST-segment. Thus, a healthy signal in which the amplitude in the ST-segment is regarded as isoelectric, preserves the low pitch components of the original stethoscope sound, whereas an ST-elevated signal is shifted by a larger frequency value, resulting in a higher pitch.

The amplitude in the ST-segment is mapped as follows (Table 4):

Sonification example S7 corresponds to the *stethoscope* sonification of a healthy ECG and example S8 to an ST-elevated signal.

6 Study design

In order to evaluate the sonification designs, we created an online survey using the Lime Survey[8] tool. The user study was approved by the ethical committee of Bielefeld University. The link to the online survey was announced through social networks and a mailing list of the sonification community. Among the study participants, a pair of studio headphones and five 16 GB USB memory sticks were raffled as a reward for participating in the online survey.

The user study was divided into four main parts: (i) a set of initial questions involving gender, age, experience with sound/music and ECG signals, among others (ii) a detection task, (iii) a classification task and (iv) a set of questions to evaluate the designs in terms of aesthetics, suitability for long-term listening, and informativeness. There were no particular requirements in terms of the playback system participants could use, however there was a calibration step where users were instructed to calibrate the playback volume at the beginning of the study.

6.1 Initial questionnaire and calibration phase

In the first part of the study participants were asked general demographics questions such as gender and age. Then, In order to get an overview about the participants' professional background, we asked about their experience with music/sound and experience regarding interpretation of ECG signals. Further questions focused on any hearing damage they were aware of by the time the survey was answered. Lastly, they were asked about the characteristics of their playback system.

Once the initial questionnaire was answered, there was a calibration step in which participants listened to an audio file of 8 s duration. The file was composed by 2 s of each sonification design, so that they would experience a good cross section of the acoustic material to be heard during the study. Participants were instructed to select a playback volume they felt comfortable with in order to continue with the study and, they were also advised to let the volume unchanged throughout the survey. However, we kept track of volume changes over the study for later analysis.

6.2 Detection task

The second part of the study was a detection task. In this part, sonifications of 10s were presented, pairing each audio file to a horizontal slider widget with a range from zero (0) to ten (10). The task was to adjust the slider to the point in time (in seconds) where they first noticed the sonification change from healthy to pathological. If there was no change, they
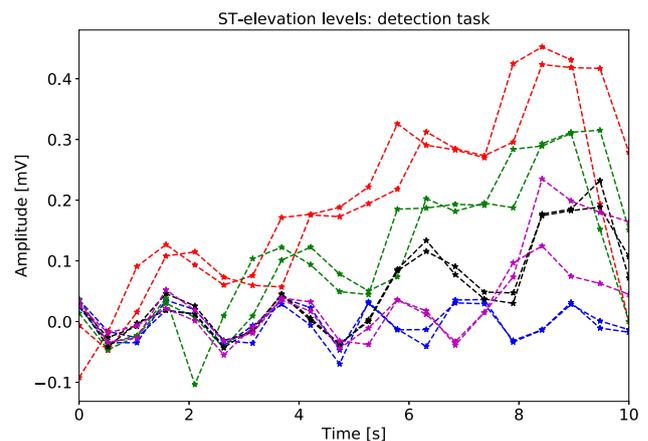

Fig. 5 ST-elevation levels of ECG files used in detection task

were instructed to set the slider to ten, meaning that after the whole audio was played there was no noticeable change.

While evaluating a sonification design, participants could always replay the sound example files corresponding to healthy and unhealthy ECG signals of that particular design. There was no limit in terms of the number of times they played a sonification or an example file. Also, there was no time limit to complete the task. Sonifications where presented to participants in a random order.

Participants where asked to evaluate ten audio files for each sonification design. All audio files started in the healthy condition where the amplitude of the ST-segment was close to isoelectricity. However, depending on the file, the amplitude of the ST-segment would either remain isoelectric or increase progressively until reaching minor to severe ST-elevation levels. Every increase in the ST-elevation levels, was produced every 2 s.

Figure 5 depicts the ST-levels of the signals used in the detection task. Note that due to the features of the ECG waveform generator, the increase from one ST-level to the next one is not a straight line, but there are fluctuations in the surrogate function. For example, there might be minor decreases in the ST-levels even when the trend indicates an increase. Overall, all signals have an increasing trend of ST-levels except for the ones intended to remain close to isoelectricity. There is also an artifact of the ECG model that generates a steep decrease in the ST-levels from the last heartbeat. However, since this change happens right at the end of the signal, it has no effect in the resulting sonification.

Duration of the audio files For sonification we compress time by a factor two, resulting in 10 s sonification for 20 s surrogate ECG data. The reason is that this allowed us to evaluate a wider range of ST-levels within the limited time available for an experiment session. Furthermore, the ability to recognize the changes in the sonifications as a result of the
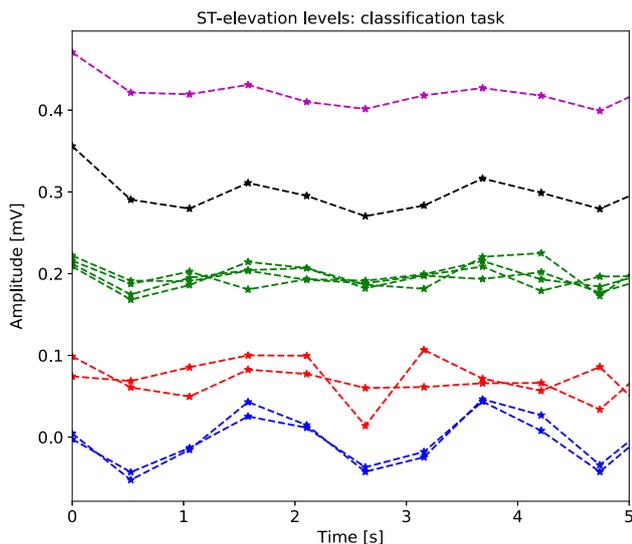

Fig. 6 ST-elevation levels of ECG files used in classification task

ST-elevation level variations is independent of rate according to our methods, as long as the rate is kept within an acceptable range that allow users to discern the main temporal patterns of the ECG signal.

6.3 Classification task

The third part of the study was a classification task. In this section, participants had to evaluate sonification examples of 5 s duration and classify them from healthy to severe ST-elevation using a 7-point Likert scale where one (1) meant healthy ECG and seven (7) meant severe ST-elevation.

Once again, at the beginning of the page for each sonification design, sound examples for a healthy and unhealthy signal were provided according to each sonification design. Again, participants could listen to these reference files as many times as they wanted.

As part of the task participants where asked to evaluate ten audio files for each sonification design. The ST-elevation levels in each audio file were kept as constant as possible, within the oscillatory limits of the ECG surrogate signal generator. Figure 6 depicts the ST-elevation levels as function of time for the ECG signals used in the classification task. Clusters according to ST-levels are represented by different colors. The blue lines show isoelectric references, while the purple line is the signal with the highest ST-elevation (approximately 0.4 mV).

6.4 Aesthetics and usability questions

After the detection and classification tasks were completed, participants were asked to answer a set of questions about each design. First, they were asked to rate on a 6-point Likert scale where one (1) meant 'Strongly Disagree' and six (6) meant 'Strongly Agree', items such as: pleasantness of the sonification, suitability for long-term listening and level of informativeness.

Additionally, they were asked to compare the proposed designs to the QRS tone commonly used in medical settings, in terms of which sound would be preferable to listen in a medical context. Sound example S9 is a typical QRS tone.[9] The rating in the comparison to the QRS tone was given using a 6-point Likert scale where one (1) meant *strongly disagree* and (6) meant (6) *strongly agree*.

At the end of the survey participants could add any additional comment they regarded as relevant.

7 Study results

A total of forty-two participants took part in the study, however, one of the participants was considered to be an outlier after analysis of results from the detection task, given that the person assigned the same detection time value to all audio across all sonification designs. After removal of the outlier forty-one surveys were considered for further analysis. *51.2%* of the participants were females and *48.8%* were males. The average time for completing the survey was 48.1 minutes. All participants reported not having a hearing loss condition.

In terms of sound/music experience *58.5%* people reported not having any prior experience, *2.4%* less than one year experience, *7.3 %* between one and three years experience and *31.7%* more than three years experience. Regarding previous knowledge about interpretation of ECG signals *14.6%* said to have experience with ECG signals while *85.4%* didn't have experience.

Lastly, when inquired about the sound system used to answer the survey, *53.7%* were using regular headphones, *17.1%* professional headphones, *19.5%* the computer's loudspeakers and *9.7%* the tablet's or smartphone's loudspeakers.

7.1 Results of detection task

To quantify results in the detection task we counted the number of heartbeats that were regarded as ST-elevated by participants. In order to determine this number we analyzed the point in time where they noticed the first change of an ECG signal from healthy to pathological. Afterwards, we counted the selected heartbeat and all subsequent heartbeats as ST-elevated. We did the previously described process for each participant across all audio files of the same sonification design.

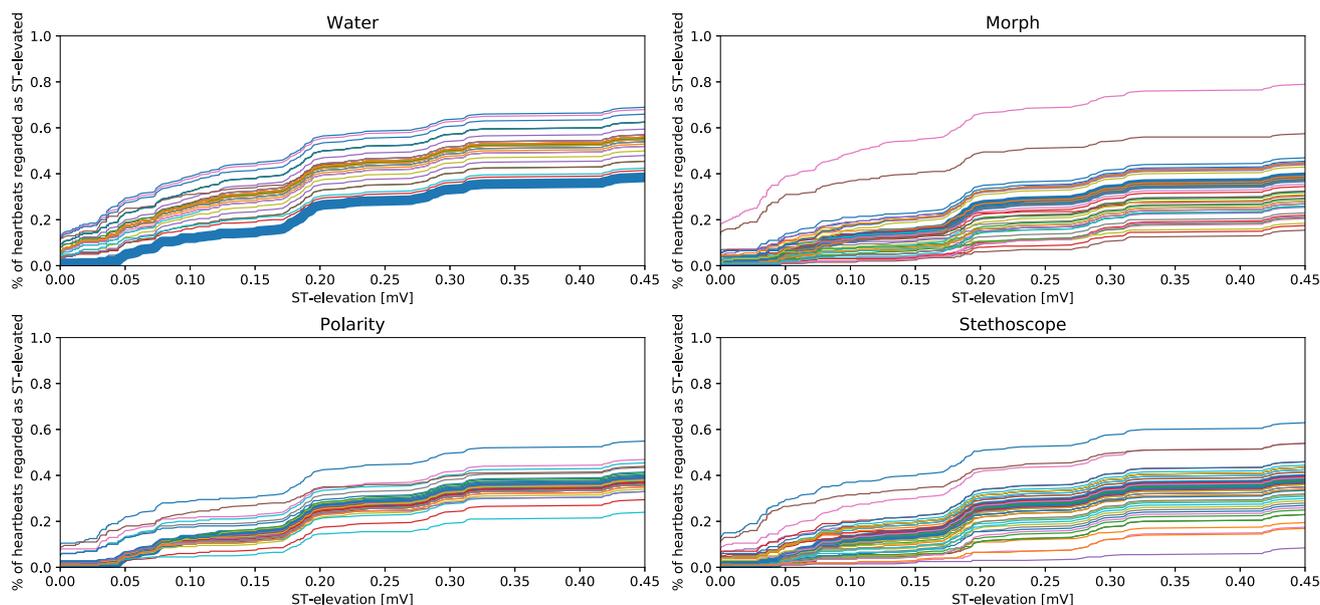[9] Retrieved from: https://freesound.org/people/samfk360/sounds/148897.

Fig. 7 Percentage of heartbeats regarded as ST-elevated in each sonification design. The thick blue line indicates the surrogate reference. Curves below the reference indicate underestimation of the ST-elevation and curves above represent overestimation of the elevation

As explained in Sect. 6.2, all ECG surrogate signals were created based on a 60 bpm rate and had a duration of 20 s. As all sonifications were rendered at a time compression of 2 resulting in a 10 s audio file, each ECG file contained a total of twenty heartbeats, one heartbeat per second. This means that a maximum of 200 heartbeats could have been regarded as ST-elevated by participants if they would have assumed all sonifications to correspond to ST-elevated from the first heartbeat on.

Figure 7 shows the percentage of heartbeats considered as ST-elevated by each participant. There is one plot for each sonification design. The thick blue line shows the target ST-elevated heartbeats percentage as a function of the ST-elevation values presented in Fig. 5, whereas the narrow lines show each participant's performance. The reference line was computed by assigning all heartbeats with an amplitude equal or higher than 0.1 mV as ST-elevated, see Sect. 3.1.

Lines above the reference correspond to an overestimation of the ST-elevation, i.e., listeners marked the beginning of the pathological sound before the signal had actually reached the minimum elevation to be considered pathological according to the medical standard. On the contrary, a line below/under the reference curve represents an underestimation of the pathology, revealing that participants did not detect the ST-elevation when it should have started to happen according to medical standards.

7.2 Results of classification task

To evaluate the performance in the classification task we estimated the error between the ST-elevation levels selected by participants and the surrogate levels shown in Fig. 6. To do so, we first applied a linear mapping from the scale values (1–7) to the surrogate levels (0 mV – 0.4 mV) to estimate to which ST-elevation level each user rating best corresponds. Subsequently, for each audio file, we calculated the root mean squared error (RMSE) between the nominal (i.e. actual) and rated ST-elevation values.

Figure 8 depicts both the RMSE obtained in the classification of each audio file (star-shaped markers) across all sonification designs and the average error by ST-elevation cluster as shown in Fig. 6.

Results show that middle ST-elevation values are harder to classify than extremes values. Generally speaking, isoelectricity and severe ST-elevation are detected with more accuracy, except for the polarity sonification where error rates are similar regardless of the ST-elevation level. Overall RMSE rates are as follows: *16.4* for the water ambience sonification, *15.5* for the morph design, *17.5* for the polarity sonification and *14.3* for the stethoscope.

7.3 Aesthetics results

Figure 9 show the results of the aesthetics group of questions: pleasantness, informativeness and long-term listening suitability.

Results are reported as $(\bar{x} \pm \sigma)$ in a scale from 1 to 6 (1 refers to *Strongly Disagree* and 6 to *Strongly agree*). Pleasantness of the sonification, was rated 5.0 ± 1.36 for the *water ambience sonification*, 3.0 ± 1.53 for the *polarity sonification*, 4.0 ± 1.65 for the *stethoscope sonification* and 3.0 ± 1.37 for the *morph sonification*. When asked about how infor-
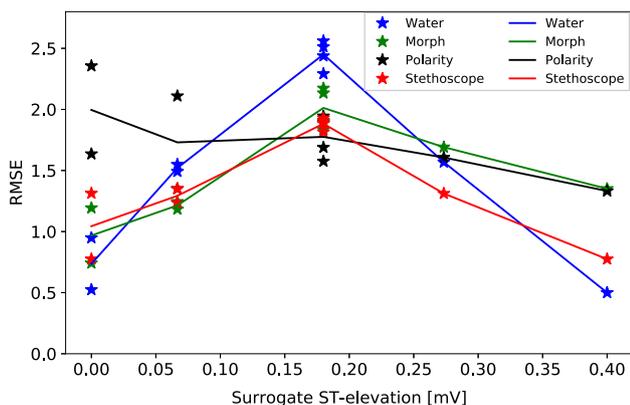
Fig. 8 Average root mean square error by ST-elevation cluster across all sonification designs. Blue color: water ambience sonification, green color: morph, black: polarity and red: stethoscope
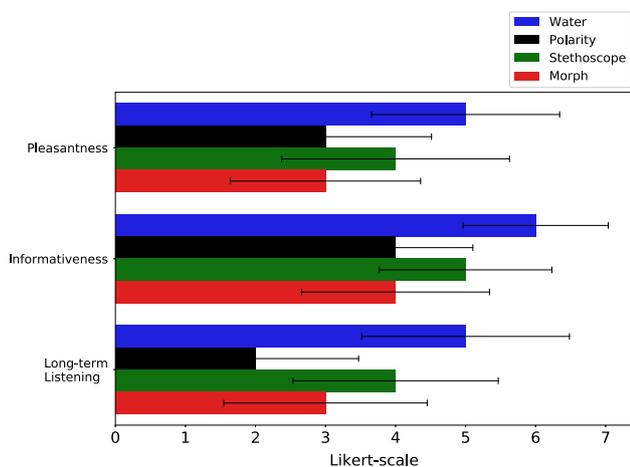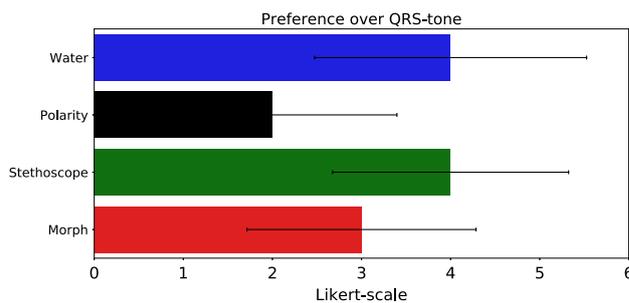

Fig. 10 Comparison to the QRS tone sound

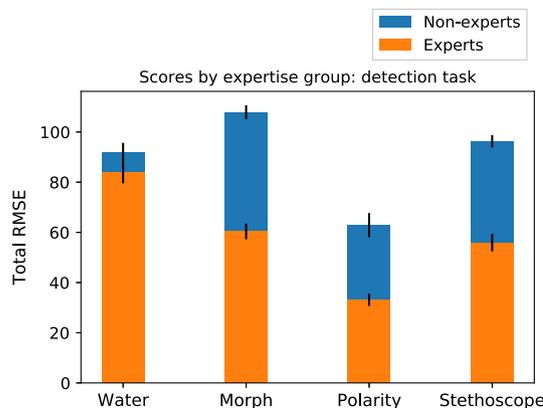**Fig. 9** Aesthetics results (median and standard deviation bars)

Fig. 11 Bar plot experts/non-experts detection task

7.4 Comparison according to level of expertise

In order to analyze how sound/music experience affected classification performance, we compared the results obtained by experts and non-experts in the detection and classification tasks. The group of experts were persons who said to have more than one year experience in sound/music related activities. On the contrary, the group of non-experts was composed by people having less or equal than one year experience.

We regarded the level of expertise based on previous sound/music experience instead of medical experience considering that the proposed study focuses on listening tasks and therefore listening abilities are more important than experience with ECG signals.

Figure 11 shows the error rates (RMSE) obtained by a sample of participants from each expertise group in the detection task. Considering that the two groups have different sample size, we randomly selected fifteen participants belonging to each expertise group in order to do the comparison in classification accuracy. Results indicate that participants from the two groups had rather a similar performance using the *water ambience* sonification. Oppositely, designs that relied mainly on pitch variations such as the *morph* and *stethoscope* sonifications, led to higher error rates from the non-experts group. The *polarity* sonification also led to higher error rates in the non-experts group, however it was the sonifi-
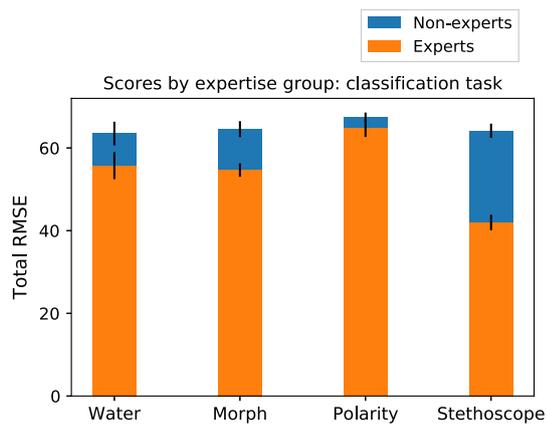
mative the sonification designs were, users gave a rating of 6.0 ± 1.05 for the *water ambience sonification*, 4.0 ± 1.12 for the *polarity sonification*, 5.0 ± 1.25 for the *stethoscope sonification* and 4.0 ± 1.36 for the *morph sonification*. Furthermore, long-term listening usability was rated as 5.0 ± 1.5 for the *water ambience sonification*, 2.0 ± 1.49 for the *polarity sonification*, 4.0 ± 1.48 for the *stethoscope sonification* and 3.0 ± 1.47 for the *morph sonification*.

Moreover, when comparing the proposed sonification designs to the QRS tone already used in medical environments, participants rated their level of agreement to the statement: "In a medical setting, the *X sonification* would be preferable to listen than the QRS tone sound". Scores were given on a scale from 1 to 6 where 1 refers to *Strongly Disagree* and 6 to *Strongly agree*). Results are depicted in Fig. 10 and reported as $(\bar{x} \pm \sigma)$ as follows: 4.0 ± 1.5 for the *water ambience sonification*, 2.0 ± 1.4 for the *polarity sonification*, 4.0 ± 1.3 for the *stethoscope sonification* and 3.0 ± 1.3 for the *morph sonification*.

Fig. 12 Bar plot experts/non-experts classification task task

cation that obtained the lowest total error rates across all sonifications designs. We found a significant difference in the detection task using the *morph sonification*, $t(40) = 3.40$, $P = 0.003$, $P < 0.01$, where experts obtained lower error rates than non-experts.

In terms of the classification task although overall lower error scores are achieved by the experts group, the difference with respect to the non-experts group is lower than in the detection task as it can be seen in Fig. 12. We found a significant difference using the *stethoscope sonification* in the classification task $t(40) = 2.64$, $P = 0.017$, $P < 0.05$

8 Discussion

This work presents the design and evaluation of four sonification methods intended to support monitoring and diagnosis of ST-elevation in ECG signals. Results from the classification task show the highest increase in detection performance with the *Polarity* sonification suggesting that a design that combines features such as pitch and loudness can make more salient variations within the ECG waves. In terms of the *water ambience* design, participants tend to overestimate the ST-elevation and regard as pathological a higher number of heartbeats, this is a consequence of the mapping, in some examples even though the elevation in the ST-segment is not yet pathological according to the medical standard, the proposed mapping results in more than one water drop triggered every heartbeat. Setting a fixed threshold in which no drops are triggered before reaching the pathological value could improve ST-elevation estimation with this design. The *Morph* and *stethoscope* designs produce similar results concerning underestimation of the ST-elevation. Considering that both methods rely mainly on pitch variations to detect ST-elevations, it is likely that participants need a more extensive training to improve their performance in the detection.

Concerning the classification task, Fig. 8 shows that using the *water ambience*, *morph* and *stethoscope* sonification designs participants obtained higher classification accuracy for signals limiting the lower and upper ST-elevation boundary (isoelectric or severe ST-elevation). On the contrary, middle ST-elevation values were harder to classify and therefore had the highest error rates. As for the *polarity sonification*, classification accuracy is rather similar across ST-elevation values, except for a minor increase in error rates when the amplitude in the ST-segment is close to isoelectricity. Interestingly, the performance achieved with each design in the detection task does not necessarily match the performance in the classification task. This suggests that some designs are better at marking transitions between healthy and unhealthy states, but are not equally suitable to convey an absolute ST-elevation. This is the case with the *polarity* sonification, which obtained the lowest error in the detection task but the highest error values across the classification task over all data, i.e., including the low and high ST-elevation examples.

Comparison between sound/music experts and non-experts show significant differences in the detection task using the *morph sonification* and in the classification task using the *stethoscope sonification*. Overall, performance in the classification and detection tasks was higher in the group of experts, which indicates that error rates can be decreased with further training.

As to the aesthetics evaluation, participants gave higher ratings to the *water ambience* sonification with regard to pleasantness and long-term listening suitability whereas the *morph* and the *polarity* design were rated with the lowest scores respectively. Regarding how informative a particular design was considered, again the *water ambience* scored higher, although in practice this was not true. The *morph* design was rated as the least informative.

After analysis of the tracked volume changes explained in Sec. 6.1 we found that participants mainly preserved the chosen volume during the calibration phase. During the classification task participants did not make changes to the volume. As for the detection task, the volume changes correspond to less than two percent of the total audio files presented. Hence, the sonification design did not affect the volume selected by participants.

Results show that not only one sonification design has the highest performance across all evaluated conditions. The task then remains to find middle points where informativeness and aesthetics meet. If we want to contribute to diminishing ear fatigue in the medical environments, it is important to consider sonifications that don't saturate the already existing soundscape but can actually contribute to monitoring and diagnosis tasks.

Our research shows promising results as participants could discern between healthy and unhealthy signals without any

medical background and after having only short training sessions. This is of great importance because for patients with acute myocardial infarction the key element is that physicians detect the ST-elevation as soon as possible.

Lastly, results obtained with the *stethoscope sonification* indicate an encouraging panorama where a combined sonification of ECG signals and stethoscope recordings could be used to provide further insight about the heart's functioning.

9 Conclusion and future work

In summary we contributed two new sonification methods for ECG sonification with a focus on assessing the ST-elevation for an diagnostic and monitoring setting, and compared four sonification approaches in an abnormality detection and classification task. We conclude that precision (i.e. the transparency and discriminability of information) and aesthetics (i.e. the acceptability of sounds as part of the auditory scene) have both to be considered and optimized when designing sonifications for practical use. The results of a online user study with 41 participants are: (i) No sonification method in particular performed better across all tasks, however, methods that depend mainly on pitch changes obtained higher detection/classification error rates among participants that have no music/sound experience. (ii) Performance achieved with the *water ambience sonification* indicates that designs that implement sounds which are rated as more pleasant (e.g. sounds of nature) could provide an interesting and effective variant to the synthesized sounds which are normally used in medical context and are linked to ear fatigue. (iii) Study results are promising for sonification research in the field of medicine as they show that even with short training sessions and no previous ECG interpretation experience, participants are able to differentiate between healthy and severe ST-elevated signals, which is an encouraging starting point for sonification to become a supporting tool in the diagnostic and monitoring of cardiac pathologies.

Based on these results there are several prospects for future research. First, we intend to continue developing combined ECG and stethoscope sonifications and evaluate how a combined method could provide further insight about the state of the heart. Moreover, we plan to extend our current sonification methods to include multi-channel ECG signals. Finally, we plan to further evaluate classification accuracy of ECG signals, also including the standard visual representation of the ECG.

References

1. Al-Kindi S, Tafrershi R (2011) Real-time detection of myocardial infarction by evaluation of ST-segment in digital ECG. J Med Imaging Health Inf 1:225–230. https://doi.org/10.1166/jmihi.2011.1032
2. Aldana Blanco AL, Steffen G, Thomas H (2016) Heart Alert: ECG Sonification for supporting the detection and diagnosis of ST segment deviations. In: Proceedings of Interactive Sonification Workshop (ISon). Bielefeld, Germany
3. Aldana Blanco AL, Steffen G, Thomas H (2017) CardioSounds: real-time auditory assistance for supporting cardiac diagnostic and monitoring. In: Proceedings of the 12th International Audio Mostly Conference on Augmented and Participatory Sound and Music Experiences. ACM, New York, USA, pp 1–4. ISBN: 978-1-4503-5373-1. https://doi.org/10.1145/3123514.3123542
4. Ballora M, Pennycook B, Ivanov PC, Glass L, Goldberger AL (2004) Heart rate sonification: a new approach to medical diagnosis. Leonardo 37(1):41–46. https://doi.org/10.1162/002409404772828094
5. Barrass S, Vickers P (2011) Sonification design and aesthetics. In: Hermann T, Hunt A, Neuhoff JG (eds) The Sonification Handbook, chapter 7, pp 145–171. Logos Publishing House, Berlin, Germany. http://sonification.de/handbook/chapters/chapter7/
6. Chandra D, Tessler MJ, Usher J (2006) Audio spectrum and sound pressure levels vary between pulse oximeters. Can J Anesth 53(1):26–32. https://doi.org/10.1007/BF03021524
7. Clifford GD, Azuaje F, Mcsharry P (2006) ECG statistics, noise, artifacts, and missing data. Adv Methods Tools ECG Data Anal 6:18
8. Di Diego J, Antzelevitch C (2014) Acute myocardial ischemia: cellular mechanisms underlying ST segment elevation. J Electrocardiol. https://doi.org/10.1016/j.jelectrocard.2014.02.005
9. Gionfrida L, Roginska A (2017) A novel sonification approach to support the diagnosis of Alzheimer's dementia. Front Neurol 8:647. https://doi.org/10.3389/fneur.2017.00647
10. Hannibal GB (2014) Interpretation of the low-voltage ECG. AACN Adv Crit Care 25(1):64–68. https://doi.org/10.1097/NCI.0000000000000001
11. Hermann T, Hildebrandt T, Langeslag P, Rinderle-Ma S (2015) Optimizing aesthetics and precision in sonification for peripheral process-monitoring. In: Proceedings of the 21st international conference on auditory display (ICAD-2015). international community for auditory display (ICAD)
12. Hildebrandt T, Hermann T, Rinderle-Ma S (2016) Continuous sonification enhances adequacy of interactions in peripheral process monitoring. Int J Hum Comput Stud 95:54–65. https://doi.org/10.1016/j.ijhcs.2016.06.002

13. Imazio M, Gaita F (2015) Diagnosis and treatment of pericarditis. Heart 101(14):1159–1168. https://doi.org/10.1136/heartjnl-2014-306362

14. Kastrati A, Caforio ALP, Bucciarelli-Ducci C, Varenhorst C, Prescott E, Crea F, Hindricks G, Bueno H, Goudevenos JA, Antunes MJ, Roffi M, Valgimigli M, Lenzen MJ, Vranckx P, Widimský P, Halvorsen S, Agewall S, Ibanez B, James S, Group ESD (2017) 2017 ESC Guidelines for themanagement of acute myocardial infarction in patients presentingwith ST-segment elevation: The Task Force for the management ofacute myocardial infarction in patients presenting with ST-segmentelevation of the European Society of Cardiology (ESC). Eur Heart J 39(2):119–177. https://doi.org/10.1093/eurheartj/ehx393

15. Kather JN, Hermann T, Bukschat Y, Kramer T, Schad LR, Zöllner FG (2017) Polyphonic sonification of electrocardiography signals for diagnosis of cardiac pathologies. Sci Rep 7:44549. https://doi.org/10.1038/srep44549

16. Klabunde RE (2017) Cardiac electrophysiology: normal and ischemic ionic currents and the ECG. Adv Physiol Educ 41(1):29–37. https://doi.org/10.1152/advan.00105.2016 PMID: 28143820

17. McSharry PE, Clifford GD, Tarassenko L, Smith LA (2003) A dynamical model for generating synthetic electrocardiogram signals. IEEE Trans Biomed Eng 50(3):289–294. https://doi.org/10.1109/TBME.2003.808805

18. Moore B (2007) Psychoacoustics. Springer, New York, pp 459–501

19. Pauletto S, Hunt A (2006) The sonification of EMG data. In: Proceedings of the 12th international conference on auditory display, pp 152–157

20. Roginska A, Mohanraj H, Keary J, Friedman K (2014) Sonification method to enhance the diagnosis of dementia. In: International conference on auditory display

21. Solet JM, Barach PR (2012) Managing alarm fatigue in cardiac care. Prog. Pediatr. Cardiol. **33**(1):85–90 (2012).https://doi.org/10.1016/j.ppedcard.2011.12.014. The Future of Pediatric and Congenital Cardiac Care (Part 2)

22. Stahl B, Thoshkahna B (2015) Real-time heart rate sonification for athletes. In: Proceedings of the 21st international conference on auditory display. University of Music and Performing Arts Graz. Institute of Electronic Music and Acoustics

23. Terasawa H, Morimoto Y, Matsubara M, Sato A, Ohara M, Kawarasaki M (2015) Guiding auditory attention toward the subtle components in electrocardiography sonification. In: Proc. ICAD. Georgia Institute of Technology

24. Vickers P (2011) Sonification for process monitoring. In: T. Hermann, A. Hunt, J.G. Neuhoff (eds.) The Sonification Handbook, chap. 18, pp. 455–491. Logos Publishing House, Berlin, Germany .http://sonification.de/handbook/chapters/chapter18/

25. Vickers P, Laing C, Debashi M, Fairfax T (2014) Sonification aesthetics and listening for network situational awareness. CoRR. http://arxiv.org/abs/1409.5282

26. Väljamäe A, Mealla S, Steffert T, Holland S, Marimon X, Benitez R, Oliveira A, Jordà S (2013) Väljamäe-steffert-2013-icad. A review of real-time EEG sonification research

27. Worrall D, Thoshkahna B, Degara N (2014) Detecting components of an ECG signal for sonification. In: Proceedings of the ICAD. Georgia Institute of Technology