Sonification enables continuous surveillance of the ST segment in the electrocardiogram

Citation for published version:

Digital Object Identifier (DOI):
10.1016/j.ajem.2022.05.016

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Peer reviewed version

Published In:
The American Journal of Emergency Medicine

General rights
Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.
Sonification enables continuous surveillance of the ST segment in the electrocardiogram

Andrea Lorena, Aldana Blanco¹, Thomas, Hermann¹, Jens, Tiesmeier², Jan, Persson³, Steffen, Grautoff⁴, Andrea L. Aldana Blanco, Steffen Grautoff, Jens Tiesmeier, Jan Persson, Thomas Hermann⁶

¹CITEC, Bielefeld University, Bielefeld, Germany
²Institute for Anesthesiology, Intensive Care- and Emergency Medicine, MKK-Hospital Luebecke, Campus OWL, Ruhr-University Bochum, Germany
³Institute for Anesthesiology, Intensive Care- and Emergency Medicine, MKK-Johannes Wesling Hospital Minden, Campus OWL, Ruhr-University Bochum, Germany
⁴Emergency Medical Services, District of Herford, Germany, and Emergency Department, Herford Hospital, Campus OWL, Ruhr-University Bochum, Germany

Abstract

Introduction ST segment elevation myocardial infarction is a common reason for out-of-hospital cardiac arrest in adult patients. The surveillance of the ST segment in the electrocardiogram is limited to visual presentation. However, the ST segment can change during the course of treatment. If ST elevation is present immediate coronary revascularization is needed, therefore detecting ST elevation changes the treatment fundamentally. Sonification of the ST segment is a new method which enables the emergency team to detect intermediate changes of the ST segment.

Material and methods We have chosen two sonification designs which were introduced to two groups, medical students and computer science students. Twenty-one participants took part in the study. The sonification was designed for evaluation of the ST segment. The user was supposed to become empowered to distinguish between no, medium–low, medium–high or extreme ST elevation by listening to the sonification. The two groups were asked to evaluate the sounds for possible ST elevation as well as for aesthetics and usability. In a second study twenty-five medical students were taking part in a medical scenario in which sonification was played during a simulated case. The patient was suffering from a myocardial infarction, ST elevation was transient and sonification sounds were changing appropriately. The students were supposed to detect these changes and act accordingly by modifying the treatment.

Results Both groups were able to classify ST segment elevation by listening to the sonification samples. The higher the ST segment, the better was the detection rate overall. In all of the three categories (pleasantness, informativeness and long-term listening) the Water Ambience sonification was rated higher compared to the Polarity sonification. Moreover, in the two groups that took part in the study, we found a significant difference when comparing classification performance using both sonification designs. For the group of medical students as $t(20) = 4.31$, $p = 3.44 \times 10^{-4}$, $p < 0.01$ and for the computer science students as $t(19) = 3.40$, $p = 9.39 \times 10^{-6}$, $p < 0.01$. In the simulated medical scenario participants indicated that 96% detected the ST elevation. 60% stated that sonification played a role whereas for 32% it did not play a role for the detection of ST elevation.

Conclusions Sonification has the potential to play an important role as a new supporting tool for the surveillance of the ST segment during the care of patients with suspicion of myocardial infarction. It
can be helpful to differentiate between ST segment elevation myocardial infarction and non-ST segment myocardial infarction especially if ST elevation is transient. Furthermore, sonification is viewed as pleasant to listen to and might not contribute to alarm fatigue.

Keywords: ECG, STEMI, sonification, monitoring, out-of-hospital cardiac arrest.

1 Introduction

Among emergency physicians it is consensus that fatal cardiac arrhythmias due to coronary heart disease are the most common reason for out-of-hospital cardiac arrest (OHCA) in adult patients. ST segment elevation myocardial infarction (STEMI) is mentioned as the most common reason for OHCA [1], [2].

Furthermore, the ERC guidelines 2021 state that parameter which support the suspicion of coronary thrombosis should be recorded [3].

The main tool to determine this diagnosis is the 12-lead electrocardiogram (ECG). Maximum permissible values of ST elevation are differing in between leads of the ECG and are dependent on age and sex of the patient according to the European Society of Cardiology (ESC) [4].

In time-critical emergencies it can be very challenging to assess the ECG accordingly adjusted to the criteria mentioned above. This is especially true if these differing thresholds are not known in detail and/or poor quality of recording is present. In these situations assessing millimeter changes makes a precise differentiation even more difficult.

<table>
<thead>
<tr>
<th>sex</th>
<th>age [y]</th>
<th>condition</th>
</tr>
</thead>
<tbody>
<tr>
<td>male</td>
<td>&lt; 40 years</td>
<td>at least two contiguous leads with ST-segment elevation ≥ 0.25 mV (2.5 mm)</td>
</tr>
<tr>
<td></td>
<td>≥ 40 years</td>
<td>at least two contiguous leads with ST-segment elevation ≥ 0.2 mV (2 mm)</td>
</tr>
<tr>
<td>female</td>
<td>any</td>
<td>≥ 0.15 mV (1.5 mm) in leads V2-V3 and/or ≥ 0.1 mV (1 mm) in the other leads</td>
</tr>
</tbody>
</table>

Table 1: Criteria for ST elevation values for STEMI diagnosis – age and sex dependency.

Additionally, it has to be mentioned that post-arrest ECG is of limited value for diagnosing acute coronary occlusion [5], [6]. However, the ST segment is subject to dynamic changes over the course of treatment. It is possible that in one patient, criteria of STEMI are met on ECG although they have not been met a few minutes earlier. ECGs which are repeatedly recorded can be of great value to detect a transient STEMI. In ambiguous cases, serial ECGs should be routinely recorded, especially if clinical symptoms are changing.

The emergency team has to focus on different tasks during patient care. Therefore, it is not possible to constantly monitor the ST segment visually and detect possible changes. In clinical practice only four of twelve leads are continuously monitored. If STEMI is documented, treatment has to be modified immediately. The patient needs to get a coronary angiogram as a procedure.

The emergency team has to process a lot of visual data during the treatment of the patient. Therefore new visual information may not be as easily processed. Actually, standard transformation of ECG signals are only visually available. Prioritizing the focus on the ECG could lead to decrease of attention on other visually received information.

If myocardial infarction is suspected, the surveillance of the ST segment is of major interest to differentiate between STEMI and non-STEMI. In the standard setting a continuous monitoring of the ST segment is difficult to provide. This monitoring gap should be closed, because it can improve patient care in the case of suspicion of myocardial infarction.
In this scenario, a different approach to monitoring, which does not require to visually focus on a given display to detect important changes, could support the timely detection of STEMI: Sonification is an option as a biosignal monitoring substitute.

Sonification is the systematic and reproducible transformation of data into audible sound for the purpose of conveying information [7]. As a matter of fact, sonification is already used in medical scenarios to monitor oxygen saturation levels as in the case of the pulse oximeter, or in cardiac monitors to provide feedback about the heart rate. Modern sonification, however, enables to create more complex sonic representations that convey much more details of a signal than alarms and warning sounds could do.

The following case report emphasizes the significance of sonification: A 47-year-old male patient had called EMS for right-sided chest pain, which he classified as 8-9 on a 10-point-scale. On arrival, the pain had reduced to 3/10 and was further waning. An ECG was obtained (Figure 1 (a)), but no significant ST segment elevation was detectable.

Although pain localization was atypical for acute coronary syndrome and pain was resolving during the course, the EMS team decided to transport the patient to a hospital with capability of acute coronary intervention. During transport right-sided pain recurred and then additionally radiated to the neck. Another ECG was recorded (cf. Fig. 1 (b)). Inferior ST segment elevation was detectable now.

The patient received pain medication (morphine), acetylsalicylic acid, and heparin. Over the next minutes, the pain resolved again. Another ECG was recorded (cf. Fig. 1 (c)) upon arrival at the hospital. ST elevation was not detectable anymore. The T wave had then become negative in lead III.

An 85% stenosis of the RCA was found and treated appropriately. The ST elevation which prompted further urgent treatment was detected only because of the EMS team obtaining serial ECGs by making extra effort. The patients’ recurrence of symptoms was recognized. If sonification had been available for continuous surveillance of the ST segment, it would have been likely to catch this brief moment of ST elevation more reliably and in a more convenient way.

To overcome the previously explained limitations of visual monitoring in the detection of STEMI in the ECG, sonification could be a well suited supporting element, as it can provide surveillance of all parts of the ECG, especially of the ST segment which is the crucial part for diagnosing STEMI. The current ERC guidelines urge scientists to work on research “developed and stimulated by ideas” [3]. Thus, this novel method might be a possibility to improve surveillance of patients with suspicion of STEMI.

To evaluate to what extent sonification could serve as a supporting tool in medical scenarios, and particularly in the detection of STEMI, we conducted a set of studies involving medical students. In the first study, participants were asked to classify a set of sonification examples among four categories that ranged from isoelectric ST segment to severe ST elevation. Moreover, they also evaluated the sonifications in terms of their aesthetic components. In order to evaluate whether medical expertise had an impact on classification performance, we compared the results obtained by medical students to those obtained by a group of computer science students who also took part in the first study. Furthermore, to gather further insight into the scope and practical issues and limitations of ECG sonification in medical settings, we conducted a second study in a simulated medical scenario involving only medical students as participants.

In this paper we first provide an overview about ECG sonification. We then present the user studies designs and introduce a number of methods for continuous ECG sonification. Finally, we present and discuss the results.

To our knowledge this is the first study to examine whether sonification can be suitable for surveillance and detection of transient ST segment elevation in the ECG.

2 Materials and methods

2.1 Sonification of ECG signals

Even though sonification is already used in medical scenarios for monitoring purposes (i.e., pulse oximeter, cardiac monitor), in recent years, researchers have further explored how sonifying other types of biosignals
could support diagnosis and monitoring processes. One of these rather novel areas of research is ECG sonification.

In general terms, research lines involving ECG sonification can be divided into two groups. On the one hand, there are methods that focus on sonifying temporal features such as the heart rate and the heart rate variability, either in medical scenarios [8], [9] or as an element to track and increase sports performance [10], [11]. On the other hand, there are research efforts that target morphology changes in the ECG signal, and thus, intend to sonify variations in the ECG signal that lead to a timely detection of pathological signs through auditory cues [12], [13].

One of the reasons why sonification is such a strong candidate for diagnosis and monitoring, is because the human auditory system has powerful capabilities in terms of its temporal and frequency [14]–[17] resolution. These auditory talents allow us to easily detect patterns and changes in auditory signals. Moreover, such hearing capabilities extend remarkably to (noisy) environments in which there are multiple sound streams present.

Because we are constantly connected to our environment via our auditory system, we are also exposed to a well-known problem in medical environments called noise fatigue. Due to a large amount of alarm sounds present, the medical staff is exposed to a higher cognitive load that can cause delayed responses or alarm desensitization. Even more, noisy environments can increase mental distress and irritability in the medical staff [18], [19]. Furthermore, noise pollution in hospitals is also a concerning factor for the well-being and recovery process of patients [20].

Even though our auditory abilities make ECG sonification a strong candidate to support diagnosis and monitoring of cardiac pathologies, it is also very important to account for noise fatigue when designing...
sonifications. The goal is that new sounds incorporated in medical environments serve the purpose of being supporting tools without contributing to noise fatigue. In other words, we want the sounds to be pleasant to listen, informative, and suitable for long-term listening.

2.2 First study: Classification task

To prepare the data set for the user study we created a set of surrogate signals. To do so, we used the ECG waveform generator\(^1\) from the Department of Engineering Science in University of Oxford and the Laboratory for Computational Physiology at the MIT \([21]\). Using surrogate signals allowed us to have a better control over the ECG features we wanted to evaluate, specifically, the ST segment amplitude. Further information about the surrogate signals and the ECG features extraction process, can be found in the Appendix A.

2.2.1 Sonification designs

In our previous user studies, sonification designs (that focus on STEMI) were evaluated in terms of aesthetics components and classification performance reached by participants \([22]\), \([23]\). In this occasion, we selected the two most representative sonification designs according to the results obtained in previous user studies. The first (1) design we selected is called Polarity sonification. We chose this sonification because it is a design that achieved high performance accuracy in the detection of transitions from healthy to pathological conditions \([23]\). The second (2) design is the Water ambience sonification. We chose this design because it is the one that received higher scores in terms of pleasantness, informativeness and suitability for long-term listening.

Before further explaining the sonifications designs, it is worth mentioning that both designs correspond to a sonification technique known as parameter mapping sonification \([24]\). In this technique, a set of features are extracted from the data and mapped to sound synthesis parameters. The mapping function linking data and its auditory representation is defined by the sonification designer according to the given use case. Moreover, it is possible to combine different sonification techniques to broaden the palette of sounds. A more extensive overview about sonification techniques can be found in \([25]\).

**Polarity sonification** The pillar of the Polarity sonification design is the mapping of the absolute voltage difference of the ECG signal to the amplitude and number of harmonics of a formant oscillator\(^2\). More specifically, we mapped the voltage to the fundamental frequency of the oscillator. As a result, the higher the voltage in the ECG signal, the higher the pitch of the resulting sound. As part of the pre-processing steps, we first added a band-pass filter in the ECG signal to remove the low-frequency respiration component. Moreover, to determine the isoelectric reference in the ECG signal, we calculated the arithmetic mean considering all samples in the ECG signal.

Additionally, we used the sign of the slope to control the panning in the stereo signal. The result is that the sound shifts to the right audio channel when the slope is positive and it shifts to the left channel when it is negative. The absolute voltage difference at \(t_i\) is given by \(v = |v_i - \mu|\), where \(\mu\) is the signal mean. The parameter-mapping used in the Polarity sonification design is shown in Table 2. Sonification examples are included as part of the supplementary material. Sonification example **Son1** corresponds to the Polarity sonification of a dataset exhibiting a healthy condition, whereas, sound example **Son2** represents an ST elevated signal.

**Water ambience sonification** The Water ambience sonification design is a mixture of two sonification techniques: Parameter-mapping Sonification and Auditory Icons. \([26]\). Similar to icons in the visual

---

1\footnote{https://physionet.org/physiotools/ecgsyn/}

2\footnote{A Formant oscillator generates a number of harmonics in relation to a fundamental frequency, in particular, by boosting harmonics around a chosen formant frequency.}
world, auditory icons are sounds that we can easily associate to a real-world element. For instance, when taking a photo with a smartphone, or a screenshot on the computer, the shutter sound of an analogue camera is played as we associate ‘something captured’ with it.

In the Water ambience design, likewise we represent higher ST levels by a higher number of drop sounds, trusting the intuition that increasing dripping sounds are associated with a problem and drops are associated with fluid, here specifically the blood circulating in the body. This also considers the aspect of noise fatigue, thus providing a noise-free environment in cases of healthy situations.

We determine the drop sound onsets within each heartbeat by linearly spacing the period of the heartbeat $T$ according to the number of drops to trigger. However, we consider a shorter heartbeat’s period given by $T - \varepsilon$ to safeguard temporal cues that reveal the heart rate. Hence, the last water drop that represents the ST segment level of a given heartbeat is triggered before the first drop in the next heartbeat. As part of the design, and in order to consider the stochastic characteristics of real water sounds, the linearly spaced water drop onsets within each heartbeat include a temporal jittering of up to 50 ms. Finally, we attenuated each drop sound within the drops of a single heart beat by a few dB, resulting in a level decay which facilitates heart rate perception even in the presence of many drops, aka high ST elevation. Table 3, shows the parameter-mapping for the Water Ambience sonification.

Table 3: Water ambience sonification: Parameter-mapping

<table>
<thead>
<tr>
<th>Data feature</th>
<th>Data Range (min, max)</th>
<th>Perceptual parameter</th>
<th>Parameter range (min, max)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ST segment amplitude</td>
<td>(0 mV, 0.4 mV)</td>
<td>number of drops</td>
<td>(1 drop, 5 drops)</td>
</tr>
</tbody>
</table>

Sonification examples Son3 and Son4 illustrate the Water ambience sonification of a healthy and ST elevated signal.

### 2.2.2 Study description

The user study consisted of two tasks: (i) a classification task, and (ii), a questionnaire to evaluate aesthetic components of the sonification designs. Before conducting the user study, we asked for ethical clearance from the ethical committee of Bielefeld University.

Two groups took part in the study: Medical students and computer science students.

**Medical experts** The first group were students from the Faculty of Medicine of the Bochum Ruhr-University (RUB) in Germany in the 9th semester of their studies. From now on, we refer to this group as the medical experts group. The day the group of medicine students took part in the study, they had a full-day training program for simulating cases in emergency medicine. All of them received an introduction to sonification and ECG sonification in particular.

---

3Supplementary material available at: tiny.cc/ssouz
Computer science students  The second group consisted of students from the computer science program at Bielefeld University that at the time of the study attended a lecture in which sonification (including ECG sonification) was taught.

Before collecting the data, we received a positive vote of the local ethics committee from Bielefeld University.

Introductory talk   Before starting the study, we gave participants a short (<15 min) introduction about sonification and the motivation of our research. For the group of medical experts, sonification was a new concept that they had heard for the first time during the introduction.

All participants belonging to one group (i.e., medical experts, computer science students) took part in the study at the same time. In both cases the study was conducted in lecture or meeting rooms with good noise level conditions.

As loudspeakers to present the sonification examples, we used a set of Genelec 8020 studio monitors⁴ to ensure that the sounds played with high sound quality and volume.

2.2.3 Classification task

The first part of the study involved a classification task in which we asked participants to evaluate sonification examples of ten seconds duration. They classified the samples from healthy to severe ST elevation using a 7-point scale, one (1) representing an isoelectric ST segment and seven (7) the highest level of ST elevation of our protocol.

First, participants evaluated the Polarity sonification. As a first step, they listened to training examples of a healthy signal (isoelectric ST segment) and then of an ECG exhibiting ST elevation. Afterwards, they started the classification task which was composed of three sets of sonifications. Each set consisted of the same audio examples but was presented in random order.

In each set participants evaluated eight sonification examples. After participants finished evaluating the Polarity sonification, they started evaluating the Water ambience design. Once again, initially they listened to the training examples and then proceeded with the classification task.

Figure 2 presents the ST elevation levels as a function of time from the surrogate data used in the classification task. Each color represents a level of ST elevation that ranged from healthy (blue dotted line) to the highest level of ST elevation (black dotted line). It can be noticed that every ST segment elevation cluster under evaluation (i.e., healthy, medium-low, medium-high, extreme) is formed by two examples. Moreover, variations in ST segment levels of the same surrogate example are a constraint of the ECG waveform generator model used.

2.2.4 Aesthetics and usability evaluation

To evaluate aesthetics and usability of the sonifications, we asked participants to rate on a 6-point Likert scale where one (1) meant Strongly Disagree and six (6) meant Strongly Agree their level of agreement to statements regarding pleasantness of the sonification designs, suitability for long-term listening and informativeness. To have an agreement about what informativeness meant in the context of the study, we explained participants that the sonification was informative if it allowed them to detect changes in the ST segment elevation.

As a last step, participants could add any comments that they considered relevant.

2.3 Second study: Simulated medical scenario

In the above described first study, participants had to complete a given task but they did not had to account for other stimuli that could affect their performance. However, that situation differs from the usual medical

⁴https://www.genelec.com/8020d
scenario, in which there are several sources competing for attention. As a result, it can be expected that it is increasingly difficult to detect subtle ECG changes of the ST segment.

For this reason, we conducted a second study in which we evaluated one of the sonifications introduced above in a simulated medical scenario. The motivation was to get closer to a more realistic situation to assess the utility of sonification for supporting medical professionals during surveillance of patients with suspicion of myocardial infarction.

2.3.1 Study description

The study was carried out as part of a collaboration with Herford Hospital in Germany and the Ruhr-Universität Bochum (RUB). Participants were medical students that attended a one-day training course as part of a lecture in emergency medicine. Over the day, they had to take part in different training scenarios; one of them being the ECG sonification setup. Before they visited the scenarios, they had an introductory session: We explained the term sonification and the background on why using sonification can improve ECG monitoring. In the introductory session, they listened to examples of ECGs with isoelectric ST segments and ST elevated ECG sonifications corresponding to the Water ambience sonification design.

Medical scenario description The medical scenario we proposed, involved a sixty-six years old female patient who presented a series of symptoms that required attention from the emergency medical service.

Figure 2: ST elevation levels of surrogate signals
The emergency teams were formed by the medical students who took part in the study. Only one team at a time took part in the scenario. All teams were formed by two persons, except for one team in which there was only one participant.

Before entering the scenario, the teams were instructed about the patient’s symptoms that originated the emergency call. These were: nausea, fatigue and light thoracic burning. Moreover, they also received basic information about the patient (i.e., name, sex, age).

Afterwards, they entered the scenario and started interacting with the patient. Once the team started attending the emergency at the patient’s house, they found out that her symptoms had temporarily improved. However, at some point during the scenario the symptoms intensified again. Overall, participants had a maximum of forty-five minutes to complete the scenario.

The patient was portrayed by the emergency medicine training dummy depicted in Figure 3. Participants interacted throughout the scenario with the dummy, but the patient’s voice was interpreted by one of our medical instructors.

The symptoms When the team entered the room, the patient started explaining the symptoms that originated the emergency call. Moreover, she also provided information about her medical history. The information given by the patient was as follows:

- History of arterial hypertension and long-term insulin-dependent type 2 diabetes.
- Taking medication: Candesartan 8 mg, Metformin 850 mg and insulin.
- Father died of myocardial infarction at the age of 70.
**Medical assessment**  In order to properly fulfill the scenario’s goals, the emergency team had to perform a series of tasks:

- Recording a 12 lead ECG.
- Recognizing the ST elevation (myocardial infarction).
- Correct selection and application of the indicated drugs.
- Referring the patient to the nearest appropriate hospital with the option of immediate cardiac catheterization.

Once the team entered the room and asked for her medical history, they immediately started measuring vital parameters such as oxygen levels, blood pressure, temperature, etc. Furthermore, they also recorded an ECG as part of the primary assessment. At the precise moment the ECG signal started on the visual display (i.e. after properly attaching all electrodes), the sonification started playing, corresponding to a healthy ECG with isoelectric ST segment. Though being manually triggered from one of the medical instructors, for the participants it created the illusion that the ECG monitor actually featured the sonification-based real-time auditory display.

The emergency equipment used by the team can be observed in Figure 4. It consisted of a medical emergency suitcase and a patient monitor/defibrillator device.

![Figure 4: Emergency equipment.](image)

**Beginning of ST elevation**  A few minutes into the medical assessment, the patient expressed that the symptoms which originated the emergency call were returning. In particular, she stated to have chest pain. At that moment, we simultaneously triggered (1), an ST elevated ECG signal on the cardiac monitor and (2) the change of the sonification into a sonification corresponding to an ST elevated ECG.

After those new events took place, we expected the medical team to recognize the ST elevation and refer the patient to an appropriate hospital.

---

5The patient monitor had a remote control that could be used to trigger ECG signals depicting different pathologies. One of them being an ST elevation.
2.3.2 Questionnaire

After finishing the simulated medical scenario, we asked participants to answer a short questionnaire. First, we collected basic demographic data (i.e., age, gender) and inquired experience in terms of music/sound.

In the second part of the questionnaire, we asked participants to evaluate using a 6-point Likert-scale where one (1) meant *Strongly Disagree* and six (6) meant *Strongly Agree* their level of agreement to the following statements:

1. The sonification is pleasant to listen to.
2. The sonification is informative, meaning that it facilitates the recognition of ST elevation changes in the ECG signal.
3. I could imagine listening to the sonification for a longer period of time.
4. I could imagine to use ECG sonification as a supporting instrument in daily practice.

Furthermore, we asked them to self-assess if they noticed the ST elevation and to provide information of whether the sonification played a role in the detection of the ST elevation. Finally, participants could add any comment they considered important. After finishing the questionnaire, we had an informal chat in which they could further comment about their overall experience with the proposed medical scenario and ECG sonification.

2.3.3 Sonification design

We decided to work with a single sonification design to avoid confusion among the participants: *Water ambience*. We chose this sonification because from our proposed designs, it received the highest scores in terms of aesthetic components. Additionally, we wanted to test a sonification design that differed from the alarm sounds that are normally heard in medical environments. Furthermore, it is important to consider that the proposed scenario was already very demanding and stressful, thus, selecting a sonification design that was regarded as unpleasant to listen, would have been counterproductive.

**QRS tone** To give participants a familiar auditory reference to detect the heart rate, we played the sonifications together with a QRS tone from the ECG monitoring system they were used to. The QRS tone was a short duration sound made using a sinusoidal waveform with a fundamental frequency of 400 Hz.

3 Results

3.1 Classification study results

We had twenty-one participants in the group of medical experts and twenty participants in the computer science group. In the results section, we refer to the group of medical experts as *MedS*, and to the computer science students as *CSS*. Information about demographic data collected can be found in the Appendix.

3.1.1 Classification task results

To evaluate the results from the classification task, we calculated the error between the ST elevation levels selected by participants and the levels from the surrogate signals (see Figure 2). However, we first did a linear mapping from the ST elevation scale values (1 to 7) to the surrogate ST elevation levels (approx. 0 mV to 0.43 mV). Moreover, taking into account that the ST elevation scale only takes discrete
values from 1 to 7, the resulting mapping is an approximation. Afterwards, we calculated the root mean squared error (RMSE) between the score assigned by participants and the expected value (considering the ST segment values measured from the surrogate data). We followed this process for each sonification example.

**Unevaluated sonification examples** We previously mentioned that the classification task included three sets of audio examples. Nonetheless, because of time constraints, the CSS group only evaluated two sets.

Furthermore, we noticed that in the MedS group, in few cases participants forgot to evaluate a given sonification example. Thus, from a total of 504 audios to be evaluated (21 participants $\times$ 24 sonifications), 4 audio examples were left unmarked.

In the CSS group, participants rated all audio examples. However, one of the participants selected in between values of the ST elevation scale. For instance, instead of crossing the number three or four, the participant put a mark in between those numbers, although this was not a valid option in the study. This indicates that either the participant did not understand the instructions given at the beginning, or was undecided about the ST elevation level of the presented sonifications. Results that we present next, exclude unrated audios and ratings given by the participant that did not complete the task according to the provided instructions.

Figures 5 and 6 present the error rates obtained by the MedS group and the CSS group. There is an error value corresponding to each evaluated sonification example. Star-shaped markers indicate the RMSE obtained in the group of MedS, while the triangle-shaped markers present the results of the CSS group.

![Figure 5: Classification task results for the MedS group.](image)

Overall, results reveal that lower error rates are achieved using the *Water ambience* sonification. In the CSS group, error rates are higher in the middle ranges of ST elevation (mid-low and mid-high). This indicates that participants can better detect extreme cases: healthy and extreme ST elevation. A similar trend can be observed in the MedS group, except for the particularity that higher error rates are found in the detection of extreme ST elevations.

To calculate the classification errors previously described, we considered data from all repetitions according to each group. Thus, in the MedS group we evaluated three classification sets whereas for the CSS we considered two sets. However, to have a comparable amount of repetitions between the two
When comparing classification performance among the two groups, we did not find a significant difference neither with the Water ambience design nor the Polarity sonification. This suggests that medical expertise did not have an effect in classification performance. We think this is a consequence of
the short exposure both groups had to ECG sonification, or sonification at all, before taking part in the study. In particular, in the case of medical experts, it is difficult to build associations between the ECG visual representation and the sonification only after the short training session they had at the beginning of the study.

3.1.2 Accuracy metrics

Besides estimating the error rates obtained by each group in the classification task, we also calculated the confusion matrix and determined the accuracy per group. To obtain the confusion matrix, we grouped all sonification files exhibiting ST elevation (mid-low, mid-high, extreme) under the ST elevated category and the healthy examples in a different category. The reason to divide the examples into these categories, is that in practice, the most critical element is to detect that there is an ST elevation as soon as possible rather than the actual level of elevation.

Table 4 show the confusion matrix for both groups after evaluating the Polarity sonification. Accuracy rates of both groups are similar, however, better classification performance is achieved by the CSS group.

<table>
<thead>
<tr>
<th>Table 4: Confusion matrix - Polarity sonification</th>
</tr>
</thead>
<tbody>
<tr>
<td>Polarity (MedS)</td>
</tr>
<tr>
<td>N = 336</td>
</tr>
<tr>
<td>Healthy</td>
</tr>
<tr>
<td>Actual Healthy</td>
</tr>
<tr>
<td>Actual ST-elevated</td>
</tr>
<tr>
<td>Accuracy</td>
</tr>
<tr>
<td>Polarity (CSS)</td>
</tr>
<tr>
<td>N = 313</td>
</tr>
<tr>
<td>Healthy</td>
</tr>
<tr>
<td>Actual Healthy</td>
</tr>
<tr>
<td>Actual ST-elevated</td>
</tr>
<tr>
<td>Accuracy</td>
</tr>
</tbody>
</table>

Furthermore, confusion matrices corresponding to the Water ambience sonification are shown in Table 5. It can be seen that accuracy rates are higher in comparison to the ones obtained with the Polarity design. Moreover, in this case the MedS group obtained better classification performance.

<table>
<thead>
<tr>
<th>Table 5: Confusion matrix - Water Ambience sonification</th>
</tr>
</thead>
<tbody>
<tr>
<td>Water ambience (MedS)</td>
</tr>
<tr>
<td>N = 335</td>
</tr>
<tr>
<td>Healthy</td>
</tr>
<tr>
<td>Actual Healthy</td>
</tr>
<tr>
<td>Actual ST-elevated</td>
</tr>
<tr>
<td>Accuracy</td>
</tr>
<tr>
<td>Water ambience (CSS)</td>
</tr>
<tr>
<td>N = 315</td>
</tr>
<tr>
<td>Healthy</td>
</tr>
<tr>
<td>Actual Healthy</td>
</tr>
<tr>
<td>Actual ST-elevated</td>
</tr>
<tr>
<td>Accuracy</td>
</tr>
</tbody>
</table>

Furthermore, we found a significant difference in the classification task after comparing the error rates obtained with both sonification designs in each of the groups. For the MedS group as $t (20) = 4.31, p =3.44 \times 10^{-4}, p < 0.01$ and for the CSS group as $t (19) = 3.40, p = 9.39 \times 10^{-6}, p < 0.01$. 

14
3.1.3 Aesthetics results

Results about aesthetics and suitability components of the sonifications, that evaluated pleasantness, informativeness and suitability for long-term listening, are presented in Figure 8. The figure on the left presents results obtained from the MedS group whereas the figure on the right exhibits results from the CSS group.

![Figure 8: Aesthetics and suitability. Mean and 95% CI. (left: MedS, right CSS)](image)

We found a significant difference among the two proposed sonification designs in terms of pleasantness. For the MedS group as $t(20) = 5.20, p = 6.35 \times 10^{-6}, p < 0.01$, and for the CSS group as $t(19) = 4.93, p = 1.70 \times 10^{-5}, p < 0.01$.

Next, we report the scores assigned by each group in the evaluation of aesthetic components as $(\bar{x} \pm \sigma)$ in a scale from 1 to 6 where 1 refers to *Strongly Disagree* and 6 to *Strongly agree*.

**MedS**  Pleasantness of the sonification, was rated 4.57 $\pm$ 1.12 for the *Water ambience* and 2.76 $\pm$ 1.14 for the *Polarity* design. Moreover, informativeness was rated as 3.95 $\pm$ 1.16 for *Water ambience* and 3.71 $\pm$ 1.49 for *Polarity*. Lastly, suitability for long-term listening obtained 3.38 $\pm$ 1.36 for *Water ambience* and 2.67 $\pm$ 1.35 for *Polarity*.

**CSS**  As for the CSS group, pleasantness was rated as 4.10 $\pm$ 1.25 for *Water ambience* and 2.25 $\pm$ 1.12 for *Polarity*. Informativeness as 4.30 $\pm$ 1.22 for *Water ambience* and 3.75 $\pm$ 1.16 for *Polarity*. Finally, suitability for long-term listening was rated as 3.45 $\pm$ 1.43 for the *Water ambience* and 2.25 $\pm$ 1.25 for the *Polarity* sonification.

3.2 Simulated medical scenario results

Twenty-five individuals took part in the simulated medical scenario. At the end of the scenario, all participants answered the questionnaire individually.

**Demographics and musical expertise**  As we previously mentioned, in the questionnaire we collected demographic data and asked participants if they were musically trained (to any level). A few of the participants decided not to answer the previously mentioned questions, however, we did collect responses from the majority of them. Further information about demographic data collected can be found in Appendix B.

3.2.1 Results of aesthetics and suitability of the sonification

Results evaluating aesthetics and suitability components of the sonification are presented in Figure 9. Neutral values are represented with lighter colors.

Overall, results indicate that participants evaluated positively aesthetic and suitability components of the *Water ambience* sonification. Furthermore, they showed a large percentage of agreement (80%) to the statement evaluating ECG sonification as a supportive instrument in medical assessment routines. This provides an encouraging panorama for ECG sonification, as results indicate that participants see a benefit in using sonification as a supporting tool in medical assessment.

3.2.2 Self-assessment and general comments about ECG sonification

Considering that participants attended the medical scenario as a team, it was also important to assess individual performance. For this reason, we included a self-assessment questionnaire to evaluate if
participants detected the ST elevation and to open a space of discussion about the role of sonification in ECG diagnosis.

Results from the self-assessment questionnaire indicate that 96% of the participants detected the ST elevation as part of the medical scenario. Hence, only one participant said to not have detected it. Moreover, when asked if sonification played any role in detecting ST elevation, 60% of the individuals stated it did play a role, 32% expressed that sonification did not play any role and 8% of participants were not sure whether it played a role.

**Additional general comments of the participants** Overall, the sonification played a very different role for participants. For instance, a group of eight (8) students expressed that they first recognized the ST elevation by looking on the ECG monitor. Among this group, three (3) participants also stated that even though they first noticed the ST elevation on the monitor, the sonification helped them to either confirm the diagnosis or to perform a faster diagnosis. Moreover, one of the participants expressed that not having disturbing noise from the commonly known QRS tone actually made the ECG detection easier, meaning that the fact that sonification differed from common alarm sounds in medical environments, favored the ST elevation detection process.

Additionally, two (2) participants expressed that they completely ignored the sounds as a result of the stressful situation. For instance, one of the participants mentioned that in the same way as it happens with the regular QRS tone of a cardiac monitor, the acoustic signal sort of vanished because he/she was so focused in solving the task. This is a common situation that results from alarm fatigue in medical environments. As explained in section 2.1, one of the consequences of alarm fatigue is that it can lead to alarm desensitization and causes underestimation of alarm sounds. The participant who did not detect the ST elevation was part of this group.

Moreover, three (3) participants stated that they first noticed a change in the sonification and therefore looked at the ECG monitor. Then, they noticed the ST elevation. Thus, in this case the sonification led them rapidly to the ECG visualization and served as a supporting tool either by providing a first hint or as a confirmation of the pathology. Interestingly, none of the participants stated that they diagnosed the ST elevation directly from sonification. In fact, they expressed that the sound made them look at the visually displayed ECG to then reach the diagnosis. This is of great importance because it means that the sonification acted as a supporting tool. In fact, considering that physicians learn to recognize changes and
patterns involving pathological behaviors from the visual signal, it is expected that they need the visual representation to reach a diagnosis.

Furthermore, twelve (12) participants did not make a particular statement on whether they first noticed the ST elevation through the ECG monitor or through the sonification. However, we collected comments expressing that sonification led to a faster diagnosis or that sonification is a good addition to recognize ECG changes timely, particularly in stressful situations.

As part of the general comments added at the end of the questionnaire, one of the participants mentioned that sonification is "hard to evaluate if you don’t have experience on the topic". Moreover, another participant stated that "the role that sonification plays, depends on whether the person has musical training or is an auditory-learning type person". These comments are in line with the importance of the training phase, as this is a fundamental step towards building associations between the data and the sonification.

Finally, one of the participants pointed out that sonification enables physicians to get information without having to turn away from the patient. In this regard, it is possible to envision use cases in which sonification can assist physicians to focus their attention on a primary task (i.e. examining the patient) while simultaneously monitoring other sets of data.

4 Discussion

Our first study shows that classification accuracy was similar in the two groups, which can be explained from the fact that both groups had no previous experience with ECG sonification. In this scenario, medical expertise did not play a role as we did not find a significant difference in terms of classification performance between the groups. Nonetheless, the sonification design did play a role in terms of classification accuracy. Overall, lower error rates were achieved using the Water ambience method. For instance, this design allowed participants from both groups to consistently determine when a sonification corresponded to an isoelectric ST segment. The CSS group was able to classify sonification examples corresponding to ST elevation cases despite not having any medical pre-knowledge.

Concerning the Polarity design, sonification examples corresponding to a healthy ECG obtained larger error rates than the middle ST elevation files (mid-low and mid-high). A reason for this could be that participants only listened to the training examples for a very short time before starting the classification task. We consider that in order to reach higher accuracy rates using the Polarity design, further training will be needed so that smaller differences in auditory references corresponding to the amount of ST elevation are better to distinguish.

In terms of the aesthetic components, the Water ambience obtained the highest score across all evaluated statements. The statement evaluating pleasantness, received a higher score from the MedS group, providing an encouraging panorama as it indicates that medical experts would be interested in using this type of sounds in their medical assessment activities. Overall, for both groups we found the same significant difference in the evaluation of pleasantness across the two sonification designs. Furthermore, informativeness obtained a higher score from the CSS group in both sonification designs. This could be an indicator that, although medical experts expressed interest about using sonifications as the Water ambience into their daily medical assessment activities, they are not sure if this type of sonification provides the information that they need in order to perform a diagnosis. Nonetheless, in this regard, it is important to consider further user studies that involve longer training phases.

Suitability for long-term listening was rated considerably higher in the Water ambience design. In particular, the CSS group assigned higher scores to this item. However, results given by the MedS group do not differ much from the scores given by the CSS group. This feedback is of great importance because medical experts know how a medical scenario normally sounds. They have a reference about how many alarm sounds are present and how often they sound. Thus, results indicate that even though they have an idea about how much exposure they would have to the presented sonification (i.e., Water ambience), they
still consider it to some level suitable for long-term listening.

In terms of the simulated medical scenario, the experience of participants about the role that sonification played in the detection of the pathology, can be generally grouped in three categories: (i) participants who completely ignored the sonification signal, either because it was a stressful situation or because they tend to do that in general with alarm sounds. (ii) Participants who first noticed the ST elevation through the ECG monitor and then attended the sonification signal. In this second group, a set of participants mentioned that the sonification helped to confirm the diagnosis. The last group were (iii), participants that first noticed a change in the sonification sound which served as a cue to look at the ECG monitor to then detect the pathology.

In general terms, results indicate that participants see the potential of ECG sonification as a supporting tool for diagnosis and monitoring of myocardial infarction. Moreover, in most cases participants expressed to be open to the idea of using sonification in their medical assessment routine.

The evaluation of the aesthetic components indicate that participants regard the proposed sonification design as pleasant to listen to. These results are in line with participants’ comments, in which they regard the sound component of the scenario as pleasant to listen as well as not distracting.

It is important to mention that none of the participants reached a diagnosis by only using sonification. In cases where they first noticed a change in the sonification, this served as a hint to look at the monitor to detect the ST elevation. Thus, the sonification served as a supporting tool. In order to test if similar results can be achieved by using ECG sonification in medical scenarios only, a follow-up user study evaluating different feedback conditions (i.e. visual, sonification, visual + sonification) needs to be made.

In our simulated case the patient had been awake and was able to mention progressive symptoms of chest pain which correlated to the ECG changes. However, there are a lot of cases in medicine in which patients can not express their symptoms, because the vigilance is restricted. This can be the case because of another vigilance affecting illness or due to narcotic medication. In these cases sudden changes of the ST segment can be the only indicator for an ongoing coronary occlusion. Continuously monitoring of the ST segment, e.g. by using sonification, can be of major importance for surveillance of potentially life threatening cardiac events.

Finally, study results indicate that the majority of medical experts that took part in the study are very open to the idea of using ECG sonification and state that it could have the potential for a supporting tool for diagnosis and monitoring of myocardial infarction. Nonetheless, we consider that the training phases are crucial so that physicians learn to better recognize patterns and changes in the sonification signal, and particularly, build associations between the visual and auditory representations of the ECG to combine the powerful capabilities of the human hearing system and their ECG interpretation expertise to assist diagnosis and monitoring of myocardial infarction.

In the simulated medical scenario in 96% of cases the pathology was properly detected and in the majority of cases sonification played a role in the detection. Therefore we consider that there is a promising outlook to use ECG sonification as a supporting tool in diagnosis and monitoring of myocardial infarction beyond the already existing sonification methods available in medical environments. Sonification could be used for prehospital care when transporting patients with suspicion of myocardial infarction. It can be helpful in being able to detect ST elevation as early as possible. By improving early detection of STEMI this method could be an additional tool to avoid following out-of-hospital cardiac arrest (OHCA) cases during prehospital emergency care. To investigate if ECG sonification could contribute to an assumed reduction of OHCA cases in this setting, further studies should be initiated.

**Funding**  Funding was provided by the German Academic Research Service “DAAD” (Grant No. 57129429) and the German Research Foundation “DFG” (Grant No. “CITEC” (EXC 277)).

**Declarations of interest:**  None
Appendix A

5 ECG surrogate signals

For the classification task we generated a set of signals at a sampling rate of 1000 Hz and constant heart rate of 60 BPM, exhibiting four distinct ST-elevation levels: healthy, medium-low, medium-high and extreme. The ECG surrogate signals are depicted in Figure 10. The surrogate data files can be found as part of the supplementary material.

![Figure 10: ST levels of surrogate signals](image)

6 Estimating the ST elevation

The next step is to perform the feature extraction. At first we removed artifacts and unwanted noise by applying a low-pass filter with a cutoff frequency of 120 Hz. We determined the cutoff frequency taking into consideration the common ranges for ECG diagnosis [27].

To calculate the ST elevation level in the ECG signal we first performed the R-peak detection. We do this by applying the method proposed by Worrall et al. [28] in a time window of 200 ms. We chose the window size considering that the average duration of the QRS complex in a 60 bpm ECG signal corresponds to 100 ms [27]).

Once we had detected the R-peaks, we estimated the ST elevation in each heartbeat. To accomplish this, we calculated the amplitude difference between the J-point and isoelectricity. The isoelectric reference can be taken either from the TP segment or the PQ segment. In our case, we selected the TP segment as the isoelectric reference considering that in cases of pathological behavior of the atrium (e.g., pericarditis), the morphology of the PQ segment can be affected.

As a means to locate the the J-point in each heartbeat, we applied the method proposed by Al-Kindi and Tafreshi [29]. First, we calculated the first derivative of the ECG signal, afterwards, we selected the first sample after the S-wave in which the derivative is zero, and picked this location as the J-point. We estimated the location and duration of the S-wave manually, considering the average QRS complex duration for an ECG signal of 60 bpm [27].
Then, to determine the amplitude at the J-point, we compute the average amplitude between the J-point and a given number of samples ahead. In practical terms we computed

$$\overline{ST}_i = \frac{1}{t_a - t_{j\text{-point}}} \sum_{t=t_{j\text{-point}}}^{t_{a-1}} g[t]$$  \hspace{1cm} (A.1)

by summing the sampled signal $h(t)$ between the segment limits $t_a$ and $t_{j\text{-point}}$. Moreover, following the aforementioned idea, we computed the amplitude in the TP segment as follows

$$\overline{TP}_i = \frac{1}{t_{\text{TPend}} - t_{\text{TPstart}}} \sum_{t=t_{\text{TPstart}}}^{t_{\text{TPend}-1}} g[t]$$  \hspace{1cm} (A.2)

Finally, we computed the ST segment level in the following way

$$\overline{ST}_{\text{amp,i}} = \overline{ST} - \overline{TP}$$  \hspace{1cm} (A.3)

The surrogate signals we generated only exhibit zero or positive ST elevation, thus, we referred to the signals as elevated and not depressed (i.e., negative ST segment amplitude).

Appendix B

7 Demographics of the classification study

7.1 Computer science students (CSS)

We only report demographic data from the CSS group considering that we did not collect this information from the MedS group. Results indicate that 30% were female and 70% were male. The average age of the group was 22.8 years. Furthermore, 20% of participants reported to have experience in terms of sound/music, for instance because they played a musical instrument. On the contrary 80% declared to not have previous experience in this field.

7.2 Demographics of the simulated medical scenario

The group of participants was formed by 72% females and 24% males. The mean age was 26.3 years old. Moreover, 68% of participants declared to have experience in terms of music, 20% said they did not have prior musical training and 12% left the question unanswered. Interestingly, the majority of participants had some level of musical expertise.

References


---

6One participant left this question unanswered.

7Results exclude six participants that left this question unanswered.


