Evaluating the Impact of Pseudo-Colour and Coordinate System on the Detection of Medication-induced ECG Changes

Alaa Alahmadi[1], Alan Davies[1], Jennifer Royle[2], Markel Vigo[1], Caroline Jay[1]
[1]School of Computer Science, The University of Manchester
[2]CRUK Manchester Institute, and The Christie NHS Foundation
Manchester, UK
alaa.alahmadi@postgrad.manchester.ac.uk,Jenny.Royle@digitalecmt.org
[alan.davies-2,markel.vigo,caroline.jay]@manchester.ac.uk

ABSTRACT

The electrocardiogram (ECG), a graphical representation of the heart's electrical activity, is used for detecting cardiac pathologies. Certain medications can produce a complication known as 'long QT syndrome', shown on the ECG as an increased gap between two parts of the waveform. Self-monitoring for this could be lifesaving, as the syndrome can result in sudden death, but detecting it on the ECG is difficult. Here we evaluate whether using pseudo-colour to highlight wave length and changing the coordinate system can support lay people in identifying increases in the QT interval. The results show that introducing colour significantly improves accuracy, and that whilst it is easier to detect a difference without colour with Cartesian coordinates, the greatest accuracy is achieved when Polar coordinates are combined with colour. The results show that applying simple visualisation techniques has the potential to improve ECG interpretation accuracy, and support people in monitoring their own ECG.

CCS CONCEPTS

• **Human-centered computing → Visualization techniques**;

KEYWORDS

Visualisation; Visual Perception; ECG; Drug-induced LQTS

CHI 2019, May 4–9, 2019, Glasgow, Scotland UK

ACM ISBN 978-1-4503-5970-2/19/05…$15.00
https://doi.org/10.1145/3290605.3300353

ACM Reference Format:
Alaa Alahmadi[1], Alan Davies[1], Jennifer Royle[2], Markel Vigo[1], Caroline Jay[1]. 2019. Evaluating the Impact of Pseudo-Colour and Coordinate System on the Detection of Medication-induced ECG Changes. In *CHI Conference on Human Factors in Computing Systems Proceedings (CHI 2019), May 4–9, 2019, Glasgow, Scotland UK*. ACM, New York, NY, USA, 13 pages. https://doi.org/10.1145/3290605.3300353

1 INTRODUCTION

A side effect of commonly prescribed medications including antihistamines, antibiotics and antidepressants is prolongation of the QT-interval, or drug-induced Long QT Syndrome (LQTS) [9, 78]. LQTS is a cardiac abnormality that can increase the risk of the life-threatening arrhythmia *torsades de pointes (TdP)*, which can lead to loss of consciousness or sudden death in young, otherwise healthy people [3, 20, 47, 78]. People may not experience symptoms and an electrocardiogram (ECG) is often the only way to identify LQTS [34, 60, 64].

ECGs are a graphical representation of the electrical activity of the heart, widely used in clinical practice to assess heart function [62]. ECG results are displayed as a line on a graph-like trace, where the 'waves' (peaks and troughs) are labelled with letters and represent different stages of the heartbeat. The duration of the QT-interval (the time period between the 'Q' and 'T' waves) represents the activity of the heart ventricles (Figure 1).

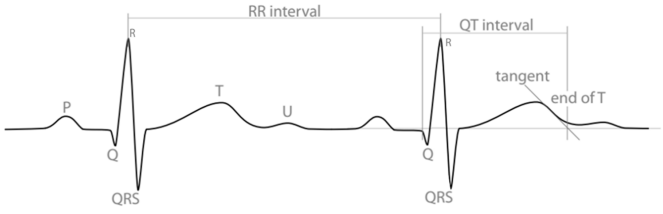

Figure 1: Measurement of the QT-interval on the ECG from the start of the Q-wave to the end of the T-wave.

Frequent ECG monitoring is advised for people at high risk of acquiring medication-induced LQTS [38, 65]. Recent technology innovations have made it possible to monitor ECGs outside of the clinical environment [55] but this approach still relies on clinician interpretation [4, 41]. This not only increases cost, but also makes it difficult (and sometimes impossible) to manage everyone who is at high risk. If lay people can interpret their own results, this may lead to a step-change in the detection and management of this potentially critical condition. However, ECG interpretation is known to be complex, even for clinicians [66, 72], and as such little work has examined self-monitoring.

Assessing the QT-interval, in particular, is known to be difficult [66, 72], and in a prior study the majority of clinicians were not able to recognise it [72]. This may be due to the fact that whilst people find it easy to perceive quantity on a vertical scale, they are poor at judging it on a horizontal scale (see e.g. [39, 40, 53, 76]). Artifacts in the ECG signal can also cause misinterpretation of QT-interval length [2].

Visualisation techniques have the potential to help highlight abnormalities within the ECG. Here we examine whether pseudo-colouring—representing continuously varying values using a sequence of colors [61, 74]—and changing the coordinate system can support lay people in identifying increases in the QT-interval. Using a psychophysical paradigm and eye tracking to systematically examine the issue, we find that:

(1) Pseudo-colouring significantly increases lay individuals' ability to identify increases in the QT-interval, even when T-wave morphology is abnormal.
(2) Coordinate system interacts with colour, such that people are most accurate in the condition where the ECG is presented using polar coordinates and pseudo-colour, and least accurate when presentation occurs with polar coordinates and no colour.
(3) According to eye tracking data, pseudo-colour helps to focus visual attention, and people are most accurate when using the polar coordinates as this concentrates colour in the center of the screen.
(4) People are significantly more satisfied when pseudo-colour is used.

2 BACKGROUND

Previous research providing a foundation for the current study is described in this section. This covers: (1) ECG interpretation; (2) ECG visualisation methods; and (3) human perception of visualised data.

ECG Interpretation

An ECG trace is a cyclic time series with each cycle representing a new heartbeat. The electrical activity is detected via leads placed on the body, where each lead produces a different electrical 'view' of heart activity. In hospital, clinicians commonly interpret short (10 second) ECGs via 12-leads [48]. This is the most comprehensive view, but useful ECG information can be gathered from a single lead 'view' (often used by mobile technologies). LQTS is detected by measuring from the beginning of the Q-wave to the end of the T-wave (identified using the tangent drawn at the maximum downslope of the T-wave) [3, 21] as shown in Figure 1.

The standard method for visualising ECG data is a Cartesian line graph showing the voltage of the heart on the Y-axis, and time in milliseconds (msec) on the X-axis [5]. A background grid supports the reader in measuring duration. To measure the QT-interval, the interpreter counts the small squares (each representing 40 msec) from the beginning of the Q-wave to the end of the T-wave [3, 21, 56].

Automated ECG interpretation was introduced in the 1950s to assist clinicians who had less training in ECG interpretation [57]. It remains far from perfect, and even the best computational methods can produce significant errors [30, 63]. Research has shown that QT-interval is underestimated or unreported by computational methods [18, 36, 45, 58, 69, 70]. The main challenge lies in identifying the end of the T-wave, especially when the morphology (shape) of the T-wave is non-standard. [17, 21, 26, 46]. This is particularly problematic, as QT-prolonging drugs often affect the morphology of the T-wave, with some drugs (e.g. quinidine and ranolazine) causing large T-wave morphology changes [71].

Each person has a unique baseline ECG that reflects their individual heart function: health status, age, gender and ethnicity all influence the ECG in general, and the QT-interval in particular [21, 27, 43]. This complicates population-level computer-derived QT calculations. Abstracting the QT interval numerically also risks masking other potentially abnormal clinically significant changes in the ECG. For instance, specific T-wave patterns can aid detection of drug-induced LQTS [11], and large T-U waves are known to precede the life-threating arrhythmia *Torsades de Pointes* [35]. As such the ECG morphology continues to provide the richest information for recognising LQTS. Current automated methods are thus a supplement to, rather than a substitute for, the human eye, and a combination of computer-visualisation methods with the gold-standard human interpretation remains the most accurate and reliable method [14, 18, 36, 45, 58, 63, 69].

ECG Visualisation

A number of studies have examined the effectiveness of visualisation techniques in supporting clinician-interpretation of ECGs. Chiang et al. [10] integrated ECG signals from the periodical and limb leads into two images of electrical heart function. This enabled clinicians to observe an overall integral heart view, which aided interpretation when viewed

alongside the 12-lead ECG. Kors et al. [37] presented a mirror image that converted the 12 leads to 24, and was shown to be effective in improving the detection-rate of heart attacks. Madias et al. [44] used a 13th, multi-use lead, which provided a further 'view' of the heart. When exploring large scale ECG data, a glyph-based interactive system has been shown to be effective in detecting arrhythmia [77]. Vectorial methods have been used to represent direction and magnitude data [16, 49] and spatial visualisations have presented the ECG on a body surface potential map [42, 68]. As these methods provide data with respect to further dimensions of the heart, they are useful as a supplement, but do not replace the standard method [7]. Furthermore, previous work in this area focused on aiding clinicians. Here, the aim is to help laypeople identify when their ECG is different (i.e. has deviated) from their normal baseline, so they know when to seek help.

Human Perception of Visualised Data

Here, we consider the problem from the perspective of the lay interpreter, rather than the data, using knowledge of visual perception to enhance the way the ECG is presented. In particular, we draw from the field of pre-attentive processing, which outlines a set of visual properties known to be detected rapidly and accurately by the human eye [50]. Examples of pre-attentive properties include colour, form, and spatial positioning. Using these properties in design can improve both the effectiveness and the efficiency of a visualisation [24, 25, 74].

Colour is a pre-attentive attribute that is noticed without conscious effort [22, 50]. Many studies have shown the effectiveness of using colour to separate visual elements from their surroundings, saving the user from having to carry out a linear visual search [23, 54, 74]. A useful technique is pseudo-colouring, which represents continuously varying values using a sequence of colors [74]. Pseudo-colouring is commonly used in geo- and time-series visualisations [61, 74]. Figure 2 shows an example of using pseudo-colouring to show changes in temperature over time.

Adnan et al. [1] have examined perception of time-series visualisations. They showed Cartesian coordinates to be most effective for detecting trends and identifying maximum and minimum values when used with positional and colour visual encodings, and Polar coordinates to be most effective for finding minimum values when using area visual encoding.

The circular layout used in the Polar coordinate system has also been employed to perceive changes in data over time. Page et al. [51, 52] proposed an "ECG Clock" generator, to visualise the changes in QT interval values automatically generated by a 24-hour Holter ECG monitor. Circular layouts have been also used to detect symmetrical patterns in data [28] and to measure symmetry in graphs [75].

3 METHOD

Measuring Visual Perception

To systematically evaluate the effectiveness of pseudo-colour and coordinate system in supporting lay people's assessment of the QT interval, we use methods from psychophysics and eye-tracking research. Psychophysical experiments investigate the relationship between physical stimuli and human perception, by varying the properties of a stimulus along one or more physical dimensions [67]. Eye-tracking is used to quantify visual behavior when performing a given task, to understand differences in locus and level of attention [13].

ECG Data Acquisition

The ECG datasets were taken from a clinical trial that assessed the effect of known QT-prolonging medication on healthy subjects [32]. As our work is motivated by supporting self-monitoring, we selected data from a single participant, whose QT-interval was seen to rise to clinically dangerous levels. The subject was a 35-year-old male who had normal QT-intervals (< 430 milliseconds) prior to taking the medication "Dofetilide" (an antiarrhythmic drug); he subsequently experienced a gradual increase in the QT-interval, which eventually reached very high levels (QT=579 milliseconds). The ECGs sampled all have a regular heart rate (HR=60 BPM) and are from lead-II, which is typically used to measure the QT-interval. The QT-values of the selected ECGs were 417, 421, 441, 447, 455, 468, 485, 537, 565 and 579 milliseconds. We categorised these values based on their clinical significance: normal (QT < 430); borderline (QT > 430 and < 470); prolonged (QT > 470 and < 500); very prolonged levels (QT > 500) [33]. The open ECG dataset is available from the PhysioNet database [19].

Visualisation Design

We used a co-design approach, creating the visualisation techniques with an expert in ECG interpretation (to ensure accuracy), and refining them with input from lay people. As a first step, R-peaks were detected in the raw ECG datasets, and a dashed vertical line used to show the halfway point of the R-R interval (Figure 1). This helps to identify the area of interest containing the QT-interval. Note that it is easy to detect the R wave in the vast majority of ECGs, as (unlike the other waveforms which vary considerably) it consistently has the greatest amplitude. We then applied pseudo-colouring, to shift the 'work' of QT interval visual encoding from perceiving distance between two waves, to perceiving colour in terms of hue and intensity.

As spectrum-approximation sequences in particular help with reading values [73], we used these as a foundation for the pseudo-colouring technique. Cool spectral colour codes (purple to blue to green) were used to indicate normal

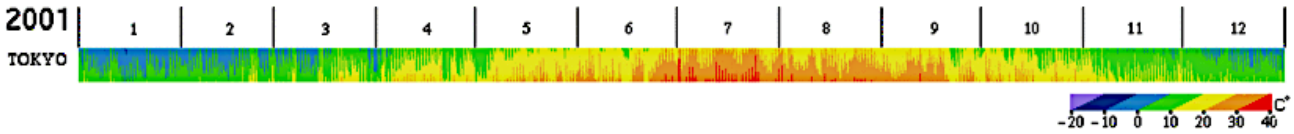

Figure 2: Using pseudo-colouring to represent temperature in time series data over the 12 months of the year [61].

QT-interval ranges, and warm colours (yellow to orange to red) to show abnormal QT-interval ranges. We applied the pseudo-colouring sequence in the area between the 0 voltage baseline up to (or down to) the signal, from the beginning of the R-wave to the R-R interval halfway point (see e.g. Figure 3). The pseudo-colouring sequence was mapped to the ECG signal such that the colour code changed every 40 milliseconds, which is equal to a small square on the standard ECG background grid.

To understand the impact of coordinate system on ECG data interpretation, we displayed the ECG signals on Cartesian and Polar coordinates with and without pseudo-colouring. We used R [31] with RStudio software version 1.1.447 to create the visualisations. Figure 3 shows ECGs with normal and very prolonged QT intervals with and without the pseudo-colouring sequence on Cartesian coordinates. Figure 4 shows the same ECGs, but on Polar coordinates. The ECGs are reduced in size for inclusion in the paper. The full size images, along with the scripts used to create them, can be found in the supplementary materials, and in our repository[1].

Experiment Design

We hypothesized that changes in the T-wave morphology (e.g. flattening of the wave, which can be caused by QT-prolonging medication [71]) might cause misperception of the QT-interval, and included this as a factor. The study thus used a counterbalanced within-subjects design with three independent variables, each with two levels:

(1) Colour-coding: no colouring; pseudo-colouring.
(2) Coordinate system: Cartesian; Polar.
(3) The T-wave morphology: normal; abnormal.

The within-subjects factorial design yielded a total of 8 (2x2x2) experimental conditions for each participant. We counterbalanced the order of visualisation presentation using a balanced Latin square to minimize practice effects. We assessed the effects of T-wave morphology in two separate tasks (described below). The order of the T-wave morphology condition (normal or abnormal) was counterbalanced across participants. The dependent variables were response correctness, reaction time, fixation location, mean fixation duration, and satisfaction.

[1] https://github.com/mbchxaa6/ECG_QT_Visualisation.

Participants

Forty two participants (22 males and 20 females) were recruited from a university campus. Eligibility for the study was determined by asking participants to rate their knowledge of ECGs/ECG interpretation, and including only people who reported no knowledge at all. Participants consisted of 34 students and 8 staff. The mean age was 30 (SD=7). The backgrounds of the participants were Computer Science (n=27), Education (n=3), Chemical Engineering (n=3), Electrical Engineering (n=3), Mathematics (n=4), History and Sociology (n=1) and Music/Violin Performance (n=1). Their sight was normal or corrected-to-normal and they reported no motor or neurological disorders.

Task and Procedure

Participants completed a 10 minute training session where they were introduced to the ECG trace and shown how to identify the QT-interval, and then shown the visualisation techniques used in the experiment. Each participant then completed an assessment task to check that they understood how to perform the measurement, where they were asked to highlight the start and end point of the QT-intervals on two different ECGs using the four visualisation techniques.

The experiment used a "two alternative forced choice" (2AFC) psychophysical discrimination task [67]. Within a trial, the participant was presented with two ECG stimuli; a baseline showing no QT-prolongation and a comparator showing an increased (or the same) QT-interval; the participant had to select the ECG that they perceived to have the longer QT-interval using the left/right arrows in the Polar condition, and up/down arrows in the Cartesian condition, according to the stimulus' position on the screen.

Participants completed all trials for one visualisation technique before moving to the next. The location of the ECG with the longer QT interval (i.e. top/bottom or left/right) was counterbalanced in the design; stimuli were then presented at random.

To determine the effects of T-wave morphology, we split the experiment into two separate 2AFC tasks, as follows.

Normal-T-wave: In this condition, participants completed a total of 20 experimental trials. In each trial, two ECGs were presented; a baseline ECG showing a normal QT-interval with a normal T-wave morphology and a comparator ECG. Two trials showed exactly the same ECG for the baseline

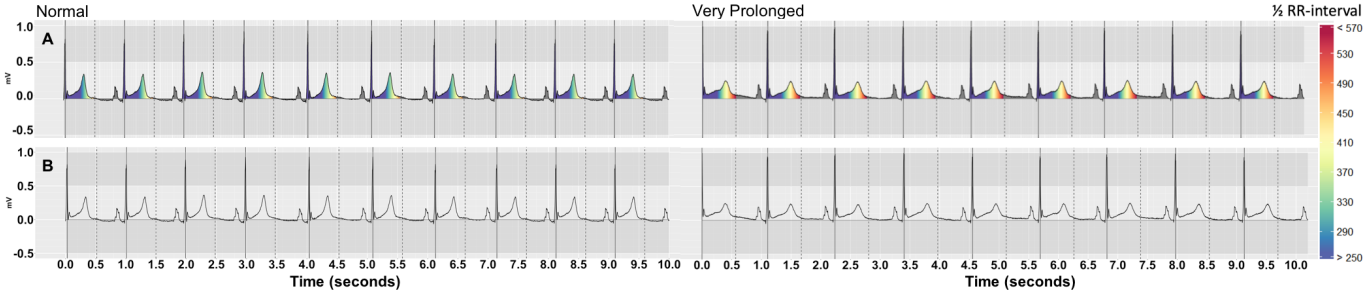

Figure 3: ECGs with normal and very prolonged QT intervals on Cartesian coordinates with (A) and without (B) pseudo-colour.

and the comparator stimuli, in order to test the validity of our method (the probability of choosing each alternative should be equal to 0.5). The other 18 trials presented the same ECG baseline (QT = 417 msec), and an ECG with a longer QT-interval that was selected from the following set of QT values, where each value was presented twice: 421, 441, 447, 455, 468, 485, 537, 565, 579 milliseconds. Figure 4 and Figures 3 show examples of the ECGs used in the normal T-wave morphology condition.

Abnormal T-wave: This condition was used to evaluate whether the visualisation techniques can help people to perceive QT-prolongation regardless of the T-wave morphology. To reduce potential fatigue, this condition contained only 8 trials. In each trial, two ECGs were presented: a baseline ECG showing a borderline QT-interval with an abnormal (flat) T-wave morphology and an ECG with an increased QT-interval. Participants had to choose the ECG with the longer QT-interval. The borderline ECG had a QT value of 447 millisecond and the comparator ECG was selected from the following set of QT values, which had either a normal or abnormal T-wave as indicated: 468 (abnormal), 485 (normal), 565 (abnormal) and 579 (normal) milliseconds. Figure 5 shows examples of the ECGs used in the abnormal T-wave morphology condition.

Apparatus

A Tobii Pro Spectrum eye-tracker and Tobii Pro lab 1.95 software were used to record eye gaze with a sampling rate of 600 HZ. Key press events were recorded to collect participants responses. The study was performed on a 23.8 inch (diagonal) Tobii Pro Spectrum eye-tracking monitor, with a resolution of 1920 x 1080 pixels. Each Cartesian coordinate ECG stimulus was 32.31cm x 6.14cm, and each Polar coordinate stimulus was 15.61cm x 12.93cm.

4 RESULTS

All anonymised raw data, along with relevant R-scripts and SPSS outputs are available in our Github repository[2].

Accuracy

Psychometric function. We used a psychometric function, which is an inferential model employed in psychophysical detection and discrimination tasks, to model the relationship between the gradual increase in the QT-interval and the correctness of participants' responses across the four visualisation techniques. The psychometric function was plotted as the percentage of correct responses (trials where the longer QT-interval stimulus was correctly identified) as a function of the QT-interval increase (Figure 6).

The results show that pseudo-colour significantly improves perception of QT-interval increases regardless of the T-wave morphology, with people able to detect smaller increases with the Polar coordinates than the Cartesian coordinates. This is important, as even these small increases are clinically significant.

When pseudo-colour is not used, T-wave morphology interacts with coordinate system in the detection of the QT-interval increases. When the T-wave morphology is normal, people perform better with Cartesian coordinates (Figure 6 (A)). However, when the T-wave morphology is abnormal (increased flattening of the T-wave), people perform better with Polar coordinates (Figure 6 (B)).

Just noticeable difference (JND) threshold. In psychophysics, the JND threshold is defined as the minimum amount of change in a stimulus necessary for it to be 'just noticeable'. In this study, we defined it as the minimum increase in the QT-interval required for it to be detectable. We estimated the 75% JND threshold as the value of the QT-interval increase from the normal baseline at which the percentage of correct responses is equal to 75%. Only the normal T-wave morphology condition was used for estimating the JND, as the abnormal condition contained insufficient trials for it

[2] https://github.com/mbchxaa6/Data_Analysis.

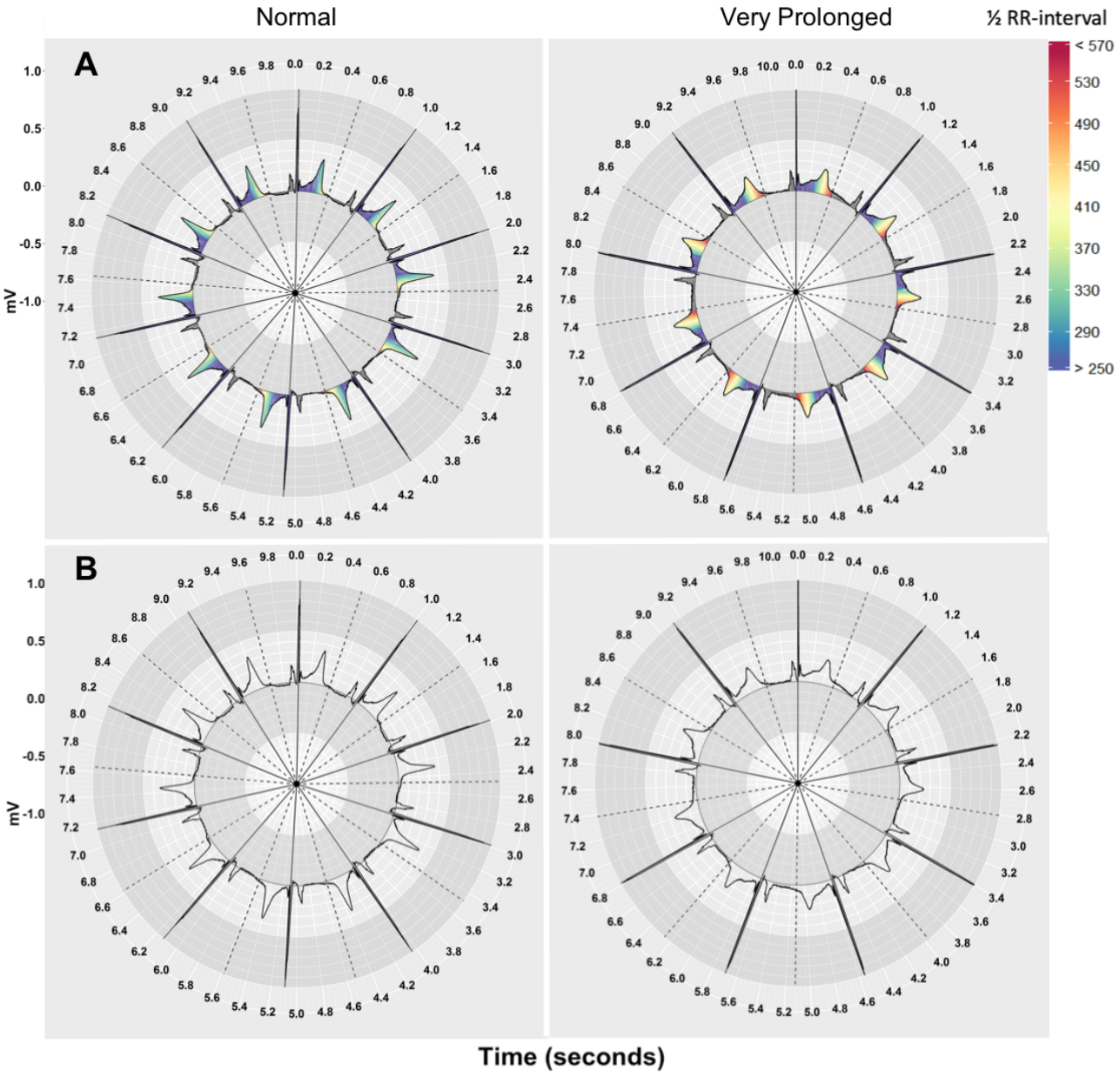

Figure 4: ECGs with normal and very prolonged QT intervals on Polar coordinates with (A) and without (B) pseudo-colour.

to be fitted with this statistical model. The JND thresholds, determined by fitting the psychometric function using a logistic function with maximum likelihood estimation (MLE) (Figure 7), were 29, 19, 65 and 9 milliseconds for *Cartesian*, *Cartesian with pseudo-colour*, *Polar* and *Polar with pseudo-colour* respectively. Pseudo-colour thus reduces the JND in both co-ordinate systems, with the effect being strongest for Polar co-ordinates.

Reaction Time

We measured reaction time as the period between the appearance of the stimuli on the screen and the key press event when people made their decision. As shown in Figure 8, pseudo-colour reduced the reaction time as the QT-interval increased in all conditions.

A Shapiro-Wilks test showed the reaction time data was not normally distributed ($p < 0.05$). We thus used a non-parametric Friedman test to compare reaction times across

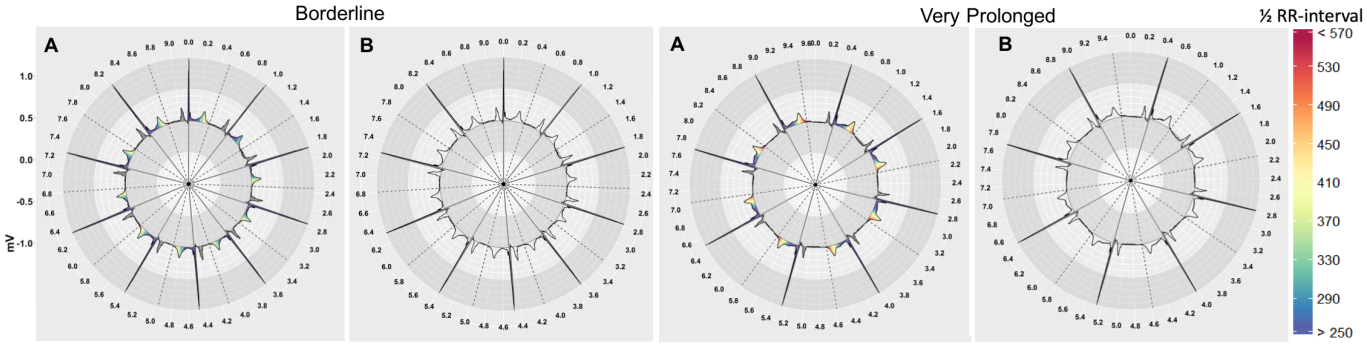

Figure 5: ECGs with abnormal T-wave morphology on Polar coordinates with (A) and without (B) pseudo-colour.

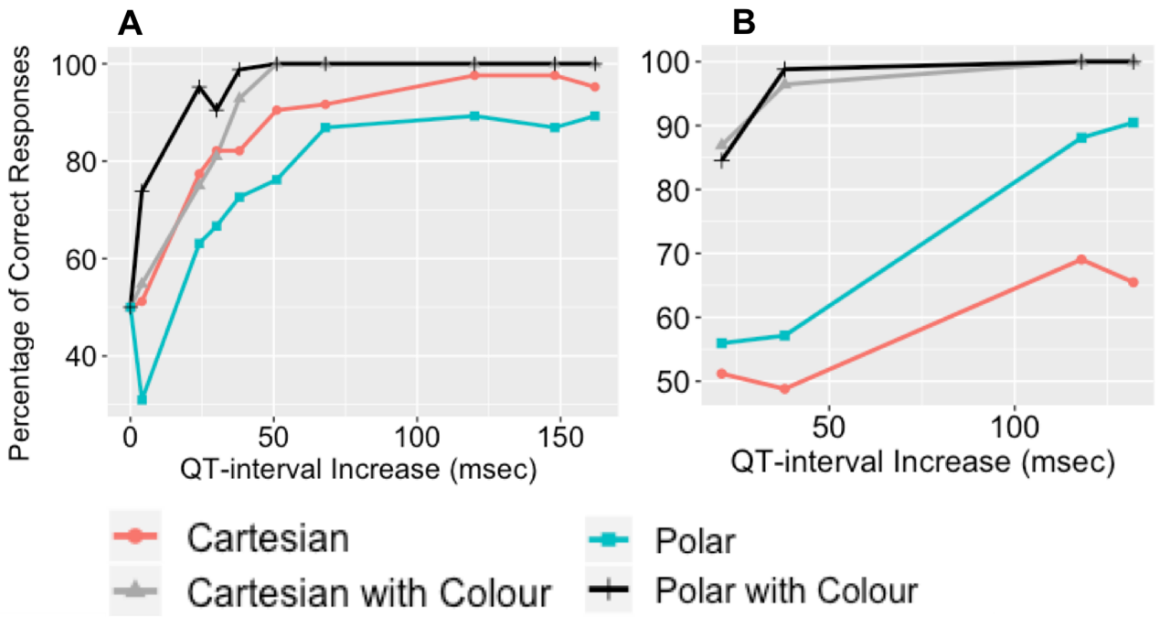

Figure 6: The psychometric function plot shows the percentage of correct responses as a function of the QT-interval increase from the baseline with (A) Normal T-wave morphology and (B) Abnormal T-wave morphology.

the four conditions. The test was conducted for each QT-interval increase and under each condition of the T-wave morphology. For all QT-interval increases, there was a statistically significant difference in reaction time according to visualisation technique, under both conditions of the T-wave morphology ($p < 0.05$) (Table 1).

To examine where the differences actually occur, *post hoc* pairwise comparisons were performed using a Wilcoxon signed-rank test with Bonferroni correction ($\alpha = 0.008$). This showed that when the QT-interval was clinically prolonged (equal to 485 msec and increased from the baseline by 68 msec) or very prolonged (greater than 500 msec and increased from the baseline by over 100 msec), reaction time was significantly faster when pseudo-colour was used, for both types of coordinate system ($p < 0.008$), regardless of T-wave morphology.

When the QT-interval was in the borderline range, the T-wave morphology and coordinate system interacted with pseudo-colour. In the trial that shows a borderline QT-interval (increased by 38 msec) with a normal T-wave morphology, reaction time when pseudo-colour was used was significantly faster for Polar coordinates than Cartesian coordinates ($Z = -3.806$, $p < 0.008$, Figure 8 (A)). However, in the trial showing a borderline QT-interval (increased by 21 msec), but with an abnormal T-wave morphology, there was the opposite effect, with people responding faster in the Cartesian with pseudo-colouring condition than in the Polar with pseudo-colouring condition ($Z = -2.870$, $p < 0.008$, Figure 8 (B)).

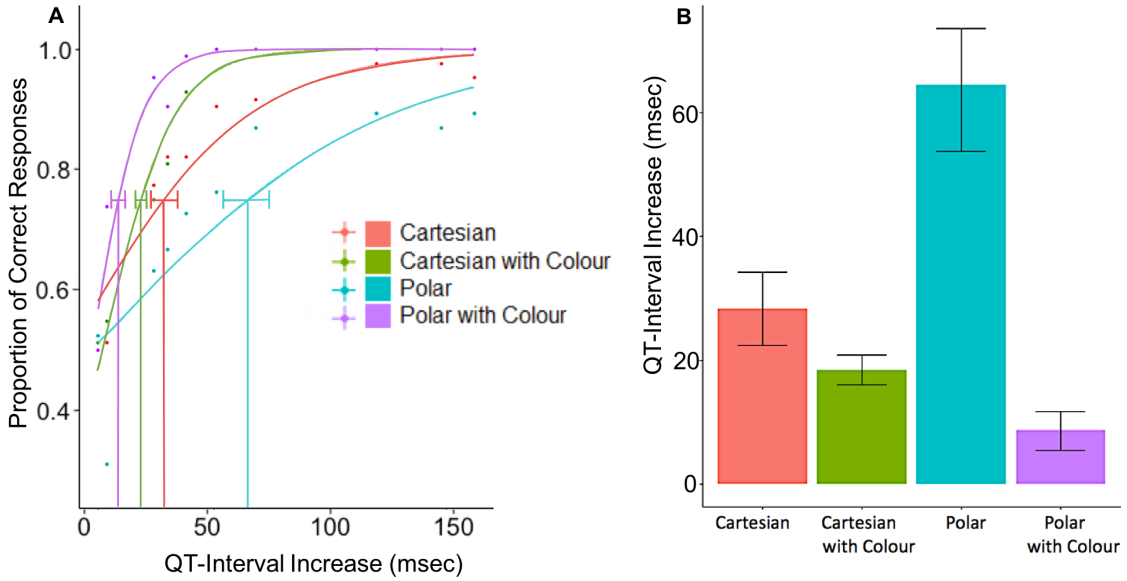

Figure 7: (A) The fitted psychometric function plot shows the proportion of correct responses as a function of the QT-interval increase from the normal baseline. (B) The just noticeable difference (JND) thresholds plot. The error bars represent bootstrap confidence intervals.

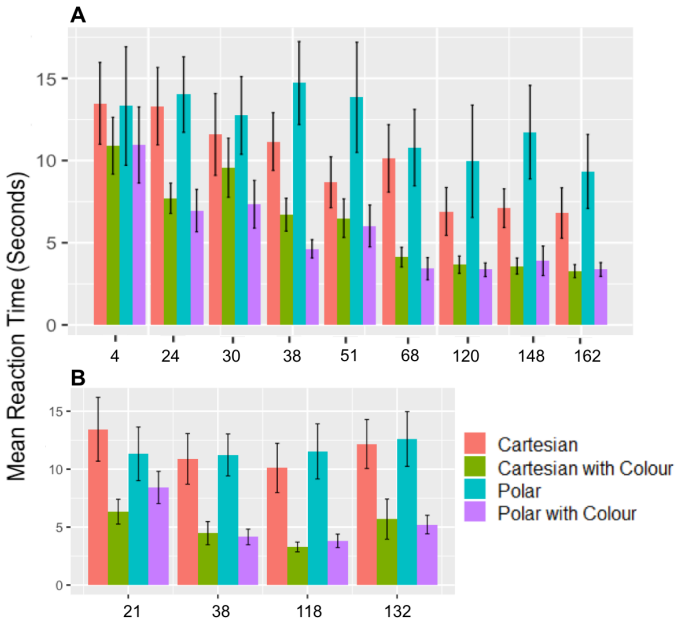

Figure 8: Mean reaction time in seconds over the QT-interval increases (msec) from the baseline with (A) Normal T-wave morphology (B) Abnormal T-wave morphology. Error bars represent 95% confidence intervals.

Eye-tracking Metrics

To calculate eye movement metrics, Tobii Pro lab software was used to create two areas of interest (AOIs) for each experimental trial: one for the baseline ECG stimulus, and one for the comparator ECG stimulus.

Mean fixation duration. The mean fixation duration metric, which is an indicator of cognitive load [12, 29], was

Table 1: Results of the Friedman test comparing reaction times in the four visualisation conditions for all QT-interval increases, and in each condition of the T-wave morphology. QT represents the value of the longer QT-interval in milliseconds. ΔQT represents the difference in milliseconds between the value of the longer QT-interval and the baseline QT-interval (i.e. the amount of QT-interval increase).

$T-wave$	QT	ΔQT	$Range$	$\chi^2(3)$	$p-value$
Normal morphology	421	4	*normal*	9.543	$p < 0.05$
	441	24	*borderline*	68.100	$p < 0.05$
	447	30	*borderline*	13.443	$p < 0.05$
	455	38	*borderline*	101.954	$p < 0.05$
	468	51	*borderline*	41.471	$p < 0.05$
	485	68	*prolonged*	101.886	$p < 0.05$
	537	120	*very prolonged*	79.225	$p < 0.05$
	565	148	*very prolonged*	90.286	$p < 0.05$
	579	162	*very prolonged*	72.265	$p < 0.05$
Abnormal morphology	468	21	*borderline*	50.671	$p < 0.05$
	485	38	*prolonged*	114.783	$p < 0.05$
	565	118	*very prolonged*	138.529	$p < 0.05$
	579	132	*very prolonged*	94.409	$p < 0.05$

used to understand whether the visualisation techniques helped people to focus on the target ECG stimulus that had the longer QT-interval. As shown in Figure 9, regardless of T-wave morphology and coordinate system, using pseudo-colour results in longer fixations on the stimulus with the longer QT-interval, compared with the baseline stimulus, and this effect becomes more pronounced as the QT-interval increases.

Satisfaction

Following the experiment, participants completed a five point Likert-type scale ranging from 'bad' (1) to 'good' (5) to rate the effectiveness of each visualisation technique in supporting the detection of increases in the QT-interval. A Friedman test showed there to be a statistically significant difference in satisfaction depending on which visualisation technique was used ($\chi^2(3) = 90.860, p < 0.05$). A *post-hoc* analysis with a Wilcoxon signed-rank test utilising a Bonferroni correction ($\alpha = 0.008$) showed a significant preference for pseudo-colour ($p < 0.008$). However, although people were faster and more accurate in the Polar coordinate condition when pseudo-colour was used, there was no difference in people's satisfaction for either coordinate system ($Z = -0.435$, $p = 0.664$).

5 DISCUSSION

Recognizing QT-interval prolongation on the standard ECG is difficult. A previous study with medical professionals has shown that accurate classification of the QT-interval as either "prolonged" or "normal" was achieved by 96% of QT experts and 62% of arrhythmia experts, but by less than 25% of cardiologists and noncardiologists [72]. The QT-interval is also underestimated or unreported by computerised methods [18, 36, 45, 58, 69, 70], and as such, human visual validation is strongly recommended [14, 63]. To support lay people, who have no experience in ECG interpretation, in detecting life-threatening changes in the ECG, we need to understand how people perceive ECG data, and the extent to which visualisation techniques can aid the interpretation process.

We used psychophysical methods to model lay people's detection of QT-interval increases when using four visualisation techniques. The results show that using pseudo-colour to represent time significantly improves accuracy in detecting increases in the QT-interval, for both coordinate systems. People are most accurate in detecting small, but clinically significant increases in the QT-interval with Polar coordinates, regardless of whether the T-wave morphology is normal or abnormal (Figure 6).

Clinical research has shown that even a small (~ 10 msec) QT-interval increase from the baseline is considered a significant side effect of a QT-prolonging drug [8, 15, 59]. When the T-wave morphology is normal, the 75% just noticeable difference (JND) thresholds were 29, 19, 65 and 9 milliseconds for *Cartesian*, *Cartesian with pseudo-colour*, *Polar* and *Polar with pseudo-colour* respectively. This shows that using a combination of Polar coordinates and pseudo-colour has the potential to support lay people in detecting the smallest clinically significant change. It also shows that colour can improve sensitivity to changes such that people can perceive increases that are much smaller than a 1mm square on the standard ECG grid (which represents 40 msec).

As well as improving accuracy, using pseudo-colour reduced reaction times and increased attention to the longer QT-interval stimulus. Eye-tracking data showed that the average fixation duration increased more on the comparator stimulus, which has the longer QT-interval, than the baseline stimulus, as the interval length increased (Figure 9). Figure 10 shows a heat map of absolute fixation duration across all participants, demonstrating that even when the interval is borderline, rather than prolonged (QT = 455 msec, increased by 38 msec), people still fixate longer on the comparator stimulus. This shows the power of the colour codes used with the spectrum-approximation pseudo-colouring sequence, where warmer colours including orange and red help to attract attention to abnormal QT-interval levels.

Visualisation Design Implications

This study shows that colour as a pre-attentive attribute can support the detection of small differences in time-series data represented along a continuous scale, that are otherwise difficult to perceive. While people find it relatively easy to perceive quantity along a vertical scale, they are known to be poor at judging size or quantity displayed along a horizontal scale (see e.g. [39, 40, 53, 76]). Time-series data are conventionally displayed horizontally. Although this study focused on a specific problem within ECG interpretation, the results may have a wider application to other forms of time-series data, for example, in supporting detection of change in seasonality in financial data, or follow-up months of survival rate among cancer treatments (e.g. in Kaplan-Meier curves).

People are able to detect the smallest differences when the ECG is presented using Polar coordinates and pseudo-colour. Eye-tracking research has shown that people's initial eye movements are more commonly located in the center of the screen [6]. According to the study's eye-tracking data, the warmer hues of the pseudo-colour helped to focus visual attention; as Polar coordinates concentrate more colour in the center of the screen than Cartesian coordinates, the increased salience may be easier to perceive in foveal vision.

6 LIMITATIONS AND FUTURE WORK

Limitations of this study include: (1) we investigated detection of QT-interval prolongation, and it is not clear whether these techniques would generalise to interpretation of other

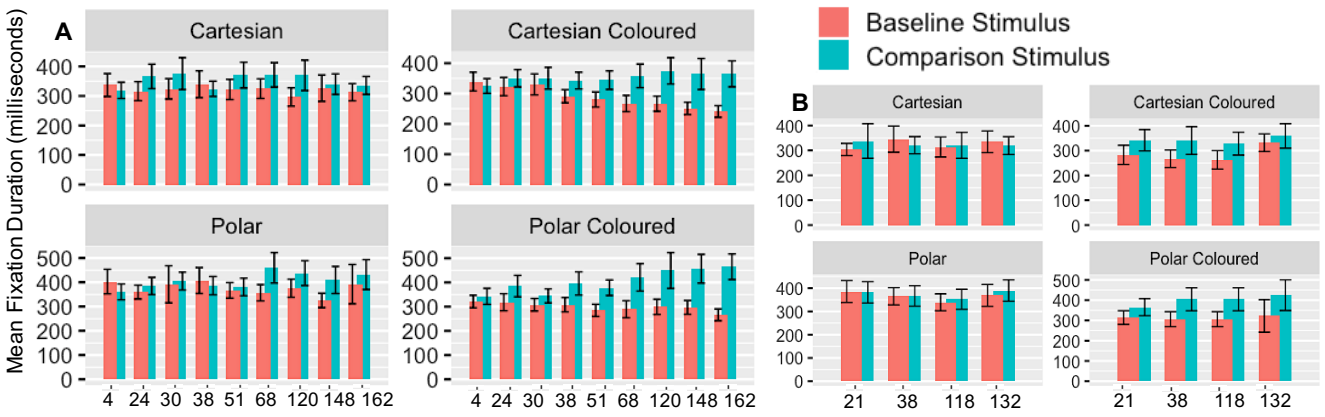

Figure 9: Mean fixation duration of the baseline and the comparator stimuli over the QT-interval increases (msec) with (A) Normal T-wave morphology and (B) Abnormal T-wave morphology. The error bars represent 95% confidence intervals.

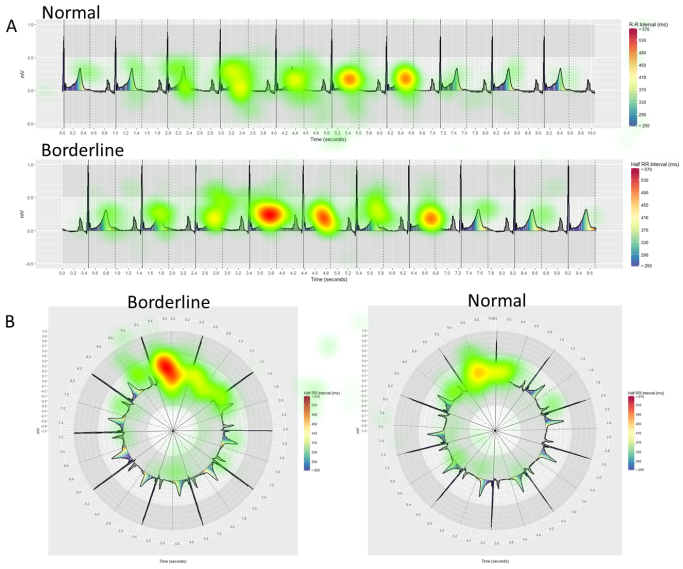

Figure 10: Heatmap of absolute fixation duration for all participants. Fixation is longer on the borderline QT-interval.

ECG abnormalities, such as changes in ST-segment elevation, or to other signal/time-series data; (2) the data used to design the stimuli were from a high quality signal with little noise; they were acquired from a 12-lead ECG, not a mobile monitoring device, where the signal is much more likely to be affected by noise; (3) we assessed an irregularity from the perspective of abnormal T-wave morphology, but this, of course, is one of many; fast or slow heart rates, abnormalities in ST-T changes and the presence of some common types of arrhythmia such as atrial fibrillation (AF) can all affect QT-interval calculation; (4) in a self-monitoring situation people may be using tablets or phones. We hypothesise that the visualisation techniques will still be beneficial, but this will need to be confirmed in a further study examining the effects of screen size and lighting setting on the visualisation techniques; and (5) our participants were highly educated, and we do not know whether the results would generalise to other demographics. Future work will include evaluating the visualisation techniques with more diverse clinical populations, particularly with low-literacy and low-income minority populations, who are taking medication that can lead to LQTS, and are using a mobile device with a wearable ECG monitor.

7 CONCLUSION

This study shows that using simple visualisation techniques significantly improves lay people's ability to accurately measure the QT-interval. This may help with self-monitoring drug-induced LQTS and enable treatment to be altered to prevent the development of life threatening complications. Whilst using a pseudo-colour sequence significantly improves people's ability to detect increases in the QT-interval when the ECG is displayed on a standard Cartesian coordinate system, the greatest accuracy is achieved when pseudo-colour is combined with Polar coordinates.

8 ACKNOWLEDGEMENT

This research is sponsored by Taibah University, Kingdom of Saudi Arabia, College of Computer Science and Engineering, Yanbu. We would like to thank the Digital Experimental Cancer Medicine Team (dgitalECMT), Manchester, UK, for their feedback and support.

REFERENCES

[1] Muhammad Adnan, Mike Just, and Lynne Baillie. 2016. Investigating time series visualisations to improve the user experience. In *Proceedings*

of the 2016 CHI Conference on Human Factors in Computing Systems. ACM, 5444–5455.

[2] Alaa Alahmadi, Alan Davies, Markel Vigo, and Caroline Jay. 2018. Can lay people identify a drug-induced QT-interval prolongation? A psychophysical and eye-tracking experiment examining the ability of non-experts to interpret an ECG. *Journal of the American Medical Informatics Association* (2018).

[3] Mark E Anderson, Sana M Al-Khatib, Dan M Roden, Robert M Califf, Duke Clinical Research Institute, et al. 2002. Cardiac repolarization: current knowledge, critical gaps, and new approaches to drug development and patient management. *American heart journal* 144, 5 (2002), 769–781.

[4] Ashish Atreja, Emamuzo Otobo, Karthik Ramireddy, and Allyssa Deorocki. 2018. Remote patient monitoring in IBD: current state and future directions. *Current gastroenterology reports* 20, 2 (2018), 6.

[5] Arlie R Barnes, Harold EB Pardee, Paul D White, Frank N Wilson, Charles C Wolferth, et al. 1938. Standardization of precordial leads: supplementary report. *American Heart Journal* 15, 2 (1938), 235–239.

[6] Markus Bindemann. 2010. Scene and screen center bias early eye movements in scene viewing. *Vision research* 50, 23 (2010), 2577–2587.

[7] Raymond R Bond, Dewar D Finlay, Chris D Nugent, George Moore, and Daniel Guldenring. 2013. Methods for presenting and visualising electrocardiographic data: from temporal signals to spatial imaging. *Journal of electrocardiology* 46, 3 (2013), 182–196.

[8] Joanna M Brell. 2010. Prolonged QTc interval in cancer therapeutic drug development: defining arrhythmic risk in malignancy. *Progress in cardiovascular diseases* 53, 2 (2010), 164–172.

[9] A John Camm, Marek Malik, and Yee Guan Yap. 2008. *Acquired long QT syndrome*. John Wiley & Sons.

[10] Huihua Kenny Chiang, Chao-Wei Chu, Gau-Yang Chen, and Cheng-Deng Kuo. 2001. A new 3-D display method for 12-lead ECG. *IEEE transactions on biomedical engineering* 48, 10 (2001), 1195–1202.

[11] Ehud Chorin, Ofer Havakuk, Arnon Adler, Arie Steinvil, Uri Rozovski, Christian van der Werf, Pieter G Postema, Guy Topaz, Arthur AM Wilde, Sami Viskin, et al. 2015. Diagnostic value of T-wave morphology changes during 'QT stretching' in patients with long QT syndrome. *Heart Rhythm* 12, 11 (2015), 2263–2271.

[12] Nicolas Debue and Cécile Van De Leemput. 2014. What does germane load mean? An empirical contribution to the cognitive load theory. *Frontiers in psychology* 5 (2014), 1099.

[13] Andrew T Duchowski. 2007. Eye tracking methodology. *Theory and practice* 328 (2007).

[14] NA Mark Estes III. 2013. Computerized interpretation of ECGs: supplement not a substitute.

[15] Food, HHS Drug Administration, et al. 2005. International Conference on Harmonisation; guidance on E14 Clinical Evaluation of QT/QTc Interval Prolongation and Proarrhythmic Potential for Non-Antiarrhythmic Drugs; availability. Notice. *Federal register* 70, 202 (2005), 61134.

[16] Ernest Frank. 1956. An accurate, clinically practical system for spatial vectorcardiography. *circulation* 13, 5 (1956), 737–749.

[17] Pieter G Postema and Arthur AM Wilde. 2014. The measurement of the QT interval. *Current cardiology reviews* 10, 3 (2014), 287–294.

[18] Anubhav Garg and Michael H Lehmann. 2013. Prolonged QT interval diagnosis suppression by a widely used computerized ECG analysis system. *Circulation: Arrhythmia and Electrophysiology* 6, 1 (2013), 76–83.

[19] Ary L Goldberger, Luis AN Amaral, Leon Glass, Jeffrey M Hausdorff, Plamen Ch Ivanov, Roger G Mark, Joseph E Mietus, George B Moody, Chung-Kang Peng, and H Eugene Stanley. 2000. PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals. *Circulation* 101, 23 (2000), e215–e220.

[20] Ilan Goldenberg and Arthur J Moss. 2008. Long QT syndrome. *Journal of the American College of Cardiology* 51, 24 (2008), 2291–2300.

[21] ILAN Goldenberg, Arthur J Moss, and Wojciech Zareba. 2006. QT interval: how to measure it and what is "normal". *Journal of cardiovascular electrophysiology* 17, 3 (2006), 333–336.

[22] Carl Gutwin, Andy Cockburn, and Ashley Coveney. 2017. Peripheral Popout: The Influence of Visual Angle and Stimulus Intensity on Popout Effects. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. ACM, 208–219.

[23] Christopher G Healey. 1996. Choosing effective colours for data visualization. In *Visualization'96. Proceedings*. IEEE, 263–270.

[24] Christopher G Healey, Kellogg S Booth, and James T Enns. 1995. Visualizing real-time multivariate data using preattentive processing. *ACM Transactions on Modeling and Computer Simulation (TOMACS)* 5, 3 (1995), 190–221.

[25] Christopher G Healey, Kellogg S Booth, and James T Enns. 1996. High-speed visual estimation using preattentive processing. *ACM Transactions on Computer-Human Interaction (TOCHI)* 3, 2 (1996), 107–135.

[26] PD Higham and RW Campbell. 1994. QT dispersion. *British heart journal* 71, 6 (1994), 508.

[27] Katerina Hnatkova, Peter Smetana, Ondrej Toman, Georg Schmidt, and Marek Malik. 2016. Sex and race differences in QRS duration. *Ep Europace* 18, 12 (2016), 1842–1849.

[28] Patrick Hoffman, Georges Grinstein, Kenneth Marx, Ivo Grosse, and Eugene Stanley. 1997. DNA visual and analytic data mining. In *Visualization'97., Proceedings*. IEEE, 437–441.

[29] Kenneth Holmqvist, Marcus Nyström, Richard Andersson, Richard Dewhurst, Halszka Jarodzka, and Joost Van de Weijer. 2011. *Eye tracking: A comprehensive guide to methods and measures*. OUP Oxford.

[30] Richard H Hongo and Nora Goldschlager. 2006. Status of computerized electrocardiography. *Cardiology clinics* 24, 3 (2006), 491–504.

[31] Ross Ihaka and Robert Gentleman. 1996. R: a language for data analysis and graphics. *Journal of computational and graphical statistics* 5, 3 (1996), 299–314.

[32] Lars Johannesen, Jose Vicente, JW Mason, Carlos Sanabria, Kristin Waite-Labott, Mira Hong, Ping Guo, John Lin, Jens Stampe Sørensen, Loriano Galeotti, et al. 2014. Differentiating drug-induced multichannel block on the electrocardiogram: randomized study of dofetilide, quinidine, ranolazine, and verapamil. *Clinical Pharmacology & Therapeutics* 96, 5 (2014), 549–558.

[33] Jonathan N Johnson and Michael J Ackerman. 2009. QTc: how long is too long? *British journal of sports medicine* 43, 9 (2009), 657–662.

[34] Ijaz A Khan. 2002. Clinical and therapeutic aspects of congenital and acquired long QT syndrome. *The American journal of medicine* 112, 1 (2002), 58–66.

[35] Paulus Kirchhof, Michael R Franz, Abdennasser Bardai, and Arthur M Wilde. 2009. Giant T–U Waves Precede Torsades de Pointes in Long QT Syndrome: A Systematic Electrocardiographic Analysis in Patients With Acquired and Congenital QT Prolongation. *Journal of the American College of Cardiology* 54, 2 (2009), 143–149.

[36] Paul Kligfield, Fabio Badilini, Isabelle Denjoy, Saeed Babaeizadeh, Elaine Clark, Johan De Bie, Brian Devine, Fabrice Extramiana, Gianluca Generali, Richard Gregg, et al. 2018. Comparison of automated interval measurements by widely used algorithms in digital electrocardiographs. *American heart journal* 200 (2018), 1–10.

[37] Jan A Kors and Gerard van Herpen. 2008. Mirror image electrocardiograms and additional electrocardiographic leads: new wine in old wineskins? *Journal of electrocardiology* 41, 3 (2008), 245–250.

[38] Teri M Kozik and Shu-Fen Wung. 2012. Acquired long QT syndrome: frequency, onset, and risk factors in intensive care patients. *Critical care nurse* 32, 5 (2012), 32–41.

[39] Gaëlle Leroux, Jeanne Spiess, Laure Zago, Sandrine Rossi, Amélie Lubin, Marie-Renée Turbelin, Bernard Mazoyer, Nathalie Tzourio-Mazoyer, Olivier Houdé, and Marc Joliot. 2009. Adult brains don't fully overcome biases that lead to incorrect performance during cognitive development: an fMRI study in young adults completing a Piaget-like task. *Developmental Science* 12, 2 (2009), 326–338.

[40] Lynn S Liben. 1991. Adults' performance on horizontality tasks: Conflicting frames of reference. *Developmental Psychology* 27, 2 (1991), 285.

[41] Emanuela T Locati. 2017. New directions for ambulatory monitoring following 2017 HRS-ISHNE expert consensus. *Journal of electrocardiology* 50, 6 (2017), 828–832.

[42] Robert L Lux, Creig R Smith, Roland F Wyatt, and JA Abildskov. 1978. Limited lead selection for estimation of body surface potential maps in electrocardiography. *IEEE Transactions on Biomedical Engineering* 3 (1978), 270–276.

[43] PW Macfarlane, SC McLaughlin, B Devine, and TF Yang. 1994. Effects of age, sex, and race on ECG interval measurements. *Journal of Electrocardiology* 27 (1994), 14–19.

[44] John E Madias. 2004. The 13th multiuse ECG lead: ShouldnâĂŹt we use it more often, and on the same hard copy or computer screen, as the other 12 leads? *Journal of electrocardiology* 37, 4 (2004), 285–287.

[45] Misha D Miller, J Porter Co-burn, and Michael J Ackerman. 2001. Diagnostic accuracy of screening electrocardiograms in long QT syndrome I. *Pediatrics* 108, 1 (2001), 8–12.

[46] J Morganroth. 2001. Focus on issues in measuring and interpreting changes in the QTc interval duration. *European Heart Journal Supplements* 3, suppl_K (2001), K105–K111.

[47] Arthur J Moss, Peter J Schwartz, Richard S Crampton, Dan Tzivoni, Emanuela H Locati, Jean MacCluer, W Jackson Hall, Lowell Weitkamp, G Michael Vincent, and A Garson Jr. 1991. The long QT syndrome. Prospective longitudinal study of 328 families. *circulation* 84, 3 (1991), 1136–1144.

[48] Stefan P Nelwan, Jan A Kors, Simon H Meij, Jan H van Bemmel, and Maarten L Simoons. 2004. Reconstruction of the 12-lead electrocardiogram from reduced lead sets. *Journal of electrocardiology* 37, 1 (2004), 11–18.

[49] Magnus O Nimmermark, John J Wang, Charles Maynard, Mauricio Cohen, Ian Gilcrist, John Heitner, Michael Hudson, Sebastian Palmeri, Galen S Wagner, and Olle Pahlm. 2011. The impact of numeric and graphic displays of ST-segment deviation levels on cardiologists' decisions of reperfusion therapy for patients with acute coronary occlusion. *Journal of electrocardiology* 44, 5 (2011), 502–508.

[50] Hans-Christoph Nothdurft. 1993. The role of features in preattentive vision: Comparison of orientation, motion and color cues. *Vision research* 33, 14 (1993), 1937–1958.

[51] Alex Page, Mehmet K Aktas, Tolga Soyata, Wojciech Zareba, and Jean-Philippe Couderc. 2016. âĂIJQT clockâĂİ to improve detection of QT prolongation in long QT syndrome patients. *Heart Rhythm* 13, 1 (2016), 190–198.

[52] Alex Page, Tolga Soyata, Jean-Philippe Couderc, and Mehmet Aktas. 2015. An open source ECG clock generator for visualization of long-term cardiac monitoring data. *IEEE Access* 3 (2015), 2704–2714.

[53] Konstantinos Papadopoulos and Eleni Koustriava. 2011. Piaget's water-level task: The impact of vision on performance. *Research in developmental disabilities* 32, 6 (2011), 2889–2893.

[54] Ronald M Pickett, Georges Grinstein, Haim Levkowitz, and Stuart Smith. 1995. Harnessing preattentive perceptual processes in visualization. In *Perceptual issues in visualization*. Springer, 33–45.

[55] Philip J Podrid. 2017. Ambulatory ECG monitoring. *UpToDate, Waltham, MA. Accessed* 6 (2017), 05–18.

[56] Dallas Price. 2010. How to read an electrocardiogram (ECG). Part 1: Basic principles of the ECG. The normal ECG. *South Sudan Medical Journal* 3, 2 (2010), 26–31.

[57] Pentti M Rautaharju. 2007. The birth of computerized electrocardiography: Hubert V. Pipberger (1920-1993). *Cardiology journal* 14, 4 (2007), 420–421.

[58] Pentti M Rautaharju, Borys Surawicz, and Leonard S Gettes. 2009. AHA/ACCF/HRS recommendations for the standardization and interpretation of the electrocardiogram: part IV: the ST segment, T and U waves, and the QT interval a scientific statement from the American Heart Association Electrocardiography and Arrhythmias Committee, Council on Clinical Cardiology; the American College of Cardiology Foundation; and the Heart Rhythm Society Endorsed by the International Society for Computerized Electrocardiology. *Journal of the American College of Cardiology* 53, 11 (2009), 982–991.

[59] Edwin P Rock, John Finkle, Howard J Fingert, Brian P Booth, Christine E Garnett, Stephen Grant, Robert L Justice, Richard J Kovacs, Peter R Kowey, Ignacio Rodriguez, et al. 2009. Assessing proarrhythmic potential of drugs when optimal studies are infeasible. *American heart journal* 157, 5 (2009), 827–836.

[60] Dan M Roden and Prakash C Viswanathan. 2005. Genetics of acquired long QT syndrome. *The Journal of clinical investigation* 115, 8 (2005), 2025–2032.

[61] Takafumi Saito, Hiroko Nakamura Miyamura, Mitsuyoshi Yamamoto, Hiroki Saito, Yuka Hoshiya, and Takumi Kaseda. 2005. Two-tone pseudo coloring: Compact visualization for one-dimensional data. (2005).

[62] Leo Schamroth, Colin Schamroth, and Leo Schamroth. 1966. An introduction to electrocardiography. (1966).

[63] Jürg Schläpfer and Hein J Wellens. 2017. Computer-interpreted electrocardiograms: benefits and limitations. *Journal of the American College of Cardiology* 70, 9 (2017), 1183–1192.

[64] Maully Shah and Christopher Carter. 2008. Long QT syndrome: A therapeutic challenge. *Annals of pediatric cardiology* 1, 1 (2008), 18.

[65] Rashmi R Shah. 2002. The significance of QT interval in drug development. *British journal of clinical pharmacology* 54, 2 (2002), 188–202.

[66] Luke Solomons, Adrian Treloar, and Ryan Noronha. 2008. Competence of psychiatric clinicians in interpreting electrocardiograms and QT intervals: can they do this? Does it matter? *Psychiatric Bulletin* 32, 8 (2008), 291–294.

[67] Stanley Smith Stevens. 2017. *Psychophysics: Introduction to its perceptual, neural and social prospects*. Routledge.

[68] Bruno Taccardi. 1963. Distribution of heart potentials on the thoracic surface of normal human subjects. *Circulation Research* 12, 4 (1963), 341–352.

[69] Soheila Talebi, Azhir Alaleh, Zuber Sam, Soman Sandeep, Ferdinand Visco, Holly Totouom-Tangho, Hossein Kalantari, and Getaw Worku Hassen. 2015. Underestimated and unreported prolonged QTc by automated ECG analysis in patients on methadone: can we rely on computer reading? *Acta cardiologica* 70, 2 (2015), 211–216.

[70] Benoît Tyl, Sara Azzam, Nathalie Blanco, and William Wheeler. 2011. Improvement and limitation of the reliability of automated QT measurement by recent algorithms. *Journal of electrocardiology* 44, 3 (2011), 320–325.

[71] Jose Vicente, Lars Johannesen, Jay W Mason, William J Crumb, Esther Pueyo, Norman Stockbridge, and David G Strauss. 2015. Comprehensive T wave morphology assessment in a randomized clinical study of dofetilide, quinidine, ranolazine, and verapamil. *Journal of the American Heart Association* 4, 4 (2015), e001615.

[72] Sami Viskin, Uri Rosovski, Andrew J Sands, Edmond Chen, Peter M Kistler, Jonathan M Kalman, Laura Rodriguez Chavez, Pedro Iturralde Torres, Osmar A Centurión, Akira Fujiki, et al. 2005. Inaccurate electrocardiographic interpretation of long QT: the majority of physicians

cannot recognize a long QT when they see one. *Heart Rhythm* 2, 6 (2005), 569–574.

[73] Colin Ware. 1988. Color sequences for univariate maps: Theory, experiments and principles. *IEEE Computer Graphics and Applications* 8, 5 (1988), 41–49.

[74] Colin Ware. 2012. *Information visualization: perception for design*. Elsevier.

[75] Eric Welch and S Kobourov. 2017. Measuring symmetry in drawings of graphs. In *Computer Graphics Forum*, Vol. 36. Wiley Online Library, 341–351.

[76] Michele Andrisin Wittig and Mary J Allen. 1984. Measurement of adult performance on Piaget's water horizontality task. *Intelligence* 8, 4 (1984), 305–313.

[77] Ke Xu, Shunan Guo, Nan Cao, David Gotz, Aiwen Xu, Huamin Qu, Zhenjie Yao, and Yixin Chen. 2018. ECGLens: Interactive Visual Exploration of Large Scale ECG Data for Arrhythmia Detection. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. ACM, 663.

[78] Yee Guan Yap and A John Camm. 2003. Drug induced QT prolongation and torsades de pointes. *Heart* 89, 11 (2003), 1363–1372.