

Experiences in digitizing and digitally measuring a paper-based ECG archive

Arttu Holkeri, MD,^{a,*} Antti Eranti, MD,^b Tuomas V. Kenttä, PhD,^c Kai Noponen, MSc,^d
M. Anette E. Haukilahti, BM,^c Tapio Seppänen, PhD,^d M. Juhani Junttila, MD,^c
Tuomas Kerola, MD,^b Harri Rissanen, MSc,^e Markku Heliövaara, MD,^e Paul Knekt, PHD,^e
Aapo L. Aro, MD,^a Heikki V. Huikuri, MD^c

^a Division of Cardiology, Heart and Lung Center, Helsinki University Central Hospital, Meilahti Tower Hospital, PL 340, Helsinki, HUS, Finland

^b Department of Internal Medicine, Päijät-Häme Central Hospital, Keskussairaalankatu 7, Lahti, Finland

^c Research Unit of Internal Medicine, Medical Research Center, Faculty of Medicine, Oulu University Hospital and University of Oulu, P.O. Box 5000, Oulu, Finland.

^d Center for Machine Vision and Signal Analysis, University of Oulu, P.O. Box 4500, Oulu, Finland.

^e National Institute for Health and Welfare, P.O. Box 30, Helsinki, Finland

Abstract

Background: No established method for digitizing and digital measuring of paper electrocardiograms (ECG) exists. We describe a paper ECG digitizing and digital measuring process, and report comparability to manual measurements.

Methods: A paper ECG was recorded from 7203 health survey participants in 1978–1980. With specific software, the ECGs were digitized (ECG Trace Tool), and measured digitally (EASE). A sub-sample of 100 ECGs was selected for manual measurements.

Results: The measurement methods showed good agreement. The mean global (EASE)-(manual) differences were 1.4 ms (95% CI 0.5–2.2) for PR interval, –1.0 ms (95% CI –1.5–[–0.5]) for QRS duration, and 11.6 ms (95% CI 10.5–12.7) for QT interval. The mean inter-method amplitude differences of RampV5, RampV6, SampV1, TampII and TampV5 ranged from –0.03 mV to 0.01 mV.

Conclusions: The presented paper-to-digital conversion and digital measurement process is an accurate and reliable method, enabling efficient storing and analysis of paper ECGs.

© 2017 Elsevier Inc. All rights reserved.

Keywords:

Electrocardiography; Conversion; Digitization

Introduction

For over a century, the standard electrocardiogram (ECG) has been the most important tool in clinical medicine for detecting abnormalities in the electrical activity of the heart [1], and still plays a key role in diagnosing ischemic events and arrhythmias. Due to its non-invasiveness, low cost, and wide-scale availability, the 12-lead ECG as a prognostic tool has been a focus of great interest [2].

Until the digital format was introduced at the end of the 20th century, ECGs were recorded solely on paper [3]. Today, the main format for the recording and storage of ECGs is digital, which enables superior storing, transmission and retrieval potentials, compared to paper ECGs [4]. Furthermore, ECG

measuring and interpretation is nowadays mainly done concurrently with the recording using automated digital software [4], with the capability of fast and highly reproducible electrocardiographic analysis, which also allows the measurement of more complicated ECG parameters [5].

Still, a great number of archived ECGs exist only in paper format, which has spurred continuous efforts to find methods to digitize them. In the process of paper-to-digital conversion, paper ECGs are first scanned into digital images, which are then converted into digital signals, which in turn can be subsequently measured digitally. There have been propositions and requirements for ECG digitization and digital measuring [4,6,7]; however, no single method is currently in universal use, and all previously proposed methods have their own advantages and disadvantages [8,9]. An accurate and reliable paper ECG digitizing and digital measuring method would enable efficient studying of archived paper ECGs of patient cohorts.

* Corresponding author at: Heart and Lung Center, Meilahti Tower Hospital, PL 340, 00029 HUS, Finland.

E-mail address: arttu.holkeri@helsinki.fi

We developed and tested a novel method for scanning, digitizing and digital measuring a paper ECG archive, using the custom-made ECG Trace Tool and EASE computer software. The focus of this study was to describe the used paper-to-digital conversion and digital measuring process, and to assess the reliability of the process, and the correlation between manual and final digital ECG measurements. We also analyzed the ability of the method to differentiate between normal and abnormal intervals.

Methods

Study population, Mini-Finland Health Survey

The study population consisted of subjects from the Mini-Finland Health Survey, which was a part of the Social Insurance Institution's Finnish Mobile Clinic Survey (<https://www.thl.fi/en/web/thlfi-en/research-and-expertwork/population-studies/finnish-mobile-clinic/mini-finland-health-survey>). The survey took place in 1978–1980. The purpose of the survey was to obtain information about the health status of the Finnish population, focusing on prevalent

diseases. Fig. 1 summarizes the study design of the Mini-Finland Health Survey as well as the present ECG digitization and digital measurement process. The study protocol has been published in detail elsewhere [10,11].

Electrocardiographic recordings

The ECGs were recorded in paper format in 1978–1980 by specifically trained and experienced technicians with Kone Oy's Olli 308 ECG device with a paper speed of 50 mm/s and calibration of 10 mm/mV. The device contained an analogue fourth order Butterworth low-pass filter with the corner frequency of 150 Hz (−3 dB) and the bandwidth of the ink plotter (Siemens-Elema) was 0 Hz–650 Hz (−3 dB). The standard 12-lead ECG recordings were approximately 5 s long and recorded three leads at a time (limb leads, augmented limb leads, V1–V3 and V4–V6). In addition to the standard leads, orthogonal leads (X, Y, Z) [12] were recorded on a separate paper sheet for approximately 10 s.

ECG digitization with the ECG Trace Tool

Approximately 35 years later, the five sheets of each ECG were scanned by trained technicians, with a Fujitsu fi-6670 image scanner, into a single Multi-Page TIFF file. The scanner had specific firmware installed to allow long document scanning with a 600 dpi resolution. To enable the digitization of the scanned ECGs into a digital signal format, a proprietary application, the ECG Trace Tool (ETT), was developed by K.N. at the Center for Machine Vision and Signal Analysis (CMVS) at the University of Oulu on the MATLAB platform (MathWorks, Natick, MA). ETT works as a stand-alone graphical-user-interface (GUI) application on Windows, Mac, and Linux operating systems. The core functionality of ETT is the fast, automated image processing and analysis capacity that reduces the amount of user interaction, facilitating the processing of large ECG archives. ETT also provides tailored manual tools for image retouching — e.g., to quickly remove pencil markings or text in the scanned images.

In this study, the digitizing was performed by specifically trained technicians. First, the scanned image was rotated to align the printed grid on the sheet with the digital image pixel coordinate system in ETT. Furthermore, the user specified (cropped) image regions in which to search for a specific lead of the ECG. The aim was to have a minimum of three normal QRS complexes from each of the standard 12 leads. All available data from leads X, Y, and Z were digitized.

Secondly, the ECG lead was automatically traced out from the image with ETT, and the result was further smoothed using a cubic spline to add robustness to noise and tracing errors with both global and local controls for the amount of smoothing. In this work, more smoothing was applied in the more vertical parts of the ECG. The optimization of the several user-specified parameters was performed by A.E. and K.N.

Thirdly, in some cases making a distinction between ECG ink tracings and other markings proved too hard for the automated algorithms, and the user could force the tracing to proceed through certain pixels without editing the image

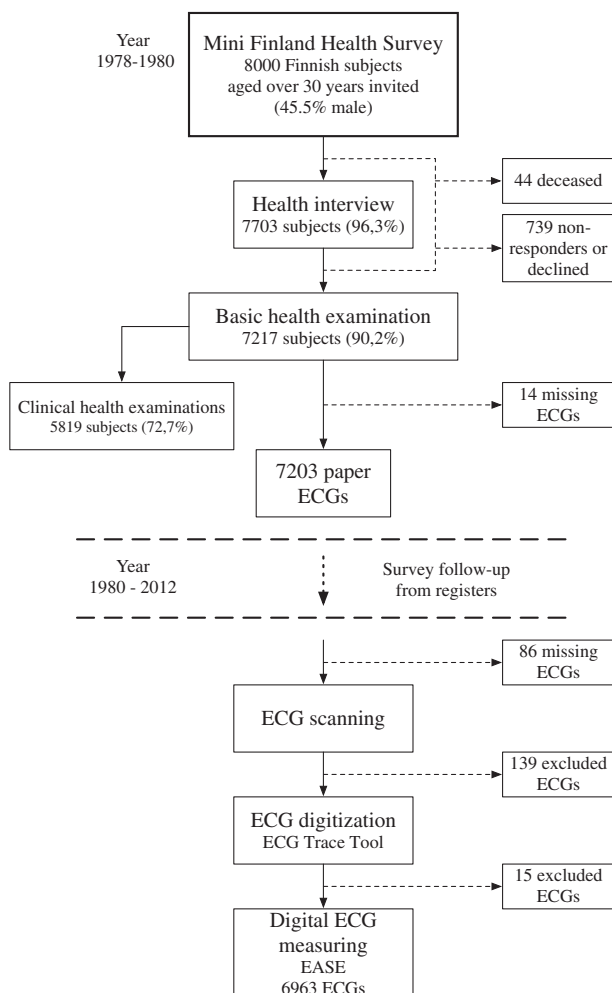


Fig. 1. Flow chart of Mini-Finland Health Survey study design. Caption: A flow chart showing the study design of the Mini-Finland Health Survey, the follow-up registers, and the digitization and digital measurement process of the ECGs.

content. Typically, user intervention was needed if ECG leads intersected, or if there were notes from previous manual measurements or stamps indicating which leads were on the given paper. On average, user intervention was needed in 0–3 leads per ECG, and a single intervention consumed approximately 10–20 s. After corrective measures, the users tried to reprocess the ECG signal as an iterative process.

Finally, the smoothed spline was sampled with output sampling rate of 1181.1 Hz, and the result was band-pass-filtered using a zero-phase Butterworth IIR filtering technique with the corner frequencies of 0.25 Hz and 150 Hz to remove any slow baseline wander and high-frequency noise. The sampling rate and filter corner frequencies were user definable. The horizontal one sample per pixel rate was chosen for this study to avoid aliasing problems. Also, at 600 DPI the width of the blurred ink trace of the ECG's typically spans several pixels. Thus, the quantization noise caused by the vertical pixel resolution should be well below the limits set by any interpretational ambiguity.

After the results for one lead were satisfactory, the user proceeded to the next lead. In this manner, the ECG was processed one lead at a time until all of them had been addressed. Lastly, the result was saved as a digital ECG signal file. Overall, the digitizing speed using this method was approximately 6 min per ECG.

Digital measurement with the electrocardiogram analysis software (EASE)

The digitized signals were analyzed by three examiners (A. Ho., A. E., and A. Ha.) with custom-made ECG analysis software (EASE), which is an interactive program for research-oriented digital ECG analysis. The software detects R peaks from the digitized ECG lead signals and creates low-noise median beats for each of the twelve leads and trimmed mean beats for the XYZ leads; the trimming is done by discarding 20% of the lowest and highest values before calculating the mean. Subsequently, the software locates the P, QRS, and T waves as well as their boundaries, which are marked with red vertical lines as in Figs. 2 and 3.

The analysis in EASE is performed in three stages. In the ECG strip view, the analyst reviews the digitized ECG complexes from each lead set and edits the selection of beats for representative beat formation by removing or adding beats, if necessary—e.g., by removing premature beats and complexes with excessive artifacts. The ECG strip view is also used in the manual grading of the ECG, where abnormal changes in ECG complexes and arrhythmias are identified and coded similarly to the Minnesota code classification system. In the second part of the analysis, the analyst reviews the constructed median beats and edits the automatically detected ECG intervals and baselines, if necessary. The grading of early repolarization (ER), QRS fragmentation, P terminal force, and U waves is also carried out at this point. In the measurement view, leads are displayed in separate windows, which are tiled in a 3×4 matrix. These windows can be zoomed collectively, or each individual lead window

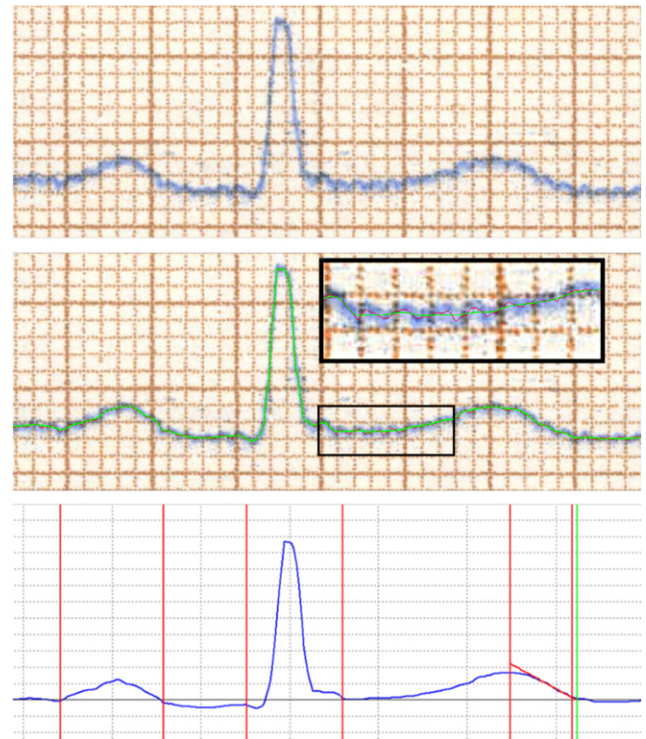


Fig. 2. Demonstration of the digitization and digital measurement process with ECG Trace Tool and EASE. Caption: The top image shows a complex from lead II of a scanned ECG. The middle image shows the same complex with a red tracing line and green smoothing line of ETT. The red tracing line is better seen in the zoomed box of the lead in the image. The bottom image displays the result of the same lead after digitizing and averaging of the complexes in EASE. From left to right, red vertical lines marking the beginning of the P wave, the end of the P wave, the beginning of the QRS complex, the end of the QRS complex, the peak of the T wave, and the end of the T wave. The oblique red line indicates the tangent of the steepest downslope of the T wave, and the green vertical line indicates the end of the T wave with the tangent method.

can be enlarged to a full-screen view. In this full-screen view, all other leads or leads from the same lead set can be superimposed on the background, as demonstrated in Fig. 3. In the final stage, the XYZ data is analyzed in a separate window that combines the functionality of the ECG strip view and the measurement view. Here, the analyst can edit the beat selection for averaging as well as the automatically detected intervals for the orthogonal XYZ leads. The digital measurement accuracy of intervals was 2 ms, and of amplitudes was 0.01 mV. After completing the analysis of the recording, EASE saves the results in spreadsheet format for subsequent statistical analyses. These results contain the manual ECG grading as well as several ECG and vectorcardiographic (VCG) features that are automatically calculated based on the user-verified interval measurements.

Missing ECGs and exclusions

From the 7217 subjects who participated in the basic health examinations, a total of 254 ECGs were missing, not recorded, not scanned, not measured digitally, or excluded. Of these, 14 subjects' ECGs were not recorded for unknown reasons, 86 ECGs were not found for the digitization, 25 ECGs were excluded because of pacemaker rhythm, 2nd or

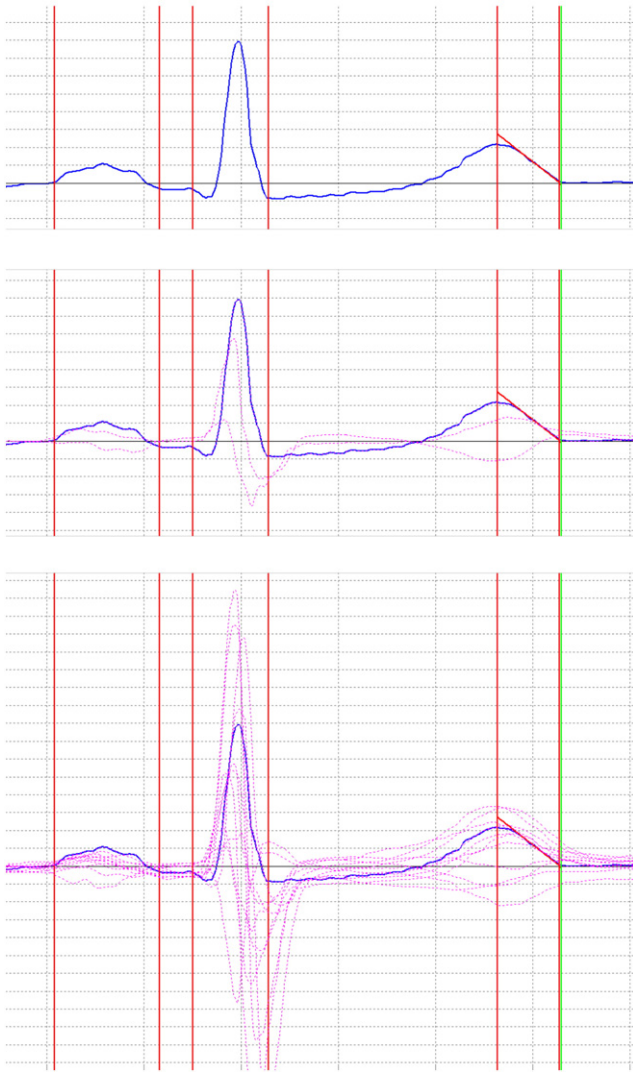


Fig. 3. A demonstration of EASE's feature comparing the lead being currently measured with other leads. Caption: In the top image, a complex from lead I. Notice the end of the QRS complex and the beginning of the ST segment having some disturbances. In the middle image is a comparison with leads from the same paper, and in the bottom image, a comparison with all leads. When compared with other leads, the QRS duration is actually substantially longer than when measured from a single lead. EASE also visualizes numeric values for current interval durations in the measurement window.

3rd degree AV block, rare findings not representing the general population (such as subacute myocardial infarction or probable long QT syndrome), and 129 were excluded due to technical problem with original ECG recording or EASE.

Validation of the digitization and digital measurement process

A total of 6963 participants' ECGs were digitally measured with EASE. One hundred of these served as a sample to validate the used ECG digitizing and digital measuring method. ECGs were selected randomly using a computer algorithm, with 50 ECGs fulfilling criteria of abnormal interval lengths (PR interval ≥ 220 ms, QRS duration ≥ 110 ms or QTc interval ≥ 450 ms in males and ≥ 460 ms in females, both in limb or augmented limb leads

and in precordial leads) or inferior or lateral ER pattern, and 50 ECGs not fulfilling these criteria in EASE measurements. Minimum of at least 10 ECGs from each abnormality group was set in the algorithm. Algorithm used threshold method for the QT interval for the selection of subjects. When re-classified according to tangent method QT interval, this resulted the final sample to contain 48 ECGs with abnormal interval length or lateral or inferior ER pattern and 52 ECGs not fulfilling these criteria.

Blinded from the possible selection criteria of ECGs, one examiner (A. Ho.) manually measured the PR interval, QRS duration, and QT interval from each lead, as well as heart rate from lead X, R wave amplitudes from leads V5 and V6, S wave amplitude from lead V1, and T wave amplitudes from leads II and V5 from the paper ECGs. The manual interval measurements were rounded to the nearest 10 ms and amplitude measurements to the nearest 0.05 mV, due to the capability of reliable eyesight accuracy. For the manual interval measurements, the most distinct complex was used. Manual amplitude measurements were taken using the median of the representative complexes. QT interval was measured using the tangent method in both digital and manual measurements. The ECGs containing the 10 largest inter-method measurement differences in PR intervals, QRS durations, and QT intervals underwent visual re-analysis to find an explanation for the difference.

Statistical analysis

Descriptive statistical analysis provided mean values and 95% confidence intervals. The inter-method differences were calculated as (EASE measurement)–(manual measurement). In addition to individual leads, descriptive analyses of digital and manual measurements were performed globally on PR interval, QRS duration, and QT interval, as a mean of all measurements. Visual assessment of frequency histograms served as a basis for distribution analyses. The Pearson correlation coefficient was used to estimate inter-method correlation, and the statistical significance of the difference between the correlation coefficient and zero was calculated. The inter-method agreement of abnormal intervals was analyzed using Cohen's kappa coefficient and McNemar's test. Statistical significance was set at $p < 0.05$ for all analyses. All analyses were performed using IBM SPSS Statistics version 23.

Results

Comparison between manual and digital measurements

There was a strong agreement between the manual and digitized interval and amplitude measurements, as demonstrated in Tables 1 and 2.

For PR interval and QRS durations, the differences between digital and manual measurements were close to zero, suggesting no systematic over- or under-measuring with either method. The differences in QT intervals were systematically positive, indicating repetitively longer EASE measurements. In individual leads, the mean inter-method differences ranged from -4.3 ms to 5.4 ms for PR interval,

Table 1

Name: Descriptive data on the manual and digital (EASE) interval measurements for the representative leads I, II, aVF, V2, V5, and Z.

	Manual mean (95% CI) (ms)	EASE mean (95% CI) (ms)	Difference mean (95% CI) (ms)	Correlation (Pearson)
PR global	163.7 (161.9–165.4)	165.0 (163.3–166.7)	1.4 (0.5–2.2)	0.881*
QRS global	92.5 (91.6–93.3)	91.4 (90.5–92.3)	–1.0 (–1.5(–0.5))	0.830*
QT global	369.6 (367.7–371.5)	381.0 (379.0–383.1)	11.6 (10.5–12.7)	0.853*
PR I	156.7 (150.2–163.2)	158.5 (151.9–165.0)	1.2 (–1.7–4.0)	0.907*
PR II	166.3 (159.4–173.1)	171.9 (165.0–178.8)	5.2 (3.2–7.3)	0.956*
PR aVF	168.4 (161.1–175.6)	171.3 (165.0–177.6)	2.9 (–0.3–6.2)	0.896*
PR V2	164.2 (157.3–171.2)	166.9 (161.0–172.8)	2.6 (–0.6–5.8)	0.889*
PR V5	160.8 (154.1–167.6)	161.2 (154.8–167.6)	0.5 (–3.5–4.4)	0.824*
PR Z	167.1 (160.1–174.1)	172.5 (165.5–179.5)	5.4 (3.5–7.3)	0.963*
QRS I	86.4 (82.9–89.9)	86.3 (82.8–89.9)	–0.1 (–2.5–2.4)	0.766*
QRS II	91.1 (87.8–94.4)	87.8 (84.1–91.4)	–3.3 (–5.1(–1.6))	0.874*
QRS aVF	93.5 (90.1–96.9)	91.5 (88.1–94.9)	–2.0 (–4.2–0.1)	0.798*
QRS V2	98.1 (95.3–100.9)	96.1 (93.0–99.2)	–2.0 (–3.7(–0.4))	0.841*
QRS V5	91.8 (88.4–95.2)	90.6 (87.2–94.0)	–1.2 (–2.9–0.5)	0.822*
QRS Z	97.2 (94.2–100.2)	96.9 (93.7–100)	–0.3 (–2.0–1.3)	0.849*
QT I	366.5 (359.6–373.4)	380.1 (372.8–387.5)	13.6 (9.4–17.9)	0.824*
QT II	370.2 (363.4–377)	382.9 (375.3–390.6)	11.7 (7.9–15.5)	0.860*
QT aVF	369.5 (361.9–377.1)	379.1 (371.3–386.9)	10.7 (7.1–14.3)	0.889*
QT V2	370.9 (363.4–378.4)	380.8 (373.0–388.6)	9.9 (6.7–13.1)	0.914*
QT V5	372.6 (365.1–380.1)	387.6 (379.8–395.4)	15.3 (11.0–19.6)	0.843*
QT Z	374.8 (367.1–382.5)	384.5 (376.8–392.3)	9.8 (6.2–13.3)	0.895*
Heart rate	69.8 (67.3–72.4) □	69.0 (66.6–71.4) □	–0.8 (–1.4(–0.3)) □	0.978*

Description: Difference calculated as (EASE measurement)–(manual measurement). * = Correlation is significant at the 0.01 level (2-tailed). Heart rate as beats per minute (bpm).

from –3.3 ms to 0.3 ms for QRS duration, and from 4.1 ms to 18.6 ms for QT interval. The digital and manual amplitude measurements of R waves in leads V5 and V6, of the S wave in lead V1, and of T waves in leads II and V5 were similar, with only slight differences.

The proportions of major inter-method measurement differences were small, as demonstrated in Table 3. Assessed as a mean of all leads, the percentages of subjects with ≤ 20 ms inter-method differences were 88.3% for PR intervals, 95.7% for QRS durations, and 76.4% for QT intervals.

Table 4 illustrates the inter-method agreement in prolonged intervals. Both prolonged PR interval and QRS duration had good inter-method agreement, with kappa coefficients of >0.7 . Prolonged QTc had weaker inter-method agreement.

Reasons for major measurement differences

The ECGs with the largest inter-method interval differences underwent visual reanalysis to determine the reason for the difference. Common causes were low-amplitude waves, artifacts, and baseline disturbances, causing difficulty in determining the beginning and end of the waveforms. The

merging of waveforms was also a frequent source of inter-method differences, especially with T or U waves merging with P waves with tachycardic heart rates. A single ECG contained an extremely long PR interval which was unmeasurable with EASE. A diverse ECG measurement practice of a biphasic T wave followed by a prominent discordant U wave was another distinct reason for a major inter-method difference in a single ECG.

Discussion

By means of a comprehensive analysis of 100 subjects' ECGs recorded in paper format over three decades ago, we have demonstrated the feasibility and reliability of a novel process for digitizing and analyzing archived paper ECGs, using custom made software, the ECG Trace Tool and EASE. In estimating heart rate, PR interval, QRS duration, as well as the amplitudes of R, S, and T waves, the novel process demonstrated excellent agreement with manual measurements. However, in estimating QT interval, there were some discrepancies between the two methods, with manual measurements yielding shorter intervals. Nevertheless, the number of subjects with a >20 ms inter-method QT

Table 2

Name: Descriptive data on manual and digital (EASE) amplitude measurements.

	Manual mean (mV) (CI 95%)	EASE mean (mV) (CI 95%)	Difference mean (mV) (95% CI)	Correlation (Pearson)
Lead II, T-wave	0.17 (0.15–0.20)	0.18 (0.16–0.21)	0.01 (0.00–0.02)	0.951*
Lead V1, S-peak	1.12 (1.02–1.23)	1.14 (1.01–1.26)	–0.03 (–0.06–0.00)	0.962*
Lead V5, R-peak	1.64 (1.51–1.77)	1.64 (1.51–1.77)	0.00 (–0.02–0.02)	0.990*
Lead V5, T-wave	0.28 (0.24–0.33)	0.29 (0.25–0.33)	0.01 (0.00–0.02)	0.967*
Lead V6, R-peak	1.32 (1.22–1.42)	1.32 (1.23–1.42)	0.00 (–0.01–0.01)	0.993*

Description: Difference calculated as (EASE measurement)–(manual measurement). * = Correlation is significant at the 0.01 level (2-tailed).

Table 3

Name: Percentage of subjects with >10 ms or >20 ms measurement differences between digital (EASE) and manual measurements of PR intervals, QRS durations, and QT intervals from the representative leads, and mean of all leads.

	Difference > 10 ms [%]	Difference > 20 ms [%]
PR, mean of all leads	34.1	11.7
QRS, mean of all leads	23.6	4.3
QT, mean of all leads	55.4	23.6
PR I	38.3	9.6
PR II	29.2	7.3
PR aVF	27.8	7.2
PR V2	28.9	9.3
PR V5	45.8	14.6
PR Z	27.8	4.1
QRS I	29.0	10.0
QRS II	23.0	7.0
QRS aVF	26.0	5.0
QRS V2	21.0	3.0
QRS V5	20.0	3.0
QRS Z	17.0	2.0
QT I	55.0	23.0
QT II	59.6	22.2
QT aVF	58.6	33.3
QT V2	50.0	13.0
QT V5	55.6	22.2
QT Z	45.9	22.4

interval difference was only 23.6%.

Similar paper-to-digital ECG conversion methods have been reviewed recently [8]. The reviewed methods differed in sample sizes, algorithms, and functions. We used an enriched sample of 100 ECGs for the validation. Our sample size was relatively large compared to the 9 studies reviewed by Waits and Soliman out of which only 3 studies had 100 or more subjects. Different studies also applied different statistical validation methods, ranging from a comparison of interval lengths and amplitudes [13–16] to comparisons of areas and signal overlays [13,17]. Most studies used only the leads with the most distinguishable complexes for the validation [13,15,17,18] or studied only specific parameters [17]. In the present study, all basic intervals from all basic 12 leads as well as the orthogonal X, Y, and Z leads were used in the validation, in addition to the R, S, and T wave amplitudes from the representative leads.

The digital ECG measuring process consists of wave recognition and measurements of durations and amplitudes [19]. Different digital measuring methods are currently widely available and many are included as features in the digital ECG recording devices. Recommendations for the standardization and interpretations of ECGs have been made [4], including formation of a representative single-lead

complex and comparison of simultaneously acquired global complexes to detect accurate onset and offset of waveforms. In EASE, representative single lead complexes are formed, with the ability to modify the selected complexes, if the automated wave recognition doesn't select all the complexes or selects un-representative complexes. In addition, EASE enables comparison of the lead being measured to other leads of the ECG, if needed.

One focus of our interest was to assess how well the digital measurements discriminate between subjects with normal and abnormal intervals as opposed to manual measurements. The kappa-values of inter-method agreement for prolonged intervals were satisfactory for PR intervals and QRS durations; for QTc intervals, the kappa value was lower, which resulted from EASE's over-measurement of QT intervals.

The differences in QT interval measurements between the methods can have several causes. The sources of error in manual measurements include measurement accuracy, the determination of baseline level, and the determination of the T wave downslope tangent. Potential influential factors with EASE are the semi-manual adjustment of the isoelectric line and the digital baseline correction due to artifacts and distortions, as even a minor transition in the baseline affects the total QT interval measured with the tangent method. In addition, measuring the end of a small-amplitude T wave is difficult regardless of the method, as the point of the maximum downslope is difficult to detect. The variability of QT interval measurements with low-amplitude T waves has been shown to be particularly problematic [20]. Furthermore, the difference between the complexes used for the measurements may in part explain the discrepancy between the results, as EASE measured the QT interval from the averaged complex, whereas the manual measurements applied the most distinct complex. Another common cause for QT interval measurement difference is the inter- and intra-examiner difference, as the inter-examiner difference can be >20 ms in manual measurements [21], and the inter-manufacturer QT interval difference between different commercial automatic ECG measurement systems can be >10 ms [22], indicating a universal challenge in analogous QT interval measurements.

The process of ECG digitization and digital measurement with the method employed herein have several advantages over manual measurement of paper ECGs. The paper-to-digital conversion of ECGs enables the storage of vast amounts of patients' old paper ECGs, with minimal physical storage space required for digital ECGs compared to paper ECGs. Digital measurement enables the studying of

Table 4

Name: Inter-method agreement for prolonged PR interval, QRS duration, and QTc interval in the enriched sample of 100 subjects.

Manual prolonged prevalence	EASE prolonged prevalence	Difference of prevalences	p-value (McNemar's test)	Cohen's Kappa	95% Confidence Interval of Kappa
PR 0:08	0:11	0:03	0:25	0.83	0.64–1.02
QRS 0:14	0:20	0:06	0:07	0.72	0.54–0.90
QTc 0:04	0:07	0:03	0:45	0.33	–0.04–0.70

Description: Difference calculated as (EASE measurement)–(manual measurement). Intervals were considered prolonged with a PR interval of ≥ 220 ms, a QRS duration of ≥ 110 ms, or a QTc interval of ≥ 450 ms in males and ≥ 460 ms in females.

more complicated ECG parameters using sophisticated algorithms, which would be difficult or virtually impossible to measure manually [23], and computer software facilitates the automatic re-measurement of the ECGs afterwards. Furthermore, the ECG signal is shown with better resolution on the computer screen compared to a paper ECG, yielding better measuring accuracy, as seen in this study with the 10-ms accuracy in the manual and 2-ms accuracy in the digital measurement of intervals. The smoothening of the ECG signal in ETT and the averaging of the ECG complexes in EASE reduce the amount of baseline disturbance, isoelectric artifacts, and the effect of deviant complexes. Another significant advantage of EASE is the ability to zoom and to superimpose leads in full-screen mode, as demonstrated in Fig. 3, which is especially useful for determining accurate boundaries for low-amplitude waveforms.

Limitations

In the present study, we compared the entire process of scanning, digitization and digital measuring of paper ECGs to manual measurements. Although the accuracy and reliability of the whole process is important for the utilization of the method, the inter-method difference can be influenced by either the digitizing or digital measuring method. The visual re-analysis of the inter-method measurement differences was performed to examine possible factors and flaws to consider while using this method in the future. Many of the major differences resulted from the actual paper ECG signal, including baseline disturbances, artifacts, and difficult or nearly isoelectrical complex morphologies. ECG recordings, whether in paper or digital format, are prone to several signal disturbances and artifacts, which disturbs the analysis and measurement of the ECG [24]. The analysis of a difficult ECG is challenging regardless of the measurement method—however, digital measurement methods can improve the accuracy substantially. Smoothening removes most of the baseline disturbances, and the averaging of the complexes as well as superimposed real-time comparison with other leads helps in the measurement of the durations of difficult complexes. The ECGs of the Mini-Finland Survey were particularly challenging, as they were recorded in 1978–1980 with relatively thick signal ink tracing and with abundant baseline disturbances. A better paper ECG quality could possibly improve the agreement between manual and EASE measurements. In addition to factors concerning the paper ECGs, inter-examiner errors also affected the measurement results, as three examiners carried out the digital measurements. This type of error is possible in all multi-examiner ECG measurements, rather than being specific to this method. With QT intervals, which showed the most inter-method discrepancy in the present study, the inter-observer measurement difference has been shown to be over 20 ms between various cardiologists [21].

Conclusion

In conclusion, using archived ECGs from the large Mini-Finland Health Survey, we demonstrated that the ECG

Trace Tool and EASE software are accurate and promising methods for digitizing and digitally measuring old paper-format ECGs. The use of these tools for the analysis of non-digitized ECGs from different populations and patient cohorts will be helpful in further efforts to explore the prognostic significance of different patterns in the 12-lead ECG.

Funding

This work was supported by the Finnish Foundation for Cardiovascular Research. Arttu Holkeri received funding from Ida Montin's Foundation and the Finnish Medical Foundation. Antti Eranti received funding from Paavo Nurmi's Foundation, Onni and Hilja Tuovinen's Foundation, the Finnish Medical Foundation, the Orion Research Foundation, and the Veritas Foundation.

Funding sources had no involvement in the study design.

Declaration of interest

Conflicts of interest: none.

Acknowledgements

National Institute for Health and Welfare.

The persons participating in the ECG digitization were: Julius Rissanen, Klaus Rissanen, Pihla Martinmäki, Oona Holkeri, Crista Hämäläinen, Natalia Hämäläinen, Tiina Saarinen, Markus Nieminen, and Samuli Koho.

References

- [1] Barold SS. Willem Einthoven and the birth of clinical electrocardiography a hundred years ago. *Card Electrophysiol Rev* 2003;7:99–104.
- [2] Narayanan K, Chugh SS. The 12-lead electrocardiogram and risk of sudden death: current utility and future prospects. *Europace* 2015;17(Suppl 2):ii7–13. <https://doi.org/10.1093/europace/euv121>.
- [3] Pettis KS, Savona MR, Leibrandt PN, Maynard C, Lawson WT, Gates KB, et al. Evaluation of the efficacy of hand-held computer screens for cardiologists' interpretations of 12-lead electrocardiograms. *Am Heart J* 1999;138:765–70.
- [4] Kligfield P, Gettes LS, Bailey JJ, Childers R, Deal BJ, Hancock EW, et al. Recommendations for the standardization and interpretation of the electrocardiogram: part I: the electrocardiogram and its technology 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. *J Am Coll Cardiol* 2007;49:1109–27. <https://doi.org/10.1016/j.jacc.2007.01.024>.
- [5] Reichlin T, Abächerli R, Twerenbold R, Kühne M, Schaer B, Müller C, et al. Advanced ECG in 2016: is there more than just a tracing? *Swiss Med Wkly* 2016;146:w14303. <https://doi.org/10.4414/smw.2016.14303>.
- [6] Baeza R. Electrocardiogram digitization: a practical perspective on the usefulness of a new tool to convert paper electrocardiograms into digital waveform. *J Electrocardiol* 2005;38:321–3. <https://doi.org/10.1016/j.jelectrocard.2005.06.100>.
- [7] Stockbridge N. Points to consider in electrocardiogram waveform extraction. *J Electrocardiol* 2005;38:319–20. <https://doi.org/10.1016/j.jelectrocard.2005.06.090>.
- [8] Waits GS, Soliman EZ. Digitizing paper electrocardiograms: status and challenges. *J Electrocardiol* 2017;38:123–30.

- [9] Martis RJ, Acharya UR, Adeli H. Current methods in electrocardiogram characterization. *Comput Biol Med* 2014;48:133–49, <https://doi.org/10.1016/j.combiomed.2014.02.012>.
- [10] Aromaa A, Heliövaara M, Impivaara O, Knekt P, Maatela J. Aims, methods and study population. Part 1. In: Aromaa A, Heliövaara M, Impivaara O, Knekt P, Maatela J, editors. *The execution of the Mini-Finland Health Survey*; 1989.
- [11] Knekt P, Rissanen H, Järvinen R, Heliövaara M. Cohort profile: the Finnish Mobile Clinic Health Surveys FMC, FMCF and >MFS. *Int J Epidemiol* 2017, <https://doi.org/10.1093/ije/dyx092>.
- [12] Frank E. An accurate, clinically practical system for spatial vectorcardiography. *Circulation* 1956;13:737–49.
- [13] Badilini F, Erdem T, Zareba W, Moss AJ. ECGScan: a method for conversion of paper electrocardiographic printouts to digital electrocardiographic files. *J Electrocardiol* 2005;38:310–8, <https://doi.org/10.1016/j.jelectrocard.2005.04.003>.
- [14] Bhullar HK, Fothergill JC, Goddard WP, de Bono DP. Automated measurement of QT interval dispersion from hard-copy ECGs. *J Electrocardiol* 1993;26:321–31.
- [15] Ravichandran L, Harless C, Shah AJ, Wick CA, McClellan JH, Tridandapani S. Novel tool for complete digitization of paper electrocardiography data. *Transl Eng Health Med* 2013;1, <https://doi.org/10.1109/JTEHM.2013.2262024>.
- [16] Widman LE, Hines LS. Digitization of electrocardiograms by desktop optical scanner. *J Electrocardiol* 1991;24:325–38.
- [17] Karsikas M, Huikuri H, Perkiömäki JS, Lehtola L, Seppänen T. Influence of paper electrocardiogram digitizing on T wave and QRS complex morphology parameters. *Ann Noninvasive Electrocardiol* 2007;12:282–90, <https://doi.org/10.1111/j.1542-474X.2007.00176.x>.
- [18] Mitra S, Mitra M, Chaudhuri BB. Generation of digital time database from paper ECG records and Fourier transform-based analysis for disease identification. *Comput Biol Med* 2004;34:551–60, <https://doi.org/10.1016/j.combiomed.2003.08.001>.
- [19] Hongo RH, Goldschlager N. Status of computerized electrocardiography. *Cardiol Clin* 2006;24:491–504, <https://doi.org/10.1016/j.ccl.2006.03.005> [x].
- [20] McLaughlin NB, Campbell RW, Murray A. Accuracy of four automatic QT measurement techniques in cardiac patients and healthy subjects. *Heart* 1996;76:422–6.
- [21] Murray A, McLaughlin NB, Bourke JP, Doig JC, Furniss SS, Campbell RW. Errors in manual measurement of QT intervals. *Br Heart J* 1994;71:386–90.
- [22] Sano M, Aizawa Y, Katsumata Y, Nishiyama N, Takatsuki S, Kamitsuji S, et al. Evaluation of differences in automated QT/QTc measurements between Fukuda Denshi and Nihon Koden systems. *PLoS One* 2014;9, <https://doi.org/10.1371/journal.pone.0106947>.
- [23] Waks JW, Sitlani CM, Soliman EZ, Kabir M, Ghafoori E, Biggs ML, et al. Global electric heterogeneity risk score for prediction of sudden cardiac death in the general population: the Atherosclerosis Risk in Communities (ARIC) and Cardiovascular Health (CHS) studies. *Circulation* 2016;133:2222–34, <https://doi.org/10.1161/CIRCULATIONAHA.116.021306>.
- [24] Garcia-Niebla J, Llontop-García P, Valle-Racero JI, Serra-Autonell G, Batchvarov VN, de Luna AB. Technical mistakes during the acquisition of the electrocardiogram. *Ann Noninvasive Electrocardiol* 2009;14:389–403, <https://doi.org/10.1111/j.1542-474X.2009.00328.x>.