smw · swiss medical weekly

Original article | Published 01 September 2023 | doi: https://doi.org/10.57187/smw.2023.40096
Cite this as: Swiss Med Wkly. 2023;153:40096

Accuracy in detecting atrial fibrillation in single-lead ECGs: an online survey comparing the influence of clinical expertise and smart devices

Simon Weidlich[ab*], Diego Mannhart[ab*], Teodor Serban[ab], Philipp Krisai[ab], Sven Knecht[ab], Jeanne du Fay de Lavallaz[ab], Tatjana Müller[ab], Beat Schaer[ab], Stefan Osswald[ab], Michael Kühne[ab], Christian Sticherling[ab], Patrick Badertscher[ab]

[a] Department of Cardiology, University Hospital Basel, Basel, Switzerland
[b] Cardiovascular Research Institute Basel, University Hospital Basel, University of Basel, Basel, Switzerland
[*] These authors contributed equally to this work and share first authorship

Patrick Badertscher, MD
University Hospital Basel
Petersgraben 4
CH-4031 Basel
patrick.badertscher[at]
ubs.ch

Summary

BACKGROUND: Manual interpretation of single-lead ECGs (SL-ECGs) is often required to confirm a diagnosis of atrial fibrillation. However accuracy in detecting atrial fibrillation via SL-ECGs may vary according to clinical expertise and choice of smart device.

AIMS: To compare the accuracy of cardiologists, internal medicine residents and medical students in detecting atrial fibrillation via SL-ECGs from five different smart devices (Apple Watch, Fitbit Sense, KardiaMobile, Samsung Galaxy Watch, Withings ScanWatch). Participants were also asked to assess the quality and readability of SL-ECGs.

METHODS: In this prospective study (BaselWearableStudy, NCT04809922), electronic invitations to participate in an online survey were sent to physicians at major Swiss hospitals and to medical students at Swiss universities. Participants were asked to classify up to 50 SL-ECGs (from ten patients and five devices) into three categories: sinus rhythm, atrial fibrillation or inconclusive. This classification was compared to the diagnosis via a near-simultaneous 12-lead ECG recording interpreted by two independent cardiologists. In addition, participants were asked their preference of each manufacturer's SL-ECG.

RESULTS: Overall, 450 participants interpreted 10,865 SL-ECGs. Sensitivity and specificity for the detection of atrial fibrillation via SL-ECG were 72% and 92% for cardiologists, 68% and 86% for internal medicine residents, 54% and 65% for medical students in year 4–6 and 44% and 58% for medical students in year 1–3; p <0.001. Participants who stated prior experience in interpreting SL-ECGs demonstrated a sensitivity and specificity of 63% and 81% compared to a sensitivity and specificity of 54% and 67% for participants with no prior experience in interpreting SL-ECGs (p <0.001). Of all participants, 107 interpreted all 50 SL-ECGs. Diagnostic accuracy for the first five interpreted SL-ECGs was 60% (IQR 40–80%) and diagnostic accuracy for the last five interpreted SL-ECGs was 80% (IQR 60–90%); p <0.001. No significant difference in the accuracy of atrial fibrillation detection was seen between the five smart devices; p = 0.33. SL-ECGs from the Apple Watch were considered as having the best quality and readability by 203 (45%) and 226 (50%) participants, respectively.

CONCLUSION: SL-ECGs can be challenging to interpret. Accuracy in correctly identifying atrial fibrillation depends on clinical expertise, while the choice of smart device seems to have no impact.

Introduction

There are several FDA-cleared and CE-marked smart devices on the market capable of recording a 30-second single-lead ECG (SL-ECG) and interpreting it via automated algorithms for rhythm classification. According to 2020 ESC guidelines, smart devices are approved for the diagnosis of atrial fibrillation by manual interpretation [1]. The current manufacturers' algorithms are able to classify tracings into "signs of atrial fibrillation", "sinus rhythm" or "inconclusive". Numerous studies have shown a high rate (19–33%) of inconclusive tracings [2–9]. The data surge caused by more and more (inconclusive) patient-initiated SL-ECG tracings is challenging for today's healthcare system [10, 11]. Physicians and medical personnel need to be advised on how to assess and interpret these recordings [12–14], especially since a majority of healthcare professionals use or intend to use smart devices [15, 16].

It is still being determined how the performance of manually interpreted tracings varies between different smart devices. While a frequency range between 0.5 Hz and 150 Hz has been recommended for standard 12-lead ECG [4], the technical specifications of the acquired ECG signal for the currently available smartwatches need to be clarified. There are major differences to be observed in the PDF exports of SL-ECGs from different manufacturers including the orientation of the PDF (landscape or portrait), the thickness and colouring of the waveform and the grid design. Other differences may include the maximal amplitude to be shown on the export, as well as the possibility of enhancement (vectorised or rasterised export). Figure 1 provides a brief overview of the differences between the exportable tracings of five commercially available smart devices. Compared to the standardised 12-lead ECG tracing, with its thin waveform and light grid behind it, these SL-ECG tracings differ in various aspects to what clinicians are used to in their everyday practice and this might impact the diagnostic accuracy of manual interpretation. In addition, the quality and readability of SL-ECGs might differ between different smart devices. This topic has, to our knowledge, not been systematically assessed.

Although the first SL-ECG device was introduced over 10 years ago [17] and new uses of smart devices are currently being explored [18–20], the amount of training and experience required to reliably interpret recordings generated by this novel technology remain uncertain. Currently there are no data or studies addressing this topic.

This study aimed to assess the impact of the smart devices and level of clinical expertise affect accuracy in detecting atrial fibrillation in single-lead ECGs based on exports of PDF tracings from the same patient cohort. In addition, the quality and readability of SL-ECGs were assessed by all participants.

Materials and methods

Study design

In this prospective comparative study, we sent email messages with links to medical students at Swiss universities and to physicians of varying training levels at major Swiss hospitals inviting them to interpret SL-ECGs in an online survey (REDCap electronic data capture tools [RRID:SCR_003445] hosted at University Hospital Basel, Switzerland) [21, 22]. The deans' offices from the universities of Basel, Bern, Zürich, Lausanne, Geneva and Università Svizzera italiana were contacted via email. The survey links were distributed among all enrolled medical students once between February 2022 and November 2022. Medical students were invited to share the survey link with colleagues at other Swiss universities (snowball sampling technique). Regarding clinical experts, the survey link was sent once between February 2022 and November 2022 to the cardiology departments of University Hospital Basel, Zürich, Bern and Lausanne and to the department of internal medicine of University Hospital Basel. The survey was accompanied by a cover letter explaining the study's purposes.

The study was carried out according to the principles of the Declaration of Helsinki, was preregistered (clinicaltrial.gov, NCT04809922) and approved by the local ethics committees (BASEC ID 2020-02425). All survey participants included in the analysis provided digital informed consent. The authors designed the study, gathered and analysed the data according to STROBE guidelines [23] for observational studies, vouched for the data and analysis, wrote the paper and decided to submit the manuscript for publication.

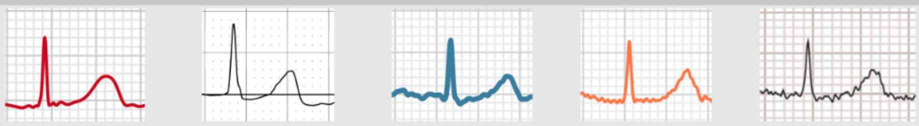
Inclusion criteria

Participants had to be aged 18 or older at the time of their participation and to have provided online confirmed consent. In addition, medical students had to provide an institutional email address. Participants who failed to provide online confirmed consent, failed to complete the baseline questionnaire or interpreted fewer than five SL-ECGs were excluded from the final analysis.

Study set-up

Single-lead ECGs were recorded as part of the BASEL Wearable Study (NCT04809922). Patients who were scheduled for catheter ablation procedures, electrical cardioversions, pacemaker (PM) or implantable cardioverter defibrillator (ICD) implantation and provided written con-

Figure 1: Overview of PDF export characteristics by manufacturer based on tracings recorded in June 2021.

	Apple	AliveCor	Fitbit	Samsung	Withings
PDF-orientation	Landscape	Portrait	Portrait	Landscape	Landscape
Thickness of waveform (relative to gridline)	2.5	1	3	2.5	1
Coloring of waveform	Yes (red)	No (black)	Yes (blue)	Yes (orange)	No (black)
Background	white	white	white	white	white
Grid design	5mm lines (dark grey) 1mm lines (same thickness as 5mm lines, light grey)	Second lines (black) 5mm lines (0.5 thickness of second lines, grey) 1mm dots (thickness of grid/second lines, light grey)	5mm lines (dark grey) Two 1mm lines (0.5 thickness of 5mm lines, lighter grey) with slight offset to one another to generate 3d-impression	5mm lines (dark grey) 1mm lines (same thickness as 5mm lines, light grey)	5mm lines (red) 1mm lines (same thickness as 5mm grid, light red)
Limitation of amplitude	±1.5mV	≥ ± 3.0mV	±1.5mV	±1.5mV	±1.5mV
Vectorised waveform	yes	yes	yes	yes	no
Example					

sent were included. From this cohort, five patients with atrial fibrillation as the underlying heart rhythm and five patients with sinus rhythm were chosen. Five immediate sequential SL-ECGs were recorded from five smart devices: Apple Watch 6® (Apple Inc., Cupertino, California, USA), Fitbit Sense® (Fitbit, San Francisco, California, USA), AliveCor KardiaMobile® (AliveCor, Mountain View, USA), Samsung Galaxy Watch3® (Samsung, Seoul, South Korea), Withings ScanWatch® (Withings, Issy-les-Moulineaux, France). This resulted in 25 SL-ECGs with sinus rhythm and 25 SL-ECGs with atrial fibrillation. The first complete SL-ECG measurements were taken in each case. The patient's rhythm was determined on an almost simultaneously recorded 12-lead ECG interpreted by two independent cardiac electrophysiologists acting as the gold standard. The PDF exports were then cropped to only show the tracing without revealing the automated measurements such as the heart rate, the proposed diagnosis or the manufacturer's details, as these all could influence the interpretation.

All modified PDF exports were then randomised and reallocated, so that all participants had the same order of SL-ECGs in the survey. Participants could perform the survey using whichever electronic device they preferred (smartphone, tablet, laptop, desktop computer). The baseline questionnaire collected information on age, sex, medical expertise (student/resident/specialist), whether participants had previous experience with at least five SL-ECGs and the average number of ECGs viewed and interpreted per week. After interpreting 5 SL-ECGs as atrial fibrillation, sinus rhythm or inconclusive, the participants were asked whether they wanted to finish the questionnaire or continue to another 5 SL-ECGs. Participants could choose to continue an additional nine times after the initial 5 SL-ECGs (resulting in a maximum of 50 SL-ECG interpretations). At the end of the survey, participants were asked which devices provided the best and worst quality and the best and worst readability; this question was illustrated with a figure showing artefact-free SL-ECG recordings from the five smart devices side by side.

Statistical analysis

Continuous variables are presented as means and standard deviations (SD) or medians and interquartile ranges (IQR); categorical variables as numbers and percentages. The T-test was used for continuous, normally distributed data, the Wilcoxon test for skewed variables, paired comparisons were conducted if indicated. For comparison of multiple groups, the Kruskal-Wallis test was performed for the overall comparisons and the pairwise Wilcoxon rank-sum tests for pairwise comparisons in the analysed group. Categorical variables were compared using Chi-square or Fisher's exact test as appropriate. No recordings had to be excluded from the analysis. The accuracy for detection of atrial fibrillation, sensitivity and specificity of each participant were calculated and compared to the cardiac electrophysiologist-interpreted 12-lead ECG as the reference standard. The primary outcome was the accuracy, sensitivity and specificity of each smart device stratified by the level of medical expertise (specialist cardiology; resident; master's student [year 4–6]; bachelor's student [year 1–3]) for the detection of atrial fibrillation. Secondary outcomes were comparisons of previous experience with SL-ECG and general ECG interpretations per week, as well as first and last five SL-ECGs among participants who interpreted all 50 SL-ECGs. Inconclusive answers were deemed false for the accuracy calculation and either false-positive if the gold standard was sinus rhythm or false-negative if the gold standard was atrial fibrillation, for the calculation of the sensitivity and specificity. All analyses were two-tailed and p values of <0.05 were considered statistically significant. All statistical analyses were performed using R version 4.0.2 (R Foundation for statistical computing, Vienna, Austria) with RStudio (version 2022.12.0+353). All software programs used were open-source and publicly accessible.

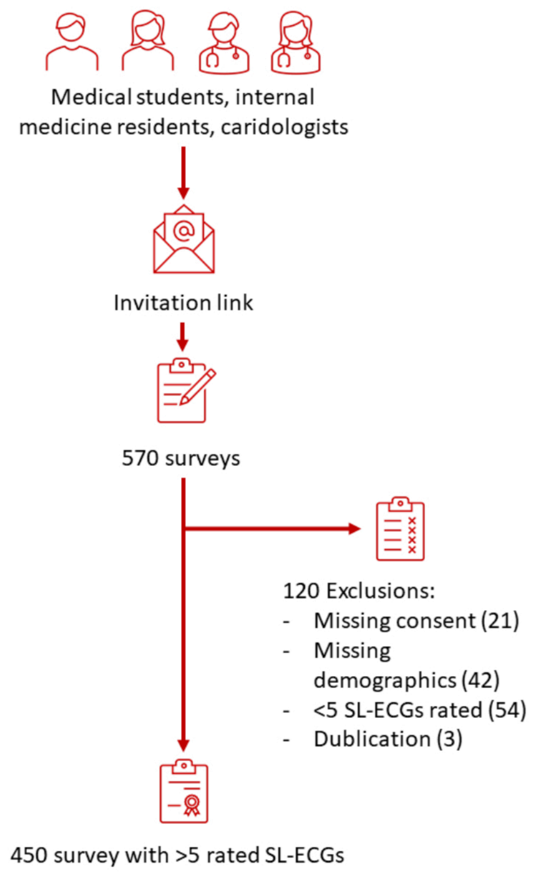
Results

Baseline data

In this prospective, multicentre study, 570 participants started the survey. From these, 21 were excluded due to missing consent, 42 due to missing demographic information, 54 due to fewer than five interpreted SL-ECGs and 3 due to duplication (figure 2), leaving a total of 450 participants who were included between February and November 2022.

Their mean age was 27 years (SD 6.9 years); 55% were female; 70% were medical students (15% bachelor's students and 55% master's students), 21% were internal medicine residents, 6% were board-certified cardiologists and 3% were other specialists or medical professionals; 63%

Figure 2: Study flowchart.



stated that they assessed ECGs of some sort weekly. Previous smartwatch ECG experience was present in 21% of participants. We included students from eight universities in Switzerland, all offering a master's degree course in medicine (table 1).

Overall 10,865 SL-ECGs were interpreted with an overall median accuracy of 64% (IQR 52–78%). Of these, 2337 (22%) were deemed 'inconclusive'. The breakdown of the number of SL-ECGs interpreted per participant is: 50 by 107 (24%) participants; 45 by 2; 40 by 9; 35 by 18; 30 by 16; 25 by 31; 20 by 48; 15 by 69; 10 by 86; and 5 by 64. On average, participants interpreted 20 SL-ECGs (IQR 10–40). The best-interpreted SL-ECG had an accuracy of 97% and the worst-interpreted SL-ECG had an accuracy of 10%.

– 1869 SL-ECGs from the Apple Watch 6 were interpreted with a median accuracy of 65% (IQR 48–90%), sensitivity of 48% (CI 45–51%) and specificity of 86% (CI 84–89%).
– 2128 SL-ECGs from the Fitbit Sense were interpreted with a median accuracy of 69% (IQR 63–81%), sensitivity of 68% (CI 65–71%) and specificity of 73% (CI 71–76%).
– 2212 SL-ECGs from AliveCor KardiaMobile were interpreted with a median accuracy of 74% (IQR 58–85%), sensitivity of 62% (CI 59–64%) and specificity of 83% (CI 80–85%).
– 1955 SL-ECGs from Samsung Galaxy Watch3 were interpreted with a median accuracy of 50% (IQR 41–80%), sensitivity of 51% (CI 48–54%) and specificity of 66% (CI 63–69%).
– 2701 SL-ECGs from Withings ScanWatch were interpreted with a median accuracy of 49% (IQR 40–70%), sensitivity of 54% (CI 51–56%) and specificity of 51% (CI 49–54%).

There was no significant difference in the accuracy of these devices, p = 0.33 (figure 3).

The automated algorithms of each smart device resulted in an accuracy of 50–70%, a sensitivity of 40–75% and a specificity of 40–67%.

Comparison between different smart devices regarding quality and readability

Single-lead ECGs from the Apple Watch 6 were ranked as having the best quality and readability by 203 (45%) and 226 (50%) participants, respectively (figure 3). The 226 participants who ranked the SL-ECGs from Apple Watch 6 as having the best readability between them interpreted 1148 SL-ECGs with a median accuracy of 67% (IQR 50–83%) from the Apple Watch 6 and 5162 SL-ECGs with a median accuracy of 63% (IQR 52–77%) from the other smart devices, p = 0.003. Inversely, 182 and 172 participants ranked the SL-ECGs from Withings ScanWatch as having the worst quality and worst readability (figure 4). A total of 1103 SL-ECGs from Withings ScanWatch were rated with the worst readability and interpreted with a median accuracy of 50% (IQR 40–70%). In the other 3277 smart devices rated by these participants, the median accuracy was 70% (IQR 58–83%); p <0.001.

Smartwatch ECG interpretation between groups of different expertise

Board-certified cardiologists interpreted a total of 590 ECGs with a median accuracy of 81% (IQR 75–90%), sensitivity of 72% (CI 67–77%) and specificity of 92% (CI 88–95%). Internal medicine residents interpreted 1930 SL-ECGs with a median accuracy of 80% (IQR 70–80%), sen-

Table 1:
Baseline characteristics of the groups with different level of expertise. "Other" consisted of: other specialist (n = 8) and other professional in the medical field (n = 5).

	Cardiology specialist	Resident	Master's student	Bachelor's student	Other	Total
	(n = 26)	(n = 95)	(n = 250)	(n = 66)	(n = 13)	(n = 450)
Sex (%)						
Female	5 (19%)	43 (45%)	151 (60%)	43 (65%)	6 (46%)	248 (55%)
Indifferent	1 (4%)	0 (0%)	3 (1%)	1 (2%)	0 (0%)	5 (1%)
Male	20 (77%)	52 (55%)	96 (38%)	22 (33%)	7 (54%)	197 (44%)
Age (years) (mean [SD])	45 (12)	30 (3.4)	25 (2.3)	22 (2.4)	38 (11)	27 (6.9)
Experience with smart device ECGs = Yes (%)	18 (69%)	23 (24%)	38 (15%)	8 (12%)	6 (46%)	93 (21%)
University (%)						
Basel	0 (0%)	0 (0%)	111 (44%)	12 (18%)	0 (0%)	123 (27%)
Bern	0 (0%)	0 (0%)	55 (22%)	29 (44%)	0 (0%)	84 (19%)
Zurich	0 (0%)	0 (0%)	32 (13%)	10 (15%)	0 (0%)	42 (9%)
Lausanne	0 (0%)	0 (0%)	28 (11%)	14 (21%)	0 (0%)	42 (9%)
Geneva	0 (0%)	0 (0%)	18 (7%)	0 (0%)	0 (0%)	18 (4%)
USI	0 (0%)	0 (0%)	2 (1%)	0 (0%)	0 (0%)	2 (0%)
Lucerne	0 (0%)	0 (0%)	1 (0%)	0 (0%)	0 (0%)	1 (0%)
St. Gallen	0 (0%)	0 (0%)	1 (0%)	0 (0%)	0 (0%)	1 (0%)
Missing			2 (0.8%)	1 (1.5%)		
Average ECGs looked at per week (%)						
<1	7 (27%)	6 (6%)	2 (1%)	0 (0%)	0 (0%)	15 (3%)
1–5	1 (4%)	8 (8%)	109 (44%)	47 (71%)	1 (8%)	166 (37%)
5–25	1 (4%)	24 (25%)	118 (47%)	18 (27%)	3 (23%)	164 (36%)
25–50	10 (38%)	19 (20%)	4 (2%)	0 (0%)	1 (8%)	34 (8%)
>50	7 (27%)	38 (40%)	17 (7%)	1 (2%)	8 (62%)	71 (16%)

SD: standard deviation; ECG: electrocardiogram; USI: Università della Svizzera italiana.

sitivity of 68% (CI 65–71%) and specificity of 86% (CI 83–88%). Master's medical students interpreted 6515 SL-ECGs with a median accuracy of 60% (IQR 50–70%), sensitivity of 54% (CI 52–56%) and specificity of 65% (CI 63–67%). Bachelor's medical students interpreted 1590 SL-ECGs with a median accuracy of 50% (IQR 40–60%), sensitivity of 44% (CI 40–47%) and specificity of 58% (CI 54–61%) (table 2 and figure 5).

Difference in groups with prior single-lead ECG experience and number of ECGs interpreted per week

Participants who stated they had prior SL-ECG experience interpreted 2210 SL-ECGs with an accuracy of 73% (IQR 55–80%), a sensitivity of 63% (CI 60–66%) and a specificity of 81% (CI 78–83%). In comparison, participants without prior experience interpreted 8655 SL-ECGs with an accuracy of 60% (IQR 50–74%), p <0.001, a sensitivity of 54% (CI 53–56%) and a specificity of 67% (CI 65–68%).

Comparison between groups with different estimated ECGs interpreted per week (<1, 1–5, 5–25, 25–50, >50)

Figure 3: Boxplots illustrating the diagnostic accuracy of all participants in correctly identifying atrial fibrillation for each of the five smart devices evaluated. Overall comparison with the Kruskal-Wallis test, p = 0.33. The purple line indicates the median values for each smart-device.

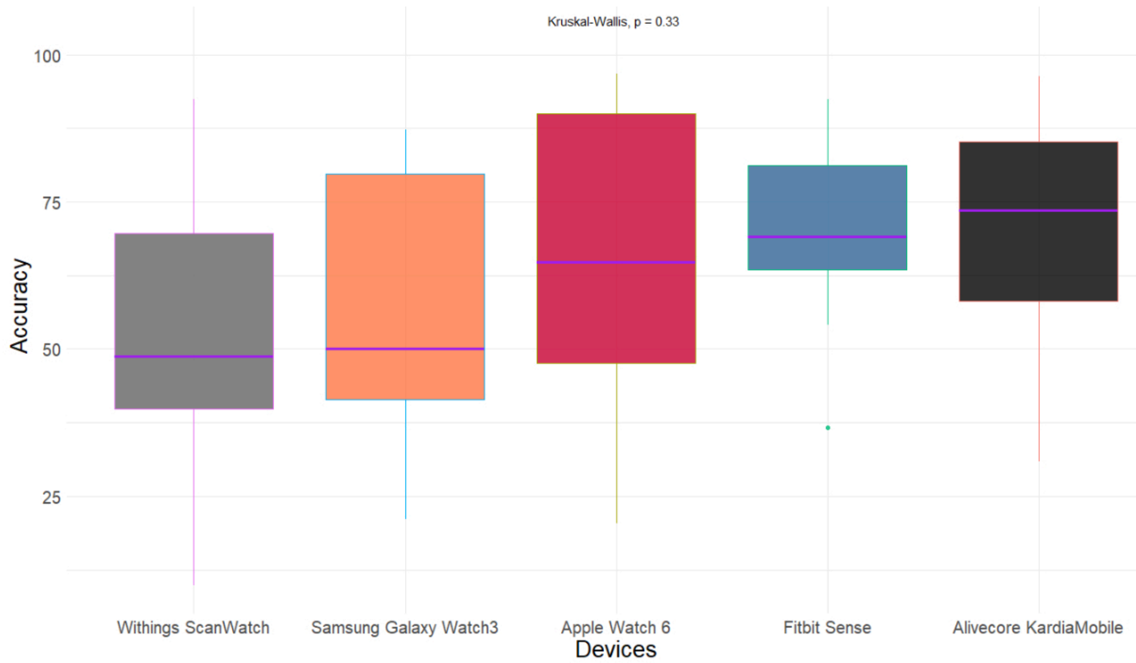
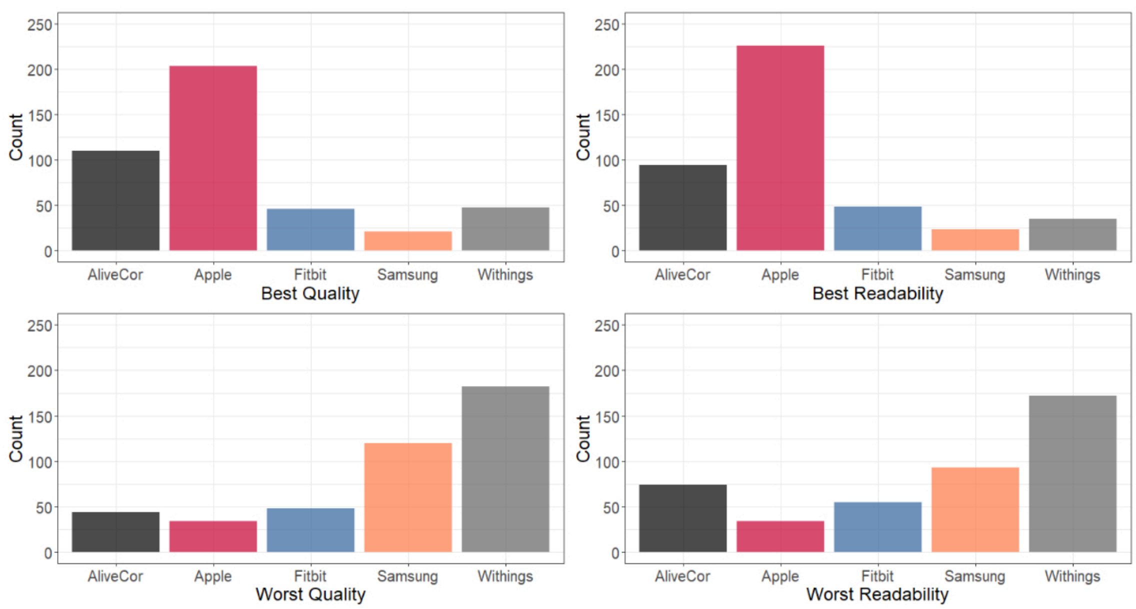


Figure 4: All five smart devices were rated by all participants for best quality (top left), best readability (top right), worst quality (bottom left), and worst readability (bottom right).



showed that the accuracy increased with the number of ECGs seen per week: 56%, 64%, 73%, 80% and 78%, respectively; p <0.001.

Learning curve throughout the single-lead ECG questionnaire

107 participants interpreted all 50 SL-ECGs. The median accuracy in the first 5 ECGs was 60% (IQR 40–80%) and increased in the last 5 SL-ECGs to 80% (IQR 60–90%); p <0.001. Sensitivity and specificity in the first 5 SL-ECGs were 59% (CI 54–65%) and 71% (CI 64–77%), respectively. In the last 5 SL-ECGs, sensitivity was 65% (CI 59–72%) and specificity was 77% (CI 72–82%).

Discussion

In this multicentre prospective study, the accuracy of manual interpretation was assessed for 10,865 SL-ECGs from 5 different smart devices and medical personnel with varying levels of expertise. In addition, the quality and readability of the SL-ECGs were assessed by all participants. As a result, we report the following main findings:

First, single-lead ECGs were recorded with a short instruction and without repetition, which resulted in SL-ECGs with various qualities. Among 10,865 SL-ECGs, 2337 (22%) were deemed inconclusive by participants, decreasing accuracy, sensitivity and specificity. Second, overall accuracy, sensitivity and specificity in correctly classifying SL-ECGs did not differ between the smart devices evaluated. Third, we found significant differences when the five manufacturers' smart devices were rated for readability and quality. SL-ECGs from the Apple Watch 6 were ranked best, while SL-ECGs from the Withings ScanWatch were ranked worst for quality and readability. Fourth, participants' accuracy was better for highly-ranked SL-ECGs regarding quality and readability, and worse for SL-ECGs with lower-ranked quality and readability. Fifth, accuracy, sensitivity, and specificity in correctly classifying SL-ECGs highly correlated with the level of expertise. Sixth, previous experience in interpreting SL-ECGs increased accuracy significantly. Similarly, a learning curve was seen throughout the SL-ECG questionnaire with better accuracy, sensitivity and specificity in the last 5 SL-ECGs compared to the first 5 SL-ECGs of the questionnaire.

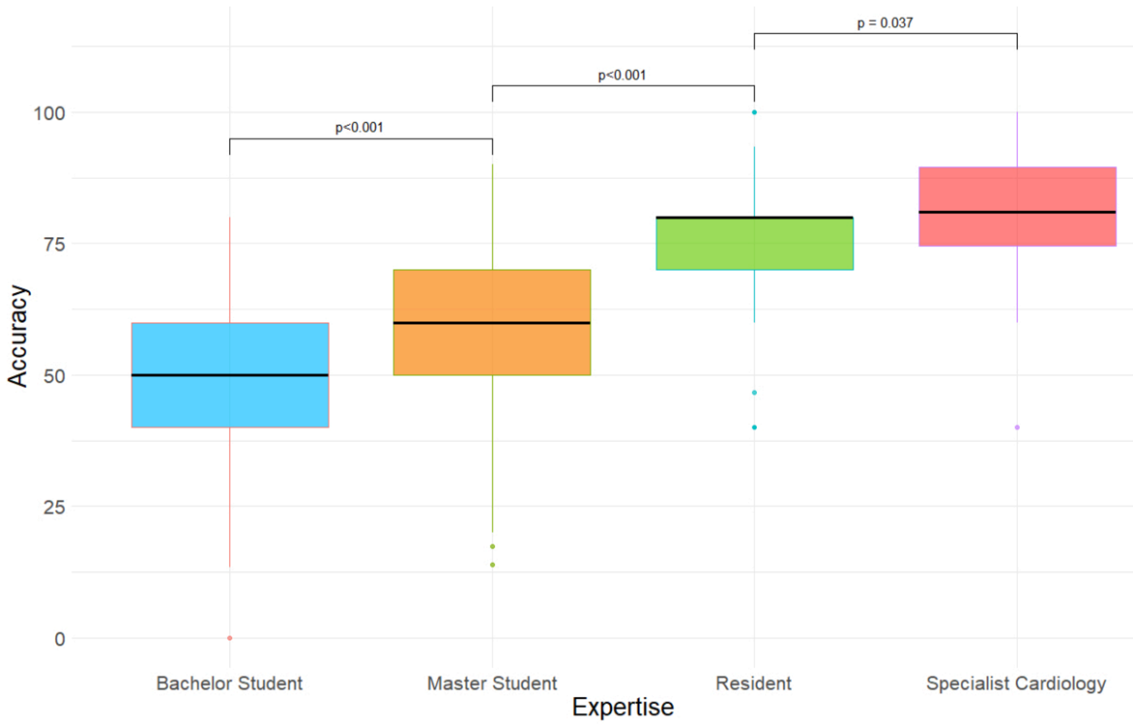
Our findings corroborate and extend the findings of previous literature assessing the accuracy of interpreting SL-ECGs from wearable smart devices. Furthermore, to the best of our knowledge, our study is the first to assess diagnostic accuracy of manual interpretation in correctly classifying rhythm via SL-ECGs from 5 different smart devices and the first to assess the quality and readability of SL-

Table 2:
Overall median accuracy in detecting atrial fibrillation, by level of expertise and overall.

	Cardiology specialist (n = 26)	Resident (n = 95)	MA medical student (n = 250)	BA medical student (n = 66)	Total (n = 437)
Accuracy in detecting AF (median [IQR])	81 (75–90)	80 (70–80)	60 (50–70)	50 (40–60)	62 (52–77)
Sensitivity (CI)	72% (67–77%)	68% (65–71%)	54% (52–56%)	44% (40–47%)	56% (55–57%)
Specificity (CI)	92% (88–95%)	86% (83–88%)	65% (63–67%)	58% (54–61%)	69% (68–70%)

AF: atrial fibrillation; IQR: interquartile range; CI: confidence interval.

Figure 5: Boxplot illustrating diagnostic accuracy in correctly identifying atrial fibrillation according to level of clinical expertise. Overall comparison with the Kruskal-Wallis test showed p <0.001. The black line indicates the median values for each smart device.



ECGs from 5 different smart devices in a large group of participants.

Our study highlights that SL-ECG tracing quality can differ between models with a possible direct impact on their diagnostic value. Our findings might help to understand better the strengths and limitations of each of these smart devices. Since previous studies mainly used only one smart device like the Apple Watch [5, 6] or the KardiaMobile [2, 3, 24–27] a direct comparison between different smart devices regarding the quality and readability of the SL-ECGs and their impact on the diagnostic value was not undertaken. Abu-Alrub et al. [28] compared three smart devices (Apple Watch, Samsung Galaxy Watch and Withings ScanWatch) and found the highest rating of difficult or uninterpretable SL-ECGs for the Samsung Galaxy Watch. Similarly to this study, the lower quality rendered SL-ECG interpretation more difficult and resulted in decreased sensitivity and specificity for manual interpretation of SL-ECGs from the Samsung Galaxy Watch [28].

Even with no significant difference between the smart devices, some questions for further research emerged. The Withings ScanWatch is the only smart device with a rasterised SL-ECG. It is possible that this influenced subjective ratings of quality and readability in our survey. Similarly, while the Apple Watch 6 SL-ECG showed the best rating for quality and readability, it demonstrated the worst sensitivity of the assessed smart devices. It is unclear whether this is coincidental or whether the underlying algorithm loses/alters information for a smoother and more appealing recording. On the other hand, the AliveCor KardiaMobile, showed the highest accuracy, although not to a statistically significant extent. It is worth considering whether this is due to its different approach in recording the SL-ECG with a handheld smart device rather than a smartwatch, but further studies with more recordings are needed.

The reported sensitivity and specificity for SL-ECG interpretation of residents and cardiologists is in the range previously reported for general practitioners and cardiologists [2–4, 6–8, 24–28]. Mant and Fitzmaurice et al. [29]found a sensitivity and specificity of 84% and 86% for GPs and 68% and 82% for practice nurses, respectively. Of note, the SL-ECGs were taken from a 12-lead ECG and not from a smart device.

Only one-fifth of participants stated prior SL-ECG experience, reflecting the fact that SL-ECGs are not part of medical teaching. For participants who interpreted all 50 SL-ECGs, the accuracy, sensitivity and specificity in the last five SL-ECGs were better than in the first five. This implies that even a tiny effort has an impact on SL-ECG interpretation. This raises the question of what proper teaching could achieve. Maybe in the future, with the upcoming smart devices and rising quantity of SL-ECGs, consideration should be given to training medical personnel on SL-ECGs and even on photoplethysmography tracings, tachograms and Poincaré plots [24].

Limitations

We acknowledge several limitations in this survey-based study. First the study is based on ten patients, five in sinus rhythm and five in atrial fibrillation, resulting in 50 SL-ECGs from five different smart devices. This leads to a low number of different individual recordings with few considerations to different body builds, possibly impacting assessment of quality and readability. This approach also resulted in a prevalence of 50% of atrial fibrillation, which is higher than in the general population. Future studies should include more SL-ECG recordings from a higher number of individuals. Second, SL-ECGs were recorded nearly simultaneously, meaning sequentially with a very short time; however difference between recordings and therefore possible variation in quality could occur, especially in recordings with atrial fibrillation. Third, we did not differentiate the internal medicine residents' working experience and specialty, limiting further findings. Fourth, best/worst quality/readability was not further explained, leaving room for different interpretations. However, the answers seem consistent in positive and negative ratings, suggesting a consensus. Fifth, SL-ECGs were randomly chosen from a database of 166 patients with recorded SL-ECGs from the five different smart devices, resulting in different quality, readability and difficulty, thereby impacting the diagnostic accuracy. This may be resolved with a repetition of the SL-ECG recording. Nevertheless, physicians could come across such SL-ECGs in their future work. Last, we only compared five smart devices in a market that is starting to be inundated with more and cheaper smart devices capable of recording SL-ECGs, making it almost impossible to make general statements.

Conclusion

Single-lead ECGs from five different manufacturers have various qualities and can be challenging in interpretation. Diagnostic accuracy was better for highly ranked smart device SL-ECGs regarding quality and readability, and worse for SL-ECGs with lower ranked quality and readability. Accuracy for correct atrial fibrillation diagnosis varied by level of expertise. Previous SL-ECG experience constitutes a benefit in interpreting SL-ECGs and a learning curve is achieved with little effort implying that SL-ECG teaching should be implemented in future training.

Data availability statement

The datasets generated during and/or analysed during the current study are available from the corresponding author on reasonable request.

Acknowledgments
The authors thank the patients who participated in the study and Corinne Isenegger, Claudius Vernier and David Vögeli for contributing in patient recruiting and data collection.

Author contributions: DM and PB contributed to conception and design of the Study, SW, DM and PB organized the database, performed the statistical analysis, and wrote the first draft of the manuscript. TS, PK, SK, JdF, TM, BS, SO, MK and CS revised it critically for important intellectual content. PB supervised the whole work. SW finalized the manuscript. All authors contributed to manuscript revision, read, and approved the submitted version.

Financial disclosure
The study was supported by the Swiss Heart Foundation and the University of Basel.

Potential competing interests
All authors have completed and submitted the International Committee of Medical Journal Editors form for disclosure of potential conflicts

of interest. PB has received research funding from the “University of Basel“, the “Stiftung für Herzschrittmacher und Elektrophysiologie”, the “Freiwillige Akademische Gesellschaft Basel”, the Swiss Heart Foundation, and Johnson&Johnson, all outside the submitted work, and reports personal fees from Abbott, Boston Scientific and Pfizer BMS. SK has received funding of the “Stiftung für kardiovaskuläre Forschung”. CS: Member of Medtronic Advisory Board Europe and Boston Scientific Advisory Board Europe, received educational grants from Biosense Webster and Biotronik and a research grant from the European Union’s FP7 program, and Biosense Webster and lecture and consulting fees from Abbott, Medtronic, Biosense-Webster, Boston Scientific, Microport, and Biotronik all outside the submitted work. MK reports grants from the Swiss National Science Foundation (Grant numbers 33CS30_148474, 33CS30_177520, 32473B_176178, 32003B_197524), the Swiss Heart Foundation, the Foundation for Cardiovascular Research Basel and the University of Basel, grants from Bayer, grants from Pfizer, grants from Boston Scientific, grants from BMS, grants from Biotronik, grants and personal fees from Daiichi Sankyo. all outside the submitted work. BS reports speaker’s bureau for Medtronic and Zoll, both outside the submitted work. JdFdL has received research funding from the Swiss Heart Foundation and a personal grant from the Goldschmidt Jacobson Foundation, both outside the submitted work. TS has received research funding from the “Gottfried & Julia Bangerter-Rhyner Foundation” and the Swiss Academy for Medical Sciences.

References

1. Hindricks G, Potpara T, Dagres N, Arbelo E, Bax JJ, Blomström-Lundqvist C, et al.; ESC Scientific Document Group. 2020 ESC Guidelines for the diagnosis and management of atrial fibrillation developed in collaboration with the European Association for Cardio-Thoracic Surgery (EACTS): The Task Force for the diagnosis and management of atrial fibrillation of the European Society of Cardiology (ESC) Developed with the special contribution of the European Heart Rhythm Association (EHRA) of the ESC. Eur Heart J. 2021 Feb;42(5):373–498. http://dx.doi.org/10.1093/eurheartj/ehaa612. PubMed. 1522-9645
2. Selder JL, Breukel L, Blok S, van Rossum AC, Tulevski II, Allaart CP. A mobile one-lead ECG device incorporated in a symptom-driven remote arrhythmia monitoring program. The first 5,982 Hartwacht ECGs. Neth Heart J. 2019 Jan;27(1):38–45. http://dx.doi.org/10.1007/s12471-018-1203-4. PubMed. 1568-5888
3. William AD, Kanbour M, Callahan T, Bhargava M, Varma N, Rickard J, et al. Assessing the accuracy of an automated atrial fibrillation detection algorithm using smartphone technology: the iREAD Study. Heart Rhythm. 2018 Oct;15(10):1561–5. http://dx.doi.org/10.1016/j.hrthm.2018.06.037. PubMed. 1556-3871
4. Avram R, Ramsis M, Cristal AD, Nathan V, Zhu L, Kim J, et al. Validation of an algorithm for continuous monitoring of atrial fibrillation using a consumer smartwatch. Heart Rhythm. 2021 Sep;18(9):1482–90. http://dx.doi.org/10.1016/j.hrthm.2021.03.044. PubMed. 1556-3871
5. Seshadri DR, Bittel B, Browsky D, Houghtaling P, Drummond CK, Desai MY, et al. Accuracy of Apple Watch for Detection of Atrial Fibrillation. Circulation. 2020 Feb;141(8):702–3. http://dx.doi.org/10.1161/CIRCULATIONAHA.119.044126. PubMed. 1524-4539
6. Ford C, Xie CX, Low A, Rajakariar K, Koshy AN, Sajeev JK, et al. Comparison of 2 Smart Watch Algorithms for Detection of Atrial Fibrillation and the Benefit of Clinician Interpretation: SMART WARS Study. JACC Clin Electrophysiol. 2022 Jun;8(6):782–91. http://dx.doi.org/10.1016/j.jacep.2022.02.013. PubMed. 2405-5018
7. Bumgarner JM, Lambert CT, Hussein AA, Cantillon DJ, Baranowski B, Wolski K, et al. Smartwatch Algorithm for Automated Detection of Atrial Fibrillation. J Am Coll Cardiol. 2018 May;71(21):2381–8. http://dx.doi.org/10.1016/j.jacc.2018.03.003. PubMed. 1558-3597
8. Badertscher P, Lischer M, Mannhart D, Knecht S, Isenegger C, Du Fay de Lavallaz J, et al. Clinical validation of a novel smartwatch for automated detection of atrial fibrillation. Heart Rhythm O2. 2022 Feb;3(2):208–10. http://dx.doi.org/10.1016/j.hroo.2022.02.004. PubMed. 2666-5018
9. Mannhart D, Lischer M, Knecht S, du Fay de Lavallaz J, Strebel I, Serban T, et al. Clinical Validation of 5 Direct-to-Consumer Wearable Smart Devices to Detect Atrial Fibrillation: BASEL Wearable Study. JACC Clin Electrophysiol. 2023 Feb;9(2):232–42. http://dx.doi.org/10.1016/j.jacep.2022.09.011. PubMed. 2405-5018
10. Schnabel RB, Marinelli EA, Arbelo E, Boriani G, Boveda S, Buckley CM, et al. Early diagnosis and better rhythm management to improve outcomes in patients with atrial fibrillation: the 8th AFNET/EHRA consensus conference. EP Europace 2022. doi: http://dx.doi.org/10.1093/europace/euac062.
11. Leclercq C, Witt H, Hindricks G, Katra RP, Albert D, Belliger A, et al. Wearables, telemedicine, and artificial intelligence in arrhythmias and heart failure: Proceedings of the European Society of Cardiology Cardiovascular Round Table. EP Europace. 2022;24:1372–83. doi: http://dx.doi.org/10.1093/europace/euac052.
12. Kalarus Z, Mairesse GH, Sokal A, Boriani G, Średniawa B, Arroyo RC, et al. Searching for atrial fibrillation: looking harder, looking longer, and in increasingly sophisticated ways. An EHRA position paper. EP Europace; 2022. http://dx.doi.org/10.1093/europace/euac144.
13. Svennberg E, Tjong F, Goette A, Akoum N, Di Biase L, Bordachar P, et al. How to use digital devices to detect and manage arrhythmias: an EHRA practical guide. Europace. 2022 Jul;24(6):979–1005. http://dx.doi.org/10.1093/europace/euac038. PubMed. 1532-2092
14. Linz D, Hermans A, Tieleman RG. Early atrial fibrillation detection and the transition to comprehensive management. Europace. 2021 Apr;23(23 Suppl 2):ii46–51. http://dx.doi.org/10.1093/europace/euaa424. PubMed. 1532-2092
15. Boriani G, Svennberg E, Guerra F, Linz D, Casado-Arroyo R, Malaczynska-Rajpold K, et al. Reimbursement practices for use of digital devices in atrial fibrillation and other arrhythmias: a European Heart Rhythm Association survey. Europace. 2022 Nov;24(11):1834–43. http://dx.doi.org/10.1093/europace/euac142. PubMed. 1532-2092
16. Manninger M, Zweiker D, Svennberg E, Chatzikyriakou S, Pavlovic N, Zaman JA, et al. Current perspectives on wearable rhythm recordings for clinical decision-making: the wEHRAbles 2 survey. Europace. 2021 Jul;23(7):1106–13. http://dx.doi.org/10.1093/europace/euab064. PubMed. 1532-2092
17. Saxon LA. Ubiquitous wireless ECG recording: a powerful tool physicians should embrace. J Cardiovasc Electrophysiol. 2013 Apr;24(4):480–3. http://dx.doi.org/10.1111/jce.12097. PubMed. 1540-8167
18. Nasarre M, Strik M, Daniel Ramirez F, Buliard S, Marchand H, Abu-Alrub S, et al. Using a smartwatch electrocardiogram to detect abnormalities associated with sudden cardiac arrest in young adults. Europace. 2022 Mar;24(3):406–12. http://dx.doi.org/10.1093/europace/euab192. PubMed. 1532-2092
19. Mannhart D, Hennings E, Lischer M, Vernier C, Du Fay de Lavallaz J, Knecht S, et al. Clinical Validation of Automated Corrected QT-Interval Measurements From a Single Lead Electrocardiogram Using a Novel Smartwatch. Front Cardiovasc Med. 2022 Jun;9:906079. http://dx.doi.org/10.3389/fcvm.2022.906079. PubMed. 2297-055X
20. Stark K, Czermak T, Massberg S, Orban M. Watch out for ST-elevation myocardial infarction: a case report of ST-elevation in single-lead electrocardiogram tracing of a smartwatch. Eur Heart J Case Rep. 2020 Nov;4(6):1–4. http://dx.doi.org/10.1093/ehjcr/ytaa353. PubMed. 2514-2119
21. Harris PA, Taylor R, Minor BL, Elliott V, Fernandez M, O’Neal L, et al.; REDCap Consortium. The REDCap consortium: building an international community of software platform partners. J Biomed Inform. 2019 Jul;95:103208. http://dx.doi.org/10.1016/j.jbi.2019.103208. PubMed. 1532-0480
22. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)—a metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform. 2009 Apr;42(2):377–81. http://dx.doi.org/10.1016/j.jbi.2008.08.010. PubMed. 1532-0480
23. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP; STROBE Initiative. Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. BMJ. 2007 Oct;335(7624):806–8. http://dx.doi.org/10.1136/bmj.39335.541782.AD. PubMed. 1756-1833
24. Gruwez H, Evens S, Proesmans T, Duncker D, Linz D, Heidbuchel H, et al. Accuracy of Physicians Interpreting Photoplethysmography and Electrocardiography Tracings to Detect Atrial Fibrillation: INTER-PRET-AF. Front Cardiovasc Med. 2021 Sep;8:734737. http://dx.doi.org/10.3389/fcvm.2021.734737. PubMed. 2297-055X
25. Himmelreich JC, Karregat EP, Lucassen WA, van Weert HC, de Groot JR, Handoko ML, et al. Diagnostic Accuracy of a Smartphone-Operated, Single-Lead Electrocardiography Device for Detection of Rhythm and Conduction Abnormalities in Primary Care. Ann Fam Med. 2019 Sep;17(5):403–11. http://dx.doi.org/10.1370/afm.2438. PubMed. 1544-1717
26. Karregat EP, Himmelreich JC, Lucassen WA, Busschers WB, van Weert HC, Harskamp RE. Evaluation of general practitioners’ single-lead electrocardiogram interpretation skills: a case-vignette study. Fam Pract. 2021 Mar;38(2):70–5. http://dx.doi.org/10.1093/fampra/cmaa076. PubMed. 1460-2229

27. Koshy AN, Sajeev JK, Negishi K, Wong MC, Pham CB, Cooray SP, et al. Accuracy of blinded clinician interpretation of single-lead smartphone electrocardiograms and a proposed clinical workflow. Am Heart J. 2018 Nov;205:149–53. http://dx.doi.org/10.1016/j.ahj.2018.08.001. PubMed. 1097-6744
28. Abu-Alrub S, Strik M, Ramirez FD, Moussaoui N, Racine HP, Marchand H, et al. Smartwatch Electrocardiograms for Automated and Manual Diagnosis of Atrial Fibrillation: A Comparative Analysis of Three Models. Front Cardiovasc Med. 2022 Feb;9:836375. http://dx.doi.org/10.3389/fcvm.2022.836375. PubMed. 2297-055X
29. Mant J, Fitzmaurice DA, Hobbs FD, Jowett S, Murray ET, Holder R, et al. Accuracy of diagnosing atrial fibrillation on electrocardiogram by primary care practitioners and interpretative diagnostic software: analysis of data from screening for atrial fibrillation in the elderly (SAFE) trial. BMJ. 2007 Aug;335(7616):380. http://dx.doi.org/10.1136/bmj.39227.551713.AE. PubMed. 1756-1833