UHASSELT
KNOWLEDGE IN ACTION

Maastricht University

Faculteit Geneeskunde en Levenswetenschappen
School voor Levenswetenschappen

master in de biomedische wetenschappen

Masterthesis

Evaluation of FibriCheck against the standard of care in detecting atrial fibrillation in cryptogenic ischemic stroke patients

Heike Blockx

Scriptie ingediend tot het behalen van de graad van master in de biomedische wetenschappen, afstudeerrichting klinische biomedische wetenschappen

PROMOTOR :

Prof. Dr. Pieter VANDERVOORT

BEGELEIDER :

Mevrouw Femke WOUTERS

Mevrouw Myrte BARTHELS

2023
2024

De transnationale Universiteit Limburg is een uniek samenwerkingsverband van twee universiteiten in twee landen: de Universiteit Hasselt en Maastricht University.

UHASSELT
KNOWLEDGE IN ACTION

www.uhasselt.be
Universiteit Hasselt
Campus Hasselt:
Martelarenlaan 42 | 3500 Hasselt
Campus Diepenbeek:
Agoralaan Gebouw D | 3590 Diepenbeek

UHASSELT
KNOWLEDGE IN ACTION

Maastricht University

Faculteit Geneeskunde en Levenswetenschappen

School voor Levenswetenschappen

master in de biomedische wetenschappen

Masterthesis

Evaluation of FibriCheck against the standard of care in detecting atrial fibrillation in cryptogenic ischemic stroke patients

Heike Blockx

Scriptie ingediend tot het behalen van de graad van master in de biomedische wetenschappen, afstudeerrichting klinische biomedische wetenschappen

PROMOTOR :

Prof. Dr. Pieter VANDERVOORT

BEGELEIDER :

Mevrouw Femke WOUTERS

Mevrouw Myrte BARTHELS

Evaluation of FibriCheck® against the standard of care in detecting atrial fibrillation in cryptogenic ischemic stroke patients*

Heike Blockx[1], Femke Wouters[1,2,3], Myrte Barthels[1,2,3,4], David Verhaert[5], Thomas Phlips[6] and Pieter Vandervoort[1,2,3,5]

[1]Faculty of Medicine and Life Sciences, Hasselt University, Agoralaan, 3590 Diepenbeek, Belgium
[2]Limburg Clinical Research Center/Mobile Health Unit, Faculty of Medicine and Life Sciences, Hasselt University, Martelarenlaan 42, 3500 Hasselt, Belgium
[3]Department of Future Health, Ziekenhuis Oost Limburg, Synaps Park 1, 3600 Genk, Belgium
[4]Qompium NV, Hasselt, Belgium
[5]Department of Cardiology, Ziekenhuis Oost-Limburg, Synaps Park 1, 3600 Genk, Belgium
[6]Department of Cardiology, Jessa Ziekenhuis, Stadsomvaart 11, 3500 Hasselt, Belgium

*Running title: *REMOTE study*

To whom correspondence should be addressed: Pieter Vandervoort, Tel: +32 (0) 89 32 70 85; Email: pieter.vandervoort@uhasselt.be

Keywords: Atrial fibrillation, Cryptogenic stroke, Implantable loop recorders, mHealth, Photoplethysmography

ABSTRACT

BACKGROUND – Atrial fibrillation (AF) is a significant risk factor for ischemic strokes. Identifying AF in cryptogenic ischemic stroke patients is crucial for preventing recurrence. Traditional diagnostic tools like short-term electrocardiograms and long-term implantable loop recorders (ILRs) are costly, time-consuming, and invasive, presenting significant drawbacks.

METHOD – This multicentric, prospective, randomized, blinded, and interventional study in Ziekenhuis Oost Limburg, Genk, and Jessa Ziekenhuis, Hasselt aims to compare ILR with FibriCheck® on a smartphone or smartwatch in detecting AF in cryptogenic stroke patients over a six-month period.

RESULTS – A total of 190 patients were included and randomized in the smartphone (n=102) or smartwatch group (n=88). In the smartphone group, the sensitivity, specificity, and accuracy of FibriCheck® were 94.44%, 99.61%, and 99.61%, respectively, and the specificity was 99.81% in the smartwatch group. Furthermore, FibriCheck® shows significantly higher detection (962 vs 81 AF episodes, $p<0.001$) and longer duration (23 [9–71.5] minutes vs 2 [2–16.80] minutes, $p<0.001$) of the first AF episodes than ILR. Furthermore, there is a significant difference in AF duration between smartphones and smartwatches, with smartwatches showing a shorter AF duration (334 [22.90–879] minutes vs 12 [9–27] minutes, $p<0.001$). The percentage of low-quality measurements during the day is significantly higher compared to the percentage during the night (respectively, 58.46% and 7.46%, $p<0.001$).

CONCLUSION – FibriCheck® has promising potential for AF detection in cryptogenic stroke patients. Further research and larger-scale studies are needed to confirm these findings and investigate the integration of FibriCheck® into standard clinical protocols for secondary stroke prevention.

INTRODUCTION

Ischemic stroke – Cerebrovascular accidents (CVA), better known as strokes, are the second leading cause of death worldwide (1). These CVAs can be indicated as ischemic or hemorrhagic, whereby ischemic strokes are the most prevalent types, at approximately 85% (2, 3). Hemorrhagic strokes occur as a consequence of cerebral bleeding, which is usually caused by a ruptured blood vessel (4). Ischemic strokes, on the other hand, occur when a blood clot prevents or reduces the blood flow to certain parts of the brain. These ischemic strokes can be classified into subtypes using the Trial of Org 10172 in Acute Stroke Treatment (TOAST) classification. These subtypes consist of large-artery atherosclerosis, cardioembolism,

small-vessel occlusion, stroke of other determined etiology, and stroke of undetermined etiology (5). This last subtype is frequently referred to as a cryptogenic ischemic stroke (6).

Timely recognition of stroke is essential since "time is brain". The National Institutes of Health Stroke Scale (NIHSS) is a simple stroke recognition tool. This scale is an observation list that identifies a stroke patient's neurological characteristics and symptoms. This instrument includes domains such as consciousness and attention level, vision, motor and sensory functions, neglect, as well as language and speech production. A high score indicates a more severe form of CVA (7). To confirm the diagnosis of stroke, a Computed Tomography (CT) and Magnetic Resonance Imaging (MRI) scan will be performed, revealing the bleeding and ischemic areas (8). In contrast to a stroke, transient ischemic accidents (TIAs) usually resolve in less than one hour. However, it causes the same symptoms associated with a stroke (9).

Ischemic stroke treatments consist of medication, such as thrombolytics, or mechanical procedures like thrombectomy to restore the blood flow. To reduce the risk of clot formation and, thus, stroke recurrence, other therapeutical treatments can be used, including antiplatelets and antithrombotics (3). Hypertension, atrial fibrillation, cigarette smoking, diabetes mellitus, and hypercholesterolemia are risk factors that are associated with the development of an ischemic stroke, of which atrial fibrillation (AF) is one of the most important. Moreover, patients with AF have a five times higher risk of developing a stroke and, in addition, these AF-related strokes have a higher morbidity and mortality than non-AF-related strokes (10, 11).

Atrial fibrillation – According to the European Society of Cardiology (ESC) guidelines, AF is a supraventricular tachyarrhythmia with uncoordinated atrial electrical activation and ineffective atrial contraction (10). Atrial fibrillation can be classified into four different types. Firstly, paroxysmal AF, which terminates spontaneously or within seven days after intervention. Secondly, persistent AF is referred to as continuous AF that persists for more than seven days. Thirdly, long-lasting AF is continuous AF for over 12 months. Lastly, in permanent AF, the decision has been made to accept the presence of AF and to stop further efforts to treat the cardiac arrhythmia (10, 12).

Furthermore, AF is one of the most common types of cardiac arrhythmias worldwide, with a global estimation of over 40 million in 2016 and still increasing (10). The occurrence of AF is elevated with age, underlying heart diseases, hypertension, endocrine disorders, and others. The standard therapy for AF includes rhythm and rate control and managing stroke prevention (13). Under normal circumstances, the cardiac conduction system regulates the rhythmic contraction and relaxation of the heart through a coordinated interplay of electrical signals generated in the sinoatrial node (14, 15). However, the coordination of atrial activation is disrupted in AF, resulting in disorganized and rapid atrial contractions (16). This dysregulation leads to incomplete emptying of the atria, fostering blood stasis and subsequent formation of blood clots (17, 18). Over time, these thrombi may dislodge and traverse to the brain, where they can occlude or restrict blood flow in cerebral vessels, causing inadequate oxygen and nutrient supply. Without timely intervention, this series of events can lead to irreversible brain damage or mortality due to oxygen and nutrient deprivation (2). Additional factors that may contribute to AF-related strokes include endothelial dysfunction, fibrosis, and impaired myocyte function (18, 19). Furthermore, AF is seen as a complication after cardiac surgery, including coronary artery bypass grafting and valve replacement. In patients with AF, determination of Congestive heart failure, Hypertension, Age, Diabetes mellitus, prior Stroke or TIA or thromboembolism, Vascular disease, Age, Sex category (CHA_2DS_2-VASc) score is recommended since it will predict the risk of developing stroke (20).

Detection procedures – Identifying the underlying cause and risk factors of a stroke and, consequently, detecting potential AF episodes are crucial for preventing stroke recurrence. The standard AF detection procedures involve an electrocardiogram (ECG), which assesses cardiac rhythm over a short period of time (10). If AF is not evident on the initial ECG, the patient is equipped with a Holter for continuous cardiac rhythm monitoring. European Stroke Organization guidelines suggest Holter monitoring of more

than 48 hours post-stroke, while ESC guidelines recommend monitoring at least 72 hours post-stroke (10, 21). However, these first two methods are short-term detection procedures, limiting the chances of detecting AF, as AF may not be consistently present and may be overlooked. In cases where AF is not detected during Holter monitoring, ESC guidelines recommend long-term monitoring. The patient will undergo the insertion of an implantable loop recorder (ILR) to monitor heart rhythm for an extended period of time (10). Although, the invasive nature of the procedure, high costs, and time-consuming work pose significant drawbacks. Hence, there is a need for an alternative approach to long-term monitoring.

Mobile health – One potential and reliable alternative could involve leveraging mobile health (mHealth), which refers to integrating mobile devices, such as smartphones, patient monitoring tools, and other wireless devices, into medical health practices (22). These devices are classified into two main categories: photoplethysmography (PPG)- and ECG-based mobile devices (23). PPG technology allows volumetric blood flow variations to be measured via a light source and photodetector, resulting in the heart rate. Furthermore, embedded algorithms are used to determine the heart rhythm (23, 24). On the other hand, the ECG detects the electrical activity of the heart, provides detailed insight into the heart rhythm, and is used to diagnose cardiac arrhythmias (23). In this study, FibriCheck®, a PPG-based mobile application, will be investigated and compared to ILR. FibriCheck® is a non-invasive mHealth application used to measure cardiac rhythm. This method was investigated in previous studies for its ability to detect arrhythmias. While Proesmans *et al.* demonstrated a sensitivity and specificity of 96% and 98.5%, respectively, Wouters *et al.* illustrated the capability of FibriCheck® in a case report, demonstrating detection of AF using the mHealth application with similar precision as an ILR in cryptogenic ischemic stroke patients (24, 25). Preliminary results from the REMOTE trial illustrated the potential of mHealth to detect AF in stroke patients. This study also detected confirmed AF episodes on the ILR by PPG-based mHealth on the smartphone (26). This REMOTE trial, of which this project is a part, is still ongoing. This study aims to compare ILR, the standard of care, with mHealth (FibriCheck®) on a smartphone or smartwatch in detecting AF in cryptogenic ischemic stroke patients over a six-month period. Additionally, the quality of the smartwatch FibriCheck® measurements will be investigated. It is expected that the detection of AF in patients with cryptogenic ischemic stroke using a PPG-based method (FibriCheck®) will be non-inferior to the detection of AF using ILR. Furthermore, the quality of smartwatch measurements at night is expected to be higher than those measured during the day because of reduced patient activity.

METHODS

Study design – This multicentric, prospective, randomized, blinded, and interventional study in Ziekenhuis Oost Limburg (ZOL), Genk and Jessa Ziekenhuis, Hasselt aims to compare ILR, the standard of care, with mHealth (FibriCheck®) on a smartphone or smartwatch in detecting AF in cryptogenic ischemic stroke patients over a six-month period. Ethical approval from the Ethics Committee of ZOL Genk was granted on the 24th of June 2020. Patients were included starting in September 2020. This study was performed in accordance with the Declaration of Helsinki and registered on clinicaltrials.gov (NCT05006105).

Study population – All patients planned to undergo ILR implantation were subjected to participation in the study after screening. Patients older than 18 years old, who were diagnosed with cryptogenic ischemic stroke or TIA, who could give their informed consent, and who spoke Dutch could be included in the study. Patients who had a history of AF or atrial flutter, who had a life expectancy of less than one year, who were not suitable for an ILR insertion, who had an indication or contra-indication for permanent oral anticoagulants (OAC) during inclusion, who had untreated hyperthyroidism, who had had myocardial infarction or CABG less than one month before stroke or TIA, who had a patent foramen ovale (PFO) and this is or was an indication to start OAC, who were included in other clinical trials that may affect the results of this study and who did not own a smartphone could not be included in the study.

Randomization – After inclusion and signing informed consent, patients were

randomized into the smartphone group or the smartwatch group. Randomization was performed in a 2:2 ratio.

Study procedure and intervention – All study patients received an ILR after inclusion, with batteries designed to last around three years. The ILR was surgically implanted subcutaneously at cardiac level through a minimally invasive procedure. Both study groups used the FibriCheck® application. The participants in the smartphone group were asked to perform two manual (spread over the day, e.g., in the morning and in the evening) one-minute measurements daily on the smartphone over a six-month period. On the other hand, the participants in the smartwatch group used FibriCheck® on the smartwatch (Fitbit® Versa 2, Fitbit Inc., San Francisco, United States), which measured the cardiac rhythm semi-continuously (automatically every nine minutes) over a six-month period. In case of symptoms such as palpitations or angina, all patients could perform additional measurements. The FibriCheck® measurements were analyzed by the FibriCheck® algorithm and annotated as normal/sinus rhythm (green), warning/other arrhythmias (orange), urgent/possible AF (red), and insufficient quality (blue). On the other hand, ILR measurements were annotated as sinus rhythm, AF appropriate, no annotation, or AF inappropriate. However, only AF-appropriate measurements were included in the data analysis to represent the AF episodes detected by ILR. An overview of the study procedure and intervention is presented in Figure 1.

Study outcomes – The primary outcome was to determine the detection of AF, using PPG-based mHealth (FibriCheck®) compared to ILR, in cryptogenic stroke or TIA patients, over a six-month period. The secondary outcome was to assess the measurement quality of FibriCheck® on the smartwatch and compare this quality between day and night in patients with cryptogenic ischemic stroke and TIA over a six-month period.

Data collection – Demographic and clinical data were collected from the electronic health records (HiX, Chipsoft, Amsterdam, The Netherlands, and KWS, NexuzHealth, Belgium), ILR dashboards (Biotronik, SE & Co. KG, Berlin, Germany, Medtronic, Ireland and Abbott, Illinois, United States), and FibriCheck® dashboard (Qompium nv, Hasselt, Belgium). All study data were captured into an electronic case report form (eCRF) (Castor EDC, New York, United States). All data and information were collected according to the General Data Protection Regulation (GDPR). Further, measurements were divided into smartphone and smartwatch based on the device used to perform the measurement, not on which study group the patient was in. This was due to the ability of patients in the smartwatch group to perform measurements using a smartphone as well.

Statistical analysis – To analyze the quality of the FibriCheck® measurements in the smartwatch group, data were divided into daytime (between 8 a.m. and 10 p.m.) and nighttime (between 12 p.m. and 6 a.m.) based on the prospective observational study of Hermans *et al.* (27). Insufficient quality FibriCheck® measurements (blue) were referred to as low-quality measurements and normal, warning and urgent FibriCheck® measurements (green, orange and red, respectively) were referred to as high-quality measurements. To describe the precision of the FibriCheck® application, the sensitivity, specificity, and accuracy were calculated using the following formula:

$$\text{Sensitivity} = \frac{\text{True positive}}{\text{True positive} + \text{False negative}}$$

$$\text{Specificity} = \frac{\text{True negative}}{\text{True negative} + \text{False positive}}$$

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FP} + \text{TN} + \text{FN}}$$

FN, False Negative; FP, False Positive; TN, True Negative; TP, True Positive

The start and end times were estimated to determine the duration of the AF episode detected by FibriCheck®. The start of the AF episode is calculated by identifying the time of the last green measurement (i.e., no AF) and the first red measurement (i.e., AF). The start time is determined by calculating the midpoint between these two measurements. Similarly, the end of the AF episode is determined by identifying the time of the last red measurement and the next green measurement. The midpoint between these measurements is used to

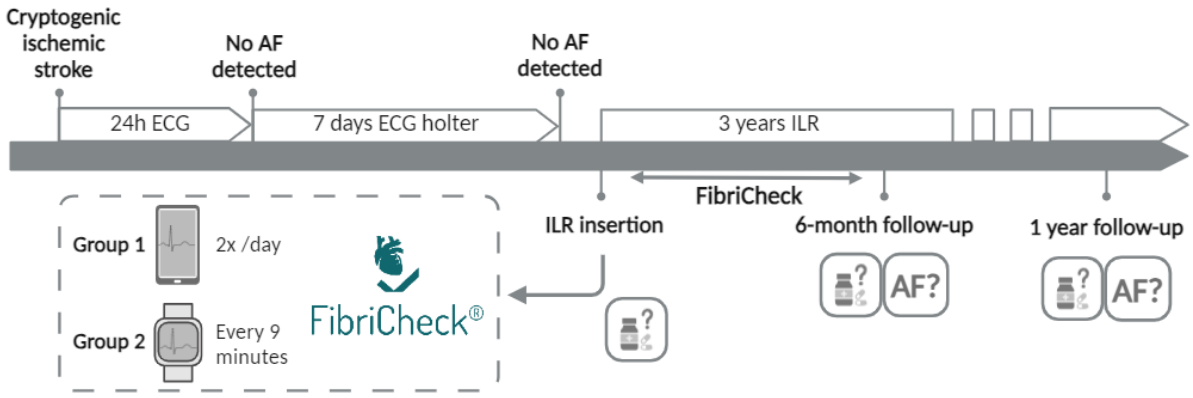

Figure 1. Schematic overview of the study procedure and intervention.

AF, atrial fibrillation; ECG, electrocardiogram; ILR, implantable loop recorder.

determine the end time of the AF episode. This method was described in a case report by Wouters et al. (25). The duration of the AF episodes detected with FibriCheck® and ILR was defined as the duration of the first AF episode detected.

Statistical analysis was performed using IBM SPSS statistics version 29.0.2 (Armonk, United States) for the demographic data and R version 23.03.0 (Posit PBC, Boston, Massachusetts, United States) for the digital device data. Normality was tested by means of the Shapiro-Wilk test. Continuous data were expressed as median with interquartile ranges ([Q1 - Q3]), and categorical data were expressed as frequency (n) and percentages (%). When data was not normally distributed, a Mann-Whitney U test was used to compare the two study groups, and a Wilcoxon signed-rank test was used to compare within one group. Categorical data were compared using Pearson's Chi-squared test for unpaired data and a McNemar Chi-squared test for paired data. Data were considered statistically significant when p-value < 0.05.

RESULTS

Study population - During a study period from October 2020 until February 2024, 1009 patients were assessed for eligibility, of which 224 participated in this trial, as presented in Figure 2. These 224 patients were randomized into one of the two study groups: 115 in the smartphone group and 109 in the smartwatch group. In the smartphone group, 103 patients were followed up, and 102 were analyzed. In the smartwatch group, 100 patients were followed up and 88 patients were analyzed. The most important reasons for drop-out were too much effort to perform the FibriCheck® measurements, anxiety, and allergic reaction to the wristband of the smartwatch.

Demographic data – Overall, demographical data were similar in both groups. However, congestive heart failure (CHF) and antihypertensive medication use were significantly different between both groups (6 (5.90%) vs 0, p=0.021 and 67 (67.0%) vs 70 (80.50%), p=0.031 respectively). Patient demographics are shown in Table 1.

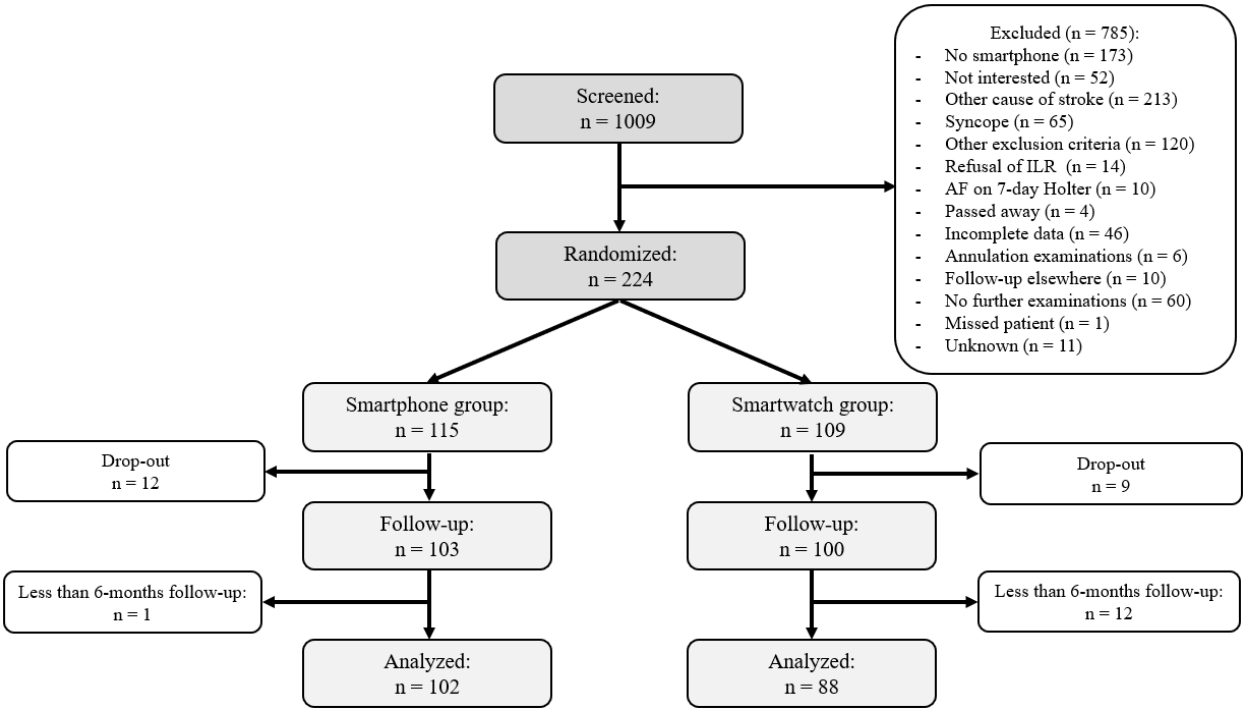

Figure 2. Flow chart of patient recruitment, analysis, follow-up, and exclusion.

AF, atrial fibrillation; ILR, implantable loop recorder.

Table 1. Demographic data.

	Smartphone group (n = 102)	**Smartwatch group (n = 88)**	**Total (n = 190)**	
Patient characteristics	**n (%) or median [Q1-Q3]**	**n (%) or median [Q1-Q3]**	**n (%) or median [Q1-Q3]**	**P-value**
Age, years	64 [57 – 73.25]	65 [58 – 71]	65 [57 – 72]	0.684
Sex, male	55 (53.92)	69 (78.41)	124 (65.26)	0.457
BMI, kg/m²	26.20 [24.16 – 28.68]	27.76 [24.29 – 29.39]	26.83 [24.21 – 29.27]	0.283
Smoking				0.153
No	43 (42.2)	46 (52.3)	89 (46.8)	
Former	27 (26.5)	25 (28.4)	52 (27.4)	
Current	32 (31.4)	17 (19.3)	49 (25.8)	
Heavy drinking	4 (3.9)	2 (2.3)	6 (3.2)	0.517
Diagnosis:				0.284
Stroke	82 (80.4)	65 (73.9)	147 (77.4)	
TIA	20 (19.6)	23 (26.1)	43 (22.6)	

Medical history				
History of stroke	8 (7.8)	14 (15.9)	22 (11.6)	0.083
History of TIA	7 (6.9)	7 (7.9)	14 (7.4)	0.774
Arterial hypertension	60 (58.8)	57 (64.8)	117 (61.6)	0.401
Diabetes	9 (8.8)	13 (14.8)	22 (11.6)	0.201
Hypercholesterolemia	63 (61.8)	54 (61.4)	117 (61.6)	0.955
Dyslipidemia	8 (7.8)	9 (10.1)	17 (8.9)	0.566
Family history of cardiovascular conditions	40 (39.2)	32 (36.4)	72 (37.9)	0.686
Congestive heart failure	6 (5.9)	0	6 (3.2)	**0.021**
Chronic kidney disease	6 (5.9)	4 (4.5)	10 (5.3)	0.681
Peripheral artery disease	8 (7.8)	6 (6.8)	14 (7.4)	0.787
Angina pectoris	6 (5.9)	5 (5.7)	11 (5.8)	0.953
Myocardial infarction	7 (6.9)	6 (6.8)	13 (6.8)	0.990
Atherosclerosis	30 (29.4)	23 (26.1)	53 (27.9)	0.616
Coronary artery disease	15 (14.7)	9 (10.2)	24 (12.6)	0.354
CHA_2DS_2-VASc	4 [3 – 5]	4 [3 – 5]	4 [3 – 5]	0.790
Pre-MRS	0	0 [0 – 0.5]	0	0.254
NIHSS	1 [0 – 4]	1 [0 – 3.5]	1 [0 – 4]	0.848
PFO	34 (33.0)	11 (12.6)	45 (23.7)	0.750
Medication during hospitalization:				
Antiplatelet	97 (98.0)	87 (98.9)	184 (98.4)	0.631
Anticoagulantia	8 (8.1)	16 (16.0)	15 (8.0)	0.975
Thrombolytics	16 (16.0)	11 (12.6)	27 (14.4)	0.477
Thrombectomy	4 (4.0)	7 (8.0)	11 (5.9)	0.256
Antihypertensiva	67 (67.0)	70 (80.5)	137 (73.3)	**0.031**
Antidiabetics	16 (16.0)	13 (41.9)	29 (15.5)	0.793
Hypolipidemic	83 (83.0)	74 (85.1)	157 (84.0)	0.655

Statistical analysis was done using the Pearson Chi-Squared test for categorical data and the Mann-Whitney U test for continuous data with a significance level of $p<0.05$.

BMI, Body Mass Index; CHA_2DS_2-VASc, Congestive heart failure, Hypertension, Age, Diabetes, Stroke/TIA, Vascular disease, Age and Sex category; MRS, Modified Rankin Scale; NIHSS, National Institutes of Health Stroke Scale; TIA, Transient Ischemic Attack; PFO, Patent Foramen Ovale.

FibriCheck® compared to ILR – A total of 1,047,808 FibriCheck® measurements were performed, including 25,415 with a smartphone and 1,022,393 with a smartwatch. Of this total of FibriCheck® measurements, 602,538 (57.50%) were normal (i.e., sinus rhythm), 414,142 (39.50%) were of insufficient quality, 30,252 (2.90%) were warning (i.e., other arrhythmias), and 876 (0.10%) were urgent (i.e., AF episode). Besides this, a total of 1,830 AF detections were done by ILR, of which 1,088 (59.50%) had no annotation, 81 (4.40%) were annotated as AF approved, 6 (0.30%) were indeterminate, 383 (20.90%) were inappropriate, 211 (11.50%) were annotated as sinus rhythm, 56 (3.10%) were extrasystoles or premature beats, 1 (0.10%) was overseeing, and 2 (0.10%) were noise.

To compare the detection of AF between FibriCheck® and ILR, the percentages of FibriCheck® measurements annotated by ILR were calculated. Each FibriCheck® measurement was compared and linked to the ILR-approved and telemonitored annotation (sinus rhythm or AF appropriate), as illustrated in Figures 3A and B. Furthermore, the detection of sinus rhythm and other arrhythmias by both detection methods was evaluated and shown in Figures 3B and C.

In Figure 3A, the urgent (red) FibriCheck® measurements (i.e., AF episodes) were compared and linked to the approved ILR annotations (sinus rhythm or AF appropriate). A comparison between the two groups shows that the smartphone group has a significantly higher percentage of true-positive measurements (AF detected by FibriCheck® and annotated as AF appropriate by ILR) than the smartwatch group (17.52% vs. 0%, $p<0.001$). Within each group, the percentage of urgent measurements annotated as AF appropriate by ILR was substantially lower than those annotated as sinus rhythm. In the smartphone group, 82.47% of all urgent measurements were false positives (AF detected by FibriCheck® and annotated as sinus rhythm by ILR), compared with 17.52% true positives. Moreover, the smartwatch group had 0% true positives compared to 100% false positives ($p<0.001$). Notably, one patient in the smartwatch group had an AF episode detected by FibriCheck® and annotated as AF-appropriate. However, this case is outside the analyzed sample size due to the incomplete six-month follow-up.

The patient in this case was a 72-year-old male who was diagnosed with a TIA, had a BMI of 23.59 kg/m², had hypercholesterolemia, and was a former smoker. The ILR recorded an AF episode from 02:34 a.m. until 07:20 a.m. (more than 4 hours). During this period, multiple smartwatch FibriCheck® measurements were taken, identifying three AF episodes separated by a green and blue measurement: the first from 02:29 a.m. until 05:47 a.m. (more than 3 hours), the second from 05:56 a.m. until 06:23 a.m. (approximately half an hour), and the third from 06:32 a.m. until 07:08 a.m. (more than 30 minutes). A screenshot of the AF episode is shown in Figure 1S.

In Figure 3B, normal (green) FibriCheck® measurements (i.e., sinus rhythm) were linked to ILR-approved annotation (sinus rhythm or AF appropriate). Within each group, a significantly higher percentage of normal FibriCheck® measurements were annotated by ILR as sinus rhythm compared with AF-appropriate. These sinus rhythm annotated measurements were nearly 100% in both groups (99.99% vs 0.01% in the smartphone group and 99.999% vs 0.001% in the smartwatch group, $p<0.001$).

In Figure 3C, the warning (orange) FibriCheck® measurements (i.e., other arrhythmia) were linked to the ILR-approved annotations (AF appropriate or no AF), with 99.99% and 100% of all warning measurements annotated as no AF in the smartphone and smartwatch groups, respectively.

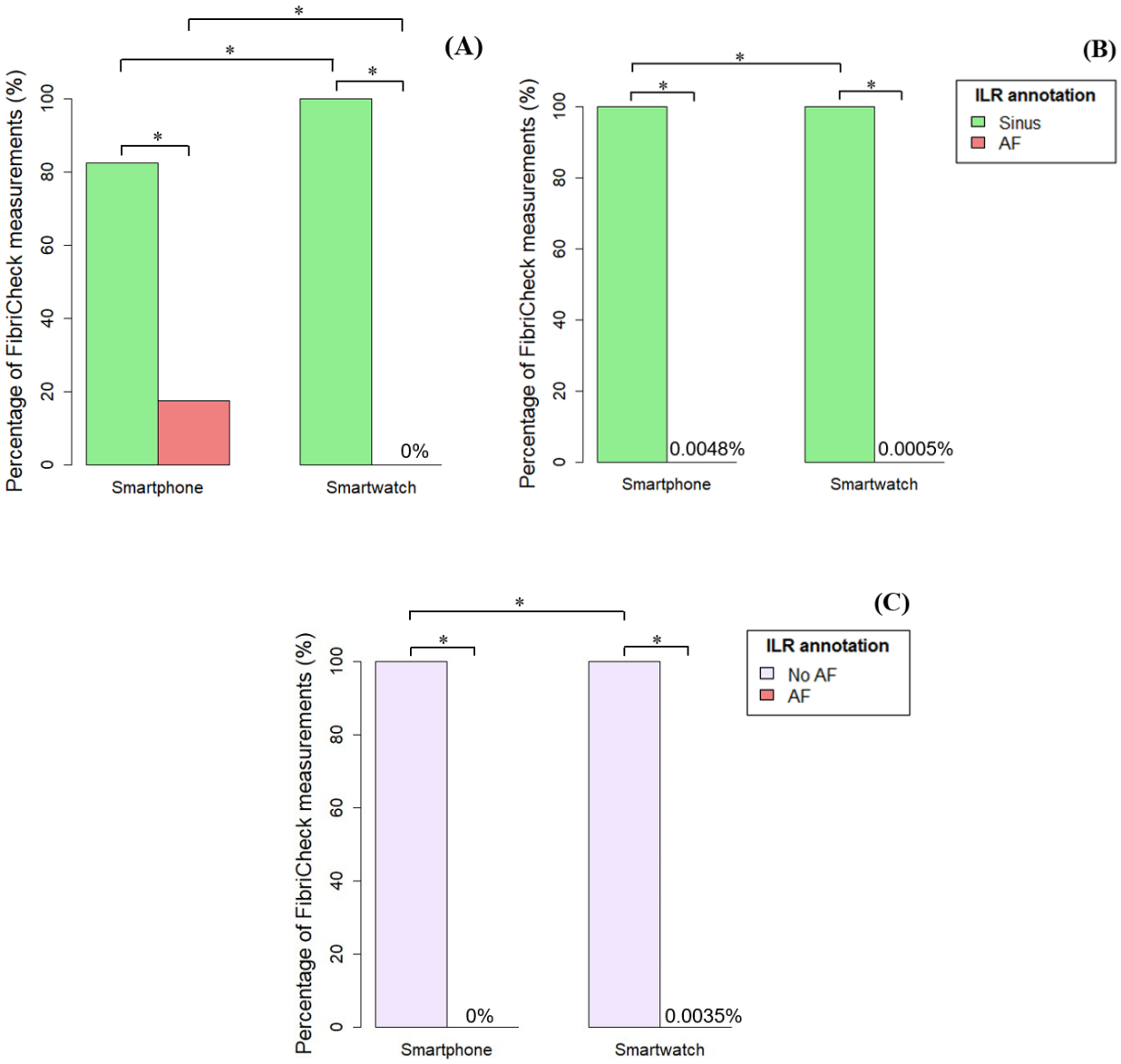

Figure 3. Number of FibriCheck®measurements per device compared to the annotation of ILR.
(A) Percentage of urgent/red FibriCheck® measurements annotated by the ILR: sinus rhythm (green) and AF appropriate (red) **(B)** Percentage of normal/green FibriCheck® measurements annotated by the ILR: sinus rhythm (green) and AF appropriate (red) and **(C)** Percentage of warning/orange FibriCheck® measurements annotated by the ILR: sinus rhythm (green) and no AF (light purple).

Statistical analysis was done using the McNemar Chi-squared test for paired data and the Pearson Chi-squared test for unpaired data.
*p<0.001.

ILR, implantable loop recorder

Additionally, the true-positive, false-positive, true-negative, and false-negative recordings were obtained for FibriCheck® measurements in both the smartphone and smartwatch groups on measurement level. This allowed the determination of sensitivity, specificity, and accuracy on a measurement level, which are presented in Table 2 and Table 3. In the smartphone group, the sensitivity, specificity, and accuracy of FibriCheck® measurements were 94.44%, 99.61%, and 99.61%, respectively. Specifically, there were 17 true-

positive, 1 false-negative, 20,438 true-negative, and 80 false-positive recordings. For the smartwatch group, the specificity of FibriCheck® measurements was 99.87% with 580,421 true-negative, 3 true-positives, and 779 false-positive. However, the calculation of sensitivity was not possible due to the absence of true-positive cases. Furthermore, the true-positive, false-positive, true-negative, and false-negative measurements were obtained at the patient level, allowing the sensitivity, specificity, and accuracy of the FibriCheck® application to be calculated, as presented in Table 4. The sensitivity was 68.75%, the specificity was 61.49%, and the accuracy was 62.11%.

Table 2. Number of smartphone measurements by FibriCheck® and ILR.

		FibriCheck	
		AF	**Sinus**
ILR	**AF**	17	1
	Sinus	80	20 438
		Sensitivity	94.44%
		Specificity	99.61%
		Accuracy	99.61%

AF, Atrial Fibrillation; ILR, implantable loop recorder

Table 3. Number of smartwatch measurements by FibriCheck® and ILR.

		FibriCheck	
		AF	**Sinus**
ILR	**AF**	0	3
	Sinus	779	580 421
		Sensitivity	N/A
		Specificity	99.87%

AF, Atrial Fibrillation; ILR, implantable loop recorder; NA, not applicable

Table 4. Number of patients with AF or sinus rhythm annotated as AF approved or sinus rhythm by ILR.

		FibriCheck	
		AF	**Sinus**
ILR	**AF**	11	5
	Sinus	67	107
		Sensitivity	68.75%
		Specificity	61.49%
		Accuracy	62.11%

AF, Atrial Fibrillation; ILR, implantable loop recorder

Quality of smartwatch FibriCheck® measurements – There were a total of 412,742 insufficient quality smartwatch measurements and 3088 insufficient quality smartphone measurements, representing 47.6% and 12.15% of all smartwatch and smartphone measurements, respectively.

To determine the quality differences between FibriCheck® measurements in the smartwatch group during night vs day, the percentages of low-quality and high-quality FibriCheck® measurements were calculated and presented in Figure 4. The percentage of low-quality measurements during the day is significantly higher compared to the percentage during the night (respectively, 58.46% and 7.46%, $p<0.001$). The percentage of high-quality measurements is significantly lower during the day compared to night (42.54% and 92.54%, $p<0.001$).

Characteristics of AF detection methods – Table 5 shows significant differences in the detection and duration of AF episodes between different detection methods and devices (FibriCheck® vs ILR and smartphone vs smartwatch). FibriCheck® shows significantly higher detection (962 vs 81 AF episodes, $p<0.001$) and longer duration (23 [9 – 71.5] minutes vs 2 [2 – 16.80] minutes, $p<0.001$) of AF episodes than ILR. When using FibriCheck®, there is a significant difference in the duration of AF episodes between smartphones and smartwatches, with smartwatches showing a shorter duration of AF episodes (334 [22.90 – 879] minutes vs 12 [9 – 27] minutes, $p<0.001$). The number of reminder

notifications to perform a FibriCheck® measurement was higher in the smartwatch group compared with the smartphone group, although these results did not differ significantly.

To gain more insight into the duration of AF episodes detected by ILR and FibriCheck®, the number of the first detected and approved AF episodes within a defined duration category is compared between the two detection methods and presented in Figure 2S. The duration of the AF episodes is divided into 6 categories. The number of AF episodes with a duration between 6 minutes and 1 hour was significantly higher with FibriCheck® compared with ILR (43 vs. 3, respectively, p <0.001). There were no significant differences in the number of AF episodes detected by the two detection methods in the other 5 categories.

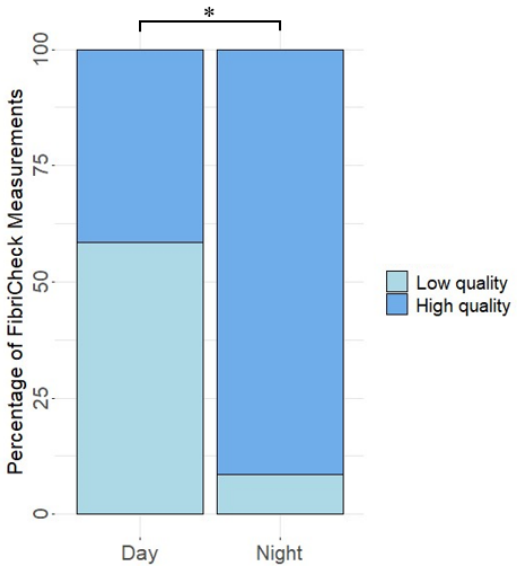

Figure 4. Quality of smartwatch FibriCheck® measurements during day and night.
The quality annotation is presented in percentages of all smartwatch FibriCheck® measurements. Low-quality measurements (light blue) indicate the measurements annotated as insufficient quality (blue measurement) by the FibriCheck® algorithm. High-quality measurements (dark blue) indicate the measurements annotated as normal, warning, and urgent (green, orange, and red measurements) by the FibriCheck® algorithm.
Daytime was defined from 8 a.m. until 10 p.m., and nighttime was defined from 12 a.m. until 6 a.m (27).

Statistical analysis was done using the McNemar Chi-squared test.
*p<0.001.

Table 5. Characteristics of the measurements.

	FibriCheck	ILR	p-value
Number of AF episodes (n)	962	81	**<0.001**
Duration of first detected AF episodes, minutes (median [IQR])	23 [9 – 71.50]	2 [2 – 16.80]	**<0.001**

	FibriCheck		
	Smartphone	**Smartwatch**	**p-value**
Duration of first detected AF episodes, minutes (median [IQR])	334 [22.90 – 879]	12 [9 – 27]	**<0.001**
Reminder notifications (median [IQR])	5 [2 – 10]	6 [1 – 19]	0.273

Statistical analysis was done using the Pearson Chi-squared test for categoric data, and Wilcoxon signed rank or Mann Whitney U test for continuous data with a significance level of p<0.05.

AF, atrial fibrillation; ILR implantable loop recorder; IQR, interquartile range.

DISCUSSION

The primary outcome was to determine the detection of AF using PPG-based mHealth (FibriCheck®) on a smartphone or smartwatch compared to ILR in cryptogenic stroke or TIA patients over a six-month period. The secondary outcome was to determine the quality of smartwatch FibriCheck® measurements between day and night in cryptogenic ischemic stroke and TIA patients over a six-month period.

FibriCheck® compared to ILR – The percentage of true positive AF detections was significantly higher in the smartphone group compared to the smartwatch group. On the other hand, the percentage of true-negative AF detections was similar in both groups. Further, the number of false-negative AF detections at the patient level was 5, which are the patients that FibriCheck® failed to detect. When assessing the sensitivity, specificity, and accuracy on measurements level, it resulted in, respectively, 94.44%, 99.61%, and 99.61% in the smartphone group and a specificity of 99.81% in the smartwatch group. This is consistent with a study by Proesmans *et al.* (24). This study tested the diagnostic accuracy of the FibriCheck® mobile phone app in comparison to the 12-lead ECG. The study population was patients with known paroxysmal or persistent AF. According to this study, FibriCheck® had a sensitivity of 95.3%, a specificity of 96.20%, and an accuracy of 95.8% on the measurements level. On the other hand, when assessing the sensitivity, specificity, and accuracy on the patient level, it resulted in a sensitivity of 68.75%, a specificity of 61.49%, and an accuracy of 62.11%. These results are contradictory to the results from the study by Proesmans *et al.*, which had a significantly higher sensitivity, specificity, and accuracy (respectively 95.6%, 96.6%, and 96.1%) compared to this study. The lower sensitivity, specificity, and accuracy observed in this study can be attributed to the patient-level categorization for AF. In this context, a patient is classified as an AF patient if at least one red measurement was detected (i.e., AF episode). Consequently, a patient with mainly green and only one or two red measurements is still considered an AF patient. This is the case for 24 patients. As mentioned in the results, there is a relatively high number of false-positives at measurement level. These large numbers of false-positives can affect the overall categorization of patients, leading to the potential for incorrect categorization of patients and, thus, possibly an overestimation of the AF patients. Moreover, the proportion of data collected at the measurement level is much higher than the collected data at the patient level. Both factors could affect the sensitivity, specificity, and accuracy of FibriCheck®. However, further research is needed to confirm this assumption.

A study by Selder *et al.* aimed to assess the accuracy of AF detection with FibriCheck® on the smartwatch. Here, the specificity was 99% on the measurement level, which is in line with our results (specificity of 99.81%). Further, Selder *et al.* found a sensitivity of 95% and an accuracy of 97%. In our study, it was impossible to determine sensitivity and accuracy because there were no true-positive measurements (28). Further follow-up of patients may provide this result in the future, as there was one case reported with a true-positive value, but it was not part of the analyzed patients of this study.

Quality of smartwatch FibriCheck® measurements – Since not everyone has the same bedtime and wake-up time, the periods between 10 p.m. and 12 a.m. and between 6 a.m. and 8 a.m. were excluded from the analysis (27). As assumed, the number of low-quality smartwatch measurements was significantly higher during the day, whereas the high-quality measurements were substantially lower during the day compared to the night. In a case report by Wouters *et al*., the quality of FibriCheck® measurements was compared between smartphones and smartwatches. In this report, the low-quality measurements were significantly more present in the smartwatch group compared to the smartphone group. This could be explained by the fact that patients in the smartphone group intentionally took the measurements without motion when performing the measurements. In our study, we focused exclusively on smartwatch measurements. The higher frequency of low-quality measurements during the day can probably be attributed to increased physical activity. Additionally, it was noted that measurement quality was significantly higher at night, likely due to reduced movement compared to daytime (25).

Characteristics of AF detection methods – The duration of the AF episodes was significantly longer in the smartphone group compared with the smartwatch group (334 [22.90 – 879] minutes vs 12 [9 – 27] minutes, p<0.001), which was expected. This may be due to the method used to estimate the start and end time of the AF episode, as described in the methods. The smartphone group was required to take at least two FibriCheck® measurements a day and spread them throughout the day (e.g., in the morning and in the evening). These were sometimes forgotten by the participants, resulting in a longer period between the last normal measurements and the first AF measurement and between the last AF measurement and the next measurement of regular rhythm. This may indicate that the AF episode is longer, but this is probably due to the compliance of the measurements. When comparing the duration of AF episodes between FibriCheck® and ILR, it resulted in a longer duration of AF when using FibriCheck® against ILR (23 [9 – 71.50] vs 2 [2 – 16.80] minutes, p<0.001). This might also be due to the method used to estimate the duration of the AF episode.

Furthermore, FibriCheck® shows a significantly higher detection rate of AF episodes compared to ILR (962 vs. 81, p<0.001), and a significantly higher number of AF episodes with a duration between 6 minutes and 1 hour was detected by FibriCheck® compared to ILR (43 vs. 3, respectively, p<0.001). However, it is important to note that this study showed a large number of false-positives, making these findings uninterpretable. In many cases, only one red measurement was observed among mainly green measurements.

Limitations – There are several challenges in ensuring the correct use of FibriCheck® in patients with cryptogenic ischemic stroke. Since the study group consists mainly of older patients, age-related issues play an important role. For this population, the use of technology can be challenging, leading to a lower number of measurements. Technological problems further complicate data collection. Problems with Bluetooth connectivity, malfunctioning smartphones, or smartwatches can interfere with the use of the FibriCheck® app and inhibit data transfer. In addition, post-stroke symptoms add another level of challenge. Patients with cryptogenic ischemic stroke may have cognitive or physical problems that interfere with their ability to use FibriCheck® consistently, resulting in inconsistent monitoring and reduced data reliability. Further, the study is ongoing and not all patients were fully followed. The analysis did not include these patients, resulting in unequal group sizes.

Lastly, only the AF-approved episodes by ILR were used in the analysis. However, it is possible that the episodes marked as 'no annotation' could also be relevant. This is because in patients with numerous episodes, not all episodes may have been approved and, therefore, not included in the analysis. In 28 red measurements, there was an AF without annotation detected by the ILR. These cases were possibly appropriate AF episodes detected by ILR.

CONCLUSION

In conclusion, FibriCheck® has promising potential for AF detection in patients with cryptogenic stroke or TIA. Further research and larger-scale studies are needed to confirm these findings and investigate the integration of FibriCheck® into standard clinical protocols for secondary stroke prevention. These conclusions are based on comparing and analyzing long-term follow-up for AF detection with FibriCheck® versus the standard method, ILR.

REFERENCES

1. The top 10 causes of death: World Health Organization 2020 [Available from: https://www.who.int/news-room/fact-sheets/detail/the-top-10-causes-of-death.
2. Khaku AS, Tadi P. Cerebrovascular Disease. StatPearls. Treasure Island (FL)2023.
3. Hui C, Tadi P, Patti L. Ischemic Stroke. StatPearls. Treasure Island (FL)2023.
4. Unnithan AKA, J MD, Mehta P. Hemorrhagic Stroke. StatPearls. Treasure Island (FL)2023.
5. Adams HP, Jr., Bendixen BH, Kappelle LJ, Biller J, Love BB, Gordon DL, Marsh EE, 3rd. Classification of subtype of acute ischemic stroke. Definitions for use in a multicenter clinical trial. TOAST. Trial of Org 10172 in Acute Stroke Treatment. Stroke. 1993;24(1):35-41.
6. Yaghi S, Elkind MS. Cryptogenic stroke: A diagnostic challenge. Neurol Clin Pract. 2014;4(5):386-93.
7. Chalos V, van der Ende NAM, Lingsma HF, Mulder M, Venema E, Dijkland SA, et al. National Institutes of Health Stroke Scale: An Alternative Primary Outcome Measure for Trials of Acute Treatment for Ischemic Stroke. Stroke. 2020;51(1):282-90.
8. Birenbaum D, Bancroft LW, Felsberg GJ. Imaging in acute stroke. West J Emerg Med. 2011;12(1):67-76.
9. Panuganti KK, Tadi P, Lui F. Transient Ischemic Attack. StatPearls. Treasure Island (FL)2023.
10. Hindricks G, Potpara T, Dagres N, Arbelo E, Bax JJ, Blomstrom-Lundqvist C, et al. 2023 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. 2023;42(5):373-498.
11. Elsheikh S, Hill A, Irving G, Lip GYH, Abdul-Rahim AH. Atrial fibrillation and stroke: State-of-the-art and future directions. Curr Probl Cardiol. 2024;49(1 Pt C):102181.
12. January CT, Wann LS, Alpert JS, Calkins H, Cigarroa JE, Cleveland JC, Jr., et al. 2014 AHA/ACC/HRS guideline for the management of patients with atrial fibrillation: executive summary: a report of the American College of Cardiology/American Heart Association Task Force on practice guidelines and the Heart Rhythm Society. Circulation. 2014;130(23):2071-104.
13. Nesheiwat Z, Goyal A, Jagtap M. Atrial Fibrillation. StatPearls. Treasure Island (FL)2023.
14. Oberman R, Shumway KR, Bhardwaj A. Physiology, Cardiac. StatPearls. Treasure Island (FL)2023.
15. Mohan R, Boukens BJ, Christoffels VM. Lineages of the Cardiac Conduction System. J Cardiovasc Dev Dis. 2017;4(2).
16. Moyer CB, Norton PT, Ferguson JD, Holmes JW. Changes in Global and Regional Mechanics Due to Atrial Fibrillation: Insights from a Coupled Finite-Element and Circulation Model. Ann Biomed Eng. 2015;43(7):1600-13.
17. Ibrahim F, Murr N. Embolic Stroke. StatPearls. Treasure Island (FL)2023.
18. Escudero-Martinez I, Morales-Caba L, Segura T. Atrial fibrillation and stroke: A review and new insights. Trends Cardiovasc Med. 2023;33(1):23-9.
19. Kamel H, Okin PM, Elkind MS, Iadecola C. Atrial Fibrillation and Mechanisms of Stroke: Time for a New Model. Stroke. 2016;47(3):895-900.
20. Gazova A, Leddy JJ, Rexova M, Hlivak P, Hatala R, Kyselovic J. Predictive value of CHA2DS2-VASc scores regarding the risk of stroke and all-cause mortality in patients with atrial fibrillation (CONSORT compliant). Medicine (Baltimore). 2019;98(31):e16560.
21. Rubiera M, Aires A, Antonenko K, Lemeret S, Nolte CH, Putaala J, et al. European Stroke Organisation (ESO) guideline on screening for subclinical atrial fibrillation after stroke or transient ischaemic attack of undetermined origin. Eur Stroke J. 2022;7(3):VI.
22. mHealth programmes are sponsored by government WHO World Health Organization [Available from: https://www.who.int/data/gho/indicator-metadata-registry/imr-details/4774.
23. 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;24(6):979-1005.
24. Proesmans T, Mortelmans C, Van Haelst R, Verbrugge F, Vandervoort P, Vaes B. Mobile Phone-Based Use of the Photoplethysmography Technique to Detect Atrial Fibrillation in Primary Care: Diagnostic Accuracy Study of the FibriCheck App. JMIR Mhealth Uhealth. 2019;7(3):e12284.

25. Wouters F, Gruwez H, Vranken J, Ernon L, Mesotten D, Vandervoort P, Verhaert D. Will Smartphone Applications Replace the Insertable Cardiac Monitor in the Detection of Atrial Fibrillation? The First Comparison in a Case Report of a Cryptogenic Stroke Patient. Front Cardiovasc Med. 2022;9:839853.
26. Wouters F, Gruwez H, Vranken J, Vanhaen D, Daelman B, Ernon L, et al. The Potential and Limitations of Mobile Health and Insertable Cardiac Monitors in the Detection of Atrial Fibrillation in Cryptogenic Stroke Patients: Preliminary Results From the REMOTE Trial. Front Cardiovasc Med. 2022;9:848914.
27. Hermans ANL, Isaksen JL, Gawalko M, Pluymaekers N, van der Velden RMJ, Snippe H, et al. Accuracy of continuous photoplethysmography-based 1 min mean heart rate assessment during atrial fibrillation. Europace. 2023;25(3):835-44.
28. Selder JL, Te Kolste HJ, Twisk J, Schijven M, Gielen W, Allaart CP. Accuracy of a Standalone Atrial Fibrillation Detection Algorithm Added to a Popular Wristband and Smartwatch: Prospective Diagnostic Accuracy Study. J Med Internet Res. 2023;25:e44642.

Acknowledgments – First of all, I would like to thank my supervisors, Femke Wouters and Myrte Barthels, for the internship as well as the valuable feedback and support I received. Furthermore, I would like to thank my fellow students in Future Health, especially Anouk, Jelena, Marie, Ruben, Janne, and Imke, for the enjoyable period, support, and feedback.

Author contributions – Femke Wouters conceived and designed the research. Heike Blockx and Femke Wouters performed the study procedure and data analysis. Heike Blockx wrote the paper. All authors carefully reviewed the paper.

SUPPLEMENTARY FIGURES

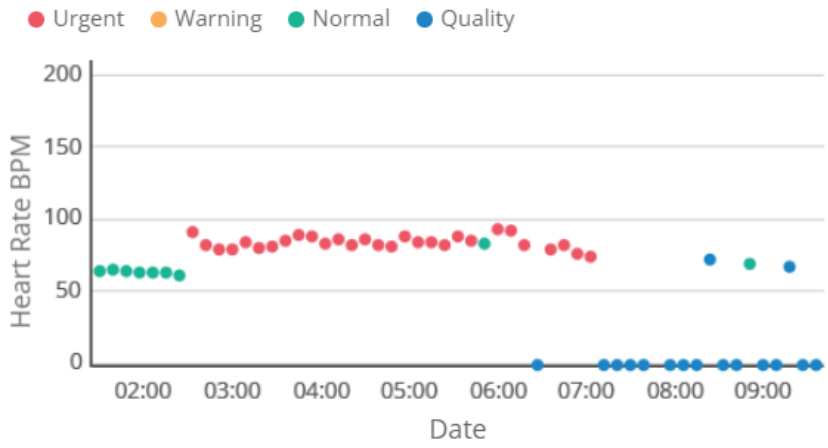

Figure 1S: print screen of the FibriCheck® measurements from the FibriCheck® dashboard of the study patient

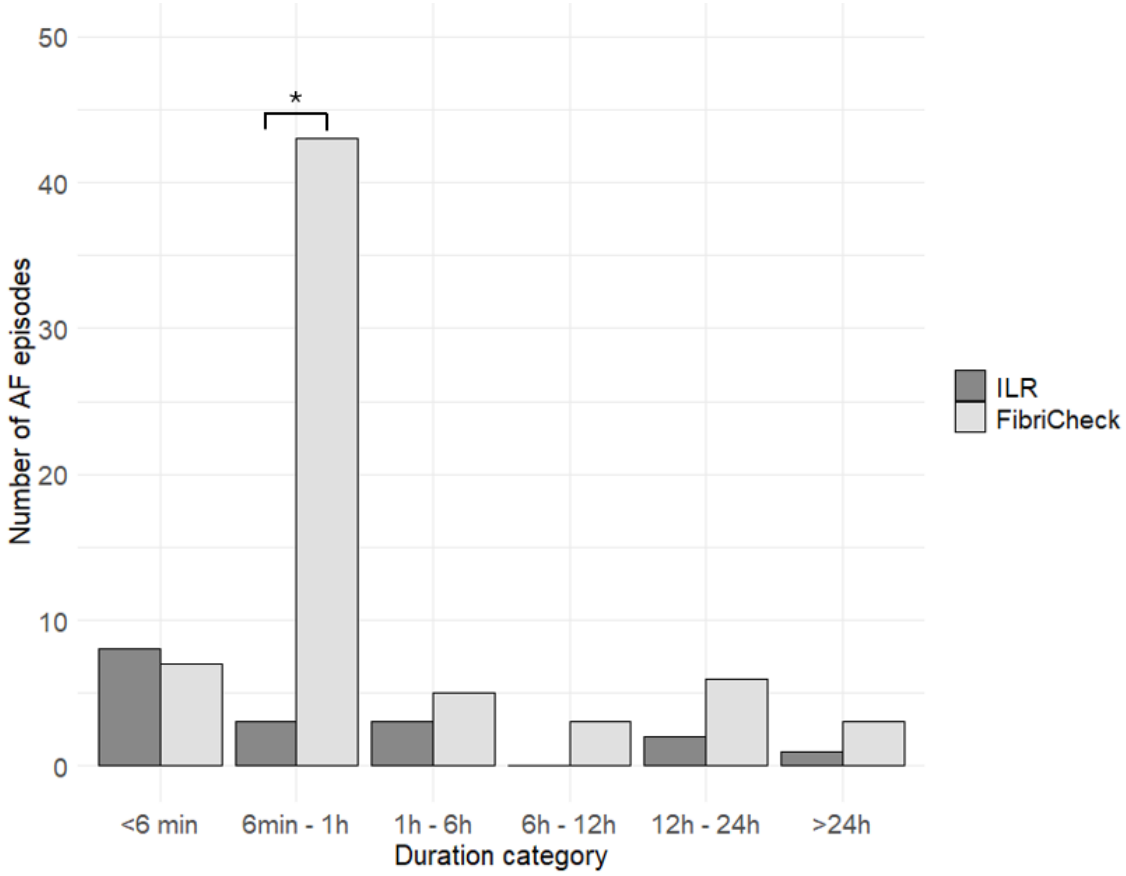

Figure 2S. Number of approved AF episodes detected by ILR vs FibriCheck® per duration category.
Statistical analysis was done using the Pearson Chi-squared test and Fisher exact test.
*p<0.001.

AF, atrial fibrillation; ILR, implantable loop recorder.