ORIGINAL ARTICLE

Large-Scale Assessment of a Smartwatch to Identify Atrial Fibrillation

Marco V. Perez, M.D., Kenneth W. Mahaffey, M.D., Haley Hedlin, Ph.D., John S. Rumsfeld, M.D., Ph.D., Ariadna Garcia, M.S., Todd Ferris, M.D., Vidhya Balasubramanian, M.S., Andrea M. Russo, M.D., Amol Rajmane, M.D., Lauren Cheung, M.D., Grace Hung, M.S., Justin Lee, M.P.H., Peter Kowey, M.D., Nisha Talati, M.B.A., Divya Nag, Santosh E. Gummidipundi, M.S., Alexis Beatty, M.D., M.A.S., Mellanie True Hills, B.S., Sumbul Desai, M.D., Christopher B. Granger, M.D., Manisha Desai, Ph.D., and Mintu P. Turakhia, M.D., M.A.S., for the Apple Heart Study Investigators*

ABSTRACT

BACKGROUND

Optical sensors on wearable devices can detect irregular pulses. The ability of a smartwatch application (app) to identify atrial fibrillation during typical use is unknown.

METHODS

Participants without atrial fibrillation (as reported by the participants themselves) used a smartphone (Apple iPhone) app to consent to monitoring. If a smartwatch-based irregular pulse notification algorithm identified possible atrial fibrillation, a telemedicine visit was initiated and an electrocardiography (ECG) patch was mailed to the participant, to be worn for up to 7 days. Surveys were administered 90 days after notification of the irregular pulse and at the end of the study. The main objectives were to estimate the proportion of notified participants with atrial fibrillation shown on an ECG patch and the positive predictive value of irregular pulse intervals with a targeted confidence interval width of 0.10.

RESULTS

We recruited 419,297 participants over 8 months. Over a median of 117 days of monitoring, 2161 participants (0.52%) received notifications of irregular pulse. Among the 450 participants who returned ECG patches containing data that could be analyzed — which had been applied, on average, 13 days after notification — atrial fibrillation was present in 34% (97.5% confidence interval [CI], 29 to 39) overall and in 35% (97.5% CI, 27 to 43) of participants 65 years of age or older. Among participants who were notified of an irregular pulse, the positive predictive value was 0.84 (95% CI, 0.76 to 0.92) for observing atrial fibrillation on the ECG simultaneously with a subsequent irregular pulse notification and 0.71 (97.5% CI, 0.69 to 0.74) for observing atrial fibrillation on the ECG simultaneously with a subsequent irregular tachogram. Of 1376 notified participants who returned a 90-day survey, 57% contacted health care providers outside the study. There were no reports of serious app-related adverse events.

CONCLUSIONS

The probability of receiving an irregular pulse notification was low. Among participants who received notification of an irregular pulse, 34% had atrial fibrillation on subsequent ECG patch readings and 84% of notifications were concordant with atrial fibrillation. This siteless (no on-site visits were required for the participants), pragmatic study design provides a foundation for large-scale pragmatic studies in which outcomes or adherence can be reliably assessed with user-owned devices. (Funded by Apple; Apple Heart Study ClinicalTrials.gov number, NCT03335800.)

From the Division of Cardiovascular Medicine (M.V.P.), Stanford Center for Clinical Research (K.W.M., A.R., N.T.), the Quantitative Sciences Unit (H.H., A.G., V.B., J.L., S.E.G., M.D.), Information Resources and Technology (T.F., G.H.), Department of Medicine (S.D.), and the Center for Digital Health (M.P.T.), Stanford University, Stanford, Apple, Cupertino (L.C., D.N., A.B., S.D.), and the Veterans Affairs Palo Alto Health Care System, Palo Alto (M.P.T.) — all in California; the University of Colorado School of Medicine, Aurora (J.S.R.); the Division of Cardiovascular Disease, Cooper Medical School of Rowan University, Camden, NJ (A.M.R.); the Lankenau Heart Institute and Jefferson Medical College, Philadelphia (P.K.); StopAfib.org, American Foundation for Women's Health, Decatur, TX (M.T.H.); and the Duke Clinical Research Institute, Duke University, Durham, NC (C.B.G.). Address reprint requests to Dr. Perez or Dr. Turakhia at Stanford Center for Clinical Research, 1070 Arastradero Rd., Palo Alto, CA 94304, or at mvperez@stanford.edu or mintu@stanford.edu.

*A complete list of the Apple Heart Study Investigators is provided in the Supplementary Appendix, available at NEJM.org.

Drs. M. Desai and Turakhia contributed equally to this article.

N Engl J Med 2019;381:1909-17.
DOI: 10.1056/NEJMoa1901183

Wearable devices with optical sensors, such as smartwatches, are commonly used to measure wearers' pulse rates.[1] Algorithms that use pulse wave data to detect atrial fibrillation and atrial flutter have been developed.[1,2] An Apple Watch application (app) can use intermittent, passively detected pulse rate data in an algorithm that identifies episodes suggestive of atrial fibrillation.[3]

Atrial fibrillation (which in this article also refers to atrial flutter) is the most commonly diagnosed clinically significant cardiac arrhythmia and affects approximately 6 million people in the United States,[4] with a lifetime risk as high as 1 in 3.[5] Atrial fibrillation is associated with a quintupling of the risk of stroke.[6] The paroxysmal nature of atrial fibrillation may result in diagnostic delays since the electrocardiogram (ECG) can appear normal between episodes. In addition, atrial fibrillation can be minimally symptomatic or clinically silent.[7] Approximately 700,000 people in the United States may have undiagnosed atrial fibrillation.[8] Continuous traditional heart monitors or implantable devices increase the detection of atrial fibrillation in populations at high risk[7,9-12] but have limited monitoring periods and require either invasive procedures or activation by the user.

The widespread use of Internet-connected devices provides an opportunity to conduct large, siteless, pragmatic trials at a lower cost. The goal of the Apple Heart Study was to evaluate the ability of an irregular pulse notification algorithm to identify atrial fibrillation with the use of an Apple Watch app by consumers.

Methods

Study Design and Oversight

Details of the study have been described previously.[13] This was a prospective, single-group, open-label, siteless, pragmatic study. The research protocol was approved by the institutional review board at Stanford University and by a central institutional review board (Advarra).

Apple sponsored the study and owns the data. All study data are stored at Stanford on Stanford data platforms. The analyses presented here were performed by Stanford quantitative scientists independent of the sponsor. Stanford has the right to publish regardless of the outcome. All the authors, including authors employed by the sponsor, reviewed and approved the manuscript and vouch for the accuracy and completeness of the data.

Study Population

The app, which used the irregular pulse notification algorithm, was available in the United States for download from the Apple App Store from the time the study launched on November 29, 2017, until August 1, 2018. Major eligibility criteria included possession of a compatible Apple iPhone and Apple Watch, an age of 22 years or older, United States residency, and proficiency in English, as reported by the participant. Participants who reported previous atrial fibrillation or current use of oral anticoagulation agents were not eligible. All participants provided electronically signed informed consent. (The consent form, along with a description of the algorithm, telemedicine visit protocol, and methods used for tachogram sampling, are provided in the Supplementary Appendix, available with the full text of this article at NEJM.org.)

Monitoring and Study Intervention

The study app was used to verify eligibility, obtain participants' consent, provide study education, and direct participants through the study procedures. After a participant provided consent, the irregular pulse notification algorithm was activated. The study used the Apple Watch photoplethysmography sensor, which used light-emitting and light-sensitive diodes to intermittently and passively measure changes in blood flow while participants were at rest. These signals were used to generate pulse intervals (tachograms) over 1 minute, which were classified as regular or irregular on the basis of the variation in the pulse interval (Fig. S1 in the Supplementary Appendix). Participants were prompted to initiate a telemedicine visit directly from the app. After the initial notification of an irregular pulse, subsequent tachograms and notifications were recorded but were not provided to the participant. The notification feature was active until September 1, 2018.

Study visits were conducted by physicians from a national telehealth servicer (American Well) with the use of a standardized protocol. Participants with urgent symptoms were directed to go to an urgent care clinic or emergency department. Participants whose eligibility was confirmed and whose symptoms were not urgent were mailed an ECG patch (ePatch) to wear for up to 7 days. The

ECG patches were returned by mail and initially examined by trained technicians. Participants with serious arrhythmias were contacted immediately and directed to seek urgent medical care. The ECG patch reports were read by two clinicians, and discrepant interpretations were then reconciled by a committee of clinicians coordinated by the Stanford Center for Clinical Research. In addition, 3-minute ECG strips from each patch, time-aligned to sampled tachograms, were separately read by two clinicians, with disagreements resolved by a third clinician and then a committee, if necessary.

Participants were prompted to initiate a second telemedicine visit to discuss the ambulatory ECG findings and were directed to subsequent care. Study-visit physicians did not initiate treatments. Participants who received irregular pulse notifications were asked to complete a survey, included in the study app, 90 days after notification. All enrolled participants, regardless of notification status, were directed to a Web-based end-of-study survey to be completed by January 31, 2019. All adverse events were reviewed by personnel at the study safety monitoring desk at the Stanford Center for Clinical Research.

PRIMARY AND SECONDARY OUTCOMES

There were two coprimary outcomes: atrial fibrillation of greater than 30 seconds' duration on ECG patch monitoring in a participant who received an irregular pulse notification, and simultaneous atrial fibrillation on ECG patch monitoring during intervals when the participant had an irregular tachogram. Key secondary outcomes were simultaneous atrial fibrillation on ECG patch monitoring when the pulse notification algorithm detected an irregular pulse and participant report of contact with a health care provider outside the study within 3 months after notification of an irregular pulse.

STATISTICAL ANALYSIS

We calculated the minimum number of participants with analyzable data from ECG patches that would ensure sufficient precision to estimate both the proportion of atrial fibrillation detected in participants 65 years of age or older and the positive predictive value of the tachogram. We targeted 503 ECG patches in each age group (<65 and ≥65 years), yielding 97.5% confidence intervals around the atrial fibrillation yield for participants 65 years of age or older and a positive predictive value of tachograms no wider than 0.10. Given the large volume and diversity of data, methods were applied so that key statistics presented were arrived at independently by at least two members of the data team to ensure reproducibility.

Statistical analyses are described in the statistical analysis plan, which is included in the Supplementary Appendix. Means and standard deviations are provided for continuous characteristics, and frequency distributions with percentages are presented for binary and categorical characteristics. For participants for whom key data were missing, we compared observed participant-level characteristics of participants with missing values with characteristics of participants without missing values to help in interpreting key relationships.

The ECG patch subgroup included participants who received a notification, reported no history of atrial fibrillation before enrollment, were not receiving anticoagulant therapy, had no urgent symptoms at the first study visit, and wore their ECG patch within 14 days after shipment for at least 1 hour and returned it within 45 days after the first study visit. We estimated the yield of atrial fibrillation for participants 65 years of age or older as the proportion who had confirmed atrial fibrillation on subsequent ECG patches. We characterized the positive predictive value of the tachograms by calculating the proportion of sampled irregular tachograms for which atrial fibrillation was confirmed on simultaneous ECG patch strips. We estimated the positive predictive value of the notification by calculating the proportion of participants with atrial fibrillation confirmed on at least one ECG strip that was simultaneous with the tachograms that led to the notification. Finally, we estimated the proportion of participants who reported contact with a health care provider among those who were notified and responded to the 90-day questionnaire. Confidence intervals for relevant quantities were provided on the basis of Gaussian assumptions or, if the Gaussian assumption yielded bounds that crossed 0 or 1, on the bias-corrected and accelerated bootstrap interval; 97.5% confidence intervals were provided for two key quantities (the yield of atrial fibrillation for participants 65 years of age or older and the positive predictive value of the tachogram) and 95% confidence intervals were provided for other estimates.

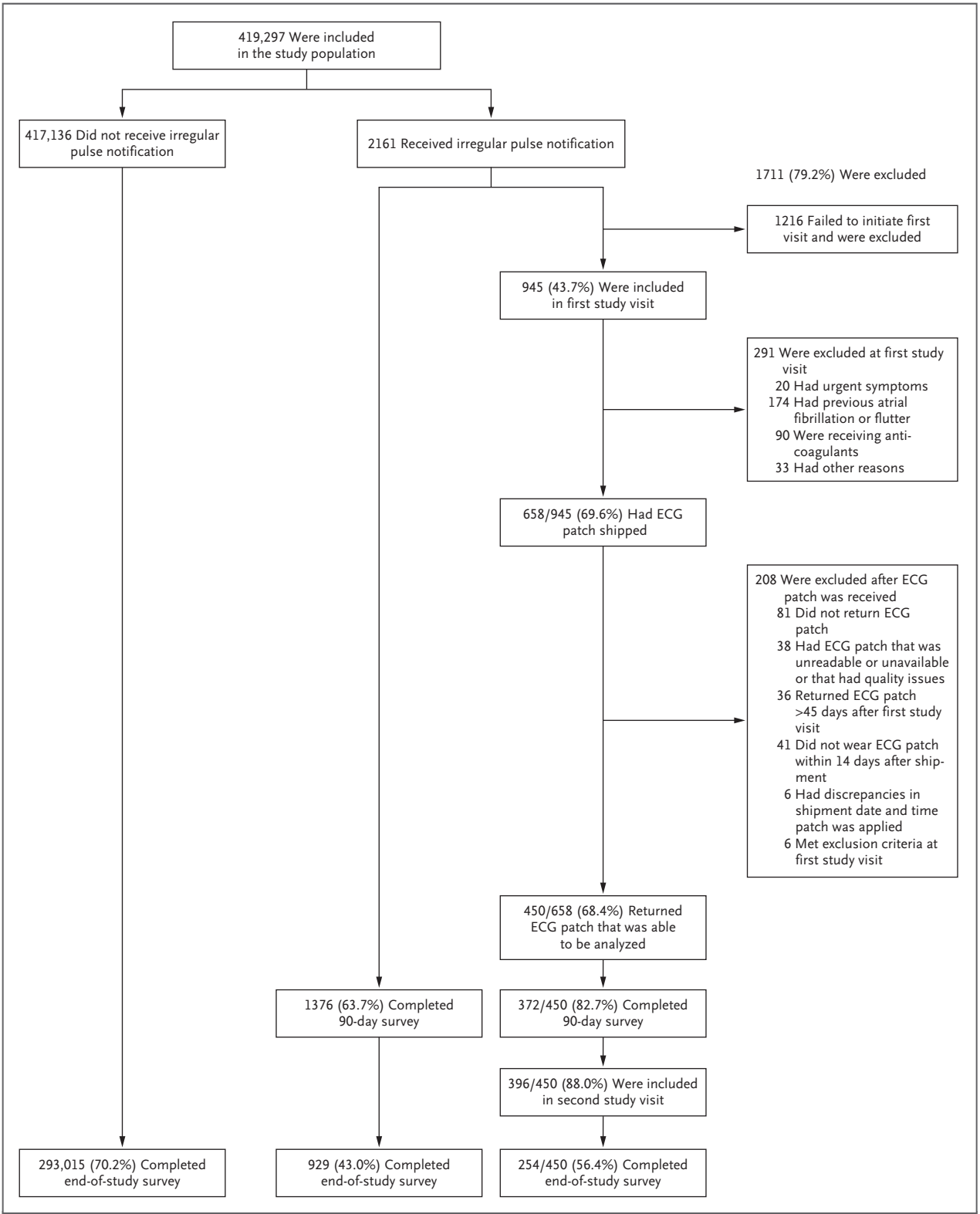

Figure 1. Participant Selection.

Of the 1216 potential participants who were excluded because they did not have a first study visit, 4 received an ECG patch. AF denotes atrial fibrillation, and ECG electrocardiography.

RESULTS

BASELINE CHARACTERISTICS

During an 8-month period, 419,297 participants were recruited from 50 states and the District of Columbia (Fig. 1). Baseline characteristics of the participants are summarized in Table 1.

IRREGULAR PULSE NOTIFICATIONS

Over a median monitoring time of 117 days (interquartile range, 113 to 186), irregular pulse notifications were received by 2161 participants (0.52%), ranging from 3.1% of those 65 years of age or older to 0.16% of those 22 to 40 years of age (Fig. 2). Participants who received irregular pulse notifications were older, less likely to be female, more likely to be white, and more likely to have a CHA_2DS_2-VASc score of 2 or higher than the overall cohort (Table 1). (Scores on the CHA_2DS_2-VASc, which is a measure of the risk of stroke among persons with atrial fibrillation, range from 0 to 9, with higher scores indicating a greater risk.) Among participants who received a notification, 50% received notification by day 38 after enrollment, and 90% by day 133 after enrollment (Fig. S2).

ATRIAL FIBRILLATION ON SUBSEQUENT AMBULATORY ECG MONITORING

Among participants who received a notification, 450 (20.8% of all notified) returned an ECG patch, which was applied a mean (±SD) of 13±16 days after initial notification. ECG patches were worn for an average of 6.3 days. Participants who returned ECG patches had baseline characteristics similar to the full cohort of participants who received a notification.

Among the 450 participants who returned ECG patches, atrial fibrillation was identified in 153, resulting in a diagnostic yield of atrial fibrillation on ECG patches of 34% (97.5% confidence interval [CI], 29 to 39) (Fig. 3). The ECG patches worn by participants 65 years of age or older had a diagnostic yield of atrial fibrillation of 35% (97.5% CI, 27 to 43), whereas among participants younger than 40 years of age, the diagnostic yield of atrial fibrillation was 18% (95% CI, 6 to 31).

Among the 153 participants with atrial fibrillation confirmed on ambulatory ECG, 20% had continuous atrial fibrillation, whereas most of the remaining participants with atrial fibrillation had atrial fibrillation less than 50% of the time they were monitored and 89% had an episode that lasted at least 1 hour (Figs. S3 and S4). Of the 20 participants who were urgently contacted, 18 had atrial fibrillation with ventricular rates greater than 200 beats per minute for more than 30 seconds, 1 had a pause lasting more than 6 seconds, and 1 had nonsustained ventricular tachycardia lasting more than 6 seconds.

POSITIVE PREDICTIVE VALUES

Of the 6968 tachograms sampled for adjudication, 270 were excluded because they were not of sufficient quality to be read. Of the 2089 irregular tachograms sampled from participants who had received a notification for analysis, 1489 showed simultaneous atrial fibrillation on ECG patch monitoring, resulting in a positive predictive value of the individual tachogram of 0.71 (97.5% CI, 0.69 to 0.74). For tachograms in the subgroup of participants 65 years of age or older, the positive predictive value was 0.60 (97.5% CI, 0.56 to 0.64). In the 600 irregular tachograms without simultaneous atrial fibrillation on ECG patch monitoring, frequent premature atrial contractions (6 or more over a 3-minute period) were identified in 77%, frequent premature ventricular contractions (6 or more over a 3-minute period) in 16%, and atrial tachycardias (3 or more consecutive beats) in 38%. The identification of these arrhythmias was not mutually exclusive. Sinus arrhythmia alone was found in 28 (4.7%) of the 600 irregular tachograms without atrial fibrillation.

Of the 86 participants who had irregular pulse notifications during simultaneous use of an ECG patch, 72 showed evidence of atrial fibrillation on concurrent ECG patch strips. This resulted in a positive predictive value for the irregular pulse notification of 0.84 (95% CI, 0.76 to 0.92) among participants who had received an irregular pulse notification. For irregular pulse notifications in participants 65 years of age or older, the positive predictive value was 0.78 (95% CI, 0.64 to 0.92).

90-DAY SURVEY

Of the 2161 participants who received an irregular pulse notification, 1376 (64%) returned a 90-day survey. Of these, 787 (57%) reported contact with a health care provider outside the study, 28% were prescribed a new medication, 33% were recommended to see a specialist (e.g., a cardiologist), and 36% were recommended to have additional testing. In total, 1041 (76%) stated that they had contacted the study visit doctor, a health care provider outside the study, or both.

Table 1. Characteristics of Participants Enrolled in the Apple Heart Study at Baseline.*

Characteristic	Total Cohort (N=419,297)	Notification Subgroup (N=2161)	ECG Patch Subgroup (N=450)
Sex — no. (%)†			
Female	177,087 (42)	461 (21)	102 (23)
Male	238,700 (57)	1672 (77)	335 (74)
Other	396 (0.1)	0	0
Not reported	3,114 (0.7)	28 (1.3)	13 (2.9)
Age — yr	41±13	57±15	59±14
Age distribution — no. (%)			
≥65 yr	24,626 (5.9)	775 (36)	181 (40)
55–64 yr	42,633 (10)	556 (26)	114 (25)
40–54 yr	132,696 (32)	488 (23)	106 (24)
22–39 yr	219,179 (52)	341 (16)	49 (11)
Not reported	163 (<0.1)	1 (<0.1)	0
Race or ethnic group — no. (%)†			
White	286,190 (68)	1747 (81)	379 (84)
Hispanic	48,775 (12)	104 (4.8)	20 (4.4)
Black	32,275 (7.7)	106 (4.9)	16 (3.6)
Asian	26,156 (6.2)	87 (4.0)	8 (1.8)
American Indian	4,696 (1.1)	20 (0.9)	3 (0.7)
Pacific Islander	1,493 (0.4)	6 (0.3)	0
Middle Eastern	3,652 (0.9)	9 (0.4)	2 (0.4)
Other or mixed race	7,958 (1.9)	32 (1.5)	6 (1.3)
Not reported	8,102 (1.9)	50 (2.3)	16 (3.6)
Medical history — no. (%)			
Obesity	160,197 (38)	984 (46)	192 (43)
Hypertension	86,338 (21)	917 (42)	200 (44)
Diabetes	20,443 (4.9)	255 (12)	53 (12)
Heart failure	2,511 (0.6)	72 (3.3)	10 (2.2)
Stroke or TIA	4,153 (1.0)	66 (3.1)	10 (2.2)
Peripheral artery disease	2,596 (0.6)	52 (2.4)	10 (2.2)
CHA_2DS_2-VASc score ≥2‡	55,277 (13)	713 (33)	171 (38)
Current smoking — no. (%)	25,458 (6.1)	88 (4.1)	10 (2.2)
Alcohol: ≥1 drink/wk — no. (%)	190,463 (45)	1092 (51)	227 (50)

* Plus–minus values are means ±SD. Percentages may not total 100 because of rounding. ECG denotes electrocardiography, and TIA transient ischemic attack.
† Sex and race or ethnic group were reported by the participants.
‡ Scores on the CHA_2DS_2-VASc, which is a measure of the risk of stroke among persons with atrial fibrillation, range from 0 to 9, with higher scores indicating a greater risk.

END-OF-STUDY SURVEY

Of the 2161 participants who received a notification, 929 (43%) completed an end-of-study survey; among 417,136 participants who never received a notification, 293,015 (70%) completed the survey (Table 2). Of those notified, 404 (44%) reported a new atrial fibrillation diagnosis, whereas among those who received no notification, 3070 (1.0%) reported a new atrial fibrillation diagnosis. The notification subgroup reported a greater incidence

of strokes, heart failure, and myocardial infarctions than did the non-notification group. The notification subgroup was also more likely to start receiving anticoagulant therapy or aspirin. Of the 404 notified participants who reported new atrial fibrillation, 95 (24%) reported undergoing cardioversion, 12 (3%) received an implantable loop recorder, 82 (20%) started antiarrhythmic therapy, and 71 (18%) underwent catheter ablation.

Of 1038 adverse events reviewed (Fig. S5), 16 (1.5%) were related to the app; of those, 15 were anxiety-related. None of the adverse events related to the app resulted in hospitalization or urgent medical attention.

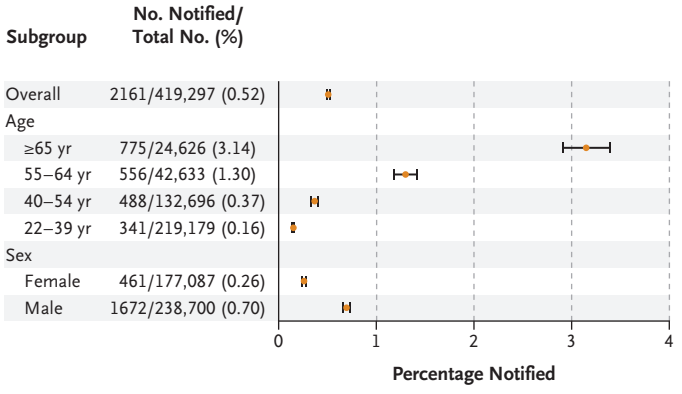

Figure 2. Irregular Pulse Notifications, According to Age and Sex.
Horizontal bars indicate 97.5% confidence intervals.

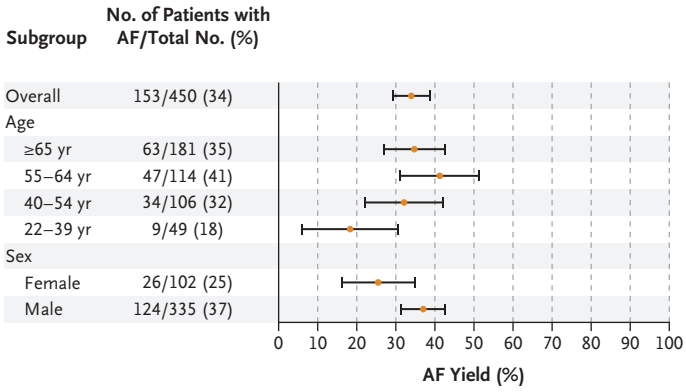

Figure 3. Yield of Atrial Fibrillation on ECG Patch Monitoring.
Horizontal bars indicate 97.5% confidence intervals.

DISCUSSION

The Apple Heart Study was a prospective, single-group study that was based on a siteless, pragmatic design. Of the 419,297 participants enrolled, only 0.52% received an irregular pulse notification, and among those with an initial notification who returned an ECG patch, 84% (95% CI, 76 to 92) of their subsequent notifications were confirmed to be atrial fibrillation. Of participants 65 years of age and older, 3.2% received notifications. These estimates may help providers better understand the implications of irregular pulse notifications when patients present for clinical care.

The overall yield of atrial fibrillation on an ECG patch was 34% among those who received notifications. This finding is clinically relevant because these participants had a relatively high burden of atrial fibrillation, with a majority of episodes lasting more than 1 hour. The absence of atrial fibrillation on a subsequent ECG patch does not imply that the initial notification was a false positive. Rather, atrial fibrillation may have been paroxysmal and infrequent, which is the most common pattern in early-stage atrial fibrillation. The index atrial fibrillation episode may have ended by the time the ECG patch was worn, which was, on average, 13 days after the initial notification.

Although the percentage of participants younger than 40 years of age who received notifications (0.16%) was low, the atrial fibrillation yield on ECG patch monitoring in this group was also lower (18%) than in other age groups. This may be a reflection of the paroxysmal nature of atrial fibrillation at the earlier stages of disease, but further studies are needed to better understand the public health implications of identifying irregular pulse in persons younger than 40 years of age.

The positive predictive value of an individual tachogram was 0.71 (97.5% CI, 0.69 to 0.74) and the positive predictive value of an irregular pulse notification was 0.84 (95% CI, 0.76 to 0.92), which suggests that algorithms that rely on confirmation of multiple irregular tachograms before triggering a notification improve accuracy. Many of the irregular tachograms not adjudicated as atrial fibrillation were instead concordant with rhythms that may warrant further clinical attention and require additional study. The positive predictive values were measured for participants who had already received an irregular pulse no-

Table 2. End-of-Study Survey.

Variable	Notification Subgroup (N=929)	Non-notification Subgroup (N=293,015)
New diagnosis — no. (%)		
Atrial fibrillation	404 (43)	3070 (1.0)
Stroke	7 (0.8)	321 (0.1)
TIA	12 (1.3)	498 (0.2)
Heart failure	30 (3.2)	648 (0.2)
Myocardial infarction	10 (1.1)	574 (0.2)
Major bleeding	7 (0.8)	842 (0.3)
Medication use — no. (%)*		
Warfarin	20 (2.2)	265 (0.1)
Direct oral anticoagulant	202 (22)	996 (0.3)
Aspirin	338 (36)	40,774 (14)

* This category refers to medication use since enrollment in the study, as reported by the participants.

tification and are therefore only an estimate of the positive predictive value of an initial notification in the overall cohort.

This study also provides insight into the way digital alerts result in engagement with the health care system. That 76% of notified participants who returned a survey contacted either the telemedicine provider or a nonstudy provider suggests that many actively sought medical attention. The remaining may have ignored the notification because they knew they had atrial fibrillation, were asymptomatic, did not trust the notification, or did not feel that the notification, even if true, required follow-up.

There are several limitations to the study. Participants did not initiate contact with the study provider after notification and fewer returned ECG patches (450 of 2161 notified) than anticipated. As a result, the targeted statistical precision for estimating the yield of atrial fibrillation on patch monitoring, which was one of our primary end points, was not met. The reported confidence intervals appropriately reflect the uncertainty of our key quantities of interest among participants who returned their ECG patches; however, the generalizability of these estimates to participants who did not return ECG patches remains uncertain. Nevertheless, no qualitative differences were observed between those notified and excluded from the analysis and those notified who provided ECG patches with data that could be analyzed. The study was not designed to assess the algorithm as a screening tool or to measure sensitivity, specificity, or false positive results. The algorithm was designed to minimize false positive findings,[3] and the low incidence of notifications reflects this intent. Furthermore, the algorithm was not designed to detect short episodes of atrial fibrillation, and participants with a low burden of atrial fibrillation could have been missed. The study objective was not to address the use of the Apple Watch as a population screening tool. Patients using this technology should be aware that the absence of an irregular pulse notification does not exclude possible arrhythmias. Conversely, notification based on an irregular pulse from a photoplethysmography signal should not be used for a definitive diagnosis of atrial fibrillation. Since rhythm-detection technologies are rapidly evolving, additional studies using features such as wearable ECG monitoring devices will need to be performed as the technology becomes available. Nevertheless, uncertainty remains about the benefits of diagnosing and treating asymptomatic atrial fibrillation, particularly in persons whose episodes of atrial fibrillation are of 6 hours' duration or less.

There was no direct physical contact with participants from the time of enrollment and consent to interaction with the telemedicine provider and ECG patch monitoring. Although our siteless, pragmatic study design allowed us to enroll more than 400,000 participants in 8 months, we relied on the participants' assessments regarding their eligibility for inclusion and regarding outcomes. Substantial loss to follow-up results in uncertain validity and generalizability inherent to this design. At enrollment, persons with previous atrial fibrillation were asked not to participate, but several participants who received notifications later reported a history of atrial fibrillation. Although we mitigated this misclassification by verifying enrollment criteria at the study visit, this kind of misclassification illustrates the challenges of relying on the participants themselves to assess enrollment eligibility and outcomes. In the future, studies may be able to leverage health record data directly from smartphones. As the number of app-based studies grows, development of methods to maximize engagement and the accuracy of data reported by participants is an important area of investigation. Although the participants we enrolled were geographically, racially, and ethnically diverse, the cohort was skewed

toward a younger demographic, reflective of smartwatch owners. Studies using similar designs will need to consider these factors to ensure that all affected age and socioeconomic groups are represented.

We found that the probability that a participant was notified of an irregular pulse was low, but among participants who were notified of an irregular pulse, more than one third had atrial fibrillation identified on a subsequently worn ECG patch monitor, and among those notified who returned an ECG patch, positive notifications were concordant with atrial fibrillation 84% (95% CI, 76 to 92) of the time. We believe that these data support the ability of the algorithm to correctly identify atrial fibrillation in users whom it notifies of irregular pulses. Rigorous investigation of this technology and of its use in a clinical setting is needed, including the ways this technology can guide further evaluation and treatment to improve clinical outcomes. Finally, this study provides a foundation on which further research in digital health can be conducted.

Supported by Apple.

Disclosure forms provided by the authors are available with the full text of this article at NEJM.org.

A data sharing statement provided by the authors is available with the full text of this article at NEJM.org.

We thank Rebecca Gardner from the Quantitative Sciences Unit for critical analyses and review of data; and Alaina Mitsch, Ph.D., of MedErgy HealthGroup (Yardley, PA) for technical editorial assistance with an earlier version of the manuscript, funded by Stanford University.

REFERENCES

1. Strain T, Wijndaele K, Brage S. Physical activity surveillance through smartphone apps and wearable trackers: examining the UK potential for nationally representative sampling. JMIR Mhealth Uhealth 2019;7(1):e11898.
2. Tison GH, Sanchez JM, Ballinger B, et al. Passive detection of atrial fibrillation using a commercially available smartwatch. JAMA Cardiol 2018;3:409-16.
3. Apple. Using Apple Watch for arrhythmia detection. December 2018 (https://www.apple.com/healthcare/site/docs/Apple_Watch_Arrhythmia_Detection.pdf).
4. Colilla S, Crow A, Petkun W, Singer DE, Simon T, Liu X. Estimates of current and future incidence and prevalence of atrial fibrillation in the U.S. adult population. Am J Cardiol 2013;112:1142-7.
5. Weng LC, Preis SR, Hulme OL, et al. Genetic predisposition, clinical risk factor burden, and lifetime risk of atrial fibrillation. Circulation 2018;137:1027-38.
6. Wolf PA, Abbott RD, Kannel WB. Atrial fibrillation as an independent risk factor for stroke: the Framingham Study. Stroke 1991;22:983-8.
7. Sanna T, Diener H-C, Passman RS, et al. Cryptogenic stroke and underlying atrial fibrillation. N Engl J Med 2014;370:2478-86.
8. Turakhia MP, Shafrin J, Bognar K, et al. Estimated prevalence of undiagnosed atrial fibrillation in the United States. PLoS One 2018;13(4):e0195088.
9. Healey JS, Connolly SJ, Gold MR, et al. Subclinical atrial fibrillation and the risk of stroke. N Engl J Med 2012;366:120-9.
10. Belkin MN, Soria CE, Waldo AL, et al. Incidence and clinical significance of new-onset device-detected atrial tachyarrhythmia: a meta-analysis. Circ Arrhythm Electrophysiol 2018;11(3):e005393.
11. Steinhubl SR, Waalen J, Edwards AM, et al. Effect of a home-based wearable continuous ECG monitoring patch on detection of undiagnosed atrial fibrillation: the mSToPS randomized clinical trial. JAMA 2018;320:146-55.
12. Halcox JPJ, Wareham K. Response by Halcox and Wareham to letter regarding article, "Assessment of remote heart rhythm sampling using the AliveCor heart monitor to screen for atrial fibrillation: the REHEARSE-AF study." Circulation 2018;137:2193-4.
13. Turakhia MP, Desai M, Hedlin H, et al. Rationale and design of a large-scale, app-based study to identify cardiac arrhythmias using a smartwatch: the Apple Heart Study. Am Heart J 2019;207:66-75.