Journal of the American Medical Informatics Association, 27(9), 2020, 1359–1363
doi: 10.1093/jamia/ocaa137
Research and Applications

Research and Applications

Clinical evaluation and diagnostic yield following evaluation of abnormal pulse detected using Apple Watch

Kirk D. Wyatt,[1] Lisa R. Poole,[2] Aidan F. Mullan,[3] Stephen L. Kopecky,[4] and Heather A. Heaton[2]

[1]Division of Pediatric Hematology/Oncology, Department of Pediatric and Adolescent Medicine, Mayo Clinic, Rochester, Minnesota, USA, [2]Department of Emergency Medicine, Mayo Clinic, Rochester, Minnesota, USA, [3]Division of Biomedical Statistics and Informatics, Mayo Clinic, Rochester, Minnesota, USA, and [4]Division of Preventive Cardiology, Department of Cardiovascular Medicine, Mayo Clinic, Rochester, Minnesota, USA

Corresponding Author: Heather A. Heaton, MD, Mayo Clinic, 200 First Street Southwest, Rochester, MN 55905, USA; Heaton.Heather@mayo.edu

Received 25 January 2020; Revised 9 April 2020; Editorial Decision 7 June 2020; Accepted 25 June 2020

ABSTRACT

Objective: The study sought to characterize the evaluation of patients who present following detection of an abnormal pulse using Apple Watch.
Materials and Methods: We conducted a retrospective review of patients evaluated for abnormal pulse detected using Apple Watch over a 4-month period.
Results: Among 264 included patients, clinical documentation for 41 (15.5%) explicitly noted an abnormal pulse alert. Preexisting atrial fibrillation was noted in 58 (22.0%). Most commonly performed testing included 12-lead echocardiography (n = 158; 59.8%), Holter monitor (n = 77; 29.2%), and chest x-ray (n = 64; 24.2%). A clinically actionable cardiovascular diagnosis of interest was established in only 30 (11.4%) patients, including 6 of 41 (15%) patients who received an explicit alert.
Discussion: False positive screening results may lead to overutilization of healthcare resources.
Conclusions: The Food and Drug Administration and Apple should consider the unintended consequences of widespread screening for asymptomatic ("silent") atrial fibrillation and use of the Apple Watch abnormal pulse detection functionality by populations in whom the device has not been adequately studied.

Key words: wearable electronic devices, direct-to-consumer screening and testing, mass screening, fitness trackers, monitoring, ambulatory

INTRODUCTION

The market for wearable devices that monitor physiologic parameters is expanding.[1] Apple Watch (Apple Inc, Cupertino, CA) is the most popular smartwatch on the market,[2] with newer versions having the capability to alert users of an abnormal pulse[3] and capture a single-lead electrocardiogram (ECG).[3] The U.S. Food and Drug Administration (FDA) has granted clearance for the optical abnormal pulse detection[4] and ECG[5] features. In both clearances, false positive results leading to additional unnecessary medical procedures were identified as risks. The abnormal pulse detection feature was not intended for use by individuals with atrial fibrillation or those under 22 years of age.[4]

The Apple Watch abnormal pulse detection feature was designed to address asymptomatic (so-called silent) atrial fibrillation. Atrial fibrillation is the most common cardiac arrhythmia, affecting over

30 million people worldwide; however, as many as one-third of cases may be asymptomatic.[6,7] Atrial fibrillation is associated with an increased risk for stroke and death. Current strategies for management of atrial fibrillation attempt to mitigate these risks; however; the optimal management of silent atrial fibrillation, including cases detected using Apple Watch, remains unclear.[6,8]

The Apple Heart Study[9] aimed to prospectively assess the ability of Apple Watch to detect asymptomatic atrial fibrillation. Postenrollment study exclusion and dropout rates were high. Eighteen percent of patients who initiated a study visit because of an initial irregular pulse notification reported a preexisting diagnosis of atrial fibrillation or atrial flutter, leading to study exclusion. Abnormal pulse alerts were triggered for 2161 (0.52%) participants. Among patients with an initial alert who complied with study procedures, only 34% were diagnosed with atrial fibrillation. Among patients with a previous alert, positive predictive value of a subsequent alert for detection of atrial fibrillation by single-lead ECG was 0.84. A total of 57% of participants who were alerted of an abnormal pulse reported seeking care from a local medical provider. A total of 36% of these patients were recommended to undergo additional testing, but the specific testing performed was not reported.

False positive disease screening can lead to healthcare overutilization.[10] In this study, we aimed to (1) assess current practices for evaluating patients who present with an abnormal pulse detected using Apple Watch and (2) estimate the proportion of patients in a real-world setting who are diagnosed with a clinically actionable cardiovascular diagnosis following medical evaluation in response to an Apple Watch abnormal pulse alert.

MATERIALS AND METHODS

Apple Watch abnormal pulse detection

Series 1 and later versions of Apple Watch feature abnormal pulse detection.[3] Prior to enabling abnormal pulse detection, users must confirm that they are old enough to use the functionality and have not been diagnosed with atrial fibrillation. Once the feature is enabled, Apple Watch intermittently measures a user's pulse via a light-based sensor capable of detecting changes in blood flow with each heartbeat (ie, photoplethysmography).[9,11,12] Changes in blood flow are plotted over time to generate a tachogram. The tachogram is analyzed using a proprietary algorithm on the device to determine whether the pulse is irregular. If an irregular pulse is detected, Apple Watch continues to measure the user's pulse to confirm the abnormal reading. To increase alert specificity, the user is not alerted of an abnormal pulse until 4 additional measurements confirm an abnormal pulse.[3,9,11] Patients with a pulse suggestive of atrial fibrillation are presented with an alert on the watch that indicates "Your heart has shown signs of an irregular rhythm suggestive of atrial fibrillation. If you have not been diagnosed with AFib by a physician, you should talk to your doctor." Patients with high and low heart rates similarly receive a notification on the watch when their pulse is outside of a specified range.

Series 4 and later versions of Apple Watch may be used to capture single-lead ECG.[3] After the ECG is captured, it is automatically analyzed and interpreted as atrial fibrillation, sinus rhythm, or "inconclusive."[3] Users are able to view the ECG on a connected iPhone (Apple Inc, Cupertino, CA) and can export the ECG as a PDF file that can be electronically sent to a healthcare provider.[3]

Human subjects protection

This study was approved by the Mayo Clinic Institutional Review Board.

Participant selection

Clinical notes from all Mayo Clinic sites, including those in Minnesota, Arizona, Florida, Wisconsin, and Iowa, between December 6, 2018, and April 2, 2019, were queried using an internal cohort discovery tool ("Advanced Text Explorer") to identify notes that included the term "Apple Watch." The inclusion dates represented the date the abnormal pulse detection feature was released to the general public up to the date this retrospective study was submitted to the institutional review board. Patients who declined use of their medical records for research were excluded, yielding a set of records that were manually reviewed to identify which notes referenced abnormal pulse detection. In order to capture "real-world" use of the technology, we did not exclude patients based on age or preexisting diagnoses.

Data collection

Data collection was performed by 1 physician co-author (L.R.P.). Duplicate review with assessment of interrater reliability was considered but deemed unnecessary because the variables included in the data extraction form were not expected to be subject to significant differences in interrater interpretation.

The electronic health record was further reviewed to extract patient demographics (eg, age), clinical characteristics (eg, symptoms experienced), evaluations performed (eg, diagnostic testing), and diagnoses established following evaluation, with a cutoff date for data collection of April 2, 2019. Because clinical evaluation may occur over multiple visits, all evaluations performed and diagnoses established between the index visit and the cutoff date were included. REDCap (Research Electronic Data Capture) was used for data entry.[13,14] Variables included on the data extraction form were selected by a multidisciplinary group with expertise in cardiology, emergency medicine, and clinical informatics represented. Clinically actionable cardiovascular diagnoses of interest included the following: atrial fibrillation, atrial flutter, atrioventricular block, supraventricular tachycardia, ventricular fibrillation, and ventricular tachycardia.

The primary purpose of the study was to evaluate the Apple Watch automated pulse detection alert functionality; however, owing to the retrospective nature of the study, it was not always possible to determine whether patients presented in response to an automated alert or a pulse detected using other methods (eg, manually by launching the native Heart Rate application on the watch). To limit contamination of the sample with patients who presented with an abnormal pulse detected using methods other than automated alert, manual chart review was performed. Patient records were divided into those in which clinical documentation explicitly indicated the presence of an automated alert on the device ("alert") and those in which an abnormal pulse was detected using the Apple Watch but in which an alert was not explicitly mentioned in clinical documentation ("no alert").

Statistical analysis

Continuous patient characteristics are summarized as median and interquartile range, and categorical features are summarized as count and percentage. Differences in testing based on presenting department were assessed using chi-square tests of independence. Similarly, the association between presenting symptoms and diagnostic testing, as well as between abnormal pulse alerts and diagnostic testing, were evaluated using chi-square tests. All tests were 2-sided and P values $<.05$ were considered significant. The number needed to

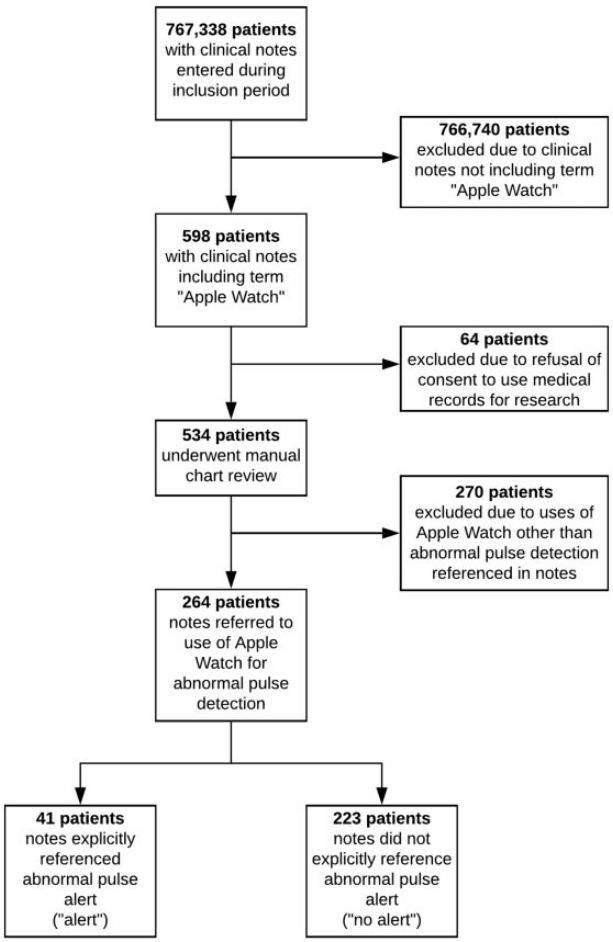

Figure 1. Patient inclusion.

diagnose was calculated as the inverse of the proportion of subjects with an abnormal pulse detected who were diagnosed with a cardiovascular condition of interest. Confidence intervals were calculated using Wilson's score interval with Yates' continuity correction. Statistical analyses were conducted using R version 3.6.1 (R Foundation for Statistical Computing, Vienna, Austria).

RESULTS

Patient characteristics

Records of 534 patients seen during the study period included the term "Apple Watch" (Figure 1). Abnormal pulse detection using Apple Watch was noted for 264 patients who were included in the analysis. The analysis included 41 (15.5%) patients whose records explicitly noted an abnormal pulse alert ("alert") and 223 (84.5%) patients whose records indicated abnormal pulse detection but did not explicitly mention of an alert ("no alert"). Median patient age was 55 (interquartile range, 37.75-69) years. Most patients were symptomatic at the time of abnormal pulse detection (26 [63.4%] alert patients and 151 [67.7%] no alert patients). Participant characteristics are shown in Table 1.

The most common department where patients initially presented was cardiology, in which the patient had a preexisting relationship (Table 1). Among patients who presented to the emergency department, 14 (33%) had received an abnormal pulse alert, only 1 of whom was asymptomatic. Approximately half of patients (n = 129, 48.9%) had a preexisting cardiovascular diagnosis, with the most common being atrial fibrillation (n = 58, 22%), followed by supraventricular tachycardia (n = 25, 9.5%) and sinus bradycardia (n = 17, 6.4%).

Table 1. Participant characteristics (N = 264)

Reference to alert	
Alert explicitly noted ("alert")	41 (15.5)
Alert not explicitly noted ("no alert")	223 (84.5)
Under 22 years of age	23 (8.7)
Sex	
Female	162 (61.4)
Male	102 (38.6)
Presence of symptoms	
Symptomatic	177 (67.0)
Asymptomatic	87 (33.0)
Specific symptoms	
Palpitations or fast heart rate	104 (39.3)
Lightheadedness, dizziness, or presyncope	51 (19.3)
Chest pain, pressure, or discomfort	32 (12)
Difficulty breathing or shortness of breath	24 (9.0)
Fatigue or weakness	18 (6.8)
Syncope	7 (2.7)
Presenting department	
Cardiology	101 (38.3)
Primary care	90 (34.1)
Emergency department	42 (15.9)

Values are n (%).

Diagnostic testing

The most commonly performed testing (Table 2) included 12-lead ECG (n = 158, 59.8%), Holter monitor (n = 77, 29.2%), and chest x-ray (n = 64, 24.2%). The proportion of patients presenting with specific symptoms who underwent each clinical evaluation is shown in Table 3. Patients seen in the emergency department were more likely to undergo 12-lead ECG, chest x-ray, or bloodwork compared with patients seen in primary care or by a cardiologist ($P < .001$ for all) (Table 2).

Only 2 (5%) patients with alert and 31 (13.9%) patients with no alert were noted to have self-recorded a single-lead ECG using Apple Watch over the 4 months following the release of this functionality. ECGs were interpreted by the Apple Watch as atrial fibrillation in 11 (33%), inconclusive in 9 (27%), and sinus rhythm in 3 (9%). For 10 (30%) patients, the interpretation was not specified in clinical notes. Among the 11 patients with a self-recorded ECG interpreted as atrial fibrillation, 9 (82%) subsequently underwent 12-lead ECG testing ordered by a physician, 7 (63%) had a preexisting diagnosis of atrial fibrillation, and 3 (27%) received a new diagnosis of atrial fibrillation.

Patients who experienced symptoms were more likely to undergo diagnostic testing than were patients who did not experience symptoms. When patients with "alert" and "no alert" are considered together, 177 (67.0%) patients experienced symptoms and 87 (33.0%) did not. A total of 139 (78.5%) patients who experienced symptoms underwent diagnostic testing, compared with 53 (60.9%) patients who did not experience symptoms ($P = .004$). There was no difference in diagnostic testing for "alert" patients compared with "no alert" patients. A total of 29 of 41 (70.7%) "alert" patients underwent diagnostic testing, compared with 163 of 223 (73.0%) "no alert" patients ($P = .9$).

Table 2. Comparison of clinical evaluations by presenting department

Evaluation	Full Cohort (n = 264)	Primary Care (n = 90)	Cardiology (n = 101)	Emergency Department (n = 42)	*P* Value[a]
12-lead ECG	158 (59.8)	51 (56.7)	56 (55.4)	38 (90.5)	<.001
Holter/event monitor	103 (39.0)	43 (47.8)	33 (32.7)	19 (45.2)	.085
Echocardiogram	60 (22.7)	19 (21.1)	25 (24.8)	12 (28.6)	.6
CV stress test	16 (6.1)	6 (6.7)	8 (7.9)	2 (4.8)	.8
Chest X-ray	64 (24.2)	16 (17.8)	18 (17.8)	26 (61.9)	<.001
Chest CT	19 (7.2)	3 (3.3)	8 (7.9)	6 (14.3)	.075
Bloodwork	61 (23.1)	12 (13.3)	15 (14.9)	30 (71.4)	<.001

CT: computed tomography; CV: cardiovascular; ECG: electrocardiography.

[a]Generated from chi-square tests comparing the usage of each assessment across the 3 departments.

Table 3. Comparison of clinical evaluations by presenting symptoms

Evaluation	Palpitations (n = 104)	Dizziness (n = 51)	Difficulty Breathing (n = 24)	Syncope (n = 7)	Fatigue (n = 18)	Chest Pain (n = 32)	None (n = 87)
12-lead ECG	77 (74)	33 (65)	16 (67)	4 (57)	14 (78)	22 (69)	39 (45)
Holter/event monitor	54 (52)	29 (57)	10 (42)	1 (14)	9 (50)	10 (31)	23 (26)
Echocardiogram	30 (29)	16 (31)	8 (33)	0 (0)	4 (22)	2 (6)	17 (20)
CV stress test	8 (8)	6 (12)	4 (17)	0 (0)	3 (17)	3 (9)	2 (2)
Chest x-ray	31 (30)	18 (35)	12 (50)	0 (0)	7 (39)	14 (44)	10 (11)
Chest CT	9 (9)	2 (4)	3 (12)	1 (14)	2 (11)	7 (22)	5 (6)
Bloodwork	35 (34)	16 (31)	9 (38)	2 (29)	9 (50)	14 (44)	6 (7)

Values are n (%).

CT: computed tomography; CV: cardiovascular; ECG: electrocardiography.

Diagnosis

A clinically actionable cardiovascular diagnosis of interest was established in 30 (11.4%) patients (atrial fibrillation in 13 [4.9%]), including 6 of 41 (15%) patients who received an abnormal pulse alert. Therefore, for patients who experienced an abnormal pulse alert and presented for medical evaluation, 7 (95% confidence interval, 3.5-14.5) patients needed to be evaluated to establish 1 diagnosis of clinically actionable cardiovascular disease. Among the 15 asymptomatic patients who presented following an abnormal pulse alert, only 1 was diagnosed with a clinically actionable cardiovascular diagnosis, yielding a number needed to diagnose of 15 (95% confidence interval, 2.9-286.5).

DISCUSSION

We describe 264 patients who sought medical attention for an abnormal pulse detected using Apple Watch. We included both patients in whom an automated alert was and was not explicitly documented in clinical notes. Explicit mention of an alert could be absent because of incomplete documentation or in cases in which patients identified an abnormal pulse in other ways, such as manually opening the native heart rate monitoring application. Most patients were symptomatic at the time an abnormal pulse was detected. Nearly half (48.9%) of patients had a preexisting cardiovascular diagnosis, and the most common department for initial evaluation was cardiology, in which the patient had a preexisting relationship. Diagnostic testing was associated with the presence of symptoms and the department to which a patient presented for initial evaluation.

In the FDA memorandum to Apple following review of the abnormal pulse detection feature, the FDA stated, "The feature has not been tested for and is not intended for use in people under 22 years of age. It is also not intended for use in individuals previously diagnosed with [atrial fibrillation]." Among patients included in this study, 8.7% were under 22 years of age, and 22% of patients had preexisting atrial fibrillation. These findings suggest that many users used the device in a manner inconsistent with FDA guidance. Stricter access controls are encouraged to ensure real-world use conforms with FDA guidance.

The observation that new clinically actionable cardiovascular diagnoses of interest were diagnosed in only 11.4% of patients following medical evaluation as directed by the treating provider suggests a high false positive rate as a screening tool for undiagnosed cardiovascular disease. False positive screening results have the potential to lead to excessive healthcare resource utilization and anxiety among the "worried well."

One study limitation is that the true rate of clinically actionable cardiovascular diseases of interest among the study population cannot be known, as atrial fibrillation can be paroxysmal and may not be detected on follow-up evaluation, especially with short-duration follow-up. Another limitation was our lack of access to claims data to assess the actual costs of evaluation. Furthermore, a standardized approach to medical evaluation was not taken. Therefore, some patients with clinically actionable cardiovascular disease may remain undiagnosed because of incomplete evaluation by the treating provider. Additionally, data were extracted by 1 reviewer. Finally, as a retrospective study, we identified included subjects based on presence of the text "Apple Watch" in clinical notes in order to assess patients using the same technology. However, this search strategy would have inadvertently excluded patients whose provider did not explicitly use the term "Apple Watch" in clinical documentation.

Very few patients in this study were noted to self-record an ECG. Although we were unable to identify which generation Apple Watch each patient used, we hypothesize that infrequent use of the ECG self-recording functionality within the study reflects a small

proportion of participants who owned a newer-generation Apple Watch that included this functionality.

Regulatory trends have shifted from reliance on large-scale prospective, randomized trials to the use of observational studies capturing "real-world evidence." The present study and the Apple Heart Study both have important limitations that constrain our ability to draw generalizable conclusions regarding the utility of the abnormal pulse detection feature of the Apple Watch. Despite these limitations, it is important to acknowledge that even false positive rates that are low in relative (ie, percentage of users) terms may translate to high absolute numbers (ie, total number of users) and yield significant healthcare overutilization.

As novel methods to screen for disease are developed, best practices for follow up and evaluation that maximize sensitivity and specificity while minimizing cost will need to be developed in tandem.[10] Since publication of the Apple Heart Study results, the authors have acknowledged the need to revise the medical community's conception of early disease detection in a direct-to-consumer era.[15] As the use of "real-world evidence" for medical device approvals becomes more mainstream, the limitations of nonrandomized study designs—including pragmatic trials—must be carefully considered and understood.[16] Only through the careful evaluation using rigorous, large-scale, prospective research studies can we comprehensively assess the risks and benefits of direct-to-consumer disease screening using novel technologies.

CONCLUSION

The FDA and Apple must carefully consider the unintended consequences of widespread direct-to-consumer screening for asymptomatic atrial fibrillation, including overutilization of healthcare resources owing to false positive screening results and use of screening tools by users in whom they have not been adequately studied.

FUNDING

This work was supported by United States Food and Drug Administration grant number U01FD005938 (Center of Excellence in Regulatory Science and Innovation grant to Yale University and Mayo Clinic). REDCap resources through the Mayo Clinic Center for Clinical and Translational Science were supported by National Institutes of Health grant number UL1TR002377. The contents of the article are solely the responsibility of the authors and do not necessarily represent the official views of the Department of Health and Human Services, Food and Drug Administration, or National Institutes of Health.

AUTHOR CONTRIBUTIONS

K.D.W. conceived of the study and designed the study with the assistance of S.L.K. and H.A.H. Data were extracted by L.R.P. and analyzed by A.F.M. The first draft of the manuscript was written by K.D.W. and revised critically for important intellectual content by all co-authors. All co-authors approve of the final version to be published and agree to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

CONFLICT OF INTEREST STATEMENT

The authors have no personal conflicts of interest to declare. All authors are employed by Mayo Clinic. Mayo Clinic has a partnership with AliveCor, which sells a direct-to-consumer electrocardiogram product. None of the authors have conducted work relating to the AliveCor product or receive funding related to the AliveCor product. Mayo Clinic and AliveCor had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

REFERENCES

1. Haghi M, Thurow K, Stoll R. Wearable devices in medical internet of things: scientific research and commercially available devices. *Healthc Inform Res* 2017; 23 (1): 4–15.
2. Reisinger D. Apple Watch series 4 outsold every other smart watch in 2018—in just one quarter. In. *Fortune* 2019. https://fortune.com/2019/06/25/apple-watch-smartwatch-sales-2018/ Accessed April 9, 2019.
3. Apple I. ECG app and irregular heart rhythm notification available today on Apple Watch. https://www.apple.com/newsroom/2018/12/ecg-app-and-irregular-heart-rhythm-notification-available-today-on-apple-watch. Published 2018. Accessed 9 April, 2019,
4. U.S. Food and Drug Administration. Letter: review of de novo request for classification of the irregular rhythm notification feature. https://www.accessdata.fda.gov/cdrh_docs/pdf18/DEN180042.pdf Accessed April 9, 2019.
5. U.S. Food and Drug Administration. Letter: review of de novo request for classification of the ECG app. https://www.accessdata.fda.gov/cdrh_docs/pdf18/den180044.pdf Accessed April 9, 2019.
6. Qiu J, Wang Y. A Smartwatch to identify atrial fibrillation. *N Engl J Med* 2020; 382 (10): 974–5.
7. Chugh SS, Havmoeller R, Narayanan K, *et al*. Worldwide epidemiology of atrial fibrillation: a Global Burden of Disease 2010 *Study. Circulation* 2014; 129 (8): 837–47.
8. Savelieva I, Camm AJ. Clinical relevance of silent atrial fibrillation: prevalence, prognosis, quality of life, and management. *J Interv Card Electrophysiol* 2000; 4 (2): 369–82.
9. Perez MV, Mahaffey KW, Hedlin H, *et al*. Large-Scale assessment of a Smartwatch to identify atrial fibrillation. *N Engl J Med* 2019; 381 (20): 1909–17.
10. Wyatt KD. A smartwatch to identify atrial fibrillation. *N Engl J Med* 2020; 382 (10): 975.
11. 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.
12. American College of Cardiology. Apple Heart Study identifies AFib in small group of Apple Watch wearers. https://www.acc.org/latest-in-cardiology/articles/2019/03/08/15/32/sat-9am-apple-heart-study-acc-2019. Accessed April 9, 2019.
13. 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; 42 (2): 377–81.
14. Harris PA, Taylor R, Minor BL, *et al*. The REDCap consortium: building an international community of software platform partners. *J Biomed Inform* 2019; 95: 103208.
15. Turakhia MP, Desai M, Perez MV. Apple Heart Study I. A smartwatch to identify atrial fibrillation. Reply. *N Engl J Med* 2020; 382 (10): 975–6.
16. Dhruva SS, Ross JS, Desai NR. Real-world evidence: promise and peril for medical product evaluation. *P T* 2018; 43 (8): 464–72.