Title

Physician responses to apple watch-detected irregular rhythm alerts.

Authors

Demkowicz, Patrick
Dhruva, Sanket
Spatz, Erica
et al.

Publication Date

2023-08-01

DOI

10.1016/j.ahj.2023.04.008

Peer reviewed

HHS Public Access

Author manuscript

Am Heart J. Author manuscript; available in PMC 2024 August 01.

Published in final edited form as:

Am Heart J. 2023 August ; 262: 29–37. doi:10.1016/j.ahj.2023.04.008.

Physician Responses to Apple Watch-Detected Irregular Rhythm Alerts

Patrick C. Demkowicz, BS[1,2,3], **Sanket S. Dhruva, MD, MHS**[4], **Erica S. Spatz, MD, MHS**[2,3], **Alexis L. Beatty, MD, MAS**[4,5], **Joseph S. Ross, MD, MHS**[1,3,6], **Rohan Khera, MD, MS**[2,3,7]

[1.]Section of General Medicine, Department of Internal Medicine, Yale School of Medicine, New Haven, CT

[2.]Section of Cardiology, Department of Internal Medicine, Yale School of Medicine, New Haven, CT

[3.]Center for Outcomes Research and Evaluation, Yale-New Haven Health System, New Haven, CT

[4.]Division of Cardiology, Department of Medicine, University of California San Francisco (UCSF), San Francisco, CA

[5.]Department of Epidemiology and Biostatistics, UCSF, San Francisco, CA

[6.]Department of Health Policy and Management, Yale School of Public Health

[7.]Section of Health Informatics, Department of Biostatistics, Yale School of Public Health, New Haven, CT

Abstract

Correspondence To: Rohan Khera, MD, MS, 195 Church St, 5th Floor, New Haven, CT 06510, rohan.khera@yale.edu; @rohan_khera.

Author Contributions

All authors were responsible for study concept and design; PD, RK, and SD for acquisition of subjects and data; PD, RK, and SD, and JR for analysis and interpretation of data; and all authors for preparation of manuscript. PD had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis

COMPETING INTERESTS

The Authors declare no Competing Non-Financial Interests but the following Competing Financial Interests: In addition to funding listed above, PD reports Roivant Sciences stock options; SD reports research funding from the Department of Veterans Affairs, from the Medical Device Innovation Consortium (MDIC) as part of the National Evaluation System for health Technology Coordinating Center (NESTcc), Greenwall Foundation, Arnold Ventures, and National Institute for Health Care Management; ES receives grant funding from the Centers for Disease Control and Prevention (20042801-Sub01), National Institute on Minority Health and Health Disparities (U54MD010711-01), the U.S. Food and Drug Administration to support projects within the Yale-Mayo Clinic Center of Excellence in Regulatory Science and Innovation (CERSI, U01FD005938), the National Institute of Biomedical Imaging and Bioengineering (R01 EB028106-01), and from the National Heart, Lung, and Blood Institute (R01HL151240); AB was employed by Apple Inc. 2018–2019 and held stock in Apple Inc. 2019–2021; JR currently receives research support through Yale University from Johnson and Johnson to develop methods of clinical trial data sharing, from the Medical Device Innovation Consortium as part of the National Evaluation System for Health Technology (NEST), from the Food and Drug Administration for the Yale-Mayo Clinic Center for Excellence in Regulatory Science and Innovation (CERSI) program (U01FD005938), from the Agency for Healthcare Research and Quality (R01HS022882), from the National Heart, Lung and Blood Institute of the National Institutes of Health (NIH) (R01HS025164, R01HL144644), and from Arnold Ventures; in addition, JR is an expert witness at the request of Relator's attorneys, the Greene Law Firm, in a qui tam suit alleging violations of the False Claims Act and Anti-Kickback Statute against Biogen Inc; in addition to funding listed above, RK also receives research support, through Yale, from Bristol-Myers Squibb and Novo Nordisk. He is a coinventor of U.S. Provisional Patent Applications 63/177,117, 63/428,569, and 63/346,610, unrelated to current work. He is also a founder of Evidence2Health, a precision health platform to improve evidence-based cardiovascular care.

Background: While the US Food and Drug Administration (FDA) has cleared smartwatch software for detecting atrial fibrillation (AF), there is lack of guidance on management by physicians. We sought to evaluate the approach to management of Apple Watch alerts for AF by physicians and assess whether respondent and case characteristics were associated with their approach.

Methods: We conducted a case-based survey of physicians practicing primary care, emergency medicine, and cardiology at 2 large academic centers (Yale and University of California San Francisco) between September to December 2021. Cases described asymptomatic patients receiving Apple Watch AF alerts; cases varied in sex, race, medical history, and notification frequency. We evaluated physician responses among prespecified diagnostic testing, referral, and treatment options.

Results: We emailed 636 physicians, of whom 95 (14.9%) completed the survey, including 39 primary care, 25 emergency medicine, and 31 cardiology physicians. Among a total of 192 cases (16 unique scenarios), physicians selected at least one diagnostic test in 191 (99.5%) cases and medications in 48 (25.0%). Physicians in primary care, emergency medicine, and cardiology reported varying preference for patient referral (14%, 30%, and 16%, respectively; P=.048), rhythm monitoring (84%, 46%, and 94%, respectively; P<.001), measurement of BNP (8%, 20%, and 2%; P=.003), and use of antiarrhythmics (16%, 4%, and 23%; P=.023). There were few physician differences in reported practices across patient demographics (sex and race), clinical complexity, and alert frequency of the clinical case.

Conclusions: In hypothetical cases of patients presenting without clinical symptoms, physicians opted for further diagnostic testing and often to medical intervention based on Apple Watch irregular rhythm notifications. There was also considerable variation across physician specialties, suggesting a need for uniform clinical practice guidelines. Additional study is required before irregular rhythm notifications should be used in clinical settings.

BACKGROUND

Smartwatches such as the Apple Watch and Fitbit have grown in popularity, with 1% of primary care patients having documentation of a device in a recent study of an academic health care system.[1] Many of these devices contain FDA-cleared software that can identify irregular rhythms, such as AF.[2,3] Despite the widespread adoption of smartwatches, there is no evidence that medical evaluation for irregular rhythms detected by smartwatches improves patient morbidity or mortality.[4–7] Conversely, healthy patients who falsely screen positive may receive unnecessary and potentially harmful testing and treatment.[8,9]

Given a lack of clinical evidence to inform decision-making, encounters where patients discuss information generated from smartwatches are challenging for physicians. Furthermore, different device companies' algorithms vary in their diagnostic yield.[10] Recent surveys have found that cardiac electrophysiologists would consider pursuing further electrophysiological evaluation as well as initiate anticoagulation in asymptomatic patients on the basis of smartwatch irregular rhythm and single-lead ECG findings despite lack of any evidence to suggest benefit of such an approach.[11,12] These findings suggest that there is substantial variation in the strategies adopted by different clinicians.

On the other hand, widespread access to affordable smartwatches may instead increase rates of AF diagnosis among marginalized patients, mitigating current disparities in care. For example, Black patients with AF are less likely than White patients to receive a formal diagnosis and receive anticoagulation.[13] Female patients with AF are less often prescribed anticoagulation therapy despite having higher stroke risk.[14] Furthermore, a study using hypothetical case histories found gender bias in attitude toward and secondary prevention of patients with coronary artery disease.[15] A study using actors portraying patients with chest pain found that women and Black adults were less likely to be referred for cardiac catheterization than men and White adults.[16]

In this study, we used hypothetical patient cases to conduct a multicenter evaluation of physicians' responses to asymptomatic patients reporting Apple Watch irregular rhythm notifications suggestive of AF. We sought to evaluate variation across physicians and assess whether specialty and personal experience with smartwatches influenced their approach to management. We also assessed whether patient characteristics influenced the approach of physicians to patient-reported episodes of Apple Watch-detected, asymptomatic AF.

METHODS

Study Population

We conducted a cross-sectional, case-based survey of attending and resident physicians practicing in primary care, emergency medicine, and cardiology departments at two geographically separated health systems, (1) Yale New Haven Hospital in New Haven, Connecticut and, (2) University of California San Francisco between September and December 2021. A convenience sample of physicians and their email addresses was gathered from institutional directory profiles, if available. The methods were performed in accordance with relevant guidelines and regulations and approved (exempted from Institutional Review Board review) by the Yale and UCSF Human Research Protection Programs (HRPPs) and San Francisco VA Health Care System Research and Development Committee. The study was funded by the National Heart, Lung and Blood Institute (under award K23HL153775 to RK and T35HL007649 to PD), and the Doris Duke Charitable Foundation (under award, 2022060 to RK). The authors are solely responsible for the design and conduct of this study, all study analyses, the drafting and editing of the paper and its final contents.

Survey Design

The survey was developed by 2 cardiologists, an internist, and a medical student, and reviewed by 2 other cardiologists. The survey is included in Supplement 1 and the Consensus-Based Checklist for Reporting of Survey Studies (CROSS) is included in Supplement 2.[17] The survey consisted of two clinical vignettes. Each vignette involved a 60-year-old person who reported recently receiving one or more irregular rhythm notifications on their Apple Watch in the absence of associated symptoms (i.e., no fatigue or racing sensation in his chest). Four variables varied across the 16 scenarios: stroke risk-factors (none or history of diabetes and hypertension), reported frequency of recent alerts (1 or >1),

sex (male or female) and race (Black or White). The atrial fibrillation stroke risk factors represent items on the $CHAD_2DS_2$-VASc score.[18]

Respondents were randomized by the survey platform to see one low stroke-risk case and one moderate-high stroke-risk case of two patients with the same race and sex. Due to an error in the survey platform settings, 5 respondents completed a third case that varied from their second case with respect to alert frequency. A respondent notified the team of the error and it was promptly corrected. Respondents were then asked to rate on a five-point Likert scale the likelihood that they would consider specific diagnostic evaluations and therapeutic interventions. Next, respondents were asked on a five-point Likert scale how important (extremely, very moderately, slightly, not at all) various factors were in determining their answer choices in the preceding cases: likelihood of AF in patient's group, risk of stroke in patient's group, strength of evidence, concern about missing a diagnosis, and concern about unnecessary testing. Finally, respondents were asked about their own backgrounds and practices. We asked respondents about their gender, race, specialty, training history, personal use of smartwatches capable of rhythm detection, and experience recommending smartwatches to their patients.

Survey Delivery

We used Qualtrics to host our survey and send email invitations. Qualtrics generated a unique invitation link for each respondent. This allowed us to track survey completion and prevent multiple participation. Completed responses were deidentified by Qualtrics. Respondents at Yale were emailed three weekly reminders and respondents at UCSF 2 reminders (1 and 3 weeks) if they did not respond. Respondents who clicked the link to the survey were then shown a study information statement and asked whether they agreed to participate (Yes/No). Respondents who clicked "No" were automatically exited from the survey. No written or verbal consent was obtained, as participation in this optional, anonymous survey posed minimal risk to the subjects. Respondents were required to select their specialty before they could advance to the case scenarios. Respondents who did not complete this step or who left all case questions unanswered were excluded. Respondents who reached the end of the survey were offered the opportunity to enter a drawing for a $100 gift card.

Outcomes

The primary outcome was the proportion of responses indicating the choice of specific interventions. These included: referring to primary care, cardiology, or electrophysiology; ordering a cardiovascular stress test, 12-lead ECG, serum brain natriuretic peptide (BNP), transthoracic echocardiogram, ambulatory rhythm monitoring (specific options included event monitor, implantable loop recorder, patch monitor, or a commercially available heart rhythm monitor such as AliveCor, which were combined into one category); and treating with aspirin, anticoagulation, or antiarrhythmics (beta blocker, calcium channel blocker, or a class IC or III antiarrhythmic). We categorized the responses as "yes" and "no", combining responses of "Extremely Likely" or "Somewhat Likely" into a yes category and responses of "Neither Likely nor Unlikely," "Somewhat Unlikely," and "Extremely Unlikely" into a no category.

Statistical Analyses

We generated descriptive statistics (specialty, sex, age, and years in practice) about our study population using percentages to report categorical variables and averages (standard deviation) to report continuous variables, stratified by respondent specialty. We also reported proportions of respondents reporting specific experiences with smartwatches (personal use, having previously recommended patients use wearable devices, and having experienced a patient reporting results from a smartwatch) stratified by respondent specialty. We used the chi-squared test to compare proportions across specialties.

We reported the proportions of cases in which respondents selected specific tests and interventions. Because respondents evaluated multiple cases, we calculated mean proportions out of all, non-unique cases. We also used a histogram distribution to report the number of diagnostic tests (stress test, 12-lead ECG, serum BNP, TTE, ambulatory rhythm monitoring), or medical interventions (aspirin, anticoagulation, antiarrhythmic) that were selected across all, non-unique cases.

We analyzed the relationship between respondent characteristics and the selection of specific tests or interventions. These included their (1) clinical background, specifically their specialty and years of practice, (2) experiences with smartwatches, including wearing a smartwatch, having had a patient report wearable findings before, and having previously recommended wearables to their patients; and (3) approach towards risk assessment and management – evaluated by questions whether the likelihood of AFib in patient's group, risk of stroke in patient's group, the strength of evidence, concern about missing a diagnosis, and concern about unnecessary testing were "very" or "extremely" important in influencing their management decisions compared with respondents who indicated these considerations were "not at all," "slightly," or "moderately" important. We used the chi-squared test to compare proportions of choosing specific tests and interventions across these characteristics.

We repeated the above analysis to evaluate the relationship between patient characteristics and respondents' likelihood of selecting each test or intervention. We compared use of each intervention by the race (Black vs. white), sex (male vs. female), and history of stroke risk factors (none vs. hypertension and diabetes) of the patient described, as well as by the number of notifications reported by the patient in the scenario (one vs. many). Analyses were conducted using R (version 4) and figures were generated using Prism (version 9). A 2-sided p value <0.05 was considered statistically significant.

Lastly, we performed sequential logistic regression to examine further how patient and clinician factors influenced the selection of interventions used at significantly different proportions across clinical specialties. For each outcome, we sequentially generated five models, each adding the following groups of independent variables to the prior model: specialty (emergency medicine, cardiology); clinician factors, including gender (male and female), years in practice (0–9, 10–19, and ≥ 20), and institution (Institution A and Institution B); prior experiences with smartwatches, including wearing a smartwatch, ever having recommended smartwatches to patients, and ever having treated a patient who reported a smartwatch alert; simulated patient factors, including race (white and Black), sex (male and female), presence of stroke risk factors (none, and both hypertension

and diabetes), and number of notifications (one and many); and the considerations that respondents scored as “very/extremely” important in influencing their responses to the cases, including the likelihood of AFib in patient’s group, the risk of stroke in patient’s group, the strength of the evidence, concern about missing a diagnosis, and concern about unnecessary testing. We calculated odds ratios with 95% confidence intervals and p-values for each variable in comparison to its reference level. We generated forest plots using these odds ratios and 95% confidence intervals.

RESULTS

Study Population

We emailed 636 physicians (excluding an additional 14 emails that bounced-back or failed), of whom 95 (14.9%) completed the survey, including 39 primary care, 25 emergency medicine, and 31 cardiology physicians (Table 1); 75 were based at Yale and 20 at UCSF. Overall, 44.2% of respondents identified as female, 55.8% were aged 44 years or younger, and 40% of respondents report being practice less than a decade, with 22.1% practicing 10–19 years, and 31.6% practicing more than 20 years (6.3% missing).

Overall, 27 (28.4%) respondents reported personally using a smartwatch, which was consistent across specialties (P=.68) (Table 2). However, respondents of different specialties reported considerably different experiences with smartwatches in their clinical practices, including the proportion who had recommended smartwatches to their patients (12.8% primary care, 12.0% emergency, and 41.9% cardiology; P=.003) and have had a patient report a smartwatch alert (30.8%, 80.0%, and 80.6%, respectively; P<.001).

Overall Approach to Case Scenarios

Respondents completed a total of 192 cases drawn from 16 unique scenarios; 3 respondents completed 1 case, 87 respondents completed 2 cases, and 5 respondents completed 3 cases. In 191 (99.5%) cases physicians selected at least one diagnostic test to work-up the asymptomatic Apple Watch irregular rhythm notification: electrocardiography (185, 96.4%), ambulatory rhythm monitoring (148, 77.1%), transthoracic echocardiography (TTE) (63, 32.8%), stress testing (19, 9.9%), and brain natriuretic peptide (BNP) evaluation (17, 8.9%) (Figure 1). Respondents selected an average of 2.3 diagnostic tests (standard deviation 0.97) (Figure 2A). In addition, referral and treatment options were commonly selected in response to the notification: in 36 (18.8%) cases physicians selected referral to a different specialty and in 48 (25.0%) new medication treatment, such as aspirin (35, 18.2%), antiarrhythmics (29, 15.1%), and/or anticoagulation (18, 9.4%). Respondents selected an average of 0.4 types of medication (standard deviation 0.8) (Figure 2B).

When asked about the factors that influenced their answers to the cases, 50.5%, 42.1%, 56.8%, 43.2%, and 48.4% of respondents rated as very important or extremely important the likelihood of AF, the risk of stroke, the strength of evidence, concern about missing a diagnosis, and concern about unnecessary testing, respectively (Supplementary Table 1).

Of 95 respondents, 41 wrote comments with additional information about their management choices (Supplementary Table 2). Ten respondents provided algorithmic descriptions of

the sequence or logic of their management approaches, and 9 indicated they wanted to review the watch tracing if available. Additional diagnostic studies were mentioned, including thyroid function tests (N=6), electrolytes (N=3), basic metabolic panel, toxicology screening, lipid profile, and sleep study. A selected intervention, such as the brand or kind of ambulatory rhythm monitor (N=6), was further specified by 8 patients. “Watchful waiting,” meaning that a patient should continue to monitor his or her watch-recorded rhythm and return following another episode, was recommended by 3 respondents. One respondent was concerned that the vignette specified the patient’s race. One respondent recommended prescribing an angiotensin-converting enzyme (ACE) inhibitor. Lastly, one respondent wished to review the evidence for Apple watch-detected atrial fibrillation.

Clinician Factors and Approach to Case Scenarios

For diagnostic testing, there were no statistically significant differences across specialties in the selection of ECG, TTE, or stress testing (Supplementary Figure 1). With respect to interventions, there were no statistically significant differences across specialties in the selection of aspirin or anticoagulation (Supplementary Figure 1). Among physicians practicing emergency medicine, primary care, and cardiology, use of referral differed by as much as 16% across specialties, 30% in emergency medicine, 14% in primary care, and 16% in cardiology (P=.048). Similarly, BNP measurement by up to 18% (20% vs. 8% vs. 2%; P=.003), ambulatory rhythm monitoring by up to 48% (46% vs. 84% vs. 94%; P<.001) and prescription of antiarrhythmic therapy by up to 19% (4% vs. 16% vs. 23%; P=.023). Years spent in practice was associated with the use of transthoracic echocardiogram (24% 0–9 years, 28% 10–19 years, 45% > 20 years; P=.023) and ambulatory rhythm monitoring (61% vs. 81% vs. 92%; P<.001) (Supplementary Figure 2).

Respondents who have previously recommended their patients to use smartwatches were more likely to order TTEs (54.8% vs. 24.8%, P<.001) and ambulatory rhythm monitoring (95.2 vs. 70.2%, P=.002) (Supplementary Table 3). There was no relationship between other respondent factors, including wearing a smartwatch, having treated patients who reported smartwatch findings, and having recommended smartwatches to patients, and selection of specific interventions.

Respondents who indicated that concern about the likelihood of AFib was “very important” or “extremely important” were more likely to order ambulatory rhythm monitoring than those who indicated it was “not at all,” “slightly,” or “moderately important” (84% vs. 67%; P=.029 (Supplementary Table 4). Those who indicated that concern about stroke risk was very/extremely important were more likely to order ambulatory rhythm monitoring (89% vs. 65%; P<.001), prescribe anticoagulation (15% vs. 4%; P=.023), and prescribe antiarrhythmics (23% vs. 6%; P=.002). Respondents who indicated that strength of evidence was very or extremely important were more likely to order TTEs (39% vs. 20%; P=.010). Respondents who ranked concern about missing a diagnosis as very/extremely important were more likely to order TTEs (41% vs. 24%; P=.022). Respondents who ranked concern about unnecessary testing as very/extremely important were less likely to make referrals (6% vs. 32%; P<.001), prescribe anticoagulation (4% vs. 14%; P=.041), and antiarrhythmics (7% vs. 21%; P=.016).

Simulated Patient Factors and Overall Approach to Case Scenarios

The distribution of case characteristics is summarized in Supplementary Table 5. Responses are reported by individual cases in Supplementary Table 6. Black race was associated with a higher receipt of antiarrhythmics (21.3% vs. 9.2%; P=.033) (Supplementary Figure 3). Selection of other diagnostics and interventions were otherwise similar between cases describing Black and White patients. Female sex was associated with lower use of serum BNP (4.0% vs. 14.1%; P=.027). Selection of other diagnostics and interventions were otherwise similar between cases describing female and male patients. Patient stroke risk (presence vs. absence of diabetes and hypertension) and reported frequency of alerts (single vs. repeated) were not associated with differences in intervention (Supplementary Table 7).

Sequential logistic regression

In sequential logistic regression we further examined the relationship between specialty and reported practices, accounting for different clinician and case characteristics. For each outcome, we generated a sequence of models adjusting for additional sets of variables: clinician factors, smartwatch experience, patient factors, and the factors that respondents reported were influential in their decision-making.

In unadjusted assessments, emergency medicine was significantly associated with higher use of referral compared with primary care (OR 2.8, 95% CI 1.2–6.8) and remained significantly associated with referral after adjustment for clinician factors, smartwatch experiences, patient factors, and respondent rationale (OR 6.0, 95% CI 1.2–34.1). Cardiology was not significantly associated with referral in unadjusted analyses compared with primary care (OR 1.2, 95% CI 0.5–3.1). After adjusting for clinician factors, smartwatch experiences, and patient factors, cardiology was associated with a significantly greater use of referral compared with primary care (OR 5.6, 95% CI 1.3–28.7) (Supplementary Figure 4A).

In unadjusted assessments, emergency medicine was significantly associated with higher use of BNP than primary care (OR 3.1, 95% CI 1.1–9.7). This relationship was no longer significant after adjustment for other factors. The large confidence interval estimates for the OR of use of BNP among cardiologists compared with primary care physicians precluded assessment. (Supplementary Figure 4B).

In unadjusted assessments, emergency medicine was significantly associated with lower use of ambulatory monitoring compared with primary care (OR 0.2, 95% CI 0.1–0.4). This relationship remained significant after adjustment for clinician factors, smartwatch experiences, and patient factors (OR 0.3, 95% CI 0.1–0.7). Cardiology was not significantly associated with use of ambulatory monitoring in unadjusted analyses (OR 2.8, 95% CI 0.9–10.4). After adjusting for clinician factors, cardiology was associated with greater use of ambulatory monitoring (OR 3.9, 95% CI 1.2–15.9) (Supplementary Figure 4C).

In unadjusted assessments, emergency medicine was significantly associated with lower use of antiarrhythmic therapy compared with primary care (OR 0.2, 95% CI 0.0–0.8). This relationship did not remain significant after adjustment. Cardiology was not significantly associated with use of antiarrhythmics in any model (Supplementary Figure 4D).

DISCUSSION

In this survey of physicians at two institutions, we used clinical vignettes to examine decision-making patterns in response to irregular rhythm notifications on an Apple Watch in patients without symptoms of AF. Notifications nearly always triggered further diagnostic workup, varying from low-cost and low-intensity tests such as ECG to higher-cost and higher-intensity procedures such as stress testing. Furthermore, in 1 out of 4 simulated encounters, respondents considered initiating medications, including antiplatelet, anticoagulant, and antiarrhythmic therapies on initial presentation. Management varied considerably by specialty, across years spent in practice, and with regard to respondent-reported experiences with wearables in their own practices. In general, differences in management among specialties persisted after multivariable adjustment for other clinician and simulated patient factors. Respondent-identified concerns, in particular concern about stroke risk and also concern about unnecessary testing, were associated with diagnostic and treatment approach. With the exception of patient sex, we found that management did not vary substantially across case characteristics such as patient race, stroke risk, and alert frequency.

Our findings are consistent with prior studies that have examined clinical decision-making based on rhythm data from smartwatches. A 4-month single-center retrospective review identified 264 patients evaluated for abnormal pulse detected using Apple Watch, 33% of whom were asymptomatic.[19] The study found that 61% of asymptomatic patients underwent diagnostic testing. The study also found variation in diagnostic testing across clinical departments, with patients seen in the emergency department more likely to undergo 12-lead ECG or bloodwork compared with patients seen in primary care or by a cardiologist. One possible explanation for a higher use of diagnostic testing in our survey is that respondents were shown only brief vignettes and not afforded the opportunity to perform a physical examination. In the retrospective study, clinicians had preexisting relationships with many patients: nearly half of all patients had a preexisting cardiovascular diagnosis and the most common department of presentation was cardiology. Clinicians may therefore be able to work up smartwatch irregular rhythm notifications more conservatively in the real world than in our survey because they have more data to inform decision-making.

Prior studies examining prescribing patterns in response to smartwatch notifications have focused on interpreting single-lead ECG tracings available on models such as the Apple Watch 4 and later. A survey of 1601 clinicians, including advanced practice providers, found that results from a 30-second single-lead ECG were sufficient for 42.7% of clinicians to recommend oral anticoagulation for patients at high risk for stroke.[11] A survey of 417 electrophysiologists worldwide found that, when presented with a single-lead ECG tracing suggesting AF, 21% would consider initiating anticoagulation in an asymptomatic patient.[12] In contrast, we found that respondents would consider anticoagulation approximately 9% of the time in response to irregular rhythm notifications. It is expected that clinicians would be less likely to consider anticoagulation based on an irregular rhythm notification derived from photoplethysmography. Prior studies of Apple Heart Study participants found that follow-up ambulatory ECG confirmed AF in only 34% of cases of irregular rhythm notification, with non-AF irregular rhythms detected in 40% of cases without AF.[20,21]

Therefore, anticoagulation may not be merited without explicit confirmation of the rhythm on an electrocardiographic modality. Manual clinician review of single-lead ECG tracings can improve the diagnostic utility of the Apple Watch automated AF-detection algorithm.[10]

It is notable that stroke risk in our study was not associated with respondent practice patterns. Some researchers have advocated for screening for AF in the high-stroke risk population with the assumption that such patients would have a greater benefit.[22] Our findings show that physicians approached patients with the least potential benefit from AF screening and treatment similarly to how they approached patients who would have the greatest potential benefit. This suggests that in actual practice, physicians might apply such hypothetical screening guidelines with a much larger scope that recommended. This finding supports the concern that smartwatch utilization may lead to greater health care expenditure on potentially unnecessary testing and treatment.[23] In patients who have true paroxysmal AF, this may also lead to disparities in access to cardiovascular care between patients who own smartwatches and patients who do not.

It is also notable that, in contrast with prior studies using simulated patients, there was little variation in management with respect to the patient's race or sex.[15,16] This study had fewer respondents and therefore may not have been sufficiently powered. Furthermore, our respondent population skewed toward younger clinicians, who may have had greater exposure to anti-bias curricula in their medical training than clinicians later than their careers.[24] This finding also only looks at one mechanism – clinician bias – of systemic racism and sexism, which may affect the outcomes of real-world patients in a myriad of ways.[25] Further research is needed to understand disparities in the treatment of smartwatch-detected irregular rhythms in real world clinical practice.

There are several limitations to our study. First, 14.9% of physicians contacted ultimately responded to our survey; respondents may have stronger attitudes toward smartwatches than non-respondents. We found in our survey that respondents with positive attitudes toward smartwatches were more likely to select certain interventions. Second, we assessed practice patterns using written clinical case vignettes. It is assumed that respondents would act similarly when managing actual patients.[16] Furthermore, assessments made on the bases of written case vignettes have been shown to correlate with those made based on in-person examinations.[26] Third, our case scenarios were limited in detail, and other potentially informative clinical cues available in the real-world were not provided. Fourth, we asked respondents to share their approach to management based solely on initial presentation, which may not reflect the sequential approach to diagnosis and treatment choices in the real world. Lastly, we surveyed physicians at two large academic centers, which may limit the generalizability of our findings.

CONCLUSIONS

Our survey demonstrates that many physicians likely have a high degree of confidence in smartwatch irregular rhythm notifications, as demonstrated by respondents' likelihood of pursuing additional diagnostic testing and interventions. This finding raises several concerns: smartwatch utilization may lead to greater healthcare expenditure on potentially

unnecessary testing and treatment; it may also lead to disparities in access to cardiovascular care between patients who own smartwatches and patients who do not. Therefore, despite the lack of recommendations by public health and professional organizations for AF screening in asymptomatic patients, our study highlights the need for further evidence to inform the development of standardized guidelines.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

DATA AVAILABILITY

Deidentified survey response data are available upon reasonable request to qualified researchers from corresponding author, RK.

REFERENCES

1. Al-Alusi MA et al. Trends in Consumer Wearable Devices With Cardiac Sensors in a Primary Care Cohort. Circ Cardiovasc Qual Outcomes 0, 10.1161/CIRCOUTCOMES.1121.008833, doi:10.1161/CIRCOUTCOMES.1121.008833.
2. Krueger AC (ed Center for Devices and Radiological Health) (2018).
3. Kozen JS (ed Center for Devices and Radiological Health) (2022).
4. Force, U. S. P. S. T. Screening for Atrial Fibrillation: US Preventive Services Task Force Recommendation Statement. JAMA 327, 360–367, doi:10.1001/jama.2021.23732 (2022). [PubMed: 35076659]
5. Svennberg E, Engdahl J & Rosenqvist M Screening for Atrial Fibrillation to Initiate Stroke-Protective Therapy. Am Fam Physician 101, 6–7 (2020). [PubMed: 31894945]
6. Noseworthy PA et al. Subclinical and Device-Detected Atrial Fibrillation: Pondering the Knowledge Gap: A Scientific Statement From the American Heart Association. Circ 140, e944–e963, doi:doi:10.1161/CIR.0000000000000740 (2019).
7. Svendsen JH et al. Implantable loop recorder detection of atrial fibrillation to prevent stroke (The LOOP Study): a randomised controlled trial. Lancet 398, 1507–1516, doi:10.1016/S0140-6736(21)01698-6 (2021). [PubMed: 34469766]
8. Varshney AS, Madias C, Kakkar R & Martin DT Watching for Disease: the Changing Paradigm of Disease Screening in the Age of Consumer Health Devices. J Gen Intern Med 35, 2173–2175, doi:10.1007/s11606-019-05626-y (2020). [PubMed: 31942673]
9. Mandrola J & Foy A Screening for Atrial Fibrillation—New Devices, Same Challenges. JAMA Intern Med 182, 251–253, doi:10.1001/jamainternmed.2021.7283 (2022). [PubMed: 35076689]
10. Ford C et al. Comparison of 2 Smartwatch Algorithms for Detection of Atrial Fibrillation and the Benefit of Clinician Interpretation. JACC Clin Electrophysiol 8, 782–791, doi:10.1016/j.jacep.2022.02.013 (2022). [PubMed: 35738855]
11. Ding EY et al. Survey of current perspectives on consumer-available digital health devices for detecting atrial fibrillation. Cardiovasc Digit Health J 1, 21–29, doi:10.1016/j.cvdhj.2020.06.002 (2020). [PubMed: 32924024]
12. Manninger M et al. Role of wearable rhythm recordings in clinical decision making—The wEHRAbles project. Clin Cardiol 43, 1032–1039, doi:10.1002/clc.23404 (2020). [PubMed: 32700414]
13. Meschia JF et al. Racial Disparities in Awareness and Treatment of Atrial Fibrillation. Stroke 41, 581–587, doi:10.1161/STROKEAHA.109.573907 (2010). [PubMed: 20190000]
14. Kassim NA, Althouse AD, Qin D, Leef G & Saba S Gender differences in management and clinical outcomes of atrial fibrillation patients. J Cardiol 69, 195–200, doi:10.1016/j.jjcc.2016.02.022 (2017). [PubMed: 27262176]

15. Abuful A, Gidron Y & Henkin Y Physicians' attitudes toward preventive therapy for coronary artery disease: Is there a gender bias? Clin Cardiol 28, 389–393, doi:10.1002/clc.4960280809 (2005). [PubMed: 16144216]
16. Schulman KA et al. The Effect of Race and Sex on Physicians' Recommendations for Cardiac Catheterization. N Engl J Med 340, 618–626, doi:10.1056/nejm199902253400806 (1999). [PubMed: 10029647]
17. Sharma A et al. A Consensus-Based Checklist for Reporting of Survey Studies (CROSS). J Gen Intern Med 36, 3179–3187, doi:10.1007/s11606-021-06737-1 (2021). [PubMed: 33886027]
18. Lip GY, Nieuwlaat R, Pisters R, Lane DA & Crijns HJ Refining clinical risk stratification for predicting stroke and thromboembolism in atrial fibrillation using a novel risk factor-based approach: the euro heart survey on atrial fibrillation. Chest 137, 263–272, doi:10.1378/chest.09-1584 (2010). [PubMed: 19762550]
19. Wyatt KD et al. Clinical evaluation and diagnostic yield following evaluation of abnormal pulse detected using Apple Watch. J Am Med Inform Assoc 27, 1359–1363, doi:10.1093/jamia/ocaa137 (2020). [PubMed: 32979046]
20. Perez MV et al. Large-Scale Assessment of a Smartwatch to Identify Atrial Fibrillation. N Engl J Med 381, 1909–1917, doi:10.1056/nejmoa1901183 (2019). [PubMed: 31722151]
21. Perino AC et al. Arrhythmias Other Than Atrial Fibrillation in Those With an Irregular Pulse Detected With a Smartwatch: Findings From the Apple Heart Study. Circ Arrhythm Electrophysiol 14, e010063, doi:10.1161/circep.121.010063 (2021). [PubMed: 34565178]
22. Lowres N, Neubeck L, Redfern J & Freedman SB Screening to identify unknown atrial fibrillation. Thromb Haemost 110, 213–222 (2013). [PubMed: 23595785]
23. Wang L et al. Association of Wearable Device Use With Pulse Rate and Health Care Use in Adults With Atrial Fibrillation. JAMA Network Open 4, e215821–e215821, doi:10.1001/jamanetworkopen.2021.5821 (2021). [PubMed: 34042996]
24. Ufomata E et al. A Policy Statement of the Society of General Internal Medicine on Tackling Racism in Medical Education: Reflections on the Past and a Call to Action for the Future. J Gen Intern Med 36, 1077–1081, doi:10.1007/s11606-020-06445-2 (2021). [PubMed: 33483823]
25. Feagin J & Bennefield Z Systemic racism and U.S. health care. Soc Sci Med 103, 7–14, doi:10.1016/j.socscimed.2013.09.006 (2014). [PubMed: 24507906]
26. Peabody JW, Luck J, Glassman P, Dresselhaus TR & Lee M Comparison of Vignettes, Standardized Patients, and Chart AbstractionA Prospective Validation Study of 3 Methods for Measuring Quality. JAMA 283, 1715–1722, doi:10.1001/jama.283.13.1715 (2000). [PubMed: 10755498]

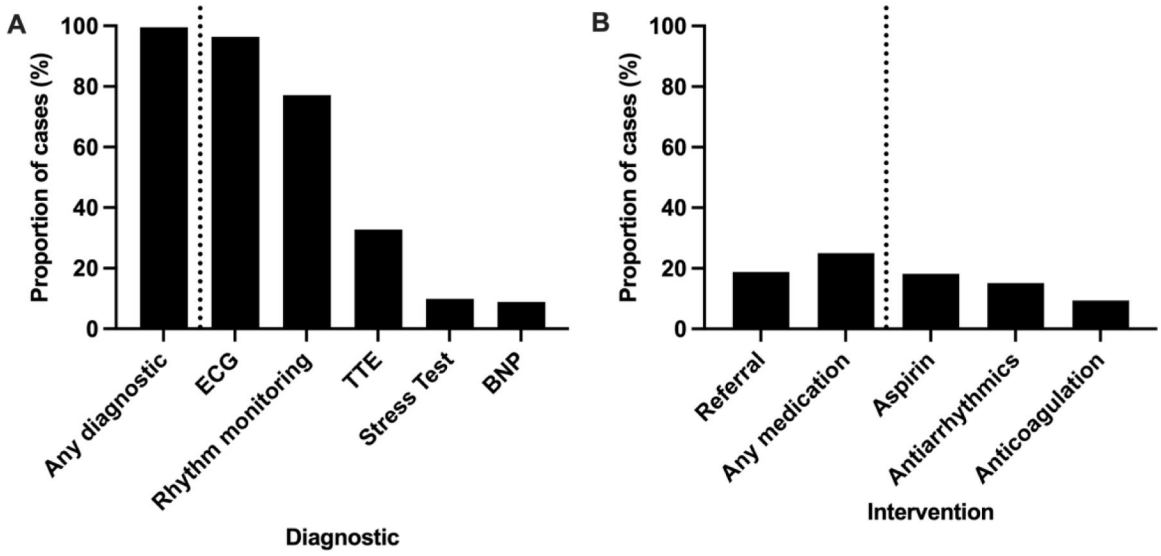

Figure 1. Testing and interventions ordered overall.

Among 95 participants, 192 hypothetical cases were completed. Here we report the average proportion of cases in which respondents indicated that they were “extremely likely” or “somewhat likely” to order specific diagnostic tests, referral/consultation, and therapeutic interventions. “Rhythm monitoring” indicates that any of the following were selected: event monitor, implantable loop recorder, patch monitor, or a commercially available heart rhythm monitor such as AliveCor. “Antiarrhythmics” included selection of a beta blocker or calcium channel blocker, or a class IC or III antiarrhythmic. ECG = electrocardiogram, TTE = transthoracic echocardiogram, and BNP = brain natriuretic peptide. Selections for referral, stress test, and aspirin were missing in 1 (0.5%) case.

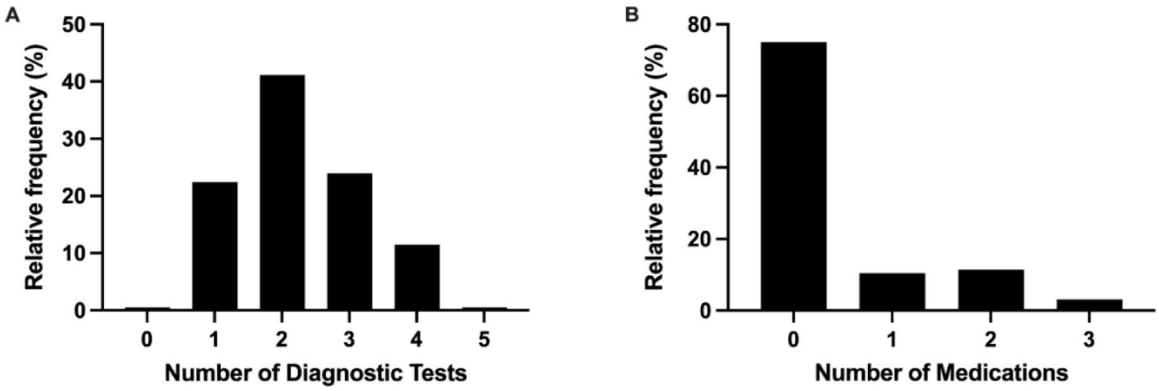

Figure 2. Frequency of (A) diagnostic testing and (B) medication treatment across cases.

The number of diagnostic tests (among ECG, rhythm monitoring, TTE, stress test, and BNP) and types of medications (among aspirin, antiarrhythmics, and anticoagulation) were tallied in each case (N = 192). The distributions of frequencies are reported here.

Table 1.

Respondent characteristics.

Respondent demographics and years in practice are presented overall and by specialty. Data were missing for 6 respondents (6.3% overall).

	Primary care (N=39)	Emergency medicine (N=25)	Cardiovascular/Electrophysiology (N=31)	Overall (N=95)
Gender				
Female	24 (61.5%)	11 (44.0%)	7 (22.6%)	42 (44.2%)
Male	13 (33.3%)	13 (52.0%)	21 (67.7%)	47 (49.5%)
Age in years				
< 45	19 (48.7%)	19 (76.0%)	15 (48.4%)	53 (55.8%)
45 – 65	17 (43.6%)	5 (20.0%)	11 (35.5%)	33 (34.7%)
> 65	1 (2.6%)	0 (0%)	2 (6.5%)	3 (3.2%)
Years in Practice				
0 – 9	13 (33.3%)	17 (68.0%)	8 (25.8%)	38 (40.0%)
10 – 19	8 (20.5%)	5 (20.0%)	8 (25.8%)	21 (22.1%)
> 20	16 (41.0%)	2 (8.0%)	12 (38.7%)	30 (31.6%)

Table 2.

Experience with smartwatches.

Respondents' experiences with smartwatches (personally wearing one, having had recommended one to a patient, and having encountered a patient reporting smartwatch alerts) are presented by specialty. Data were missing for 6 respondents (6.3% overall).

	Primary care (N=39)	Emergency medicine (N=25)	Cardiovascular/EP (N=31)	P-value
Wears a Smartwatch				0.681
Yes	13 (33.3%)	6 (24.0%)	8 (25.8%)	
No	24 (61.5%)	18 (72.0%)	20 (64.5%)	
Has Recommended Smartwatches to Patients				0.003
Yes	5 (12.8%)	3 (12.0%)	13 (41.9%)	
No	32 (82.1%)	21 (84.0%)	15 (48.4%)	
Has Experience with Patients Reporting Smartwatch Alerts				<0.001
Yes	12 (30.8%)	20 (80.0%)	25 (80.6%)	
No	25 (64.1%)	4 (16.0%)	3 (9.7%)	