Journal of the American Medical Informatics Association, 29(6), 2022, 1040–1049
https://doi.org/10.1093/jamia/ocac009
Advance Access Publication Date: 22 February 2022
Research and Applications

Research and Applications

Will Apple devices' passive atrial fibrillation detection prevent strokes? Estimating the proportion of high-risk actionable patients with real-world user data

Keith Feldman [1,2], Ray G. Duncan[3,4], An Nguyen[5], Galen Cook-Wiens[5], Yaron Elad[3,5], Teryl Nuckols[6], and Joshua M. Pevnick[3,6]

[1]Division of Health Services and Outcomes Research, Children's Mercy Kansas City, Kansas City, Missouri, USA, [2]Department of Pediatrics, University of Missouri-Kansas City School of Medicine, Kansas City, Missouri, USA, [3]Enterprise Information Services, Cedars-Sinai Health System, Los Angeles, California, USA, [4]Department of Pediatrics, Cedars-Sinai Medical Center, Los Angeles, California, USA, [5]Department of Cardiology, Cedars-Sinai Medical Center, Los Angeles, California, USA, and [6]Department of Medicine, Cedars-Sinai Medical Center, Los Angeles, California, USA

Corresponding Author: Joshua M. Pevnick, MD, MSHS, Department of Medicine, Division of General Internal Medicine, Cedars-Sinai Medical Center, Los Angeles, CA 90048, USA; joshua.pevnick@cshs.org

Received 20 September 2021; Revised 17 November 2021; Editorial Decision 7 January 2022; Accepted 27 January 2022

ABSTRACT

Objective: Utilizing integrated electronic health record (EHR) and consumer-grade wearable device data, we sought to provide real-world estimates for the proportion of wearers that would likely benefit from anticoagulation if an atrial fibrillation (AFib) diagnosis was made based on wearable device data.

Materials and Methods: This study utilized EHR and Apple Watch data from an observational cohort of 1802 patients at Cedars-Sinai Medical Center who linked devices to the EHR between April 25, 2015 and November 16, 2018. Using these data, we estimated the number of high-risk patients who would be actionable for anticoagulation based on (1) medical history, (2) Apple Watch wear patterns, and (3) AFib risk, as determined by an existing validated model.

Results: Based on the characteristics of this cohort, a mean of 0.25% ($n = 4.58$, 95% CI, 2.0–8.0) of patients would be candidates for new anticoagulation based on AFib identified by their Apple Watch. Using EHR data alone, we find that only approximately 36% of the 1802 patients ($n = 665.93$, 95% CI, 626.0–706.0) would have anticoagulation recommended even after a new AFib diagnosis.

Discussion and Conclusion: These data suggest that there is limited benefit to detect and treat AFib with anticoagulation among this cohort, but that accessing clinical and demographic data from the EHR could help target devices to the patients with the highest potential for benefit. Future research may analyze this relationship at other sites and among other wearable users, including among those who have not linked devices to their EHR.

Key words: atrial fibrillation, mobile health (mHealth), precision health, patient-generated data

BACKGROUND AND SIGNIFICANCE

Due to impaired blood flow, patients with atrial fibrillation (AFib) are known to have a 5-fold increased stroke risk over those without the diagnosis.[1] Prophylactic anticoagulation is commonly utilized as a means to reduce such risk.[2] Unfortunately, anticoagulation is known to be underutilized in this patient population, based in part on underdiagnosis of AFib.[3–5] As part of ongoing efforts to identify patients who could benefit from anticoagulation, multiple studies

have demonstrated promise in the utilization of photoplethysmography (PPG) heart rate data collected from consumer-grade wearable devices by identifying patients with suspected AFib outside of the clinical setting. To do so, machine learning and other computational approaches (eg, deep learning and time series analyses) have illustrated an ability to identify periods of AFib with high sensitivity and specificity.[6,7] However, simply identifying AFib does not necessarily warrant anticoagulation for all patients. For example, there are many patients who are already on anticoagulation, who have contraindications to anticoagulation, or for whom anticoagulation would not be indicated, even with a diagnosis of AFib, based on an otherwise low risk for stroke (eg, a young and otherwise healthy patient).

To date, it has been difficult to quantify the potential clinical benefit of PPG monitoring because most studies of this topic either use data from the electronic health record (EHR) *or* data from wearable devices. Studies using device data lack access to rich EHR data, such that demographic and clinical characteristics are usually limited to a few simple data elements (eg, age or diagnosis of diabetes). Instead, these data are typically self-reported, which may be limited by patients' knowledge of their health conditions. Similarly, EHR-based studies have lacked granular device data, such that it is impossible to know which high-risk patients are wearing devices and for how long, much less to incorporate device data (eg, observed heart rates) into a monitoring and care plan.

This study aims to directly address this gap, utilizing a health system-wide data resource that includes both consumer-grade wearable device data and EHR data. We leveraged this real-world data to estimate the proportion of wearable device users with potential to benefit from anticoagulation following an AFib diagnosis.

MATERIALS AND METHODS

Study cohort and data

This analysis utilized data from an observational cohort of CS-Link users, Cedars-Sinai Health System's (CSHS) branding of the EpicCare enterprise EHR product. CS-Link is used at Cedars-Sinai Medical Center, a large, nonprofit hospital and at many local associated provider organizations both within and outside CSHS. Together, these organizations provide multidisciplinary care across a socioeconomically diverse population.

Beginning in April 2015, to enable the EHR to provide a more complete picture of patients' health status, CSHS began inviting patients to sync consumer-directed health devices directly with the clinical EHR. While consumer interfaces initially supported only the Apple HealthKit software framework, the platform has been expanded to include applications and devices compatible with the Google Fit, Fitbit, and Withings frameworks. Once linked, passively collected data (eg, heart rate) are automatically uploaded into the EHR daily. Prior publications by the study team describe adoption of device linking, provide details around characteristics of early adopters, and discuss early clinical efforts utilizing these data.[8,9] An overview of system architecture can be found in Figure 1.

Given the extended period of possible enrollment, this study utilized only those patients who synced data through the HealthKit inter-

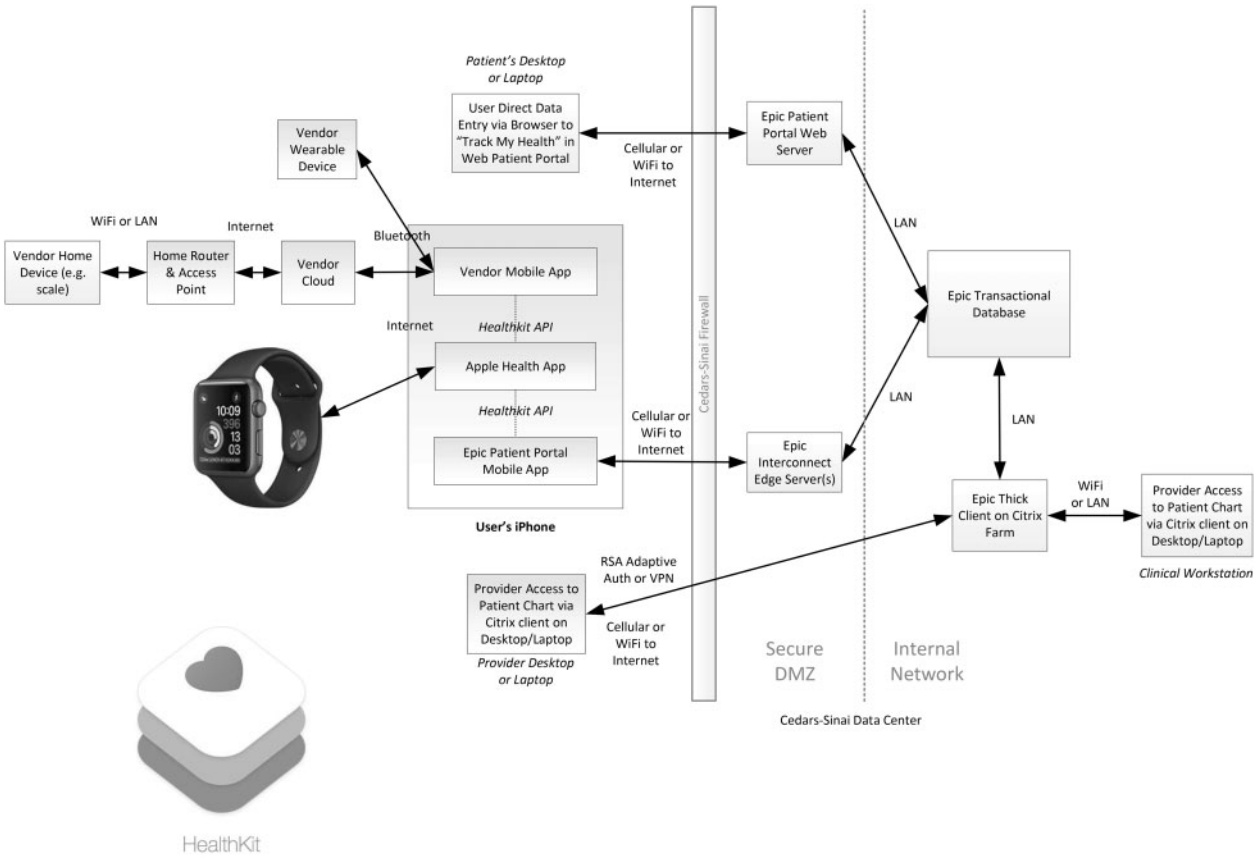

Figure 1. CS-link architecture diagram: Diagram of data flow for integration of Apple HealthKit devices with the Epic electronic medical record system. Ingested data can be viewed by patients in the Epic patient portal and by caregivers within the Epic clinical user interface.

face. Patients eligible for this study must have linked their device to the EHR between April 25, 2015 and November 16, 2018. It is important to note these data were sampled from a time before Apple Watches actively alerted users to irregular heart rhythms. This time window was selected to ensure no patients in the study had been diagnosed with AFib due to Apple Watch data. Otherwise, analysis would risk underestimating potential benefit, as patients who had already benefitted from the device would be excluded with a known diagnosis.

Patients <18 years of age were also excluded, as were any individuals who completed an explicit opt-out document excluding the use of any EHR/portal data from use in research studies. The study was approved by both the Cedars-Sinai and Children's Mercy institutional review boards.

Data elements

For all eligible patients, data were extracted from two sources, *device-level data* and *EHR data*.

Device data

In addition to the raw synced data, integration with CS-Link allowed us to leverage metadata for each datapoint including the syncing platform (eg, Withings and Apple HealthKit) and a second-level timestamp. All non-numeric data (the result of manual logging on a respective platform) and those records with missing metadata (platform and/or timestamp) were dropped prior to analysis.

Clinical data

As part of the Cedars-Sinai EHR, we were able to extract a comprehensive set of social, demographic, and clinical data on each patient. These data included: demographics: (age, sex, race, and ethnicity), anthropometric data (height, weight, and blood pressure), social variables (smoking and alcohol use), diagnosis and procedure history, and medication history. Complete definitions for each variable can be found in the Clinical data definitions section of the online Supplementary Material.

Actionability criteria

To quantify the proportion of individuals likely to benefit from detection of irregular heart rhythms through passive detection of irregular heart rhythms on the Apple Watch, we defined several criteria pertaining to aspects of actionability and risk. First, as the primary actionable course of care for patients with a recent AFib diagnosis is anticoagulation therapy, we utilized EHR data to determine if each patient would be a candidate for anticoagulation should AFib be detected (clinical candidates). Second, we identified individuals who wore their devices at sufficient intervals to allow for AFib detection based on passive detection algorithm criteria (wear patterns). Finally, we estimated the 5-year AFib risk of each patient using a previously validated model (AFib risk). Detail around each criterion can be found in the respective sections to follow.

Criteria 1: Clinical candidates for anticoagulation therapy

We first identified patients clinically ineligible for anticoagulation, and those at a sufficiently low stroke risk such that anticoagulation would not be warranted even in the case of an AFib diagnosis. Definitions for eligibility and thresholds for actionable stroke risk can be found in Table 1. Furthermore, as this work centers on identifying actionable AFib in the context of passive mHealth detection, patients with an AFib diagnosis at the time of data extraction/assessment were also considered ineligible, as it is expected action would have already been taken should it have been warranted.

Criteria 2: Device wear patterns sufficient for detection

We next assessed device wear patterns of each patient to assess if they were of sufficient length to allow for AFib detection given the Apple Watch algorithm. In recent iterations of the Apple Watch, passive monitoring from the heart rate sensor can notify users to irregular heart rhythms including AFib. Per Apple's documentation, the watch requires "five out of six sequential tachograms—including the initial one—to be classified as irregular within a 48-hour period," before alerting the user.[12] Once the first irregular pattern is detected, accelerated sampling occurs at 15-min intervals. Theorizing the watch detected an irregular rhythm immediately as the device was placed on the user and the next four tachograms were all irregular; a minimum of 60 min of wear time would be required to receive an alert. In an effort to provide conservative estimates, this duration of wear time must occur only one time after a user links their device with the EMR to be considered eligible.

Table 1. Clinical criteria to be eligible for anticoagulation

Contraindication to anticoagulation:	Patients with an absolute contraindication to anticoagulation were considered ineligible. These included history of intracranial hemorrhage, intracranial mass, or end-stage liver disease comorbid with several existing conditions including hepatic decompensation esophageal varices, ascites, hepatic coma, peritonitis, or hepatorenal syndrome as defined by Steinberg et al.[10]
Already anticoagulated:	Patients who were already anticoagulated, as evidenced by anticoagulant prescriptions within the 6 months prior to data extraction on November 16, 2018 were considered ineligible. Anticoagulants included any prescription for: warfarin, rivaroxaban, dabigatran, apixaban, edoxaban, or fondaparinux.
Insufficient stroke risk to warrant anticoagulation:	A CHA_2DS_2-VASc score was computed for each patient based on age, sex, and documentation of ICD-10 codes representing hypertension, stroke, vassal disease, and diabetes. Derived specifically for patients with AFib, CHA_2DS_2-VASc is a routinely utilized assessment of stoke risk, with the intention to start anticoagulation for high-risk patients. Specific codes to compute the score were drawn from the forward-backward ICD-10 mapping by Webster-Clark et al.[11] Males with a score <1 and females with a score <2 were considered ineligible as they would likely not be placed on anticoagulation due to their low stroke risk.

Abbreviation: AFib: atrial fibrillation; CHA_2DS_2-VASc: Congestive heart failure, Hypertension, Age ≥ 75 years, Diabetes mellitus, Stroke, Vascular disease, Age 65–74 years, Sex category (female); ICD: International Classification of Disease.

As the tachograms must be sequential, and the Apple Watch does not record pulse at set intervals (skipping measurements in the event of movement or other anomalies), it was necessary to define the period between pulse readings could be considered contiguous wear. We again take a conservative approach, allowing for up to 30 min between measurements. Recent software is known to poll for pulse data as infrequently as once per 10 min. Thus, we allow for approximately two skipped readings which we believe to be reasonable for an active individual.

Criteria 3: Future risk of developing AFib

Finally, to estimate the size of an actionable cohort, we needed an estimate of the expected incidence of AFib among the cohort. To do so, we utilized the previously validated CHARGE-AF model to compute a 5-year AFib risk for each patient.[13] Details of the estimation can be found in the online Supplementary Material.

Data analysis

Our primary analysis estimated the actionable cohort size based on each of eligibility and future risk criteria defined for each patient. Rather than arrive at a singular number by filtering ineligible patients we utilized a robust 10 000 iteration bootstrap to estimate the expected proportion of individuals that would meet each criterion.

Each iteration began by resampling, with replacement, the study cohort until the original size of the data was reached. From the resampled cohort, patient's ineligible for anticoagulation based on clinical criteria 1 were removed, followed by those with insufficient wear times as defined by criteria 2. To estimate incidence based on the 5-year AFib probability obtained from criteria 3, we draw a random value from a uniform distribution from 0 to 1 for each remaining individual, where those with values above their CHARGE-AF modeled risk were excluded. Given the bootstrap approach, this random sampling allowed us to better represent an expected proportion of patients at any stage more robustly to outliers (eg, rare individuals with very high/low future AFib risk are likely present in fewer bootstrap samples and thus less common in overall statistics).

In line with prior works, we also included an incidental identification step, to simulate a percentage of individuals expected to have an AFib diagnosis made during routine clinical care *without* the use of mHealth devices.[14,15] To do so, we sampled (without replacement) those patients remaining after criteria 3 to drop a preset percentage, initially set to the published rate of 36.09%. For completeness, we assessed a range of possible rates at 20%, 10%, and 0%.

Across the 10 000 iterations, demographic statistics and 95% confidence intervals were computed on both those patients who remain eligible as well as those dropped at each stage.

Review of high-risk patients

Taking the identification of actionable patients one step further, we concluded with a case study of individuals to understand how this framework could be used to identify high-risk patients for future monitoring. We began the case study by identifying those patients whose wear patterns made them strong candidates for detection of AFib should it develop. This was quantified using two metrics.

1. The percentage of possible days a user wore the device.
   a. Taking the first and last day a user-logged data from their device, we computed the percentage of days with *any* heart rate data to obtain at a measure of wear regularity.
2. The percentage of wear sessions in which the device was worn for >60 min.
   a. Utilizing the same 30-min threshold to define gaps in wear time, we computed the percentage of total wear intervals long enough for a watch to detect AFib under ideal circumstances.

We isolated patients at or above the 50th percentile of both measures, as they represent the most likely to benefit from passive detection algorithms. Then, using the historically recorded heart rate data, high-risk patients were ranked as follows. First a derived score was created to quantify concerning high/low pulse measurements from their device. Patients were given one-point if three occurrences of pulse data fell below 40, two of which were at least 1 h apart, and another point if five measurements were above 150, two of which were at least 1 h apart. Ties were broken by sorting on estimated 5-year AFib risk from the CHARGE-AF model, with predictions rounded to the nearest 5%. Finally, any remaining ties were broken by CHADS score indicating need for oral anticoagulation.

Demographics and pulse data for the top-five patients were reviewed together by both a cardiologist informatician and internal medicine informatician to confirm identified patients would warrant strong consideration for anticoagulation.

RESULTS

As of November 16, 2018, patients had synced 3026 devices, the majority of which were wearable devices (others included smart-scales, blood pressure monitors, etc). In total 2051 patients synced an Apple device to CS-Link during the study period. Of those, 113 opted-out from use of their data in research. As patients age and sex were drawn from the linked EHR, and recent literature has suggested maintaining unknown race/ethnicity data[16] there were no exclusions for missing demographic data. However, 136 patients were removed as incomplete anthropometric data precluded the calculation of a 5-year AFib risk score with the CHARGE-AF model. In total this resulted in a study cohort of 1802 patients representing over 64 million heart rate values across 291 464 total person days.

With respect to patient race, those who declined to answer ($n = 10$) or left the field blank ($n = 12$) were grouped with those recorded in the EHR as unknown ($n = 111$). For those with a documented race, 16 patients fell into low prevalence groupings, American Indian or Alaska Native ($n = 3$), and Native Hawaiian or Other Pacific Islander ($n = 11$), both of which were aggregated as Other. Similarly, those who declined to specify a Hispanic/Non-Hispanic ethnicity ($n = 12$) or left the field blank ($n = 11$) were grouped with those with ethnicity documented as Unknown ($n = 102$). For a small subset ($n = 105$), alcohol use was not recorded in the recent visit, and was thus not factored into the demographic summaries of use at each bootstrap iteration. A complete breakdown of cohort demographics can be found in Table 2.

Looking first to the estimation of actionability, Figure 2 presents the mean number of patients, the 95% conference interval, and the bootstrap sampling distribution of eligible patients as they are filtered from criteria 1 to 3. We find that using EHR data alone, there is significant drop to only 36.96% of patients (bootstrapped mean $n = 665.93$, 95% CI, 626.0–706.0) for whom anticoagulation would be recommended even with a positive AFib diagnosis. Of those only about 483.09 (95% CI, 447.0–520.0) are expected to wear their Apple Watch enough for detection to be possible, representing only 26.81% of the total cohort. Finally, based on character-

Table 2. Study cohort demographics

		Missing	Overall
	N		1802
	Age, years	0	45.96 (13.88)
	Sex, male (*n*, %)		1154 (64.04)
Race (*n*, %)	White	0	1223 (67.87)
	Asian		179 (9.93)
	Black or AA		153 (8.49)
	Other		138 (7.66)
	Unknown		109 (6.05)
Ethnicity (*n*, %)	Hispanic	0	239 (13.26)
	Non-Hispanic		1460 (81.02)
	Unknown		103 (5.72)
Anthropometric	Height, in	0	68.09 (4.07)
	Weight, lbs	0	186.17 (44.79)
	DBP, mmHg	0	77.44 (8.86)
	SBP, mmHg	0	124.01 (13.25)
Risk factors (*n*, %)	Hypertension	0	842 (46.73)
	Diabetes	0	166 (9.21)
	Smoking Hx	0	558 (30.97)
	Alcohol use	105	1165 (68.65)
	Surgery	0	8 (0.44)
Clinical factors	Contraindicated (*n*, %)	0	36 (2.00)
	Recent anticoagulation (*n*, %)	0	56 (3.11)
	AFib diagnosis (*n*, %)	0	136 (7.55)
	CHA_2DS_2-VASc	0	1.32 (1.44)
	CHARGE-AF risk	0	1.31 (2.99)

Abbreviations: AA: African American; AFib: atrial fibrillation; CHARGE-AF: Cohorts for Heart and Aging Research in Genomic Epidemiology Atrial Fibrillation; CHA_2DS_2-VASc: Congestive heart failure, Hypertension, Age $\geq$ 75 years, Diabetes mellitus, Stroke, Vascular disease, Age 65–74 years, Sex category (female); DBP: diastolic blood pressure; SBP: systolic blood pressure.

istics of the remaining cohort the estimated 5-year risk of AFib suggests that only an average of 4.58 (95% CI, 2.0–8.0) of these patients would be likely AFib positive and candidates for new anticoagulation based on AFib identified by their wearable device, representing only 0.25% of the total eligible cohort.

Supplementary Figure S1 presents the same analytical results with final cohorts obtained using incidental detection rates of 20%, 10%, and 0%. However, given the low AFib risk across the cohort, these values were only marginally different with the mean actionable cohort size using the published detection rate moving from 4.58 patients (36% identification rate) to 7.22 patients (0% identification rate).

Table 3 provides a detailed summary of the demographic profiles for both the included and excluded patients at each estimation stage, as well as the initial and final cohorts. For completeness, we separated the risk estimation as determined by CHARGE-AF model and the incidental detection steps into two columns (criteria 1–3: Retained and Final, respectively). Additionally, we repeated the analysis removing criteria 2 (Wear Time), allowing us to capture a potential actionable cohort for whom we could focus monitoring those most at risk. However, as our criteria for wear patterns were highly conservative (most users wear the watch for a session of sufficient length), results were only marginally different. Complete bootstrap measures of the reduced framework can be found in Supplementary Table S1.

Review of high-risk patients

Moving to the assessment of top-5 high-risk patients, Figure 3 presents an overview of their device heart rate data, while Table 4 provides detailed demographics of each. Dates were standardized to a uniform start point for privacy. Clinical review of patients indicated several notable characteristics considered worrisome for development of AFib. These included age over 60, history of several known clinical risk factors (eg, hypertension and diabetes), and social risks (eg, smoking and alcohol use). Inspection of heart rates further supports labeling these individuals as high-risk for future actionable AFib, highlighting several periods of low and high rates, in some cases becoming more common over time. In each of the five cases, physicians agreed that sufficient risk factors were present to consider anticoagulation in any of these patients if they had abnormal device data that led to a diagnosis of atrial fibrillation.

DISCUSSION

In seeking to estimate the potential for Apple Watches to benefit users by identifying AFib, we accessed data from a population of Apple device users who linked devices to the EHR. Based on the clinical and demographic characteristics of this cohort, their watch wear patterns, and their likelihood of developing AFib specified by the CHARGE-AF model, we estimate that only 0.25% of these patients would be candidates for new anticoagulation based on AFib identified by their Apple device.

Although a number needed to wear (NNW) of 400 patients (1/0.25% = 400) to detect one case of AFib with potential to benefit from anticoagulation may seem high, this is actually lower than the NNW of 2741 implied by Perez et al,[6] in which AFib was detected by Apple Watch use in only 153 (0.04%) of 419 297 participants. Because that study was limited by several real-world constraints (eg, only 21% of patients who received irregular pulse notifications continued with further testing), this may be a more accurate estimate of likely benefit.

To further improve estimates potential benefit to patients, it would also be important to consider the number needed to treat (NNT) for patients to benefit from anticoagulation in AFib. Recently a net NNT (to provide net benefit when risks are also considered) of 34 with anticoagulation was estimated for AFib.[17] Although Ding et al estimate that NNT may drop as low as 11 for patients with a baseline risk of stroke over 10%, that would still imply a NNW of around 4000 ($11\times400=4000$), or over 27 000 ($11\times2741=27\,000$) based on Perez et al, to detect a case of AFib that would actually benefit from anticoagulation (as opposed to the prior NNWs which reference the number needed to detect a case of AFib with potential to benefit from anticoagulation). Even these estimates do not account for the preference of some patients to avoid anticoagulation.

At least two important conclusions can be drawn from these estimates. First, a randomized controlled trial to study the benefit of consumer-directed heart rate monitoring devices in preventing strokes would require either a massive sample size or an enriched sample of patients very likely to experience stroke due to AFib. This is especially important in the context of initial reports showing that device users skew towards the young and healthy.[6,8] For clinical trialists or health care provider organizations seeking to identify patients who could benefit most from such devices, the simple multistep protocol we used in this study may be especially helpful.

Second, because recent work found Apple Watch use to be associated with increased health care utilization, it would be critical to estimate and minimize the extent to which false positives increased harm or utilization. For example, Wyatt et al[18] found that among 264 patients who presented for clinical care following detection of an abnormal pulse by a wearable device, only 30 (11.4%) resulted in a clinically actionable cardiovascular diagnosis. Wang et al[19]

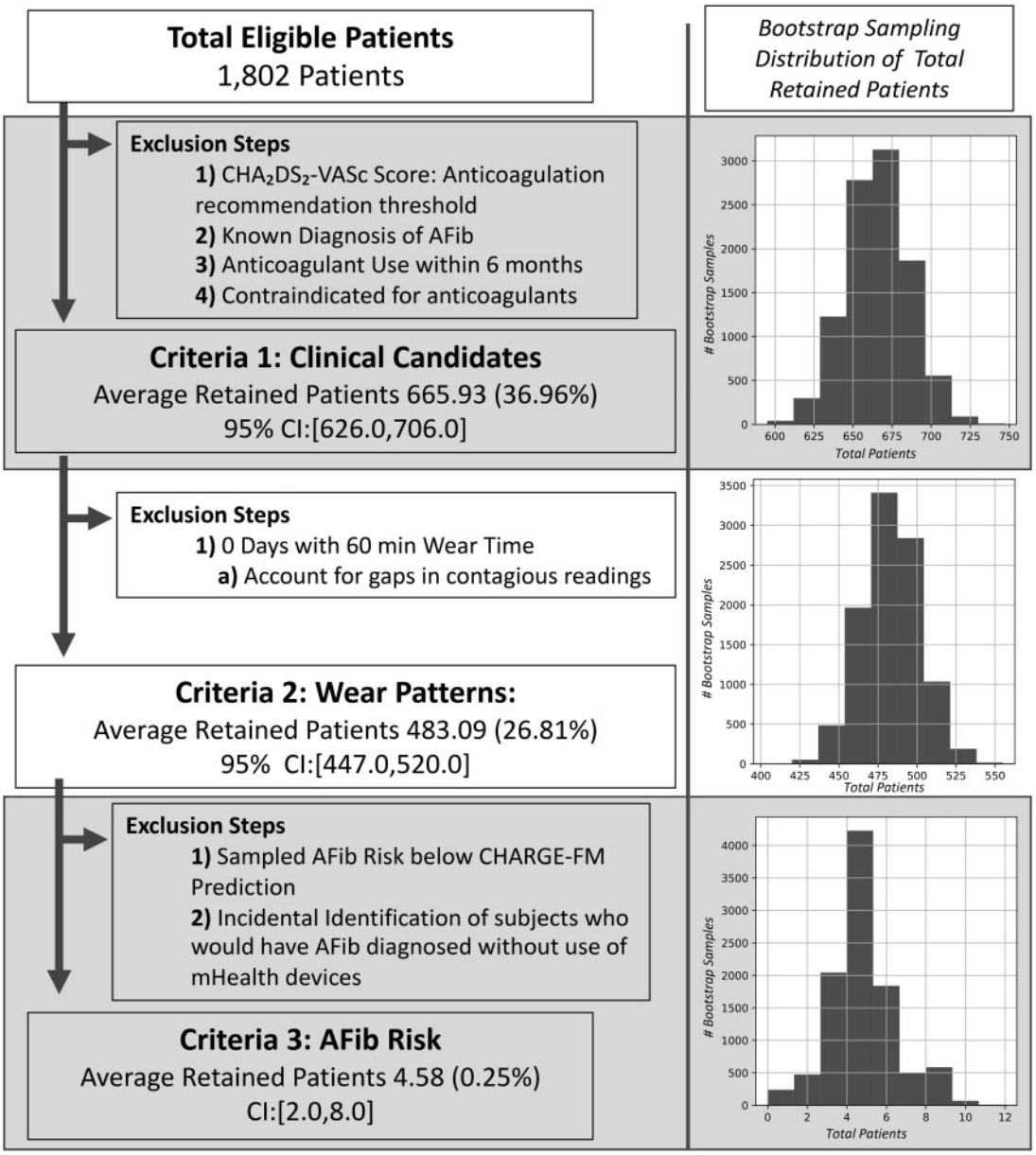

Figure 2. Cohort estimation results: the exclusion criteria for each stage of the cohort estimation bootstrap analysis are listed along with the mean patient count remaining after each step. In addition, each step is listed along with 95% confidence interval (CI) for the cohort size, proportion of total analysis cohort, and sampling distribution from the bootstrap analysis.

found that among 16 320 patients already known to have AFib, wearable use was associated with increased health care utilization, even after propensity matching for other factors, including heart rate. Because Perez et al found no AFib in 66% of Apple Watch notifications of possible AFib, it seems there would be substantial potential for false positives to detract from benefits accrued in cases of new AFib diagnosis.

In exploring whether these high NNWs might ever be reduced, it is instructive to consider which characteristics were most responsible for reducing patient eligibility. Of all the aforementioned criteria, clinical and demographic characteristics alone were sufficient to reveal the majority (63%) of patients in the study had no potential to derive benefit from heart rate monitoring, at least via new identification of AFib that would lead to anticoagulation. Although other studies of wearable devices have been able to query users about some health conditions that raise the risk of stroke,[6] many users may not be able to report their health histories in the detail available in the EHR. We would thus recommend that future studies using wearable devices to detect AFib should consider accessing EHR data to identify populations most likely to benefit before targeting them for devices.

Limitations

We recognize several limitations to this analysis. First, with respect to the cohort itself, our data are from a single site, and thus may not represent the clinical profile or wear patterns of the larger population and of those who linked their devices at later times. However, we believe datasets that include linked wearable data and complete

Table 3. Bootstrap estimates: each column presents the demographic, anthropometric, and social variables for each stage of the analysis.

Framework criteria			Criteria 1: EHR data		Criteria 1 and 2: wear patterns		Criteria 1–3: AFib risk		Incidental detection dropped
		Bootstrap sample	Drop	Retained	Drop	Retained	Drop	Retained	Final cohort
	N	1802.0 (0.0)	1136.07 (20.33)	665.93 (20.33)	182.83 (12.95)	483.09 (18.57)	475.87 (18.49)	7.23 (2.69)	4.58 (1.71)
	Age, years	45.97 (0.33)	42.36 (0.38)	52.13 (0.52)	56.77 (1.05)	50.37 (0.58)	50.18 (0.58)	62.98 (4.99)	62.98 (6.26)
	Sex, male	64.05 (1.14)	61.27 (1.46)	68.79 (1.78)	66.08 (3.47)	69.82 (2.08)	69.64 (2.09)	81.25 (15.9)	81.09 (20.17)
Race	White	67.87 (1.1)	67.42 (1.38)	68.65 (1.81)	69.44 (3.4)	68.35 (2.13)	68.1 (2.16)	84.83 (14.49)	84.78 (18.2)
	Asian	9.93 (0.7)	9.6 (0.86)	10.49 (1.2)	10.34 (2.25)	10.54 (1.4)	10.63 (1.42)	4.71 (8.57)	4.7 (10.74)
	Black or AA	8.49 (0.65)	7.49 (0.77)	10.2 (1.17)	9.28 (2.15)	10.54 (1.39)	10.61 (1.41)	5.97 (9.6)	6.07 (12.16)
	Other	7.66 (0.63)	8.62 (0.85)	6.01 (0.92)	6.01 (1.73)	6.01 (1.08)	6.06 (1.09)	2.7 (6.55)	2.74 (8.25)
	Unknown	6.05 (0.56)	6.86 (0.75)	4.66 (0.82)	4.93 (1.61)	4.56 (0.95)	4.6 (0.96)	1.8 (5.34)	1.72 (6.54)
Ethnicity	Hispanic	13.27 (0.8)	15.06 (1.06)	10.22 (1.19)	4.93 (1.6)	12.22 (1.5)	12.32 (1.52)	5.67 (9.53)	5.69 (11.95)
	Non-Hispanic	81.01 (0.93)	77.9 (1.23)	86.32 (1.35)	92.88 (1.89)	83.84 (1.7)	83.71 (1.71)	92.27 (10.92)	92.27 (13.66)
	Unknown	5.71 (0.55)	7.04 (0.76)	3.46 (0.7)	2.2 (1.08)	3.94 (0.89)	3.97 (0.9)	2.05 (5.8)	2.04 (7.25)
Anthropometric	Height, in	68.09 (0.1)	68.04 (0.12)	68.18 (0.16)	67.74 (0.29)	68.34 (0.19)	68.32 (0.19)	69.76 (1.48)	69.75 (1.87)
	Weight, lbs.	186.19 (1.06)	180.54 (1.27)	195.82 (1.8)	186.9 (3.17)	199.2 (2.15)	199.13 (2.16)	203.8 (18.54)	203.79 (23.43)
	DBP, mmHg	77.43 (0.21)	76.15 (0.24)	79.62 (0.37)	78.64 (0.71)	80.0 (0.43)	80.04 (0.43)	77.16 (3.7)	77.16 (4.64)
	SBP, mmHg	124.01 (0.31)	121.02 (0.37)	129.12 (0.5)	130.59 (0.98)	128.56 (0.58)	128.54 (0.59)	129.66 (4.73)	129.61 (5.93)
Risk factors	Hypertension	46.75 (1.16)	29.25 (1.34)	76.6 (1.61)	85.83 (2.57)	73.1 (1.99)	72.99 (2.01)	80.61 (16.16)	80.78 (20.14)
	Diabetes	9.21 (0.68)	2.47 (0.46)	20.72 (1.56)	21.84 (3.08)	20.3 (1.81)	20.25 (1.82)	23.48 (17.11)	23.5 (21.68)
	Smoking Hx	30.96 (1.08)	29.56 (1.34)	33.34 (1.82)	42.64 (3.7)	29.83 (2.05)	29.66 (2.06)	40.75 (19.93)	40.53 (25.09)
	Alcohol use	68.63 (1.13)	68.71 (1.43)	68.5 (1.87)	74.55 (3.38)	66.26 (2.22)	66.26 (2.24)	66.62 (19.91)	66.49 (24.93)
	Surgery	0.44 (0.16)	0.53 (0.21)	0.3 (0.21)	0.54 (0.54)	0.2 (0.2)	0.2 (0.21)	0.12 (1.61)	0.14 (1.99)
Clinical factors	Contraindicated	2.0 (0.33)	3.17 (0.52)	–	–	–	–	–	–
	Recent anticoagulation	3.12 (0.41)	4.94 (0.64)	–	–	–	–	–	–
	AFib diagnosis	7.56 (0.62)	11.99 (0.97)	–	–	–	–	–	–
	CHADS	1.32 (0.03)	0.81 (0.04)	2.19 (0.05)	2.6 (0.11)	2.03 (0.05)	2.03 (0.05)	2.57 (0.61)	2.57 (0.77)
	CHARGE-AF risk	1.31 (0.07)	1.0 (0.09)	1.83 (0.11)	2.73 (0.31)	1.49 (0.09)	1.45 (0.09)	4.15 (1.39)	4.14 (1.74)

At each, we include summaries of each variable for both those patients who are excluded based on the criteria (Dropped) and those who are carried forward to the next stage (Retained); also included are estimates of the full bootstrap sample at the start of each iteration and the final cohort after the incidental detection sampling. For continuous variables, data represent the mean value of the feature, while for binary data these are the average proportion of each category. In both cases the parenthetical text captures the standard deviation of the metric across the bootstrap iterations. Dashed cells represent those factors not applicable due to previous exclusion criterion.

Abbreviations: AA: African American; AFib: atrial fibrillation; DBP: diastolic blood pressure; SBP: systolic blood pressure.

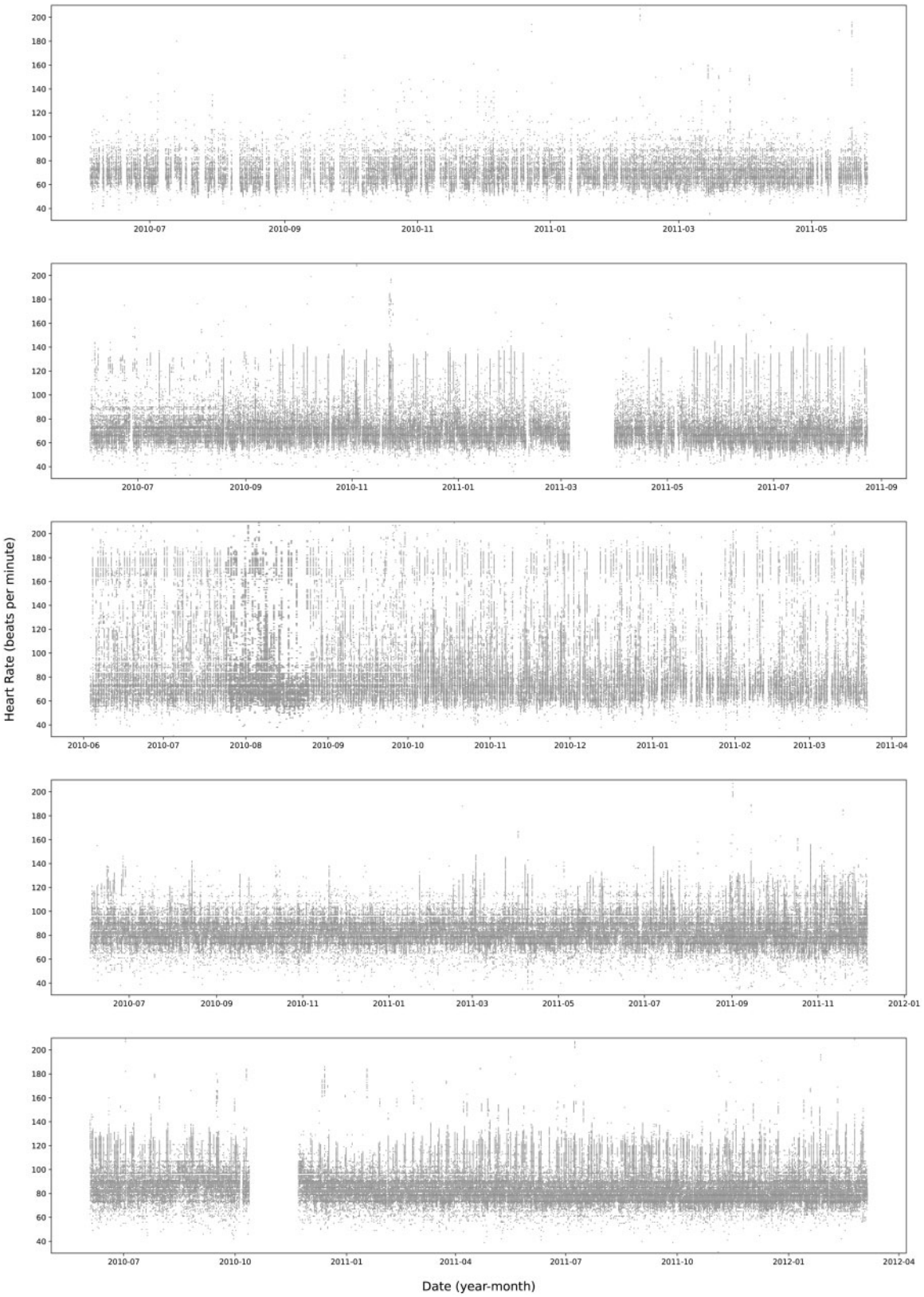

Figure 3. Validation study heart rate: overview of heart rate measurements collected from the top-5 highest risk patients. Although the dates have been shifted, the time between measurements has been maintained.

Table 4. Demographic profiles of top 5 highest risk patients making it through all cohort estimation criteria

Age (years)	Sex	Race	Ethnicity	Ht (in)	Wt. (lbs)	SBP (mmHg)	DBP (mmHg)	HTN	DM	SM	Alcohol use	CHARGE-AF (%)	CHADS
69	M	White	Non-Hispanic	73	247	118	78	Y	N	Y	Unknown	7.71	2
76	M	White	Non-Hispanic	66	151	131	70	Y	N	Y	Y	6.54	3
72	M	White	Non-Hispanic	67	184	137	78	N	Y	Y	Y	4.78	3
65	M	White	Non-Hispanic	68	269	136	75	N	Y	N	Y	3.65	3
61	M	White	Non-Hispanic	69	219	122	79	Y	Y	Y	N	3.12	3

Abbreviations: DM: diabetes; HTN: hypertension; Ht: height; M: male; *N*: No; SBP/DBP: systolic/diastolic blood pressures; SM: smoking history; Wt.: weight; Y: yes.

EHR data are scarce, and opening discussion of how their combined use can provide additional utility is valuable. Second, this analysis only estimates anticoagulation benefit of identifying AFib. Although we maintain this is the major benefit of detection, we would also acknowledge that patients might benefit in other ways from a diagnosis of AFib (eg, cardioversion, ablation, and reducing alcohol intake). Third, a larger sample size could help to reduce random error. However, even if the true percentage of patients with potential to benefit was 0.5% (near the top of our 95% CI), our conclusions regarding the importance of matching devices to high-risk patients would still hold. Lastly, as all eligible patients in our cohort must have enrolled in the CS-Link platform, is important to note potential for latent bias related to those who engage in patient-portal usage. It is well recognized that minority populations have been less likely to be offered access and/or enroll in patient portals,[20–22] and thus the cohort on which actionability was assessed may be underrepresented for the population at large. Future studies may consider more active recruitment before assessing populations at risk. Moreover, we recognize the final actionability criteria entail an estimation of future AFib. Although the CHARGE-AF model used to estimate this risk was developed and validated on large and diverse cohorts, we see how model features influence the final cohort. For example, individuals with a documented White race were found to have higher odds of future AFib (hazard ratio 1.59; 95% CI, 1.33–1.91). While finding is concordant with epidemiologic studies,[23] it fosters an increased proportion of White individuals in the final cohort which must be considered when interpreting the study results.

From a technical perspective, we attempted to use conservative criteria to determine exclusions at every stage (eg, only requiring a single day of sufficient wear time), suggesting that even this low estimate of potential for benefit may be upwardly biased. As with all ICD-based research, we cannot account for diagnoses not made or documented. We recognize that due to missing diagnoses, the CHADS score may be underestimated for some patients, for whom a more complete evaluation may be undertaken if they were thought to be at risk for AFib. Finally, we acknowledge these data were only collected from devices linked to the EHR using Apple HealthKit. Given the extended period in which Apple interfaces were synced with CS-Link, this provided the most robust sample at the time of analysis. It also allowed for a principled estimate of wear time eligibility given the company's public whitepaper.[12] However, future analyses would benefit from including data from other devices.

CONCLUSION

Analyzing data from patients who chose to connect their Apple Watches to an EHR at one site suggests that only about 0.25% of patients could derive clinical benefit from new identification of AFib, at least in terms of anticoagulation benefit to reduce the risk of stroke. Unless future research reveals higher potential for benefit at other sites and or among other wearable users, impacting the health of populations will require use of heart rate monitoring devices by patients with the greatest potential to benefit. Our analysis shows good potential for using clinical and demographic data from the EHR to identify these patients.

FUNDING

This work was supported by funding from Cedars-Sinai Precision Health.

AUTHOR CONTRIBUTIONS

JMP, TN, and KF contributed to the concept and design of the study. JMP, RGD, and AN extracted and prepared the data. JMP, RGD, AN, GCW, YE, and KF implemented the study methods. JMP, AN, YE, TN, and KF interpreted study results. All authors participated in writing of the manuscript.

SUPPLEMENTARY MATERIAL

Supplementary material is available at *JAMIA* online.

ACKNOWLEDGEMENTS

The authors thank Dr Abhi Jain for his work in organizing the clinical data used in this project and Dr Sumeet Chugh for his counsel and assistance, especially in obtaining funding.

CONFLICT OF INTEREST STATEMENT

None declared.

DATA AVAILABILITY

The data used in this analysis were aggregated from operational clinical sources with the approval of the Cedars-Sinai Medical Center Institutional Review Board (CSMC IRB) and due to concerns of patient identification are not posted publicly. The data may be shared on reasonable request to the corresponding author but note such requests may need separate approval by the CSMC IRB.

REFERENCES

1. Wolf PA, Abbott RD, Kannel WB. Atrial fibrillation as an independent risk factor for stroke: the Framingham Study. *Stroke* 1991; 22 (8): 983–8.

2. January CT, Wann LS, Alpert JS, *et al.* 2014 AHA/ACC/HRS guideline for the management of patients with atrial fibrillation: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines and the Heart Rhythm Society. *J Am Coll Cardiol* 2014; 64 (21): e1–76.
3. Xian Y, O'Brien EC, Liang L, *et al.* Association of preceding antithrombotic treatment with acute ischemic stroke severity and in-hospital outcomes among patients with atrial fibrillation. *JAMA* 2017; 317 (10): 1057–67.
4. Oldgren J, Healey JS, Ezekowitz M, *et al.* Variations in cause and management of atrial fibrillation in a prospective registry of 15,400 emergency department patients in 46 countries: the RE-LY Atrial Fibrillation Registry. *Circulation* 2014; 129 (15): 1568–76.
5. Marzec LN, Wang J, Shah ND, *et al.* Influence of direct oral anticoagulants on rates of oral anticoagulation for atrial fibrillation. *J Am Coll Cardiol* 2017; 69 (20): 2475–84.
6. Perez MV, Mahaffey KW, Hedlin H, *et al.*, Apple Heart Study Investigators. Large-scale assessment of a smartwatch to identify atrial fibrillation. *N Engl J Med* 2019; 381 (20): 1909–17.
7. Soto JT, Ashley E. DeepBeat: a multi-task deep learning approach to assess signal quality and arrhythmia detection in wearable devices. *arXiv preprint arXiv:200100155* 2020.
8. Pevnick JM, Fuller G, Duncan R, Spiegel BM. A large-scale initiative inviting patients to share personal fitness tracker data with their providers: initial results. *PLoS One* 2016; 11 (11): e0165908.
9. Pevnick JM, Elad Y, Masson LM, Riggs RV, Duncan RG. Patient-initiated data: our experience with enabling patients to initiate incorporation of heart rate data into the electronic health record. *Appl Clin Inform* 2020; 11 (4): 671–9.
10. Steinberg BA, Greiner MA, Hammill BG, *et al.* Contraindications to anticoagulation therapy and eligibility for novel anticoagulants in older patients with atrial fibrillation. *Cardiovasc Ther* 2015; 33 (4): 177–83.
11. Webster-Clark M, Huang TY, Hou L, Toh S. Translating claims-based CHA(2) DS(2) -VaSc and HAS-BLED to ICD-10-CM: impacts of mapping strategies. *Pharmacoepidemiol Drug Saf* 2020; 29 (4): 409–18.
12. Apple. *Using* Apple *Watch for Arrhythmia Detection*. Cupertino, CA: Apple, Inc.; December 2020.
13. Alonso A, Krijthe BP, Aspelund T, *et al.* Simple risk model predicts incidence of atrial fibrillation in a racially and geographically diverse population: the CHARGE-AF consortium. *J Am Heart Assoc* 2013; 2 (2): e000102.
14. Giebel GD. Use of mHealth devices to screen for atrial fibrillation: cost-effectiveness analysis. *JMIR Mhealth Uhealth* 2020; 8 (10): e20496.
15. 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 (2): 146–55.
16. Ford K, Rosinger K, Zhu Q. What do we know about "race unknown"? *Educ Res* 2020; 49 (5): 376–81.
17. Ding WY, Rivera-Caravaca JM, Marín F, Li G, Roldán V, Lip GYH. Number needed to treat for net effect of anticoagulation in atrial fibrillation: real-world vs. clinical-trial evidence. *Br J Clin Pharmacol* 2022; 88 (1): 282–9.
18. Wyatt KD, Poole LR, Mullan AF, Kopecky SL, Heaton HA. Clinical evaluation and diagnostic yield following evaluation of abnormal pulse detected using Apple Watch. *J Am Med Inform Assoc* 2020; 27 (9): 1359–63.
19. Wang L, Nielsen K, Goldberg J, *et al.* Association of wearable device use with pulse rate and health care use in adults with atrial fibrillation. *JAMA Netw Open* 2021; 4 (5): e215821.
20. Anthony DL, Campos-Castillo C, Lim PS. Who isn't using patient portals and why? Evidence and implications from a national sample of US adults. *Health Aff* 2018; 37 (12): 1948–54.
21. Grossman LV, Masterson Creber RM, Benda NC, Wright D, Vawdrey DK, Ancker JS. Interventions to increase patient portal use in vulnerable populations: a systematic review. *J Am Med Inform Assoc* 2019; 26 (8–9): 855–70.
22. Goel MS, Brown TL, Williams A, Hasnain-Wynia R, Thompson JA, Baker DW. Disparities in enrollment and use of an electronic patient portal. *J Gen Intern Med* 2011; 26 (10): 1112–6.
23. Shulman E, Kargoli F, Aagaard P, *et al.* Validation of the Framingham Heart Study and CHARGE-AF risk scores for atrial fibrillation in Hispanics, African-Americans, and non-Hispanic Whites. *Am J Cardiol* 2016; 117 (1): 76–83.