

Received 9 July 2019; revised 18 September 2019; accepted 22 September 2019.
Date of publication 13 December 2019; date of current version 25 February 2020.

Digital Object Identifier 10.1109/JTEHM.2019.2950397

Accuracy of the Apple Watch 4 to Measure Heart Rate in Patients With Atrial Fibrillation

DHRUV R. SESHADRI¹, BARB BITTEL², DALTON BROWSKY², PENNY HOUGHTALING²,
COLIN K. DRUMMOND¹, MILIND DESAI², AND A. MARC GILLINOV²

¹Department of Biomedical Engineering, Case Western Reserve University, Cleveland, OH 44106, USA

²Heart and Vascular Institute, Cleveland Clinic, Cleveland, OH 44106, USA

CORRESPONDING AUTHOR: A. MARC GILLINOV (gillinom@ccf.org)

This work was supported by the Mary Elizabeth Holdsworth Fund at the Cleveland Clinic.

ABSTRACT **Background** Wearable wrist-monitors offer an unobtrusive way to acquire heart rate data in an efficient manner. Previous work in this field has focused on studying healthy subjects during exercise but has yet to assess the efficacy of these devices in patients suffering from common cardiac arrhythmias such as atrial fibrillation. **Objective** The objective of this pilot study was to assess the accuracy of the Apple Watch heart rate monitor in fifty patients experiencing atrial fibrillation compared to telemetry. **Results** Results from this pilot clinical study demonstrated a correlation coefficient of 0.7 between all readings on the Apple Watch and telemetry. Furthermore, the Apple Watch assessed heart rate more accurately in patients who were in atrial fibrillation than in those that were not ($r_c = 0.86$, patients in AF, vs. $r_c = 0.64$, patients not in AF). **Clinical Impact** The presented data from this pilot study suggests that caution should be noted before using the Apple Watch 4 wearable wrist monitor to monitor heart rate in patients with cardiac arrhythmias such as atrial fibrillation.

INDEX TERMS Wearable sensors, Apple Watch, heart rate, cardiac arrhythmias, atrial fibrillation, accuracy, clinical trial.

I. INTRODUCTION

Atrial fibrillation and atrial flutter (AF, collectively) are the most common types of cardiac arrhythmia and affect over five million people in the United States [1]. Recent studies have suggested that ~700,000 people in the United States may have previously undiagnosed AF, thus, highlighting the need for the development and validation of accurate, non-invasive technologies to enable patients to track when they might be experiencing AF [2], [3]. The technological maturation and validity of wearable devices has recently been questioned by the medical community due to the lack of clinical data assessing their accuracy and efficacy in the identification and monitoring of human diseases [4]. Authors of this study and others have assessed the accuracy of wrist-worn monitors to measure heart rate in healthy subjects [5]–[8]; however, there is limited research in such devices to accurately monitor heart rate in individuals with cardiac arrhythmias such as AF [9]–[11]. The recent media attention surrounding the diagnosis of AF with the FDA-cleared ECG sensor on the Apple Watch (AW4) raised an important question: how reliable and accurate is this specific device in measuring heart rate in patients with AF? Building off of this hype, this study

assessed the accuracy of the AW4 HR monitor function in patients with AF to determine whether or not the AW4 should be used in the clinic.

II. STUDY DESIGN AND HUMAN SUBJECT PROTOCOL

A. NUMBER OF SUBJECTS AND RECRUITMENT

Fifty post-operative cardiac surgery patients admitted to the cardiac telemetry step-down floors at the Cleveland Clinic Main Campus from January 2019 to March 2019 were recruited by trained research personnel within five days post-surgery. All patients gave written informed consent to participate in the study. The protocol was approved by the Cleveland Clinic institutional review board (IRB) and registered at clinicaltrials.gov (NCT03798613). The study lasted from December 21st, 2018 till March 29th, 2019. The mean (SD) age of the patients was 61.4 ± 10.4 years; the mean \pm st. dev body mass index (calculated as weight in kilograms divided by height in meters squared) was 30.4 ± 4.87 ; 14 participants were women (28%), and 3 participants were African American (6%). All patients were on continuous six-lead telemetry according to standard clinical practice on the step-down floor. 50% of subjects experiencing atrial

fibrillation and 50% experiencing sinus rhythm at the time of enrollment were studied.

B. INCLUSION AND EXCLUSION CRITERIA

Inclusion criteria included patients above the age of 18 years and post-operative cardiac surgery patients on the cardiac telemetry floors. Exclusion criteria included patients with a cardiac pacemaker, use of a radial artery graft for coronary artery bypass grafting (CABG), known chronic or persistent heart rhythm disorders, and tattoos located on the skin of the wrist or forearm where the AW was placed.

C. CONFIDENTIALITY AND SAFETY PROTOCOLS

Data was inputted and analyzed from REDCAP by trained research personnel who were a part of the IRB protocol. The authors take responsibility for the data presentation and analysis in this study.

III. METHODS AND PROCEDURES

A. DATA COLLECTION

Five AW's and five iPhone 8's was purchased. The right and left wrist circumference of each patient was measured prior to placing the AW. Each patient was outfitted with an AW for no more than five minutes and the location (left or right wrist) was randomly assigned via a randomization grid. All devices were thoroughly and antiseptically cleaned by the research staff according to hospital guidelines before and after each patient's testing. Each enrolled patient had a minimum of three assessments of heart rate per day for at least two days, generating a minimum of six data points per patient.

B. DATA ANALYSIS

All statistical analyses were performed using R and SAS statistical software (SAS v9.4; SAS, Inc., Cary, NC). Graphics were constructed using SAS.

IV. RESULTS AND CLINICAL OUTCOME ANALYSIS

Across the five watches and iPhone 8's purchased, 266 heart rate values were recorded. Twenty-four data points were not obtained. Missing data can be attributed to the following reasons. For one subject, none of the five AW's was able to read an ECG rhythm at all six time points; the same watches were fine for the other subjects. In two other subjects, there was missing data due to the inability to obtain both AW and telemetry readings at certain time points due to early patient discharge. Lastly, one subject withdrew consent from study participation.

Heart rate ranges for the AW and telemetry ranged from 57 – 125 bpm (mean 88 ± 14) and 54 – 137 (mean 86 ± 15) respectively. Lin's concordance correlation coefficients (r_c) were calculated to provide a measure of agreement for the AW with telemetry. The concordance correlation coefficient (CCC) measures the degree to which the paired observations fall on the identity line. Overall there was an agreement of $r_c = 0.7$ between the AW and telemetry (0.63, 0.75, 95% confidence interval), (Figure 1).

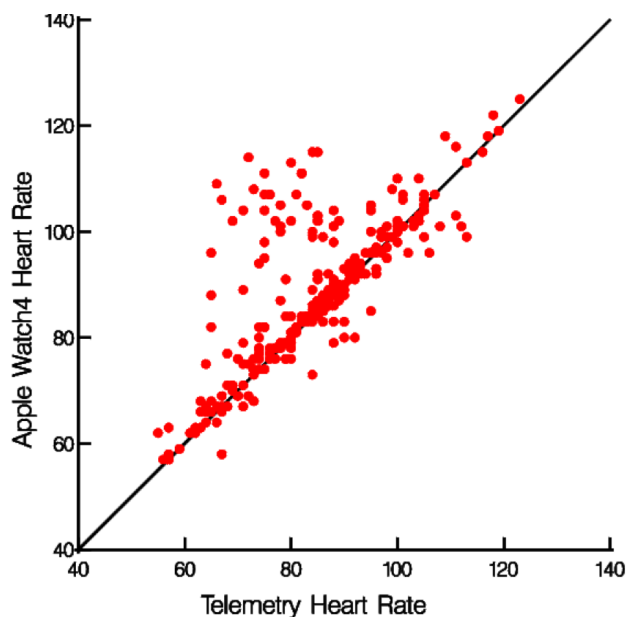


FIGURE 1. Heart rate measurements from telemetry and the Apple Watch 4.

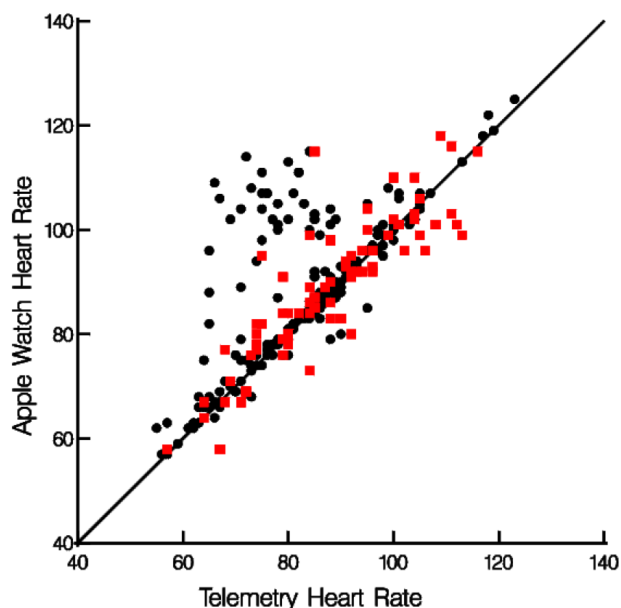


FIGURE 2. Heart rate measurements from telemetry and the Apple Watch 4 stratified by atrial fibrillation on telemetry (red = Afib, black = no Afib).

The AW assessed HR more accurately in patients who were in AF than in those that were not ($r_c = 0.86$, patients in AF, vs. $r_c = 0.64$, patient not in AF) (Figures 2 and 3).

Paired differences were calculated by subtracting the measured heart rate (HR) from the heart rate recorded on the Telemetry under each condition and at each time point. Summary data are provided for relative differences, absolute differences, and percent differences ($(HR_{tel} - HR_{watch})/HR_{tel}$). Bland-Altman analysis was performed and data was plotted to assess agreement against the mean values (Figure 4). The motivation to plot using this method was to uncover any tendency for the variation to change with the

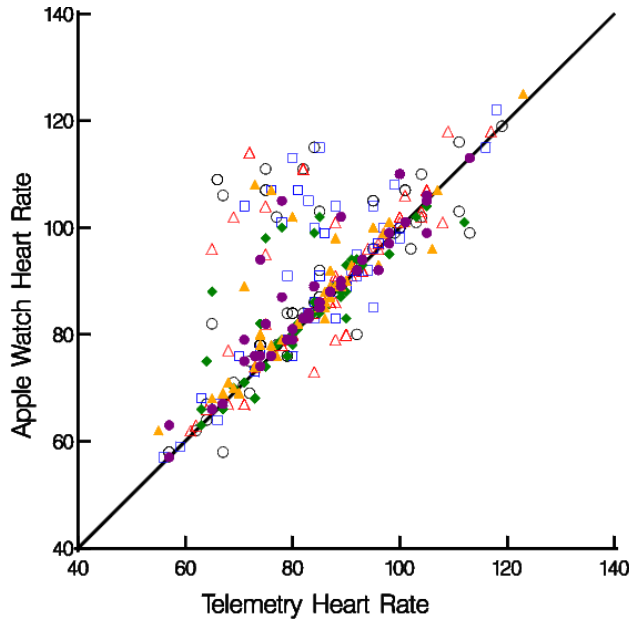


FIGURE 3. Heart rate measurements from Telemetry and Apple Watch 4 stratified by Time point (Day 1 measures 1-3 (open shapes), Day 2 measures 4-6 (filled shapes)).

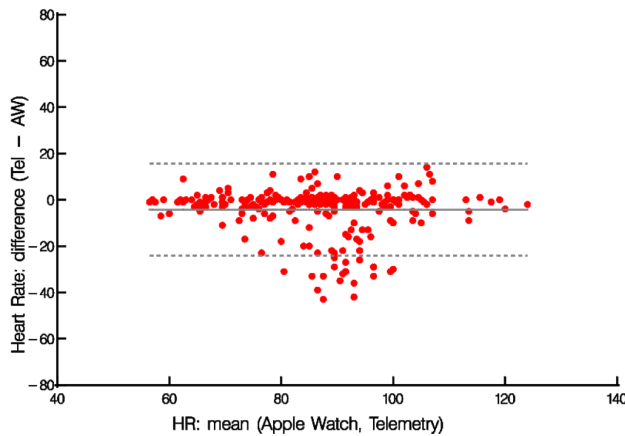


FIGURE 4. Heart rate measurements from Telemetry and Apple Watch 4 stratified by Time point (Day 1 measures 1-3 (open shapes), Day 2 measures 4-6 (filled shapes)).

magnitude of the measurement. Results suggest quite a bit of disagreement between the mean values thereby highlighting potential inaccuracies with the AW4 HR sensor compared to measurements from telemetry.

Furthermore, the study investigated whether the subjects with lower HR correlations might have had other arrhythmias, or might have had other sources of noise in the ECG traces (Table 1).

The overall goal of this pilot study was to assess the accuracy of the AW4 in measuring heart rate compared to telemetry in patients in the cardiac step-down unit following surgery. Our pilot study currently represents the first such clinical study evaluating the efficacy of this particular device to measure heart rate in a controlled clinical environment. The r_c value of 0.7 was lower than those determined from prior

TABLE 1. Association between heart rate correlations and arrhythmias as adjudicated by a cardiologist.

HR Difference Range	Rhythm from Telemetry	Rhythm from AW4
0-10	Ventricular Ectopy	Ventricular Ectopy
	None	Ventricular Ectopy
	Atrial Ectopy	Atrial Ectopy
	Noise	Ventricular Ectopy
11-20	Atrial Ectopy	Sinus
	None	Ventricular Ectopy
	Ventricular Ectopy	Ventricular Ectopy
	Sinus	Ventricular Ectopy
21-30	Sinus	Atrial Ectopy
31-40	Ventricular Ectopy	Noise

studies by members of this team. Wang et al. evaluated the efficacy of the AW 3 (and other wrist-monitors) compared to the Polar H7 monitor in 50 healthy subjects and found that the AWs had a r_c value of 0.91 [5]. Etiwy et al. assessed the accuracy of the AW 3 (and other heart rate monitors) compared to standard ECG limb leads and the Polar H7 ECG chest strap monitor in 80 patients with established cardiovascular disease enrolled in phase II or III cardiac rehabilitation (CR) [8]. The study found that overall the agreement between the AW and ECG was ($r_c = 0.8$). Our results coupled with those published by members of our team suggest the need for subsequent validation studies in the clinic to further elucidate such differences which could have enormous ramifications if heart rate readings from these devices are used to inform clinical decisions in patients with cardiac arrhythmias.

The difference in heart rate among patients with AF and sinus or other rhythms presented the following question: did the patient know they were in AF and therefore may have been more precise at getting the watch reading? A key learning from our pilot study suggested that this was one of the key limitations of the AW4. Since the patient had to manually press the button on the side of the crown to obtain a reading/tracing, there needed to be some awareness on their part that they might be in AF, unless they were in persistent/continuous AF. If the patient had no symptoms then there would not have been an incentive to obtain a reading from the AW4. The AW4 would be better served if it could automatically detect a change in the patient's rhythm and automatically obtain a reading. However, this is not the case with the current model of the AW4. In addition to the noted drawback of the device, the study noted several limitations which could have affected the accuracy of the results or long-term use-case of this technology for use in the clinic such as: 1) fluid overload and edema following surgery and 2) interference with other medical equipment on or near the patient.

V. CONCLUSION

This pilot study represents the first case in evaluating the accuracy of the AW4 to measure heart rate compared to

telemetry in a controlled clinical environment. Our results coupled with those previously published by members of this team suggest the need for further validation before wrist-worn heart rate monitors are used in the clinic to inform decision-making protocols.

ACKNOWLEDGMENT

The authors acknowledge collaboration between the Department of Biomedical Engineering, Case Western Reserve University, and the Heart and Vascular Institute, Cleveland Clinic. The Mary Elizabeth Holdsworth Fund had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication. There are no relevant conflicts of interest to disclose.

REFERENCES

- [1] E. J. Benjamin *et al.*, "Heart disease and stroke statistics—2018 update: A report from the American Heart Association," *Circulation*, vol. 137, no. 12, pp. e67–e492, Mar. 2018.
- [2] M. P. Turakhia *et al.*, "Estimated prevalence of undiagnosed atrial fibrillation in the United States," *PLoS ONE*, vol. 13, no. 4, Apr. 2018, Art. no. e0195088.
- [3] M. P. Turakhia *et al.*, "Economic burden of undiagnosed nonvalvular atrial fibrillation in the United States," *Amer. J. Cardiol.*, vol. 116, no. 5, pp. 733–739, Sep. 2015.
- [4] E. S. Izmailova, J. A. Wagner, and E. D. Perakslis, "Wearable devices in clinical trials: Hype and hypothesis," *Clin. Pharmacol. Therapeutics*, vol. 104, no. 1, pp. 42–52, Jul. 2018.
- [5] R. Wang *et al.*, "Accuracy of wrist-worn heart rate monitors," *JAMA Cardiol.*, vol. 2, no. 1, pp. 104–106, Jan. 2017.
- [6] S. Gillinov *et al.*, "Variable accuracy of wearable heart rate monitors during aerobic exercise," *Med. Sci. Sports Exerc.*, vol. 49, no. 8, pp. 1697–1703, Aug. 2017.
- [7] A. M. Gillinov *et al.*, "Variable accuracy of commercially available wearable heart rate monitors," *J. Amer. College Cardiol.*, vol. 69, no. 11, p. 336, Nov. 2017.
- [8] Etiwy, Cardiovascular Diagnosis and Therapy. *Accuracy of Wearable Heart Rate Monitors in Cardiac Rehabilitation*. Accessed: Jul. 3, 2019. [Online]. Available: <http://cdt.amegroups.com/article/view/25572>
- [9] A. N. Koshy *et al.*, "Smart watches for heart rate assessment in atrial arrhythmias," *Int. J. Cardiol.*, vol. 266, pp. 124–127, Sep. 2018.
- [10] G. D. Giebel and C. Gissel, "Accuracy of mHealth devices for atrial fibrillation screening: Systematic review," *JMIR MHealth UHealth*, vol. 7, no. 6, 2019, Art. no. e13641.
- [11] B. A. Steinberg and J. P. Piccini, "Screening for atrial fibrillation with a wearable device," *JAMA*, vol. 320, no. 2, pp. 139–141, Jul. 2018.

• • •