



Uso del Apple Watch para la detección de arritmias cardiacas

Diciembre de 2020

Índice

Descripción general	3
Introducción	3
Detección de arritmias basada en fotopletismografía	3
Descripción técnica de la función	3
Desarrollo preclínico	5
Validación clínica.....	5
Estudio Apple Heart.....	5
Diseño del subestudio del estudio Apple Heart.....	5
Resultados	6
Detección basada en el ECG	6
Descripción técnica de la función	6
Desarrollo preclínico.....	7
Validación clínica.....	7
Diseño del experimento	7
Estudio de validación clínica de la app ECG 1.0	8
Resultados	8
Estudio de validación clínica de la app ECG 2.0	12
Resultados	14
Conclusiones	17

Descripción general

Los clientes de Apple Watch tienen acceso a dos funciones de software como dispositivo médico para detectar arritmias cardíacas como la fibrilación auricular (FA): la función de notificación de ritmo irregular (FNRI) y la app ECG.

En Apple Watch Series 1 o modelos posteriores, la FNRI analiza los datos recogidos por el sensor del fotopletismógrafo (FPG) para identificar episodios de ritmo cardíaco irregular indicativos de FA y notifica al usuario si detecta un episodio. En Apple Watch Series 4, Series 5 y Series 6, la app ECG es capaz de generar un electrocardiograma (ECG) similar a un ECG de derivación I y a continuación ofrecer una clasificación de dicho electrocardiograma como ritmo sinusal (RS), fibrilación auricular (FA), frecuencia cardíaca alta o baja o no concluyentes; con ECG 2.0 se puede acceder a otras clasificaciones de la FA con frecuencia cardíaca alta o una mala grabación.

La intención de este documento es explicar más detalladamente lo que permiten hacer estas funciones, incluidas las de prueba y validación.

Introducción

La FA, un tipo de arritmia cardíaca que se caracteriza por latidos auriculares irregulares y, a veces, rápidos, es una de las principales causas de accidente cerebrovascular. No obstante, la FA es a menudo asintomática, por lo que muchas personas la padecen y no son conscientes de ello. La combinación del riesgo de accidente cerebrovascular, el cuadro clínico inicial asintomático, los tratamientos farmacológicos eficaces que reducen el riesgo de accidente cerebrovascular al mínimo y el aumento de la penetración en el mercado de dispositivos para el cliente con el potencial de detectar la fibrilación auricular ha aumentado el interés por el diagnóstico temprano de la FA fuera del ámbito clínico.

Con watchOS 5.1.2 o versiones posteriores, los Apple Watch Series 1 y modelos posteriores pueden utilizar las señales de fotopletismografía junto con un algoritmo para identificar períodos de pulso irregular indicativos de FA. Además de este algoritmo de identificación basado en la fotopletismografía, Apple Watch Series 4, Series 5 y Series 6 cuentan con un sensor eléctrico de frecuencia cardíaca que, al utilizar la app ECG, permite generar y analizar un electrocardiograma similar a un ECG de derivación I.

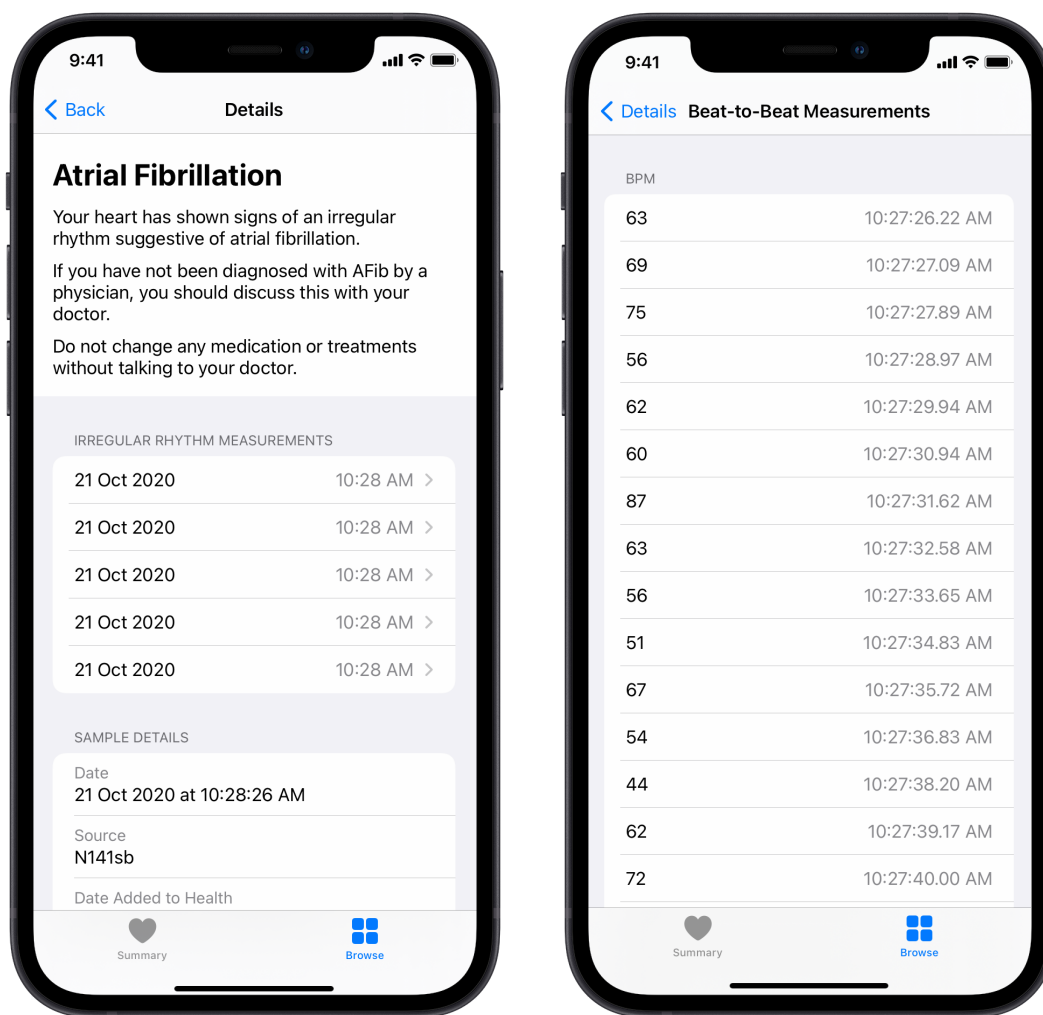
Detección de arritmias basada en fotopletismografía

Descripción técnica de la función

El Apple Watch tiene un sensor óptico de frecuencia cardíaca que usa luces LED verdes con fotodiodos fotosensibles para detectar el volumen de sangre que pasa por la muñeca del usuario mediante fotopletismografía. Estos sensores y los algoritmos subyacentes son la base de la detección de la frecuencia cardíaca (FC) y su variabilidad (HRV) habilitada en Apple Watch Series 1 y modelos posteriores. Para determinar la HRV, el Apple Watch captura un tacograma o gráfico del tiempo transcurrido entre latido y latido cada 2 o 4 horas. A partir de la versión watchOS 5.1.2, el usuario puede activar una función de detección de arritmias que utiliza estos tacogramas. Para utilizar la función de notificación de ritmo irregular en el Apple Watch, el usuario primero tendrá que completar la introducción de la app Salud en su iPhone enlazado para aprender a usarla e informarse sobre la FA. Para obtener más información sobre la experiencia de usuario, visite <https://support.apple.com/es-es/HT208931>.

Si está activada la detección de arritmias basada en fotopletismografía, cada tacograma se clasificará con un algoritmo patentado para determinar si existe un ritmo irregular. En cuanto el tacograma se clasifica como irregular, se inicia una serie de tacogramas recopilados y analizados con más frecuencia (la máxima posible, con un intervalo mínimo entre ellos de 15 minutos). Apple Watch recopila y analiza los tacogramas solo si el usuario permanece en reposo el tiempo suficiente para obtener una lectura. Por este motivo, el algoritmo no monitoriza de forma constante al usuario, sino solo cuando detecta una señal aceptable para la recopilación y el análisis. Si cinco de cada seis tacogramas secuenciales (incluido el inicial) se clasifican como irregulares durante un período de 48 horas, se notifica al usuario la posible arritmia. Además de recibir esta notificación, el usuario puede acceder a más información relacionada con los tacogramas irregulares en la app Salud (figura 1). Si dos tacogramas se clasifican como no irregulares antes de alcanzar este umbral, el ciclo se reinicia y la frecuencia de los tacogramas vuelve a ser la normal (cada dos horas).

Figura 1: Vista de la app Salud de las mediciones de ritmo irregular



En la app Salud, el usuario puede ver los momentos en los que el algoritmo ha detectado las irregularidades en los tacogramas que han desencadenado la notificación (izquierda). Si selecciona cualquiera de estas fechas y horas, podrá ver las mediciones de sus pulsaciones en ese momento calculadas en cada tacograma.

Desarrollo preclínico

Antes de los ensayos clínicos, se realizaron estudios para desarrollar el algoritmo de detección basado en fotopleletismografía y evaluar su rendimiento en diferentes afecciones y conductas del usuario; entre ellas, la respiración honda, conducir, los temblores y el movimiento de las manos, la reducción del riego sanguíneo en manos o muñecas, llevar el reloj durante la noche, la respuesta ventricular rápida en personas con FA y otras arritmias. Estos estudios se llevaron a cabo en 2300 personas sanas y más de 500 pacientes con FA.

Puesto que la fotopleletismografía se basa en la absorción de la luz, el algoritmo de detección de arritmias se probó en personas con distintos tipos y tonos de piel para garantizar que los ajustes de la plataforma del sensor correspondientes al tono de piel fueran los apropiados para detectar arritmias. La melanina presenta una gran absorción en la longitud de onda que utiliza el LED de color verde del Apple Watch, lo que puede dificultar la medición de la frecuencia cardíaca mediante fotopleletismografía en los tonos de piel más oscuros. Por este motivo, la plataforma de sensores del Apple Watch ajusta la corriente de LED (y, por tanto, la luz producida), la ganancia de los fotodiodos (la sensibilidad a la luz) y la frecuencia de muestreo para garantizar una amplitud de señal satisfactoria para todos los tonos de piel del ser humano.

A efectos de validación, se analizaron 1 300 000 tacogramas de 1124 personas (el 51 %, mujeres) con distintos tipos y tonos de piel (el fototipo según la escala de Fitzpatrick y el color según espectrofotometría de la muñeca). Puesto que lo que más preocupaba a los ingenieros era la amplitud de la señal en las personas de piel oscura, casi el 5 % de los voluntarios reclutados tenían el fototipo VI de la escala de Fitzpatrick. La validación demostró que la sensibilidad o especificidad del algoritmo no presenta una diferencia significativa en los distintos fototipos o tonos de piel.

Validación clínica

Estudio Apple Heart

El estudio Apple Heart es un estudio pragmático prospectivo de un solo grupo que se llevó a cabo a distancia con el fin de evaluar la capacidad del algoritmo de detección de pulsaciones irregulares del Apple Watch para detectar arritmias indicativas de fibrilación auricular. En este estudio, si un usuario alcanzaba el umbral de cinco tacogramas irregulares de seis, recibía una notificación en el iPhone y el Apple Watch y tenía la opción de hablar con un médico del estudio a distancia y recibir un parche de ECG ambulatorio (ePatch, BioTelemetry, Inc., Conshohocken, Pensilvania, EE. UU.). Se indicó a los participantes que llevaran puesto el ePatch durante 7 días; no obstante, se consideraban suficientes los datos recopilados de un participante con una duración mínima analizable de 1 hora.

Los resultados detallados del estudio Apple Health se publicaron en noviembre de 2019 en el *New England Journal of Medicine* (Pérez, Marco V., et al. "Large-scale assessment of a smartwatch to identify atrial fibrillation." *New England Journal of Medicine* 381.20 (2019): 1909–1917).

Diseño del subestudio del estudio Apple Heart

Se llevó a cabo un subestudio a partir de los datos recopilados en el estudio Apple Heart para determinar si el algoritmo de clasificación de tacogramas (individuales o puntuales) y el algoritmo de ciclos de confirmación (que avisa si cinco de seis tacogramas son irregulares) tienen un valor predictivo positivo aceptable en comparación con la monitorización del ePatch al identificar ritmos irregulares indicativos de FA en un subgrupo de participantes del estudio Apple Heart. Los investigadores de este estudio eran conocedores del subestudio, los posteriores análisis y el envío de los datos a la FDA (Administración de Alimentos y Medicamentos estadounidense), pero no tuvieron acceso a los resultados del subestudio mientras realizaban el estudio Apple Heart. El comité de revisión institucional (CRI) que autorizó el estudio Apple Heart determinó que este subestudio estaba exento de su supervisión. Todos los participantes del estudio Apple Heart dieron su consentimiento informado, que incluía el uso de los datos del estudio para el subestudio.

Estos datos se recopilaron de los participantes en el estudio Apple Heart que se inscribieron como voluntarios entre el 30 de noviembre de 2017 y el 22 de junio de 2018. Los voluntarios de este subestudio recibieron una notificación de ritmo irregular en la app del estudio Apple Heart y, a continuación, se les envió un parche de ECG ambulatorio (ePatch) que llevaron para que se pudieran interpretar los resultados. En este subestudio no se analizaron los tacogramas irregulares iniciales que llevaron a la primera notificación y, en teoría, el primer encuentro de teleasistencia médica, sino solamente los tacogramas irregulares y las notificaciones que se produjeron a partir de que los usuarios llevaran puesto el parche ePatch proporcionado para tales fines.

Dos validadores de ECG independientes con certificación oficial en cardiología o electrofisiología supervisaron y juzgaron los resultados de los electrocardiogramas, clasificándolos como ritmo sinusal (RS), fibrilación auricular (FA), otro ritmo irregular o ilegibles. Si no había consenso entre los validadores, los evaluaba un tercer validador con la misma cualificación. Estos validadores no tenían acceso a la clasificación de los tacogramas. La clasificación de los ECG realizada por los validadores y la clasificación de los tacogramas determinados por los algoritmos se enviaron de forma segura al estadístico del estudio para su análisis.

Resultados

De los 226 voluntarios del subestudio que recibieron una notificación inicial de arritmia y llevaron el ePatch durante aproximadamente una semana, al 41,6 % (94 de 226) se les detectó FA con el ePatch. Durante el tiempo que llevaron a la vez el Apple Watch y el ePatch, 57 de los 226 voluntarios recibieron una notificación de FA (es decir, se clasificaron como irregulares cinco de sus seis tacogramas consecutivos). De ellos, el 78,9 % (45 de 57) presentó FA también en los datos del ePatch y el 98,2 % (56 de 57) presentó fibrilación auricular u otras arritmias de importancia clínica. Estos resultados demuestran que, aunque en la mayoría de casos la notificación indicará correctamente la presencia de FA, habrá algunos en los que indicará la presencia de otro tipo de arritmia. No se observaron efectos adversos graves causados por el dispositivo.

Detección basada en el ECG

Descripción técnica de la función

Los Apple Watch Series 4, Series 5 y Series 6 incorporan un electrodo de titanio en la corona Digital Crown y una capa ultrafina de nitruro de carbono, silicio y cromo en el cristal de zafiro de la cubierta trasera. La app ECG lee y graba los impulsos eléctricos que controlan el corazón desde la punta del dedo (con la corona Digital Crown) y la muñeca (cubierta trasera del Apple Watch) del usuario, lo que crea un circuito cerrado. Para utilizar la app ECG en el Apple Watch, el usuario primero tendrá que completar la introducción de la app Salud en su iPhone enlazado para aprender a usarla e informarse sobre la fibrilación auricular. Para generar un ECG, el usuario debe abrir la app ECG instalada en el Apple Watch y poner en la Digital Crown un dedo de la mano contraria a la muñeca donde lleva el Apple Watch durante 30 segundos. La polaridad de la derivación la determina la muñeca elegida en los ajustes del Apple Watch.

Tras obtener el ECG, un algoritmo exclusivo clasifica el trazado electrocardiográfico como ritmo sinusal (RS), fibrilación auricular (FA) o no concluyente en ECG 1.0. Con ECG 2.0, en caso de estar disponible, también se accede a otras clasificaciones, como la FA con la frecuencia cardíaca alta (FC 100-150), y a una diferenciación entre una mala grabación y unos valores no concluyentes. Estas clasificaciones del ritmo (la FC media, los síntomas notificados por el usuario y la onda) se añaden a la app Salud y se almacenan en ella; el usuario puede compartirlas en forma de PDF desde la aplicación de su iPhone enlazado. Para obtener más información sobre la experiencia del usuario, visite <https://support.apple.com/es-es/HT208955>.

Desarrollo preclínico

La detección de la señal de ECG y el algoritmo de clasificación también se probaron en varios estudios antes de comenzar la validación clínica. Los sensores y el algoritmo de clasificación se probaron en personas diversas en cuanto a etnia, circunferencia de muñeca, índice de masa corporal, edad, arritmias diferentes a la FA, grado de ajuste de la correa, postura y estado/sudoración tras el ejercicio. En estas pruebas participaron aproximadamente 2500 voluntarios, el 25 % de los cuales tenía un diagnóstico previo de FA u otras arritmias cardíacas.

La principal variación en el rendimiento del algoritmo fue una mayor frecuencia de ECG ilegibles. Los factores de esta variación fueron una amplitud de señal baja (derivada de la desviación del eje derecho, detectada especialmente en aquellas personas con bajo IMC o con sudor tras sesiones de ejercicio) y los artefactos de movimiento como consecuencia de la conducta del usuario. El Apple Watch usa electrodos secos concebidos para ser mecánicamente resistentes y aguantar la corrosión, como cualquier otro dispositivo «wearable». No obstante, los electrodos secos, y especialmente los que se colocan en las extremidades, son inherentemente más propensos a introducir ruido como el descrito más arriba que los electrodos de gel temporales que se usan en los dispositivos clínicos.

Además de los factores arriba mencionados, la presencia de determinadas arritmias distintas a la FA también derivó en un rendimiento del algoritmo significativamente diferente respecto al que presentaba en los voluntarios con RS. Estas arritmias y los resultados se desglosan en la tabla 1 a continuación.

Tabla 1. Arritmias distintas de la FA que afectan al rendimiento del algoritmo (ECG 2.0)

Arritmia	Variación
Bloqueo de rama izquierda o derecha	El 7,9 % de las pruebas se clasificaron como FA
Bloqueo AV de primer grado	El 10,2 % de las pruebas no se clasificaron; de las clasificadas, el 2,4 % se clasificaron como FA
Bigeminismo	El 92,5 % no se clasificaron
Extrasístoles ventriculares frecuentes	El 48,8 % no se clasificaron; de las clasificadas, el 24,1 % se clasificaron como FA
Extrasístoles auriculares frecuentes	El 19,8 % no se clasificaron; de las clasificadas, el 23,5 % se clasificaron como FA
Taquicardia auricular	El 7,7 % se clasificaron como FA
Ritmo nodular	El 46,5 % no se clasificaron; de las clasificadas, el 29,7 % se clasificaron como FA
FC alta o baja (fuera del intervalo de 50-150 LPM)	El 94,8 % no se clasificaron

Validación clínica

Diseño del experimento

Apple financió dos estudios multicéntricos para validar la capacidad de la app ECG para (1) generar una onda similar a la de un ECG de derivación I a partir de un ECG estándar de 12 derivaciones y (2) utilizar un algoritmo para clasificar la frecuencia cardíaca de este ECG de una sola derivación como RS o FA.

Estudio de validación clínica de la app ECG 1.0

Se realizó un estudio de validación clínica para validar el rendimiento de la app ECG 1.0. Los criterios de valoración principales del estudio fueron la sensibilidad del algoritmo de clasificación del ritmo para detectar la fibrilación auricular y su grado de especificidad para detectar el ritmo sinusal. Un comité de revisión institucional externo aprobó el protocolo, el formulario de consentimiento informado y todos los materiales pertinentes antes de la inscripción de los voluntarios. Todos los voluntarios proporcionaron su consentimiento por escrito para participar en el estudio antes de inscribirse.

Se inscribieron personas con FA diagnosticada y voluntarios sin anomalías conocidas en la frecuencia cardiaca. A todos se les solicitó que se grabaran tres ECG de una sola derivación con la app ECG mientras el personal del estudio hacía al mismo tiempo tres ECG de 12 derivaciones con un dispositivo clínico aprobado por la FDA (GE Healthcare CardioSoft). Se tuvieron en cuenta las primeras pruebas para los fines de validación y análisis. Se ayudó a los participantes a ponerse correctamente el Apple Watch y se les pidió que no movieran los brazos, pudiendo apoyarlos en una mesa o en las piernas. También se les permitió hacer un electrocardiograma de prueba antes de comenzar.

Para probar (1), tres técnicos independientes certificados solaparon las tiras de resultados de 140 voluntarios elegidos al azar (70 con fibrilación auricular y 70 con ritmo sinusal) con las tiras correspondientes del electrocardiograma de derivación I realizado con el dispositivo clínico para comparar visualmente la morfología de seis complejos PQRST consecutivos. Los técnicos asignaron a cada tira de resultados una evaluación de aprobado/no aprobado decidida según la similitud morfológica visual. También se les pidió que midieran la amplitud de la onda R a partir de la referencia isoelectrica al milímetro más cercano de los dos primeros complejos QRS de las tiras del electrocardiograma de referencia y de la app ECG, y que evaluaran la coherencia entre ambas.

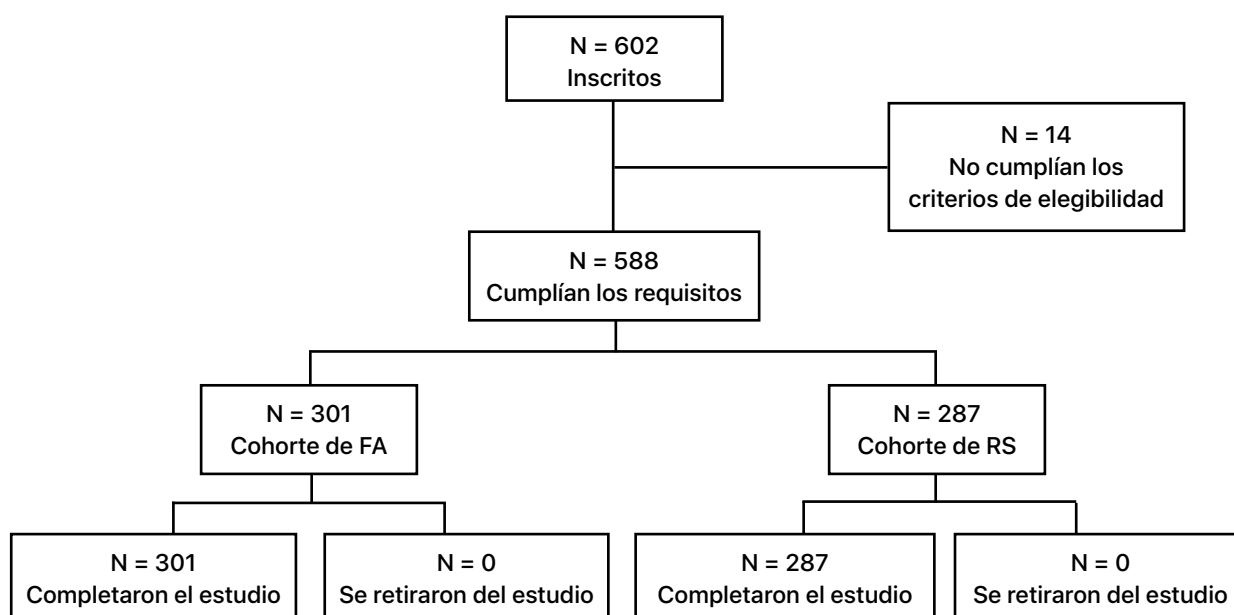
Para (2), tres cardiólogos independientes con certificación oficial y sin acceso a los demás datos revisaron cada tira de ECG de 12 derivaciones y clasificaron la frecuencia cardiaca correspondiente como ritmo sinusal, fibrilación auricular u otra arritmia (cualquier frecuencia que no fuera RS o FA dentro de los parámetros de FC), o ilegible si el resultado no se podía leer y, por tanto, no permitía el diagnóstico. El algoritmo de la app ECG clasificó los electrocardiogramas generados en esta como RS, FA, inclasificables o ilegibles. Se calculó el grado de sensibilidad y especificidad de la app ECG a la hora de clasificar el RS y la FA (en los ECG clasificables) en comparación con la interpretación de los cardiólogos de los ECG de 12 derivaciones. A continuación, se pidió a un cardiólogo independiente con certificación oficial que clasificara las tiras generadas por la app ECG en las mismas categorías.

Para los análisis del criterio de valoración principal, el límite inferior del intervalo de confianza del 97,5 % se calculó por separado para la sensibilidad y la especificidad. Si el límite inferior de la sensibilidad superaba el 90 %, la hipótesis nula se rechazaba en favor de la sensibilidad que superara el 90 %. Si el límite inferior de la especificidad superaba el 92 %, la hipótesis nula se rechazaba en favor de la especificidad que superara el 92 %.

Resultados

Se presentaron 602 sujetos para el estudio, 588 de los cuales cumplían los criterios de elegibilidad. De estos 588, se asignaron los 301 que decían padecer fibrilación auricular a la cohorte de FA, y los 287 que no decían padecerla, a la cohorte de RS. Estas asignaciones se realizaron únicamente para garantizar una inscripción satisfactoria: los evaluadores no sabían a qué cohorte pertenecían los sujetos y la presencia o ausencia de FA solo se basó en los ECG realizados en las pruebas. En cuanto a los 14 sujetos que completaron el estudio, pero no fueron asignados a ninguna cohorte, no cumplían los requisitos para participar en el estudio debido a su historial de FA paroxística sin signos de FA en el ECG en el momento de la selección. Todos los sujetos que cumplían los requisitos completaron el estudio (figura 2). No se notificaron efectos adversos durante el estudio.

Figura 2: Diagrama de clasificación de los sujetos



Tres técnicos independientes certificados detectaron equivalencia morfológica visual entre la forma de onda del ECG generado con la app ECG y la del ECG de derivación I de referencia generado con el dispositivo clínico estándar: del 98,4 % en las tiras analizadas de la cohorte de FA y del 100 % en las de la cohorte de RS (tabla 2). La proporción de voluntarios aprobados fue del 99,2 % (límite inferior del intervalo de confianza del 97,5 % = 95,7 %). Las tiras de resultados se excluían si no se podían identificar seis pulsaciones consecutivas (o complejos PQRST) sin artefactos en ningún conjunto de resultados (el de la app ECG o el de referencia).

Tabla 2. Comparación de la forma de la onda

Característica	Sujetos con FA (N = 61)	Sujetos con RS (N = 65)	Total (N = 126)	Límite inferior del intervalo de confianza*	Valor de p**
N.º de tiras comparadas (de la app ECG y de referencia) con clasificación de aprobadas	60	65	125		
N.º de tiras pareadas legibles (de la app ECG y de referencia)	61	65	126		
Proporción de tiras aprobadas	60/61 (98,4%)	65/65 (100%)	125/126 (99,2%)	95,7%	< 0,0001
N.º de tiras de resultados pareadas excluidas	8	5	13		

* Límite inferior exacto, binomial y unilateral del intervalo de confianza del 97,5 % del total.

** Prueba de la hipótesis para un éxito del sujeto > 0,8.

Síglas: FA = fibrilación auricular; RS = ritmo sinusal

Para confirmar que las ondas generadas por la app ECG y el dispositivo de referencia eran similares, la clasificación de las tiras de la app ECG del cardiólogo sin acceso a los datos se comparó con la

clasificación del cardiólogo de las tiras de referencia (tabla 3). El porcentaje de concordancia con la clasificación de las tiras del dispositivo y con los resultados de referencia de FA y RS fue del 100 % y del 99,1 %, respectivamente. Las tiras con resultados ilegibles no se incluyeron en este análisis.

Tabla 3. Clasificaciones de las tiras de la app ECG y de referencia

Característica	Total (N = 522)
Resultado final del ECG de referencia = FA	263
Clasificación de la tira de la app ECG = FA	239/263 (90,9%)
Clasificación de la tira de la app ECG = RS	0/263 (0,0%)
Clasificación de la tira de la app ECG = Otro	0/263 (0,0%)
Clasificación de la tira de la app ECG = Ilegible	24/263 (9,1%)
Porcentaje de concordancia con el resultado FA del ECG de referencia*	239/239 (100,0%)
Resultado final del ECG de referencia = RS	244
Clasificación de la tira de la app ECG = FA	0/244 (0,0%)
Clasificación de la tira de la app ECG = RS	232/244 (95,1%)
Clasificación de la tira de la app ECG = Otro	2/244 (0,8%)
Clasificación de la tira de la app ECG = Ilegible	10/244 (4,1%)
Porcentaje de concordancia con el resultado RS del ECG de referencia*	232/234 (99,1%)
Resultado final del ECG de referencia = Otro	15
Clasificación de la tira de la app ECG = FA	0/15 (0,0%)
Clasificación de la tira de la app ECG = RS	3/15 (20,0%)
Clasificación de la tira de la app ECG = Otro	12/15 (80,0%)
Clasificación de la tira de la app ECG = Ilegible	0/15 (0,0%)
Porcentaje de concordancia con el resultado Otro del ECG de referencia*	12/15 (80,0%)

* Las tiras ilegibles se excluyeron.

Siglas: FA = fibrilación auricular; RS = ritmo sinusal

En total, 485 de las 602 tiras de la app ECG y de referencia comparadas se consideraron clasificables. El resto se consideraron ilegibles o inclasificables. En la tabla 4 se desglosan las cohortes de FA y RS.

Tabla 4. Clasificación del algoritmo de la app ECG y resultado final de las tiras de referencia

Clasificación del algoritmo de la app ECG	Clasificación de las tiras de referencia				
	RS	FA	Otro	Ilegible	Total
RS	238	4	4	1	247
FA	1	236	2	2	241
Inclasificable	6	7	6	0	19
Ilegible	18	30	1	0	49
Resultado del dispositivo no facilitado*	32	13	1	0	46
Total	295	290	14	3	602

* Resultados no facilitados según los criterios preestablecidos (como la sincronización no detectada) para todos los sujetos menos uno.

Siglas: FA = fibrilación auricular; RS = ritmo sinusal

La clasificación del algoritmo de la app ECG alcanzó una sensibilidad del 98,3 % y una especificidad del 99,6 % (tabla 5). Al ampliar el análisis para incluir el 2,4 % (7 de 290) y el 2,0 % (6 de 295) de las tiras categorizadas como inclasificables por el dispositivo en las cohortes de FA y RS, la sensibilidad fue de un 95,5 % (IC del 95 %: 92,2 %, 97,8 %); y la especificidad, de un 97,1 % (IC del 95 %: 94,2 %, 98,8 %). Estos resultados cumplieron los objetivos especificados previamente en el diseño del estudio. Además, el 12,2 % (68 de 556) de las grabaciones fueron no concluyentes (por ilegibles o inclasificables) y no se clasificaron como RS ni FA. Cuando se incluyeron grabaciones no concluyentes en el análisis, la app ECG clasificó correctamente el RS en el 90,5 % (238 de 263) de los voluntarios con RS, y la FA en el 85,2 % (236 de 277) de los sujetos con FA. Los resultados de la validación clínica reflejan el uso en un entorno controlado. Al usar la app ECG en entornos reales, puede que un mayor número de tiras se consideren no concluyentes o inclasificables.

Tabla 5. Análisis de la sensibilidad y la especificidad (tiras clasificables)

Parámetro	Valor	Límite inferior del intervalo de confianza*	Valor de p**
Resultado final del ECG de referencia = FA (n)	240		
Resultado de la app ECG = FA	236/240 (98,3%)		
Resultado de la app ECG = RS	4/240 (1,7%)		
Sensibilidad	236/240 (98,3%)	95,8%	< 0,0001
Resultado final del ECG de referencia = RS (n.º)	239		
Resultado de la app ECG = FA	1/239 (0,4%)		
Resultado de la app ECG = RS	238/239 (99,6%)		
Especificidad	238/239 (99,6%)	97,7%	< 0,0001

* Límite inferior exacto, binomial y unilateral del intervalo de confianza.

** Prueba de la hipótesis para una sensibilidad > 0,9 y una especificidad > 0,92.

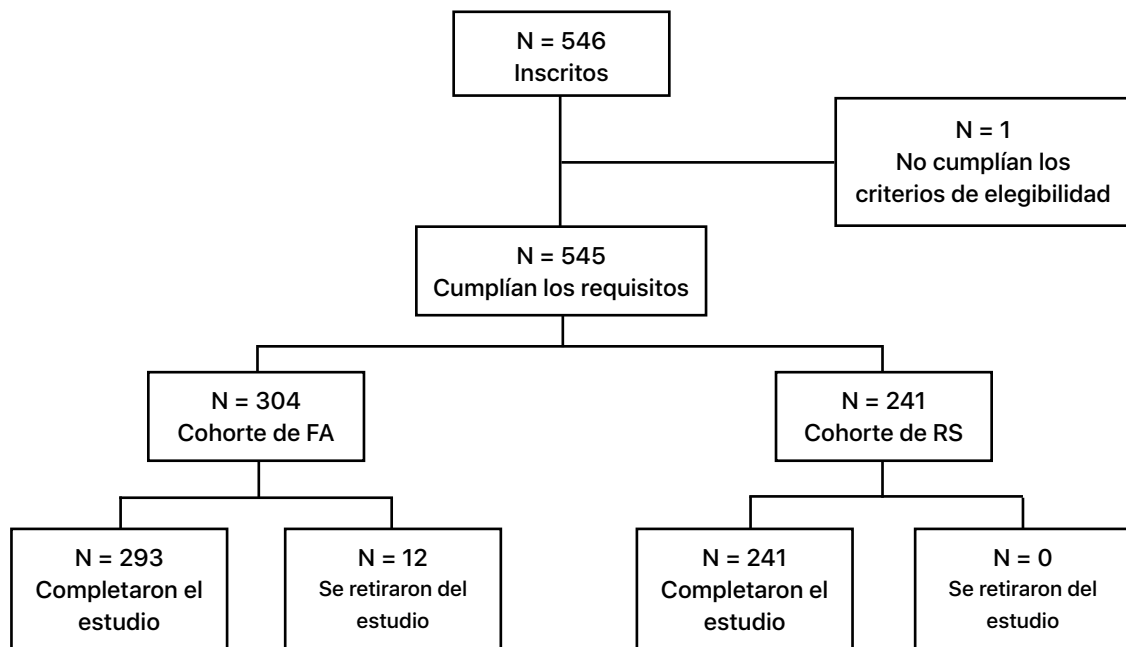
Siglas: FA = fibrilación auricular; RS = ritmo sinusal

Estudio de validación clínica de la app ECG 2.0

Se realizó un segundo estudio para respaldar y validar los algoritmos de la app ECG 2.0 en investigación (dispositivo de prueba), que amplía el intervalo de la FC clasificable (50-150 LPM) e introduce una nueva clasificación de los resultados (RS, RS con FC elevada, FA, FA con FC elevada, no concluyentes y mala grabación). El objetivo de este estudio era evaluar el rendimiento del dispositivo de prueba. Los principales criterios de evaluación fueron la especificidad y la sensibilidad. Los criterios de evaluación secundarios fueron la correcta clasificación de las siguientes categorías de sujetos en tiras específicas de ECG legibles y clasificables: RSN (FC 50-150), RS en ECG de 12 derivaciones simultáneo como RS; FA (FC 50-99), FA en ECG de 12 derivaciones simultáneo como FA, taquicardia sinusal (FC 100-150), RS en ECG de 12 derivaciones simultáneo como frecuencia cardiaca alta y FA con frecuencia cardiaca alta (FC 100-150), FA en ECG de 12 derivaciones simultáneo como FA. Además, también se evaluó la equivalencia entre la onda de la app ECG y la derivación I del electrocardiograma de 12 derivaciones, medida como la concordancia de una morfología aceptable de los complejos PQRST y la amplitud de la onda R, similar a la evaluación de la onda realizada en el estudio de validación clínica de la app ECG 1.0.

Los comités de revisión institucional apropiados de los respectivos centros de investigación del estudio prospectivo y multicéntrico con sede en EE. UU. autorizaron la realización del estudio. El investigador del centro obtuvo la aprobación del comité de revisión institucional, el formulario de consentimiento informado y cualquier material dirigido al paciente en cada centro del estudio antes de iniciar la participación.

Figura 3: Diagrama de clasificación de los sujetos



Los sujetos del estudio fueron los voluntarios con un RS normal en el momento del cribado, sin antecedentes conocidos de FA, y aquellos con FA continua, permanente o crónica conocida que presentaban FA durante el cribado. Se pidió a todos los sujetos que se hicieran un ECG de 12 derivaciones y, a la vez, un ECG de una sola derivación completo diferente con el Apple Watch; cada sujeto hizo tres pruebas. Este mismo proceso se siguió para las sesiones de ejercicio; se pidió a los voluntarios que se consideró sanos que hicieran ejercicio durante cinco minutos con una bicicleta estática para alcanzar la FC objetivo. Tanto en las sesiones de reposo como en las de ejercicio se tuvieron en cuenta las primeras pruebas para los fines de validación y análisis.

Dos cardiólogos independientes con certificación oficial y sin acceso a los demás datos analizaron en calidad de jueces los ECG de 12 derivaciones para diagnosticar la frecuencia y el ritmo cardiacos. En caso de discrepancia, un tercer juez analizó los datos. Se calculó la FC correspondiente a cada ECG de 12 derivaciones. Se grabó la FC y se seleccionó el código de diagnóstico de FC correspondiente a la frecuencia que se observó en el ECG de referencia. Se validaron los siguientes diagnósticos de ritmo cardiaco a partir de los datos de los ECG de 12 derivaciones: RS, FA, taquicardia supraventricular (TSV), otra anomalía del ritmo (extrasístoles auriculares frecuentes, extrasístoles ventriculares frecuentes, aleteo auricular, taquicardia ventricular, fibrilación ventricular, bloqueo arteriovenoso de tipo I de segundo grado, bloqueo arteriovenoso de tipo II de segundo grado, bloqueo arteriovenoso de tercer grado y otros) y no interpretables. Tres técnicos o cardiólogos sin acceso a los demás datos analizaron la onda durante la evaluación y validación de los datos del ECG de las tiras pareadas del ECG. Un revisor identificó los primeros seis complejos PQRST consecutivos nítidos legibles y sin artefactos que coincidían entre la tira del dispositivo del sujeto y la tira del dispositivo de referencia para su evaluación y otros dos revisores los utilizaron. Se excluyeron las tiras en las que no se pudieron hallar seis latidos consecutivos.

Para el análisis del criterio de evaluación principal, se puso en práctica un enfoque de muestreo con reposición para obtener unos intervalos de confianza del 95 % bilaterales para la sensibilidad y la especificidad, puesto que los datos se recopilaron de los mismos sujetos del estudio en reposo y tras el ejercicio. Los sujetos con al menos un resultado de FA (para la sensibilidad) o RS (para la especificidad) validado y con un resultado del algoritmo clasificable (RS o FA) se seleccionaron de manera aleatoria y con reemplazo. Los percentiles 2,5 y 97,5 de las

estimaciones de la distribución del muestreo con reposición representaban los límites de confianza del 95 % bilaterales. Si los límites de confianza inferiores correspondientes a la sensibilidad y la especificidad superaban el objetivo de rendimiento establecido previamente y asociado con esta métrica de rendimiento (90 % para la sensibilidad y 92 % para la especificidad), se rechazaba la hipótesis nula en favor de las hipótesis alternativas. Para la comparación de ondas se utilizaron los datos de la primera serie de pruebas de los sujetos del estudio seleccionados aleatoriamente (en reposo y después del ejercicio) para analizar las hipótesis de validación de las ondas con un enfoque de muestreo con reposición. Si el percentil 2,5 de la distribución de muestreo con reposición correspondiente a la proporción de morfologías aprobadas superaba el 80 %, se rechazaba la hipótesis nula (la proporción de morfologías aprobadas era inferior al 80 %). Se aplicó este mismo criterio a la evaluación de la proporción de concordancia con la amplitud de la onda R.

Resultados

En el estudio se inscribió a 546 sujetos; 305 se inscribieron en la cohorte de FA y 241 en la cohorte de ritmo sinusal normal (RSN). Se excluyó a un sujeto de la cohorte de FA por motivos de elegibilidad. De los 546 sujetos, 534 (293 en la cohorte de FA y 241 en la cohorte de RSN) finalizaron el estudio; 12 sujetos de la cohorte de FA no finalizaron el estudio.

Tres técnicos diferentes examinaron el conjunto de análisis de evaluación de la forma de la onda para rendir cuentas. Para la evaluación se utilizaron las tiras pareadas del reloj y del dispositivo de referencia. En total, 91 sujetos (100 %) tenía un par legible de tiras «en reposo», 87 sujetos (100 %) tenía un par de legible de tiras «después del ejercicio» y 93 sujetos (100 %) tenía al menos un par de tiras válido «en reposo» o «después del ejercicio».

El conjunto de análisis clasificable consistía en todos los sujetos con resultados legibles pareados y validados de la app ECG 2.0 y el ECG de 12 derivaciones. Se consideró como clasificable a un total de 512 sujetos (n = 279 en la cohorte de FA y n = 233 en la cohorte de RSN). El conjunto de análisis de evaluación de la onda, definido con los sujetos seleccionados de forma aleatoria con tiras pareadas legibles procedentes de la app ECG y de la derivación I del ECG de 12 derivaciones, comprendía a un total de 93 sujetos (n = 48 en la cohorte de FA y n = 45 en la cohorte de RSN).

La tabla 6 presenta los datos con clasificación cruzada de la app ECG 2.0 y los resultados validados del ECG de 12 derivaciones correspondientes a los datos en reposo y después del ejercicio. Se alcanzó un alto grado de concordancia entre los resultados validados del ECG de 12 derivaciones y el resultado del dispositivo con la app ECG en las categorías de RS y FA. De las tiras del ECG de 12 derivaciones validadas como RS (n= 470) un total de 436 resultados también fueron clasificados como RS por el resultado del dispositivo con la app ECG. Con respecto a las tiras del ECG de 12 derivaciones validadas como FA (n = 521), un total de 474 resultados fueron clasificados también como FA por el resultado del dispositivo con la app ECG. El dispositivo con la app ECG clasificó 9 resultados como ritmo sinusal, 37 resultados como fibrilación auricular, 13 resultados como no concluyentes y 9 resultados como una mala grabación que se evaluaron como otro (taquicardia supraventricular u otro) en el ECG de 12 derivaciones.

Tabla 6. Tabla de clasificación cruzada de los resultados de la app ECG y el ECG de 12 derivaciones validado (en reposo y tras el ejercicio): conjunto de análisis clasificable

Validación en terreno con resultados validados procedentes del ECG de 12 derivaciones						
Resultados del dispositivo con la app ECG	RS (50≤FC≤150)	FA (50≤FC≤150)	Otro (taquicardia supraventricular con FC 50–150; FC < 50; FC > 50)	No interpretables	No disponibles	Total
RS [RS (50–99); taq. sinusal RSN (100–150)]	436	7	9	0	2	454
FA (50≤FC≤150)	3	474	37	0	6	520
No concluyentes [incluida una FC < 50; FC > 150]	10	13	13	0	1	37
Mala grabación	14	23	9	0	3	49
No disponibles	7	4	1	1	0	13
Total	470	521	69	1	12	1073

Siglas: FA = fibrilación auricular, RS = ritmo sinusal, FC = frecuencia cardiaca, TSV = taquicardia supraventricular

Se calculó que la especificidad de la app ECG 2.0 era del 99,3 % y la sensibilidad del 98,5 % (tabla 7). Los límites inferiores del intervalo de confianza del cálculo de la especificidad (98,4 %) y la sensibilidad (97,3 %) superaban los objetivos de rendimiento establecidos previamente asociados con esta métrica de rendimiento; por tanto, las hipótesis nulas correspondientes a la especificidad y la sensibilidad se rechazaron en favor de las hipótesis alternativas y se alcanzaron los criterios de evaluación principales de especificidad y sensibilidad. Se calculó que la especificidad en reposo era del 99,1 % y que la sensibilidad en reposo era del 98,5 %. Se estimó que la especificidad después del ejercicio era del 99,5 % y que la sensibilidad después del ejercicio era del 98,6 %. Se realizó otro análisis que incorporaba la categoría de clasificación del algoritmo de «no concluyente» para los datos en reposo y después del ejercicio combinados. Se calculó que la especificidad resultante era del 97,1 %, mientras que la sensibilidad era del 96,0 %.

Se observó un alto grado de concordancia entre los resultados validados de frecuencia cardiaca procedentes del ECG de 12 derivaciones y el resultado del dispositivo con la app ECG. El grado de error de clasificación era bajo en comparación con los resultados de las tiras del ECG de 12 derivaciones de referencia. La tabla 8 ofrece un resumen de los resultados que confirma que los puntos de evaluación secundarios correspondientes a la correcta clasificación de los pacientes con RSN (FC 50-99, RS en el ECG de 12 derivaciones simultáneo) como RS, FA (FC 50-99, FA en el ECG de 12 derivaciones simultáneo) como FA, taquicardia sinusal (FC 100-150, RS en el ECG de 12 derivaciones simultáneo) como frecuencia cardiaca alta y FA con frecuencia cardiaca alta (FC 100-150, FA en el ECG de 12 derivaciones simultáneo) como FA (FC alta) en una tira de la app ECG legible y clasificable. El porcentaje de RS (FC 50-99) clasificado correctamente era del 98,3 %; el límite inferior del intervalo de confianza del 96,6 % superaba el objetivo de rendimiento establecido previamente del 90 % para este criterio de evaluación. El porcentaje de FA (FC 50-99) que se clasificó correctamente fue del 98,9 %; el límite inferior de confianza del 97,5 % superaba el objetivo de rendimiento del 81% establecido previamente. El porcentaje de RS (FC 100-150) que se clasificó correctamente fue del 90,7 %; el límite inferior del intervalo de confianza del 86,7 % superaba el objetivo de rendimiento establecido previamente del 81 %. El porcentaje de FA

(FC 100-150) que se clasificó correctamente fue del 83,0 %; el límite inferior del intervalo de confianza del 77,8 % superaba el objetivo de rendimiento establecido previamente del 75,5 %.

Tabla 7. Resultados de sensibilidad y especificidad (resultados del criterio de evaluación principal): conjunto de análisis clasificable		
Parámetro	Valor	Intervalo de confianza bilateral del 95 % del muestreo con reposición*
Combinados (en reposo y después del ejercicio)		
Resultado de referencia = FA (FC 50-150)	481	
App ECG = FA (FC 50-150)	474/481 (98,5%)	
App ECG = RS (FC 50-150)	7/481 (1,5%)	
Sensibilidad	474/481 (98,5%)	(97,3%, 99,6%)
Resultado de referencia = RS (FC 50-150)	439	
App ECG = FA (FC 50-150)	3/439 (0,7%)	
App ECG = RS (FC 50-150)	436/439 (99,3%)	
Especificidad	436/439 (99,3%)	(98,4%, 100,0%)
* Resultados obtenidos a partir del muestreo con reposición Siglas: FA = fibrilación auricular, RS = ritmo sinusal, FC = frecuencia cardiaca		

Tabla 8. Porcentaje de resultados correctos de ritmos clasificables de FA y RS		
Parámetro	Valor	Intervalo de confianza del muestreo con reposición**
Combinados (en reposo y después del ejercicio)*		
Porcentaje de resultados correctos de RS (FC 50-99) de referencia	230/234 (98,3%)	(96,6%, 99,6%)
Porcentaje de resultados correctos de FA (FC 50-99) de referencia	272/275 (98,9%)	(97,5%, 100%)
Porcentaje de resultados correctos de RS (FC 100-150) de referencia	186/205 (90,7%)	(86,7%, 94,6%)
Porcentaje de resultados correctos de FA (FC 100-150) de referencia	171/206 (83,0%)	(77,8%, 88,0%)
* El cálculo de porcentajes correctos incluye los ritmos clasificables de FA y RS obtenidos con la app ECG 2.0 ** Intervalo de confianza obtenido a partir del muestreo con reposición Siglas: FA = fibrilación auricular, FC = frecuencia cardiaca, RS = ritmo sinusal		

Además, la comparación de la onda entre el algoritmo 2.0 de la app ECG y la derivación I de un ECG de 12 derivaciones reveló que el porcentaje total de morfologías de onda en reposo y después del ejercicio combinadas era del 100 %, con la exclusión de cinco resultados en total. El límite inferior del intervalo de confianza era del 97,9 %, que superaba la proporción de morfologías aprobadas del 80 % establecida previamente. El porcentaje total de aprobados correspondiente a los resultados de la evaluación de la onda R en reposo y después del ejercicio era del 97,2 %; el límite inferior del intervalo de confianza era del 93,6 %, lo que superaba la proporción del 80 % de concordancia de la amplitud de la onda R establecida previamente. Los grupos en reposo y después del ejercicio presentaban una onda R similar.

Conclusiones

Los clientes del Apple Watch tienen acceso ahora a dos funciones opcionales que permiten detectar frecuencias cardiacas irregulares: la función de notificación de ritmo irregular (disponible en el Apple Watch Series 1 y posteriores) y la app ECG (disponible en el Apple Watch Series 4, Series 5 y Series 6).

La función de notificación de ritmo irregular es un software que clasifica de manera oportuna los tacogramas recopilados en segundo plano y notifica a los usuarios que activan la función la presencia de un ritmo cardiaco irregular. El estudio Apple Heart demostró que, de entre los participantes que recibieron una notificación mientras llevaban a la vez el Apple Watch y un parche de ECG, un 78,9 % presentó FA también en el parche de ECG y un 98,2 % presentó FA y otras arritmias de importancia clínica. Estos resultados indican que en la mayoría de los casos la notificación representará de forma precisa la presencia de FA, pero, en otros, la notificación podría indicar la presencia de una arritmia distinta a la FA.

El Apple Watch Series 4, Series 5 y Series 6 tiene un sensor eléctrico de frecuencia cardiaca que, junto con la app ECG y el algoritmo, genera un ECG similar a un ECG de una sola derivación (derivación I) para detectar la presencia de FA. Se ha demostrado que el algoritmo de código patentado que clasifica estos ECG tiene una sensibilidad de más del 98 % y una especificidad superior al 99 % para identificar la FA, al compararlos con los ECG grabados con un dispositivo de referencia y someterlos a la interpretación de expertos clínicos independientes.

Los usuarios que quieran usar estas dos funciones deben completar una introducción donde aprenderán a interpretar los resultados y a no considerarlos diagnósticos y conocerán las limitaciones de los algoritmos. Ambas funciones han obtenido la clasificación De Novo o 510K de la FDA.