



# Hartritmestoornissen detecteren met Apple Watch

December 2020

# Inhoud

<b>Overzicht</b>	<b>3</b>
Inleiding.....	3
<b>Ritmestoornissen detecteren op basis van PPG</b>	<b>3</b>
Beschrijving van de functie en technische informatie.....	3
Preklinische ontwikkeling.....	5
Klinische validatie.....	5
Apple Heart Study .....	5
Proefopzet AHS-deelonderzoek .....	5
Resultaten.....	6
<b>Detectie op basis van ECG</b>	<b>6</b>
Beschrijving van de functie en technische informatie.....	6
Preklinische ontwikkeling.....	7
Klinische validatie.....	7
Proefopzet.....	7
Klinisch validatieonderzoek ECG 1.0 .....	8
Resultaten.....	8
Klinisch validatieonderzoek ECG 2.0 .....	12
Resultaten .....	14
<b>Conclusies</b>	<b>17</b>

# Overzicht

Klanten met een Apple Watch kunnen gebruikmaken van twee softwarefuncties die dienen als medisch hulpmiddel voor de detectie van hartritmestoornissen zoals atriumfibrilleren (AF): de functie voor melding van een onregelmatig ritme en de Ecg-app.

Op Apple Watches Series 1 of nieuwer analyseert de functie voor het melden van een onregelmatig ritme de hartfrequentiegegevens die zijn verzameld door de sensor van de fotoplethysmogroaf (PPG) om episodes van een onregelmatige hartfrequentie die op AF kan duiden, te identificeren en de gebruiker te waarschuwen wanneer een episode wordt gedetecteerd. De Ecg-app op Apple Watches Series 4, Series 5 en Series 6 kan een ecg genereren dat vergelijkbaar is met een electrocardiogram met afleiding I en dat vervolgens classificeren als sinusritme (SR), atriumfibrilleren (AF), hoge of lage hartfrequentie, of onbeslist. Met ECG 2.0 zijn verdere classificaties van AF met een hoge hartfrequentie en slechte registratie beschikbaar.

In deze paper wordt uitgebreid ingegaan op de mogelijkheden van deze functies, inclusief de procedures voor testen en validatie.

## Inleiding

AF is een hartritmestoornis waarbij de boezems (atria) van het hart zich onregelmatig en soms te snel samentrekken en is een belangrijke oorzaak van beroertes. Omdat AF vaak asymptomatisch is, zijn veel mensen met AF zich er echter niet van bewust dat zij deze aandoening hebben. Door factoren als het risico op een beroerte, asymptomatische manifestatie, effectieve farmacologische behandelingen die het risico op een beroerte beperken en de groeiende marktpenetratie van consumentenapparaten die AF kunnen detecteren is er steeds meer belangstelling voor het vroegtijdig signaleren van AF buiten een klinische context.

Met watchOS 5.1.2 en nieuwer kan Apple Watch Series 1 en nieuwer perioden met een onregelmatige hartslag (een mogelijke indicatie van AF) in kaart brengen met behulp van signalen van een fotoplethysmogroaf (PPG) in combinatie met een algoritme. Naast dit identificatiealgoritme op basis van PPG hebben Apple Watches Series 4, Series 5 en Series 6 ook een elektrische hartslagsensor. Daardoor kan met de Ecg-app een ecg worden gegenereerd en geanalyseerd dat vergelijkbaar is met een ecg met afleiding I.

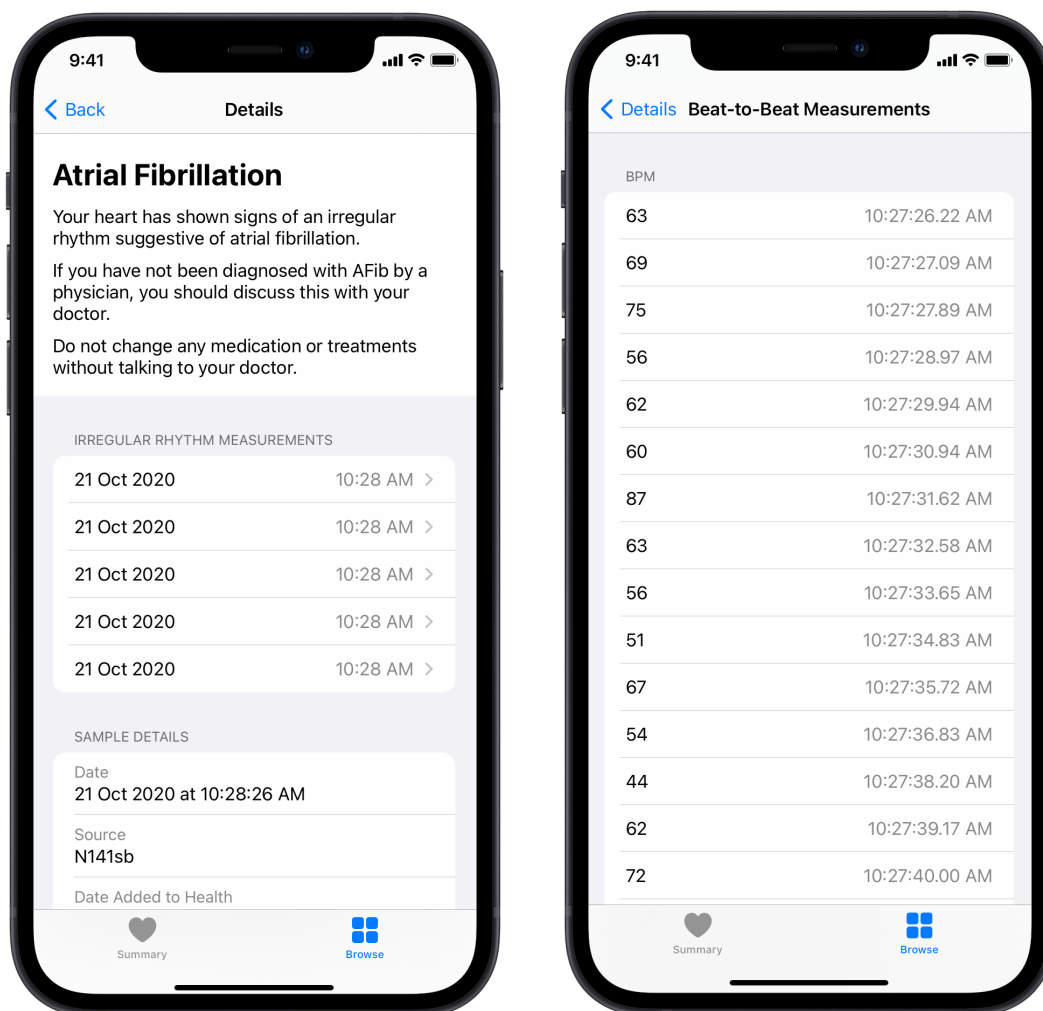
# Ritmestoornissen detecteren op basis van PPG

## Beschrijving van de functie en technische informatie

Apple Watch heeft een optische hartslagsensor die met groene ledlampjes en lichtgevoelige fotodioden vaststelt hoeveel bloed er door de pols van de gebruiker stroomt met behulp van fotoplethysmografie. Met behulp van deze sensoren en de onderliggende algoritmen worden de hartfrequentie (HF) en de hartritmevariabiliteit (HRV) gedetecteerd op Apple Watch Series 1 en nieuwer. Om de HRV vast te stellen, maakt Apple Watch om de twee tot vier uur een tachogram, een diagram van de tijd tussen de hartslagen. Vanaf watchOS 5.1.2 kan de gebruiker ook een functie voor het detecteren van ritmestoornissen inschakelen die gebruikmaakt van deze tachogrammen. Om de functie Meldingen bij onregelmatig ritme op Apple Watch te gebruiken, moet de gebruiker eerst de configuratie in de Gezondheid-app op de gekoppelde iPhone afronden. Tijdens de configuratie krijgt de gebruiker informatie over het gebruik van de functie en voorlichting over AF. Meer informatie over de gebruikerservaring vindt u op [support.apple.com/nl-nl/HT208931](https://support.apple.com/nl-nl/HT208931).

Als de functie voor het detecteren van ritmestoornissen op basis van PPG is ingeschakeld, wordt elk tachogram door Apple Watch geïnterpreteerd met een bedrijfseigen algoritme om vast te stellen of er sprake is van een onregelmatig ritme. Bij een onregelmatig tachogram wordt een proces in gang gezet waarbij tachogrammen met een hogere frequentie worden verzameld (zo vaak mogelijk, maar met tussenpozen van minimaal 15 minuten) en geanalyseerd. Apple Watch verzamelt en analyseert alleen tachogrammen als de gebruiker niet te veel beweegt om een meting te doen. Daardoor wordt de gebruiker niet constant gemonitord door het algoritme, maar gebeurt dit alleen als er een bruikbaar signaal beschikbaar is om waarden te verzamelen en analyseren. Als binnen een periode van 48 uur vijf van zes opeenvolgende tachogrammen (inclusief het eerste) als onregelmatig worden geïnterpreteerd, ontvangt de gebruiker een melding dat er mogelijk sprake is van een ritmestoornis. De gebruiker krijgt niet alleen een melding, maar kan ook meer informatie over de onregelmatige tachogrammen raadplegen in de Gezondheid-app (afbeelding 1). Als twee tachogrammen als niet onregelmatig worden geïnterpreteerd voordat de drempelwaarde is bereikt, wordt de cyclus gereset en worden tachogrammen weer met de basisfrequentie verzameld (om de twee uur).

**Afbeelding 1: Metingen van onregelmatig hartritme in Gezondheid-app**



In de Gezondheid-app kunnen gebruikers zien wanneer het algoritme een onregelmatig tachogram heeft gedetecteerd dat heeft bijgedragen aan een melding (links). Door op een datum en tijd te tikken, kunnen gebruikers de hartslagintervallmetingen bekijken die zijn berekend op basis van elk tachogram.

## Preklinische ontwikkeling

Voorafgaand aan de klinische testfase werden er onderzoeken gedaan om het detectiealgoritme op basis van PPG te ontwikkelen en de algoritmeprestaties te evalueren bij uiteenlopende omstandigheden en gebruikersactiviteiten, waaronder diep ademen, autorijden, trillende en bewegende handen, verminderde hand-/polsperfusie, het 's nachts dragen van de Watch, snelle ventriculaire reactie bij personen met AF en andere ritmestoornissen. Aan deze onderzoeken deden ruim 500 proefpersonen met AF en 2300 controleproefpersonen mee.

Omdat PPG op basis van lichtabsorptie werkt, is het detectiealgoritme voor ritmestoornissen getest met verschillende huidtypen en -tinten. Zo is ervoor gezorgd dat het sensorplatform kan worden afgestemd op verschillende huidtinten in de context van de algoritmen voor het detecteren van ritmestoornissen.

Melanine heeft een hoog absorptievermogen bij de golflengte die de groene led van Apple Watch gebruikt, waardoor het moeilijker kan zijn om hartslagmetingen met PPG uit te voeren bij donkere huidtinten. Om dit te ondervangen, past het sensorplatform van Apple Watch de ledstroom (en dus de lichtuitvoer), de versterking van de fotodiode (lichtgevoeligheid) en de bemonsteringsfrequentie aan om te zorgen voor voldoende signaalamplitude bij alle huidtinten.

Ter validatie werden 1,3 miljoen tachogrammen van 1124 proefpersonen (waarvan 51% vrouwen) met uiteenlopende huidtypen en -tinten (Fitzpatrick-huidtype en met een spectrofotometer gemeten lichtheid van de huid op de pols) geanalyseerd. Vanuit technisch oogpunt lag de focus voornamelijk op signaalamplitudes bij personen met een donkere huid. Daarom had bijna 5% van de deelnemende proefpersonen Fitzpatrick-huidtype VI. Uit validaties bleek geen significant verschil in gevoeligheid en specificiteit van het algoritme bij verschillende huidtypen/-tinten.

## Klinische validatie

### Apple Heart Study

De Apple Heart Study (AHS) is een verkennend, pragmatisch onderzoek met één arm dat virtueel werd uitgevoerd om na te gaan in hoeverre het Apple Watch-algoritme voor het melden van een onregelmatig ritme in staat is ritmestoornissen te detecteren die kunnen wijzen op AF. Wanneer tijdens het onderzoek vijf van de zes tachogrammen van een gebruiker onregelmatig waren, ontving de gebruiker via iPhone en Apple Watch een melding met de mogelijkheid om contact op te nemen met een bij het onderzoek betrokken telezorgarts en een ambulante ecg-patch (de ePatch van BioTelemetry, Inc., Conshohocken, PA) aan te vragen. De deelnemers kregen de instructie om de ePatch maximaal zeven dagen te dragen. Verzamelde gegevens voor een deelnemer met een minimale analyseerbare tijd van één uur werden als bruikbaar beschouwd.

De uitgebreide resultaten van de AHS werden in november 2019 gepubliceerd in het *New England Journal of Medicine* (Perez, Marco V., et al. "Large-scale assessment of a smartwatch to identify atrial fibrillation." *New England Journal of Medicine* 381.20 (2019): 1909–1917).

### Proefopzet AHS-deelonderzoek

Er werd een deelonderzoek uitgevoerd op gegevens die in de AHS waren verzameld om te bepalen of het classificatiealgoritme voor tachogrammen (individueel tachogram of spottachogram) en het algoritme voor de bevestigingscyclus (waarschuwniveau, vijf van de zes tachogrammen) een acceptabele positieve voorspellende waarde (PVW) hebben in vergelijking met de ePatch-controle als het gaat om het detecteren van onregelmatige hartritmen die overeenstemmen met AF in een deelgroep AHS-deelnemers. De AHS-onderzoekers waren op de hoogte van het deelonderzoek, de daaropvolgende analyses en de indiening van de gegevens bij de FDA, maar ze waren niet op de hoogte van de resultaten van het deelonderzoek terwijl de AHS nog werd uitgevoerd. De Institutional Review Board (IRB) die de AHS heeft goedgekeurd, heeft bepaald dat dit deelonderzoek niet onder toezicht van de IRB stond. Alle AHS-deelnemers hebben geïnformeerde toestemming gegeven, waarbij ze er tevens mee instemden dat hun onderzoeksgegevens voor het deelonderzoek mochten worden gebruikt.

De gegevens voor het deelonderzoek werden verzameld van AHS-deelnemers die tussen 30 november 2017 en 22 juni 2018 aan het onderzoek meededen. De proefpersonen in dit deelonderzoek ontvingen een melding van een onregelmatig hartritme via de AHS-app. Vervolgens ontvingen ze een ePatch die ze gingen dragen om de bevindingen van de ambulante ecg te interpreteren. De oorspronkelijke onregelmatige tachogrammen die leidden tot de eerste melding en mogelijk ook tot het eerste telegcontact werden niet geanalyseerd in het kader van dit deelonderzoek. Alleen onregelmatige tachogrammen en meldingen die werden gegenereerd terwijl een gebruiker de voor het onderzoek verstrekte ePatch droeg, werden geanalyseerd.

Twee onafhankelijke ecg-beoordelaars die zijn gecertificeerd door Amerikaanse beroepsraden voor cardiologie en/of elektrofysiologie, controleerden en beoordeelden de ecg-stroken en classificeerden deze als SR, AF, ander onregelmatig ritme of onleesbaar. Als de beoordelaars het niet eens waren, werd de strook geëvalueerd door een derde beoordelaar met vergelijkbare kwalificaties. Deze beoordelaars waren niet op de hoogte van de tachogramclassificatie. De ecg-classificatie van de beoordelaar en de door het algoritme bepaalde tachogramclassificatie werden beveiligd verstuurd naar de statisticus van het onderzoek voor gegevensanalyse.

## Resultaten

Van de 226 deelnemers aan het deelonderzoek die een eerste melding van een ritmestoornis ontvingen en die ongeveer een week een ePatch droegen, werd bij 41,6% (94 deelnemers) AF gedetecteerd door de ePatch. Tijdens het gelijktijdig dragen van Apple Watch en een ePatch ontvingen 57 van de 226 deelnemers een AF-melding (d.w.z. vijf van de zes opeenvolgende tachogrammen werden geclassificeerd als abnormaal). Van die groep was bij 78,9% (45 deelnemers) ook AF te zien in de ePatch-gegevens en bij 98,2% (56 deelnemers) waren AF of andere klinisch relevante ritmestoornissen te zien. Deze resultaten tonen aan dat de melding in de meeste gevallen de aanwezigheid van AF accuraat aangeeft, maar dat een melding in sommige gevallen de aanwezigheid van een andere ritmestoornis dan AF aangeeft. Er werden geen ernstige nadelige effecten van het apparaat waargenomen.

# Detectie op basis van ECG

## Beschrijving van de functie en technische informatie

Apple Watch Series 4, Series 5 en Series 6 hebben een Digital Crown met een titanium elektrode. Het saffierkristal aan de achterkant van Apple Watch is voorzien van een ultradun chroom-/silicium-/koolstofnitridelaagje. De Ecg-app leest en registreert de elektrische impulsen die het hart aansturen bij de vingertop van de gebruiker (met de Digital Crown) en bij de pols (met de achterzijde van de Apple Watch), waardoor een gesloten circuit ontstaat. Om de Ecg-app op Apple Watch te gebruiken, moet de gebruiker eerst de configuratie in de Gezondheid-app op de gekoppelde iPhone afronden. Tijdens de configuratie krijgt de gebruiker informatie over het gebruik van de functie en voorlichting over AF. Om een ecg te genereren, opent de gebruiker de Ecg-app die is geïnstalleerd op Apple Watch. Vervolgens houdt de gebruiker 30 seconden lang een vinger van de ene hand op de Digital Crown van de Apple Watch om de andere pols. De polariteit voor de afleiding wordt bepaald door de pols voor Apple Watch die is geselecteerd in Instellingen.

Nadat het ecg is gemaakt, wordt de ecg-curve met behulp van een bedrijfseigen algoritme in ECG 1.0 geclassificeerd als SR, AF of onbeslist. Met ECG 2.0 zijn mogelijk ook aanvullende classificaties beschikbaar, zoals AF met een hoge hartslag (HF 100–150) en een verdere differentiatie tussen een slechte registratie en onduidelijke metingen. Deze ritme classificaties (gemiddelde HF, door gebruiker gemelde symptomen en golfvorm) worden toegevoegd aan de Gezondheid-app en ze worden opgeslagen en kunnen door de gebruiker in PDF-indeling worden gedeeld vanuit de app op de gekoppelde iPhone van de gebruiker. Ga naar [support.apple.com/nl-nl/HT208955](https://support.apple.com/nl-nl/HT208955) voor meer informatie over de gebruikerservaring.

## Preklinische ontwikkeling

Vóór aanvang van de klinische validatie werden de algoritmen voor ecg-signaaldetectie en -classificatie in meerdere onderzoeken getest. De sensoren en het classificatiealgoritme werden getest bij verschillende etniciteiten, polsontrekken, BMI-waarden, leeftijden, ritmestoornissen anders dan AF, gradaties van strakheid van het bandje, houdingen, trainingsactiviteiten en transpiratiewaarden. Aan deze tests deden ongeveer 2500 proefpersonen mee. Bij ongeveer 25% van deze proefpersonen was eerder AF of een andere hartritmestoornis vastgesteld.

Een hogere frequentie van 'onleesbare' ecg's was de belangrijkste variatie in de algoritmeprestaties. De factoren die tot deze variatie leidden, waren een lage signaalamplitude (als gevolg van rechtersafwijking, met name bij personen met een lage BMI, of transpiratie tijdens tests na een trainingssessie) en bewegingsartefacten als gevolg van gedrag van de gebruiker. Apple Watch maakt gebruik van droge elektroden die mechanisch sterk en corrosiebestendig zijn, waardoor ze zeer geschikt zijn voor een wearable. Bij droge elektroden, vooral als deze worden aangebracht op ledematen, is de kans op ruis zoals hierboven beschreven echter groter dan bij de tijdelijke gelelektroden die worden gebruikt in klinische devices.

Naast de bovengenoemde factoren resulteerde de aanwezigheid van bepaalde niet-AF-ritmestoornissen ook in significant verschillende algoritmeprestaties in vergelijking met proefpersonen met een sinusritme. Deze voorwaarden en de resultaten worden hieronder beschreven in tabel 1.

**Tabel 1. Niet-AF-ritmestoornissen die de algoritmeprestaties beïnvloeden (Ecg 2.0)**

Ritmestoornis	Variatie
Linker- of rechterbundeltakblok	7,9% van de proeven geassocieerd als AF
Eerstegraads AV-blok	10,2% van de proeven niet geassocieerd; van de geassocieerde proeven werd 2,4% geassocieerd als AF
Bigeminie	92,5% niet geassocieerd
Frequente PVC's	48,8% niet geassocieerd; van de geassocieerde werd 24,1% geassocieerd als AF
Frequente PAC's	19,8% niet geassocieerd; van de geassocieerde werd 23,5% geassocieerd als AF
Atriale tachycardie	7,7% geassocieerd als AF
Gemengd ritme	46,5% niet geassocieerd; van de geassocieerde werd 29,7% geassocieerd als AF
Hoge/lage HF (buiten 50–150 slagen per minuut)	94,8% niet geassocieerd

## Klinische validatie

### Proefopzet

Apple heeft twee multicentrische onderzoeken gesponsord om te valideren of de Ecg-app in staat is om (1) een ecg-golfvorm te genereren die vergelijkbaar is met een ecg met afleiding I van een standaard 12-afleidingen-ecg, en (2) met behulp van een ritme classificatie algoritme hartritm te classificeren als SR of AF aan de hand van dit ecg met één afleiding.

## Klinisch validatieonderzoek ECG 1.0

Er werd een klinisch validatieonderzoek uitgevoerd om de prestaties van de ECG 1.0-app te valideren. De primaire eindpunten van het onderzoek waren de gevoeligheid van het ritme classificatie algoritme bij het detecteren van AF en de specificiteit bij het detecteren van SR. Een externe Institutional Review Board (IRB) heeft het protocol, het toestemmingsformulier en alle andere relevante materialen goedgekeurd voordat de proefpersonen aan het onderzoek gingen deelnemen. Alle proefpersonen gaven schriftelijk toestemming voor deelname voordat ze aan het onderzoek gingen deelnemen.

In het onderzoek werden deelnemers opgenomen met bekende AF en anderen bij wie geen afwijkingen van het hartritme bekend waren. Er werd hen gevraagd drie ecg's met één afleiding op te nemen met de Ecg-app terwijl onderzoekspersoneel tegelijkertijd drie 12-afleidingen-ecg's opnamen met een door de Amerikaanse FDA goedgekeurd medisch apparaat (GE Healthcare CardioSoft ECG). De eerste trials werden overwogen voor beoordeling en analyse. De deelnemers werden geholpen met het bepalen van de plaats voor Apple Watch en kregen de instructie om hun arm stil te houden, bijvoorbeeld door hun arm op een tafel of op hun been te laten rusten. Vervolgens kregen ze de gelegenheid om te oefenen met het nemen van monsters.

Voor test (1) plaatsten drie onafhankelijke, gecertificeerde cardiotechnici de gegenereerde ritmestroken van 140 willekeurig geselecteerde proefpersonen (70 AF, 70 SR) op de corresponderende strook met afleiding I van de met het medisch apparaat gegenereerde ritmestroken om de morfologie van zes opeenvolgende PQRST-complexen visueel te vergelijken. Op basis van visueel beoordeelde morfologische overeenkomsten keurden de technici elke strook goed of af. Ook werd de technici gevraagd de R-amplitude van de iso-elektrische basislijn tot de dichtstbijzijnde millimeter te meten voor de eerste twee QRS-complexen in zowel de referentieritmestroken als de met de Ecg-app gegenereerde ritmestroken en de overeenkomst tussen deze twee waarden te beoordelen.

Voor (2) werd elke referentiestrook van een 12-afleidingen-ecg blind beoordeeld door drie onafhankelijke, door Amerikaanse beroepsraden gecertificeerde cardiologen, die het ritme classificeerden als SR, AF, anders (alles wat niet SR of AF was binnen de hartslagparameters) of onleesbaar (geen diagnose mogelijk omdat de strook niet kon worden gelezen). Het algoritme van de Ecg-app classificeerde de met de Ecg-app gegenereerde ecg als SR, AF, niet classificeerbaar of onleesbaar. De gevoeligheid en specificiteit van de SR- en AF-classificaties van de Ecg-app (voor classificeerbare ecg's) vergeleken met de interpretatie van de 12-afleidingen-ecg door de cardioloog werden berekend. Eén onafhankelijke, door een Amerikaanse beroepsraad gecertificeerde cardioloog werd vervolgens gevraagd de met de Ecg-app gegenereerde ritmestroken blind te classificeren op basis van dezelfde categorieën.

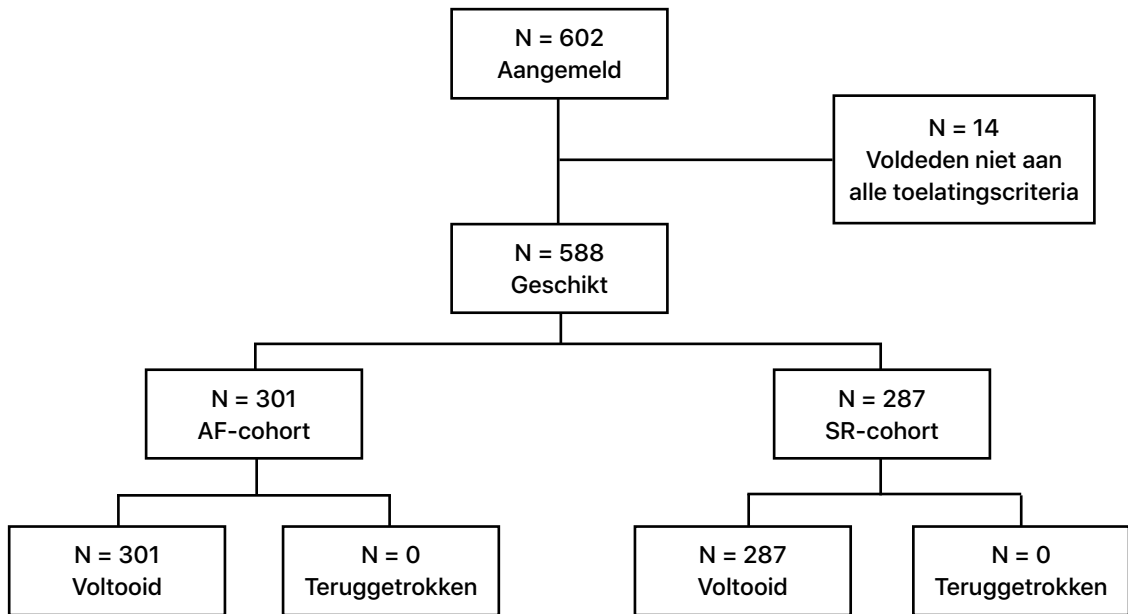
Voor de analyses van de primaire eindpunten werd afzonderlijk een eenzijdige, exacte betrouwbaarheidsondergrens van 97,5% berekend voor de gevoeligheid en specificiteit. Als de ondergrens voor gevoeligheid hoger was dan 90%, werd de nulhypothese verworpen wanneer de gevoeligheid hoger dan 90% was. Als de ondergrens voor specificiteit hoger dan 92% was, werd de nulhypothese verworpen wanneer de specificiteit hoger dan 92% was.

## Resultaten

Aan het onderzoek deden 602 proefpersonen mee, waarvan er 588 aan de toelatingscriteria voldeden. Van deze 588 proefpersonen werden er 301 die hadden aangegeven dat ze AF hadden, toegewezen aan het AF-cohort; 287 proefpersonen die niet hadden aangegeven dat ze AF hadden, werden toegewezen aan het SR-cohort. Deze cohorttoewijzingen werden alleen gebruikt om voldoende deelname te garanderen. De beoordelaars waren niet op de hoogte van de cohorttoewijzingen, en de aanwezigheid of afwezigheid van AF werd uitsluitend gebaseerd op het ecg dat tijdens de test werd gegenereerd. De 14 proefpersonen die het onderzoek wel voltooiden, maar niet werden toegewezen aan een deelnamecohort, kwamen niet in aanmerking voor deelname aan het onderzoek omdat ze in het verleden paroxismaal AF hadden gehad zonder dat AF op het ecg te zien was op het moment van screening. Alle toegelaten proefpersonen voltooiden het onderzoek (afbeelding 2). Tijdens het onderzoek werden geen ongewenste voorvallen gemeld.



**Afbeelding 2: Stroomdiagram met informatie over de proefpersonen**



Drie onafhankelijke, gecertificeerde cardiotechnici vonden visuele morfologische overeenkomsten tussen de golfvorm van de Ecg-app en het referentie-ecg met afleiding I dat werd gegenereerd met het standaard medische apparaat bij 98,4% van de geanalyseerde stroken in het AF-cohort en 100% in het SR-cohort (tabel 2). 99,2% van de stroken van alle proefpersonen werd goedgekeurd (betrouwbaarheidsondergrens van 97,5%, 95,7%). Stroken werden niet meegenomen als er niet in een van de twee sets stroken (Ecg-app of referentie) zes opeenvolgende slagen (PQRST-complexen) zonder artefact konden worden herkend.

**Tabel 2. Vergelijking van golfvormen**

Kenmerk	Proefperso- nen met AF (N = 61)	Proefperso- nen met SR (N = 65)	Totaal (N = 126)	Betrouwbaar- heidsonder- grens*	p- waarde**
Aantal gekoppelde stroken van proefpersonen (stroken van de Ecg-app en referentiestroken) dat werd goedgekeurd	60	65	125		
Aantal leesbare gekoppelde stroken van proefpersonen (stroken van de Ecg-app en referentiestroken)	61	65	126		
Percentage stroken van proefpersonen dat werd goedgekeurd	60/61 (98,4%)	65/65 (100%)	125/126 (99,2%)	95,7%	< 0,0001
Aantal gekoppelde stroken van proefpersonen dat werd uitgesloten	8	5	13		

\*Exacte binomiale, eenzijdige betrouwbaarheidsondergrens van 97,5% voor totaal.

\*\*Test van hypothese voor succes bij proefpersonen > 0,8.

Afkortingen: AF = atriumfibrilleren, SR = sinusritme

Om eens te meer te bevestigen dat de stroken die met de Ecg-app en het referentieapparaat werden gegenereerd, vergelijkbaar waren, werd de door de cardioloog blind uitgevoerde classificatie van de stroken van de Ecg-app vergeleken met de door de cardioloog uitgevoerde classificatie van de referentiestroken (tabel 3). De mate van overeenstemming tussen de classificaties van de stroken van het apparaat en de AF- en SR-referentieresultaten was respectievelijk 100% en 99,1%. Onleesbare stroken werden niet meegenomen in deze analyse.

**Tabel 3. Classificaties van stroken van Ecg-app en referentiestroken**

Kenmerk	Totaal (N = 522)
<b>Definitief ecg-referentieresultaat = AF</b>	263
Classificatie van strook van Ecg-app = AF	239/263 (90,9%)
Classificatie van strook van Ecg-app = SR	0/263 (0,0%)
Classificatie van strook van Ecg-app = Anders	0/263 (0,0%)
Classificatie van strook van Ecg-app = Onleesbaar	24/263 (9,1%)
<b>% overeenstemming met referentieresultaat AF*</b>	239/239 (100,0%)
<b>Definitief ecg-referentieresultaat = SR</b>	244
Classificatie van strook van Ecg-app = AF	0/244 (0,0%)
Classificatie van strook van Ecg-app = SR	232/244 (95,1%)
Classificatie van strook van Ecg-app = Anders	2/244 (0,8%)
Classificatie van strook van Ecg-app = Onleesbaar	10/244 (4,1%)
<b>% overeenstemming met referentieresultaat SR*</b>	232/234 (99,1%)
<b>Definitief ecg-referentieresultaat = Anders</b>	15
Classificatie van strook van Ecg-app = AF	0/15 (0,0%)
Classificatie van strook van Ecg-app = SR	3/15 (20,0%)
Classificatie van strook van Ecg-app = Anders	12/15 (80,0%)
Classificatie van strook van Ecg-app = Onleesbaar	0/15 (0,0%)
<b>% overeenstemming met referentieresultaat Anders*</b>	12/15 (80,0%)

\*Onleesbare stroken werden niet meegenomen.  
Afkortingen: AF = atriumfibrilleren, SR = sinusritme

In totaal 485 van de 602 gekoppelde stroken van de Ecg-app en referentiestroken werden als classificeerbaar beschouwd. Bij de overige paren werden stroken van de Ecg-app en/of referentiestroken als onleesbaar of niet classificeerbaar beschouwd. In tabel 4 wordt de verdeling over de AF- en SR-cohorten weergegeven.

**Tabel 4. Eindresultaat classificatie met algoritme van Ecg-app en referentiestroken**

Classificatie met algoritme van Ecg-app	Classificatie referentiestrook				
	SR	AF	Anders	Onleesbaar	Totaal
SR	238	4	4	1	247
AF	1	236	2	2	241
Niet classificeerbaar	6	7	6	0	19
Onleesbaar	18	30	1	0	49
Resultaat van apparaat niet gerapporteerd*	32	13	1	0	46
<b>Totaal</b>	295	290	14	3	602

\*Resultaten niet gerapporteerd op basis van vooraf vastgestelde criteria (zoals synchronisatie niet gedetecteerd) voor alle proefpersonen op één na  
Afkortingen: AF = atriumfibrilleren, SR = sinusritme

Bij de classificatie met het algoritme van de Ecg-app was de gevoeligheid 98,3% en de specificiteit 99,6% (tabel 5). Bij uitbreiding van de analyse met de 2,4% (7/290) en 2,0% (6/295) stroken die als niet classificeerbaar werden gecategoriseerd door het apparaat in de AF- en SR-cohorten, was de gevoeligheid 95,5% (95% BI: 92,2%, 97,8%) en de specificiteit 97,1% (95% BI: 94,2%, 98,8%). Deze resultaten voldoen aan de primaire eindpunten die vooraf zijn vastgesteld in de opzet van dit onderzoek. Bovendien was 12,2% (68/556) van de registraties onbeslist (onleesbaar of niet classificeerbaar) en was niet classificeerbaar als SR of AF. Wanneer onbesliste registraties werden meegenomen in de analyse, classificeerde de Ecg-app op de juiste wijze SR bij 90,5% (238/263) van de proefpersonen met SR, en AF bij 85,2% (236/277) van de proefpersonen met AF. De resultaten van de klinische validatie zijn gebaseerd op gebruik in een gecontroleerde omgeving. Gebruik van de Ecg-app in de praktijk kan ertoe leiden dat een groter aantal stroken als onbeslist en niet classificeerbaar wordt beschouwd.

**Tabel 5. Analyse van gevoeligheid en specificiteit – classificeerbare stroke**

Parameter	Waarde	Betrouwbaarheids- ondergrens*	p- waarde**
Definitief ecg-referentieresultaat = AF (n)	240		
Resultaat Ecg-app/apparaat = AF	236/240 (98,3%)		
Resultaat Ecg-app/apparaat = SR	4/240 (1,7%)		
<b>Gevoeligheid</b>	236/240 (98,3%)	95,8%	< 0,0001
Definitief ecg-referentieresultaat = SR (n)	239		
Resultaat Ecg-app/apparaat = AF	1/239 (0,4%)		
Resultaat Ecg-app/apparaat = SR	238/239 (99,6%)		
<b>Specificiteit</b>	238/239 (99,6%)	97,7%	< 0,0001

\*Exacte binomiale, eenzijdige betrouwbaarheidsondergrens

\*\*Test van hypothese voor gevoeligheid > 0,9 en specificiteit > 0,92

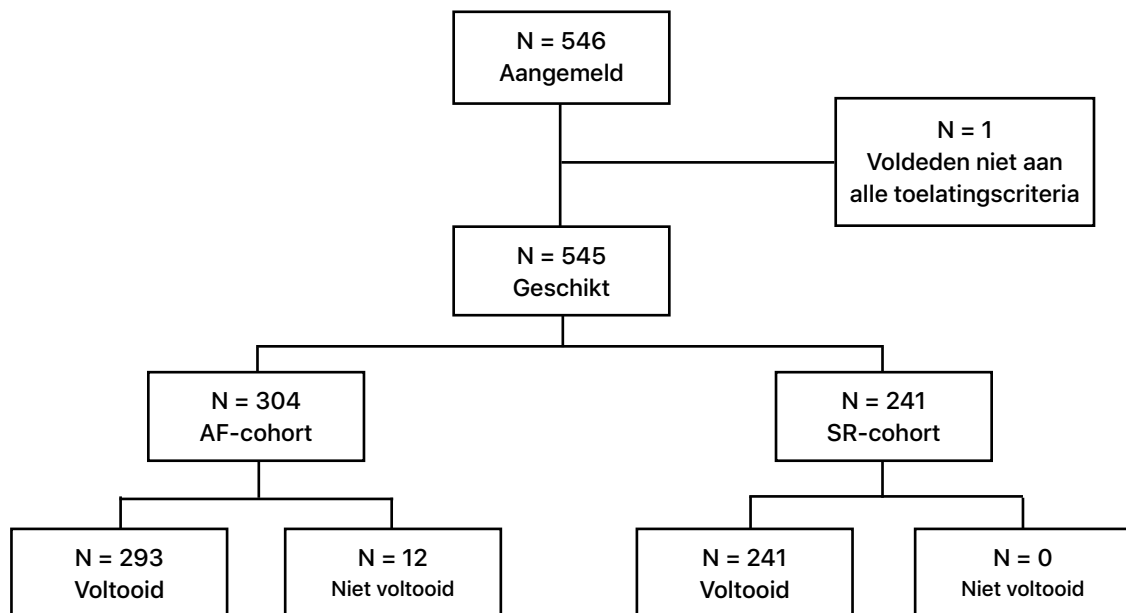
Afkortingen: AF = atriumfibrilleren, SR = sinusritme

### Klinisch validatieonderzoek ECG 2.0

Er werd een tweede onderzoek uitgevoerd om de onderzochte algoritmes van de Ecg 2.0-app (testapparaat) te ondersteunen en valideren. Met deze algoritmen wordt het classificeerbare HF-bereik (50–150 bpm) uitgebreid en zijn er nieuwe classificatieresultaten (SR, SR met hoge HF, AF, AF met hoge HF, onbeslist en slechte registratie) beschikbaar. Het doel van dit onderzoek was om de prestaties van het testapparaat te evalueren. De beoordeelde primaire eindpunten waren de specificiteit en gevoeligheid. Secundaire eindpunten waren onder andere de juiste classificatie van de volgende categorieën proefpersonen op basis van specifieke leesbare en classificeerbare ecg-teststroke: NSR (HF 50–150), SR op gelijktijdig 12-afleidingen-ecg als SR; AF (HF 50–99), AF op gelijktijdig 12-afleidingen-ecg als AF, sinustachycardie (HF 100–150), SR op gelijktijdig 12-afleidingen-ecg als hoge hartfrequentie en AF met hoge hartfrequentie (HF 100–150), AF op gelijktijdig 12-afleidingen-ecg als AF. Daarnaast werden ook de overeenkomsten tussen het ecg van de Ecg-app en dat van afleiding I van een 12-afleidingen-ecg beoordeeld op basis van een acceptabele overeenkomst van de morfologie van PQRST-complexen en de amplitude van de R-golf, op vergelijkbare wijze als de beoordeling van het ecg in klinisch validatieonderzoek ECG 1.0.

De uitvoering van dit onderzoek was goedgekeurd door de relevante institutional review boards van de desbetreffende onderzoekslocaties van het verkennende onderzoek in meerdere centra in de VS. De onderzoeker van de locatie verkreeg op elke onderzoekslocatie goedkeuring van de IRB en regelde het formulier voor geïnformeerde toestemming en alle benodigde materialen voor proefpersonen voordat het onderzoek begon.

**Afbeelding 3: Stroomdiagram met informatie over de proefpersonen**



Onder de proefpersonen waren personen met normaal SR op het moment van screening, zonder bekende voorgeschiedenis van AF, en personen met bekend persisterend, permanent of chronisch AF die tijdens de screening AF hadden. Alle proefpersonen kregen de instructie om gelijktijdig een 12-afleidingen-ecg en een afzonderlijk ecg met één afleiding met de Apple Watch te laten maken. Met elke proefpersoon werden drie proeven uitgevoerd. Bij de inspanningssessies werd hetzelfde proces gevolgd: proefpersonen die gezond geacht werden, kregen de vraag om vijf minuten te fietsen op een hometrainer tot zij de doel-HF bereikten. Bij zowel de rust- als de inspanningssessie werden de eerste proeven in overweging genomen voor beoordeling en analyse.

Twee onafhankelijke, door Amerikaanse beroepsraden gecertificeerde cardiologen beoordeelden de 12-afleidingen-ecg's blind op HF en ritmediagnose. In geval van tegenstrijdigheid keek een derde beoordelaar mee. De HF werd voor elke 12-afleidingen-ecg berekend. De HF werd geregistreerd en de diagnostische code passend bij de op de referentie-ecg vastgestelde HF werd geselecteerd. De volgende hartritmediagnoses werden toegekend op basis van de gegevens van de 12-afleidingen-ecg: SR, AF, supraventriculaire tachycardie (SVT), ander afwijkend ritme (veel premature atriale contracties, veel premature ventriculaire contracties, atriumflutter, ventriculaire tachycardie, ventrikelfibrilleren, tweedegraads AV-blok type I, tweedegraads AF-blok type II, derdegraads AV-blok en anders), en niet-interpreteerbaar. Drie onafhankelijke cardiotechnici of cardiologen bekeken het ecg tijdens de beoordeling en classificatie van ecg-gegevens van de gekoppelde ecg-stroken. De eerste zes opeenvolgende, goed afleesbare PQRST-complexen zonder artefacten die overeenkwamen op de strook van het apparaat van de proefpersoon en de strook van het referentieapparaat werden door één beoordelaar geïdentificeerd en door de twee andere beoordelaars gebruikt. Indien er geen zes opeenvolgende hartslagen konden worden gevonden, werden de stroken afgekeurd.

Voor analyse van de primaire eindpunten werd gebruikgemaakt van bootstrapping om tweezijdige 95%-betrouwbaarheidsintervallen voor gevoeligheid en specificiteit te verkrijgen, omdat de gegevens waren verzameld van dezelfde proefpersonen in rust en na inspanning. Proefpersonen met ten minste één vastgesteld AF-resultaat (voor gevoeligheid) of SR (voor specificiteit), en met een classificeerbaar algoritmeresultaat (SR of AF) werden willekeurig geselecteerd met terugleggen. Het 2,5e en 97,5e percentiel van de verdeling van bootstrapschattingen vormden de tweezijdige 95%-betrouwbaarheids grenzen. Indien de betrouwbaarheids ondergrenzen voor zowel gevoeligheid als specificiteit het vooraf vastgestelde prestatiedoel voor deze prestatie maten (90% voor gevoeligheid en 92% voor specificiteit) overschreden, werd de nulhypothese verworpen en werd de alternatieve hypothese als waar aangenomen. Voor de vergelijking van de ecg's werden gegevens van de eerste proefrun van de willekeurig geselecteerde onderzoeksproefpersonen, in rust en na inspanning, gebruikt om de hypothesen met betrekking tot de beoordeling van de ecg's te testen met behulp van bootstrapping. Indien het 2,5e percentiel van de bootstrapverdeling voor het goedkeuringspercentage van de morfologie hoger was dan 80%, werd de nulhypothese (het goedkeuringspercentage op basis van morfologie was lager dan 80%) verworpen. Hetzelfde criterium werd toegepast op de beoordeling van het overeenkomstpercentage van de amplitude van de R-golf.

## Resultaten

Er werden in totaal 546 proefpersonen ingeschreven voor het onderzoek: 305 in het AF-cohort en 241 in het cohort met normaal sinusritme (NSR). Eén proefpersoon in het AF-cohort werd uitgesloten omdat deze niet aan de toelatingscriteria voldeed. Van de 546 proefpersonen voltooiden er 534 (293 in het AF-cohort en 241 in het NSR-cohort) het onderzoek. 12 Proefpersonen in het AF-cohort voltooiden het onderzoek niet.

Ter verantwoording werd de analyseset voor ecg-beoordeling bekeken door drie afzonderlijke cardiotechnici. Voor de beoordeling werden gekoppelde stroken van het horloge en referentiestroken gebruikt. In totaal hadden 91 proefpersonen (100%) een leesbaar paar gekoppelde 'in rust'-stroken, 87 proefpersonen (100%) hadden een leesbaar paar gekoppelde 'na inspanning'-stroken en 93 proefpersonen (100%) had ten minste één leesbaar paar gekoppelde stroken voor 'in rust' of 'na inspanning'.

De classificeerbare analyseset bestond uit alle proefpersonen die een leesbaar paar gekoppelde stroken van de Ecg 2.0-app en de 12-afleidingen-ecg had. Er werden in totaal 512 proefpersonen (n = 279 in het AF-cohort en n = 233 in het NSR-cohort) als classificeerbaar beschouwd. De analyseset voor de beoordeling van ecg's, gedefinieerd als willekeurig geselecteerde proefpersonen met leesbare gekoppelde stroken van de Ecg-app en afleiding I van de 12-afleidingen-ecg, bestond uit een totaal van 93 proefpersonen (n = 48 in het AF-cohort en n = 45 in het NSR-cohort).

Tabel 6 toont de kruisgeclassificeerde gegevens van de Ecg 2.0-app en beoordeelde resultaten van 12-afleidingen-ecg voor gegevens in rust en na inspanning. Er werd een hoge mate van overeenkomst bereikt tussen de beoordeelde resultaten van de 12-afleidingen-ecg en de output van de Ecg-app voor de categorieën SR en AF. Van de stroken van de 12-afleidingen-ecg die werd beoordeeld als SR (n = 470) werden in totaal 436 resultaten ook door de Ecg-app geclassificeerd als SR. Van de stroken van de 12-afleidingen-ecg die werden beoordeeld als AF (n = 521) werden in totaal 474 resultaten ook door de Ecg-app geclassificeerd als AF. Van de resultaten die op basis van de 12-afleidingen-ecg werden ingedeeld als anders (SVT of anders), classificeerde de Ecg-app er 9 als SR, 37 als AF, 13 als onbeslist en 9 als slechte registratie.

**Tabel 6. Kruisclassificatie van Ecg-app en beoordeelde resultaten van 12-afleidingen-ecg (in rust en na inspanning) – Classificeerbare analyseset**

Ground truth op basis van resultaten beoordeelde 12-afleidingen-ecg						
Output Ecg-app/ apparaat	SR (50≤HF≤150)	AF (50≤HF≤150)	Anders (SVT of anders met HF 50–150; HF<50; HF>50)	Niet- interpreteerbaar	Niet beschikbaar	Totaal
SR [SR (50–99); NSR sinustach (100–150)]	436	7	9	0	2	454
AF (50≤HF≤150)	3	474	37	0	6	520
Onbeslist [inclusief HF<50; HF>150]	10	13	13	0	1	37
Slechte registratie	14	23	9	0	3	49
Niet beschikbaar	7	4	1	1	0	13
Totaal	470	521	69	1	12	1073

Afkortingen: AF = atriumfibrilleren, SR = sinusritme, HF = hartfrequentie, SVT = supraventriculaire tachycardie

De specificiteit van de Ecg 2.0-app werd geschat op 99,3% en de gevoeligheid op 98,5% (tabel 7). De betrouwbaarheidsondergrenzen voor de schatting van de specificiteit (98,4%) en de schatting van de gevoeligheid (97,3%) waren hoger dan de vooraf vastgestelde prestatiedoelen voor deze prestatieparameters. Om die reden werden de nulhypothese voor specificiteit en gevoeligheid verworpen en de alternatieve hypothesen werden voor waar aangenomen. Daarmee werd er voldaan aan de primaire eindpunten voor specificiteit en gevoeligheid. De specificiteit in rust werd geschat op 99,1% en de gevoeligheid in rust op 98,5%. De specificiteit na inspanning werd geschat op 99,5% en de gevoeligheid na inspanning op 98,6%. Er werd een aanvullende analyse uitgevoerd waarbij de gegevens in rust en na inspanning van de classificatiecategorie "onbeslist" van het algoritme werden meegenomen in de berekening. De daaruit voortvloeiende specificiteit werd geschat op 97,1% en de gevoeligheid op 96,0%.

Er werd een hoge mate van overeenkomst gezien tussen de beoordeelde resultaten van de HF op de 12-afleidingen-ecg en de output van de Ecg-app. Er werden weinig onjuiste classificaties gezien bij vergelijking met de resultaten van de referentiestrook van de 12-afleidingen-ecg. In tabel 8 worden de samengevatte resultaten weergegeven die de secundaire eindpunten bevestigen: een correcte classificatie van patiënten met NSR (HF 50–99, SR op gelijktijdige 12-afleidingen-ecg) als NSR, AF (HF 50–99, AF op gelijktijdige 12-afleidingen-ecg) als AF, sinustachycardie (HF 100–150, SR op gelijktijdige 12-afleidingen-ecg) als hoge hartfrequentie en AF met hoge hartfrequentie (HF 100–150, AF op gelijktijdige 12-afleidingen-ecg) als AF (hoge HF) op een leesbare en classificeerbare strook van de Ecg-app. Het percentage SR (HF 50–99) dat correct werd geclassificeerd, bedroeg 98,3%; de betrouwbaarheidsondergrens van 96,6% was hoger dan het vooraf vastgestelde prestatiedoel van 90% voor dit eindpunt. Het percentage AF (HF 50–99) dat correct werd geclassificeerd, bedroeg 98,9%; de betrouwbaarheidsondergrens van 97,5% was hoger dan het vooraf vastgestelde prestatiedoel van 81%. Het

percentage SR (HF 100–150) dat correct werd geclassificeerd, bedroeg 90,7%; de betrouwbaarheidsondergrens van 86,7% was hoger dan het vooraf vastgestelde prestatiedoel van 81%. Het percentage AF (HF 100–150) dat correct werd geclassificeerd, bedroeg 83,0%; de betrouwbaarheidsondergrens van 77,8% was hoger dan het vooraf vastgestelde prestatiedoel van 75,5%.

<b>Tabel 7. Resultaten gevoeligheid en specificiteit (resultaten primaire eindpunten) – classificeerbare analyseset</b>		
Parameter	Waarde	Tweezijdig 95%-betrouwbaarheidsinterval op basis van bootstrapping*
<b>Gecombineerd (in rust en na inspanning)</b>		
Referentieresultaat = AF (HF 50–150)	481	
Ecg-app = AF (HF 50–150)	474/481 (98,5%)	
Ecg-app = SR (HF 50–150)	7/481 (1,5%)	
<b>Gevoeligheid</b>	<b>474/481 (98,5%)</b>	<b>(97,3%, 99,6%)</b>
Referentieresultaat = SR (HF 50–150)	439	
Ecg-app = AF (HF 50–150)	3/439 (0,7%)	
Ecg-app = SR (HF 50–150)	436/439 (99,3%)	
<b>Specificiteit</b>	<b>436/439 (99,3%)</b>	<b>(98,4%, 100,0%)</b>
*Resultaten verkregen door middel van nieuwe steekproef op basis van bootstrapping Afkortingen: AF = atriumfibrilleren, SR = sinusritme, HF = hartfrequentie		

<b>Tabel 8. Percentage correct van classificeerbare AF- en SR-ritmes</b>		
Parameter	Waarde	Betrouwbaarheidsinterval op basis van bootstrapping**
<b>Gecombineerd (in rust en na inspanning)*</b>		
Percentage correct voor SR (HF 50–99) referentie	230/234 (98,3%)	(96,6%, 99,6%)
Percentage correct voor AF (HF 50–99) referentie	272/275 (98,9%)	(97,5%, 100%)
Percentage correct voor SR (HF 100–150) referentie	186/205 (90,7%)	(86,7%, 94,6%)
Percentage correct voor AF (HF 100–150) referentie	171/206 (83,0%)	(77,8%, 88,0%)
*Percentage correcte berekeningen inclusief door Ecg 2.0-app classificeerbare AF- en SR-ritmes **Betrouwbaarheidsinterval verkregen door middel van nieuwe steekproef op basis van bootstrapping Afkortingen: AF = atriumfibrilleren, HF = hartfrequentie, SR = sinusritme		

Daarnaast bleek uit de vergelijking van de ecg van het algoritme van de Ecg 2.0-app met afleiding I van een 12-afleidingen-ecg dat het totale goedkeuringspercentage voor de gecombineerde resultaten van de morfologie in rust en na inspanning 100% was, waarbij in totaal vijf resultaten werden uitgesloten. De betrouwbaarheidsondergrens was 97,9%, hoger dan het vooraf vastgestelde goedkeuringspercentage voor de morfologie van 80%. Het totale goedkeuringspercentage voor gecombineerde beoordelingsresultaten



van de R-golf in rust en na inspanning was 97,2%. De betrouwbaarheidsondergrens was 93,6%, hoger dan het vooraf vastgestelde overeenkomstpercentage van 80% voor de amplitude van de R-golf. De groepen in rust en na inspanning hadden een vergelijkbare R-golf.

## Conclusies

Apple Watch-klanten beschikken nu over twee optionele functies voor de detectie van onregelmatige hartritmes: de functie voor melding van een onregelmatig ritme (beschikbaar op Apple Watch Series 1 en nieuwer) en de Ecg-app (beschikbaar op Apple Watch Series 4, Series 5 en Series 6).

De functie voor melding van een onregelmatig ritme is software die opportunistisch verzamelde tachogrammen op de achtergrond classificeert. Gebruikers die de functie hebben geactiveerd, ontvangen een melding wanneer een onregelmatig hartritme wordt vastgesteld. De Apple Heart Study toonde aan dat 78,9% van de deelnemers die bij het gelijktijdig dragen van een Apple Watch en een ecg-patch een melding ontvingen, gelijktijdig AF vertoonden op de ecg-patch. Daarnaast vertoonde 98,2% AF en andere klinisch relevante ritmestoornissen. Deze resultaten tonen aan dat de melding in de meeste gevallen nauwkeurig de aanwezigheid van AF weergeeft, maar in sommige gevallen kan duiden op de aanwezigheid van een andere aritmie dan AF.

Apple Watch Series 4, Series 5 en Series 6 hebben een elektrische hartslagsensor die, samen met de Ecg-app en het algoritme, een ecg genereert dat vergelijkbaar is met een ecg met één afleiding (afleiding I) om AF te detecteren. Het bedrijfseigen algoritme dat deze ecg's classificeert, had bij het vaststellen van AF een gevoeligheid van meer dan 98% en een specificiteit van meer dan 99% vergeleken met ecg's die werden opgenomen met een referentieapparaat en werden geïnterpreteerd door onafhankelijke klinische experts.

Klanten die deze twee features willen gebruiken, moeten een procedure voltooien waarin ze informatie krijgen over de interpretatie en de niet-diagnostische aard van deze bevindingen en de beperkingen van de algoritmen. Beide functies zijn door de FDA goedgekeurd met De Novo- of 510K-classificatie.