



# Använda Apple Watch för arytmiidentifiering

December 2020

# Innehåll

<b>Översikt</b>	<b>3</b>
Inledning .....	3
<b>PPG-baserad identifiering av arytm</b>	<b>3</b>
Teknisk beskrivning och funktionsbeskrivning .....	3
Preklinisk utveckling.....	5
Klinisk verifiering .....	5
Apples hjärtstudie.....	5
Experimentupplägg för AHS-delstudie.....	5
Resultat.....	6
<b>EKG-baserad identifiering</b>	<b>6</b>
Teknisk beskrivning och funktionsbeskrivning .....	6
Preklinisk utveckling.....	7
Klinisk verifiering .....	8
Experimentets upplägg.....	8
Klinisk verifieringsstudie av EKG 1.0 .....	8
Resultat.....	8
Klinisk verifieringsstudie av EKG 2.0 .....	12
Resultat .....	14
<b>Slutsatser</b>	<b>17</b>

# Översikt

Apple Watch-kunder har tillgång till två program i den medicintekniska apparaten som hjälper till att identifiera hjärtarytmier som förmaksflimmer (AF): funktionen för notiser om oregelbunden rytm och EKG-appen.

På Apple Watch Series 1 eller senare analyserar funktionen för notiser om oregelbunden rytm pulldata som samlas in av fotopletysmografens (PPG) sensor för att identifiera episoder av oregelbunden puls som tyder på AF, och meddelar användaren när den identifierar en sådan episod. På Apple Watch Series 4, Series 5 och Series 6 kan EKG-appen producera en EKG-vågform (elektrokardiogram) som liknar ett avledning I-EKG, och sedan klassificera vågformen som sinusrytm (SR), förmaksflimmer (AF), hög eller låg puls, eller osäkert resultat. Med EKG 2.0 är ytterligare klassificeringar av AF med hög puls och dålig inspelning tillgängliga.

I detta blad finns en utförlig beskrivning av dessa funktioner, inklusive testning och verifiering.

## Inledning

AF, en typ av oregelbunden hjärtrytm där hjärtats förmak slår oregelbundet och ibland snabbt, är en av de vanligaste orsakerna till stroke. Men eftersom AF ofta är asymtomatiskt är många personer med AF inte medvetna om att de har detta tillstånd. Kombinationen av risken för stroke, det asymtomatiska tillståndet, effektiva farmakologiska behandlingar som minimerar risken för stroke och den ökade förekomsten av konsumentapparater som kan identifiera AF har ökat intresset för tidig identifiering av AF utanför sjukhuset.

Med watchOS 5.1.2 eller senare kan Apple Watch Series 1 och senare använda PPG-signaler kombinerat med en algoritm för att identifiera perioder av oregelbunden puls som kan tyda på AF. Utöver denna PPG-baserade algoritm har Apple Watch Series 4, Series 5 och Series 6 även en elektrisk hjärtsensor som, när EKG-appen används, gör det möjligt att producera och analysera ett EKG som liknar ett avledning I-EKG.

## PPG-baserad identifiering av arytmi

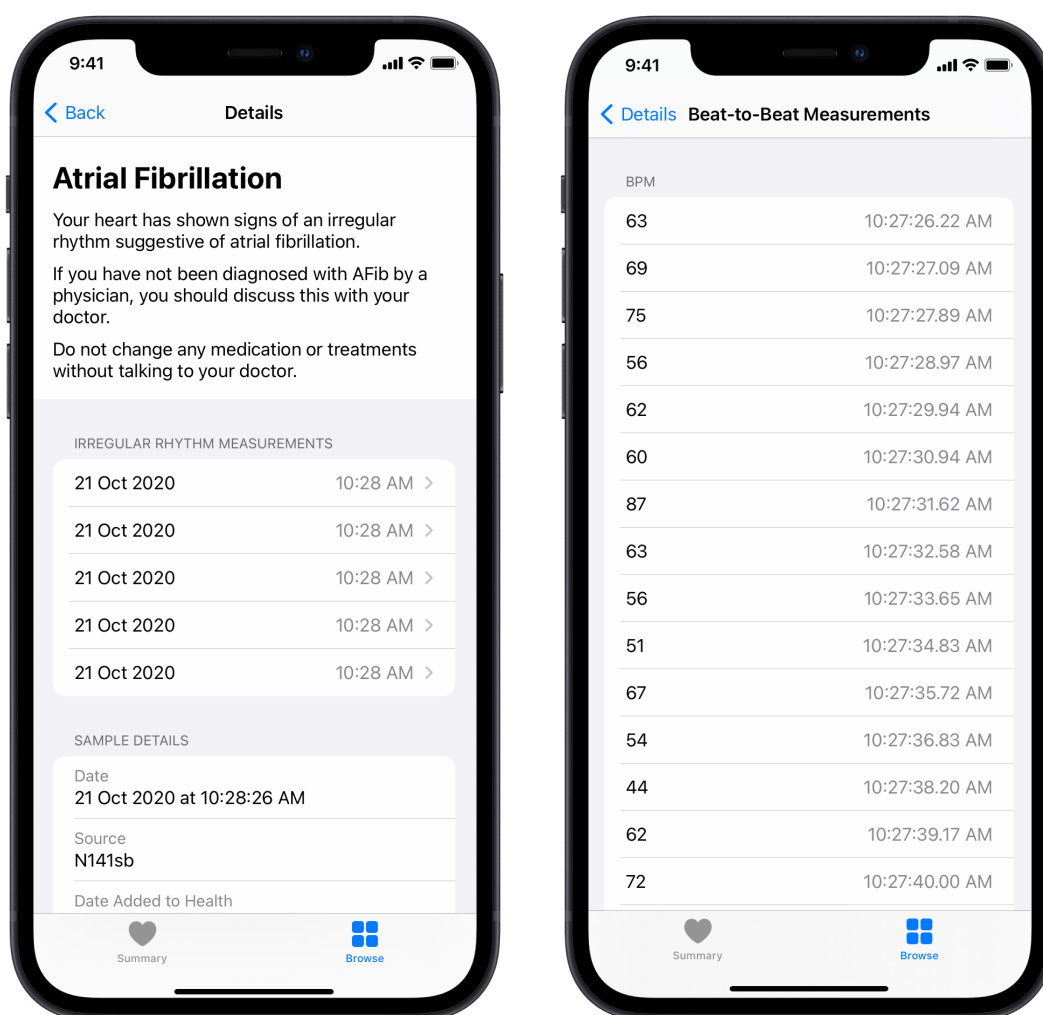
### Teknisk beskrivning och funktionsbeskrivning

Apple Watch har en optisk hjärtsensor som använder gröna LED-lampor parkopplade med ljuskänsliga fotodioder för att identifiera blodvolympulser i användarens handled med fotopletysmografi. Dessa sensorer och underliggande algoritmer är grunden för identifieringen av puls (HR) och pulsvariabilitet (HRV) som finns på Apple Watch Series 1 och senare. För att fastställa HRV tar Apple Watch ett takogram – en graf över tiden mellan hjärtslag – varannan till var fjärde timme. Från och med watchOS 5.1.2 kan användaren även välja att aktivera en funktion för arytmiidentifiering som använder dessa takogram. För att använda funktionen för notiser om oregelbunden rytm på Apple Watch måste användaren först gå igenom introduktionen till appen Hälsa på den iPhone som är parkopplad, och lära sig använda funktionen och få information om AF. Mer information om användarupplevelsen finns på [support.apple.com/sv-se/HT208931](https://support.apple.com/sv-se/HT208931).

Om den PPG-baserade arytmidentifieringen är aktiverad klassificerar Apple Watch varje takogram med en upphovsrättsskyddad algoritm för att fastställa om det finns en oregelbunden rytm. Ett oregelbundet takogram inleder en kaskad av mer frekvent takograminsamling – så ofta som möjligt, med minst

15 minuters mellanrum – och analys. Apple Watch samlar endast in och analyserar takogram om användaren är tillräckligt still för att utföra en mätning. Därför övervakar algoritmen inte användaren ständigt, utan gör det snarare opportunistiskt när det finns tillräcklig signal för insamling och analys. Om fem av sex på varandra följande takogram – inklusive det första – klassificeras som oregelbundna inom en 48-timmarsperiod meddelas användaren om den potentiella arytmien. Utöver att användaren får denna notis har han eller hon tillgång till mer information förknippad med dessa oregelbundna takogram i appen Hälsa (bild 1). Om två takogram klassificeras som icke oregelbundna innan tröskeln har nåtts återställs cykeln och återgår takograminsamlingen till den vanliga frekvensen varannan timme.

**Bild 1: Skärm med mätning av oregelbunden hjärtrytm i appen Hälsa**



I appen Hälsa kan användare se de gånger då algoritmen identifierat ett oregelbundet takogram som bidrog till en notis (vänster). Genom att trycka på ett datum och tid kan användaren se slag-till-slag-mätningarna som beräknats från varje takogram.

## Preklinisk utveckling

Före klinisk testning utfördes studier för att utveckla den PPG-baserade identifieringsalgoritmen och utvärdera algoritmens prestation under olika förhållanden och användarbeteenden. Dessa innefattade djupa andetag, bilfärd, skakningar och rörelser i händerna, minskad perfusion i handen eller handleden, användning på natten, snabb ventrikulär respons hos personer med AF samt andra arytmier. Dessa studier utfördes med 2 300 kontrollpersoner och mer än 500 personer med AF.

Eftersom PPG förlitar sig på förmågan att absorbera ljus testades algoritmen för arytmidentifiering på flera olika hudtyper och hudtoner för att säkerställa att justeringar av sensorplattformen baserat på hudton var tillräckliga för att identifiera arytmier. Melanin har en hög absorptionsförmåga vid våglängden för den gröna LED-lampan på Apple Watch, vilket potentiellt försvårar PPG-pulsmätningen på mörkare hudtoner. För att korrigera för detta justerar Apple Watch-sensorplattformen LED-strömmen (och på så sätt ljusflödet), fotodiodförstärkningen (känslighet för ljus) och samplingsfrekvensen för att säkerställa adekvat signalamplitud för alla hudtoner.

I verifieringssyfte analyserades 1,3 miljoner takogram från 1 124 personer (51 % kvinnor) med olika hudtyper och hudtoner (baserat på Fitzpatrick-hudtyper och spektrofotometermätningar av hudfärgen på handleden). Eftersom det var viktigt att verifiera signalamplituden hos personer med mörk hud hade nästan 5 % av de medverkande personerna Fitzpatrick-hudtyp VI. Verifieringen uppvisade ingen signifikant skillnad i algoritmens sensitivitet eller specificitet för olika hudtyper eller hudtoner.

## Klinisk verifiering

### Apples hjärtstudie

Apples hjärtstudie (AHS) är en prospektiv, enarmad pragmatisk studie som genomfördes virtuellt för att utvärdera förmågan hos den Apple Watch-baserade algoritmen att identifiera arytmier som tyder på AF. Om en deltagare i studien uppfyllde fem-av-sex-tröskeln fick han eller hon en iPhone- och Apple Watch-notis och hade möjligheten att kontakta en studieläkare via telefon och få ePatch, ett ambulans EKG-plåster (från BioTelemetry, Inc. i Conshohocken, Pennsylvania, USA). Deltagarna fick anvisningen att bära ePatch i upp till sju dagar, men data som samlades in från en deltagare ansågs adekvata om den analyserbara tiden var minst en timme.

De detaljerade resultaten från AHS publicerades i november 2019 i *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).

### Experimentupplägg för AHS-delstudie

En delstudie med data som samlades in i AHS genomfördes för att fastställa om takogramklassificeringsalgoritmen (individuellt eller punkttakogram) och bekräftelsecykelalgoritmen (varningsnivå, fem av sex takogram) har ett godtagbart positivt prediktivt värde (PPV) jämfört med ePatch-övervakningen beträffande identifiering av oregelbunden hjärtrytm som överensstämmer med AF hos en undergrupp av AHS-deltagarna. AHS-prövarna var kända till delstudien, efterföljande analyser och datainlämningen till FDA, men var blindade för delstudiens resultat medan AHS pågick. Den institutionella granskningsnämnden (IRB) som godkände AHS fastställde att denna delstudie var befriad från IRB-granskning. Alla AHS-deltagare gav informerat samtycke som innefattade användningen av deras studiedata i delstudien.

Data till delstudien samlades in från AHS-deltagare som skrevs in mellan 30 november 2017 och 22 juni 2018. Personerna i delstudien fick en notis om oregelbunden hjärtrytm från AHS-appen och fick därefter en ePatch som de bar för tolkning av de ambulanta EKG-fynden. De initiala oregelbundna takogrammen som ledde till den första notisen och potentiellt initierade det första hälsomötet via telefon analyserades inte som en del av delstudien. Endast oregelbundna takogram och notiser som uppträdde medan användaren bar den ePatch som tillhandahållits av studien analyserades.

Två oberoende EKG-granskare med legitimation inom kardiologi och elektrofysiologi i USA granskade och utvärderade EKG-remsorna och klassificerade dem som SR, AF, en annan oregelbunden rytm eller oläsbara. Om granskarna inte var överens utvärderade en tredje granskare med liknande kvalifikationer remsan. Dessa granskare var blindade för takogramklassificeringen. Granskarnas EKG-klassificering och takogramklassificeringen som fastställts av algoritmen skickades på ett säkert sätt till studiens statistiker för dataanalys.

### Resultat

Av delstudiens 226 deltagare som fick en initial arytmnotis och bar ePatch i ungefär en vecka hade 41,6 % (94 deltagare) AF som identifierades av ePatch. När en Apple Watch och ePatch bars samtidigt fick 57 av 226 deltagare en notis om AF – det vill säga, fem av sex på varandra följande takogram klassificerades som oregelbundna. Av dessa 57 visade 78,9 % (45 deltagare) överensstämmande AF på ePatch-data och 98,2 % (56 deltagare) AF eller andra kliniskt relevanta arytmier. Dessa resultat visar att även om notisen i majoriteten av fallen korrekt anger förekomsten av AF kan en notis i vissa fall ange förekomsten av en annan arytm än AF. Inga allvarliga biverkningar observerades.

## EKG-baserad identifiering

### Teknisk beskrivning och funktionsbeskrivning

Apple Watch Series 4, Series 5 och Series 6 har en titanelektrod i Digital Crown och ett lager av ultratunn krom-kisel-kolnitrid på safirkristallen på baksidan av Apple Watch. EKG-appen avläser och registrerar de elektriska impulserna som styr hjärtat från användarens fingerspets (med Digital Crown) och handleden (med baksidan av Apple Watch), vilket skapar en sluten krets. För att använda EKG-appen på Apple Watch måste användaren först gå igenom introduktionen till appen Hälsa på den iPhone som är parkopplad, och lära sig använda funktionen och få information om förmaksflimmer. För att producera ett EKG öppnar användaren EKG-appen som är installerad på Apple Watch och lägger sedan ett finger – från handen på den sida som inte bär Apple Watch – på Digital Crown i 30 sekunder. Avledningspolariteten fastställs av placeringen av Apple Watch på handleden och väljs i Inställningar.

När EKG:t har tagits klassificerar en upphovsrättsskyddad algoritm EKG-kurvan som SR, AF eller osäkert resultat i EKG 1.0. Där EKG 2.0 är tillgänglig är även ytterligare klassificeringar som AF med hög puls (HR 100–150) och ytterligare differentiering mellan dålig registrering och osäkra resultat tillgängliga. Dessa rytmklassificeringar – genomsnittlig HR, användarrapporterade symtom och vågform – läggs till i appen Hälsa och sparas och kan delas av användaren som en PDF-fil från appen på användarens parkopplade iPhone. Mer information om användarupplevelsen finns på [support.apple.com/sv-se/HT208955](https://support.apple.com/sv-se/HT208955).

## Preklinisk utveckling

Detekteringen av EKG-signalen och klassificeringsalgoritmen testades även i flera studier innan klinisk verifiering påbörjades. Sensorerna och klassificeringsalgoritmen testades för olika etniciteter, handledsomkretsar, BMI-intervall, åldrar, arytmier som inte var AF, åtdragning av armbandet, hållningar och motionstillstånd och svettning. Cirka 2 500 personer deltog i dessa tester. Ungefär 25 procent av dem hade tidigare fått diagnosen AF eller andra oregelbundna hjärtrytmer.

Ökad frekvens av "oläsbara" EKG:n var den huvudsakliga variationen i algoritmprestationen. Faktorerna som ledde till denna variation var låg signalamplitud (som ett resultat av högeraxelavvikelse – i synnerhet observerad hos personer med lågt BMI – eller svettning som observerades under testningen efter motion) och rörelseartefakter orsakade av användarens beteende. Apple Watch använder torra elektroder som är mekaniskt starka och korrosionsbeständiga på ett sätt som är adekvat för bärbara enheter. Men torra elektroder, särskilt de som placeras på armar eller ben, är mer benägna att producera brus – till exempel sådant brus som beskrivs ovan – jämfört med tillfälliga gelelektroder som används i kliniska enheter.

Utöver faktorerna som nämns ovan resulterade förekomsten av vissa arytmier som inte var AF även i signifikant annorlunda algoritmprestation jämfört med personer med SR. Dessa tillstånd och resultaten anges i tabell 1 nedan.

**Tabell 1. Arytmier som inte är AF som påverkar algoritmens prestation (EKG 2.0)**

Arytmi	Variation
Vänster- eller högergrenblock	7,9 % av försök klassificerade som AF
AV-block första graden	10,2 % av försöken oklassificerade, 2,4 % klassificerade som AF av klassificerade försök
Bigemini	92,5 % oklassificerade
Frekventa PVC-slag	48,8 % oklassificerade, 24,1 % klassificerade som AF av klassificerade
Frekventa PAC-slag	19,8 % oklassificerade, 23,5 % klassificerade som AF av klassificerade
Förmakstakykardi	7,7 % klassificerade som AF
Blandad puls	46,5 % oklassificerade, 29,7 % klassificerade som AF av klassificerade
Hög/låg HR (utanför 50–150 bpm)	94,8 % oklassificerade

# Klinisk verifiering

## Experimentets upplägg

Apple finansierade två multicenterstudier för att verifiera EKG-appens förmåga att (1) producera en EKG-vågform som liknar ett avledning I-EKG från ett vanligt 12-avlednings-EKG och (2) använda en rytmklassificeringsalgoritm för att klassificera detta 1-avlednings-EKG som antingen SR eller AF.

## Klinisk verifieringsstudie av EKG 1.0

En klinisk verifieringsstudie genomfördes för att verifiera prestationen för EKG-appen 1.0. Studiens primära effektmått var rytmklassificeringsalgoritmens sensitivitet beträffande identifiering av AF och specificitet beträffande identifiering av SR. En extern IRB godkände protokollet, formuläret för informerat samtycke och allt annat relevant material före inskrivning av deltagare, och alla deltagare gav skriftligt samtycke till att delta innan de skrevs in i studien.

Både personer med känt AF och utan några kända rytmavvikelser deltog i studien. De registrerade tre 1-avlednings-EKG:n med EKG-appen samtidigt som studiepersonalen registrerade tre 12-avlednings-EKG:n med en klinisk apparat (GE Healthcare CardioSoft EKG-enhet) som godkännts av FDA. De första försöken utvärderades och analyserades. Deltagarna fick hjälp med placeringen av Apple Watch och fick anvisningen att hålla armarna stilla – eventuellt genom att vila armarna på ett bord eller på benen – och fick öva på att registrera ett EKG innan testet påbörjades.

För att testa (1) la tre oberoende legitimerade EKG-tekniker de producerade rytmremorna från 140 slumpmässigt utvalda personer (70 med AF och 70 med SR) över motsvarande avledning I-remsor från rytmremorna som producerats med den kliniska apparaten för att visuellt jämföra morfologin hos sex på varandra följande PQRST-komplex. EKG-teknikerna tilldelade varje remsa beteckningen "godkänd" eller "underkänd" baserat på visuellt bedömda morfologiska likheter. De mätte även R-vågsamplituden från isoelektrisk baslinje till närmaste millimeter för de första två QRS-komplexen i både referensremorna och remorna som producerats av EKG-appen, och utvärderade överensstämmelsen mellan dem.

För (2) granskade tre blindade, oberoende legitimerade kardiologer varje referensremsa från 12-avlednings-EKG och klassificerade rytmen som SR, AF, annan (allt som inte var SR eller AF inom HR-parametrarna) eller oläsbar (en diagnos kunde inte ställas eftersom remsan inte kunde avläsas). EKG-appens algoritm klassificerade EKG:t som producerats av EKG-appen som SR, AF, ej klassificerbar eller oläsbar. Sensitiviteten och specificiteten för EKG-appens klassificering av SR och AF (för klassificerbara EKG:n) jämförts med kardiologernas tolkningar av 12-avlednings-EKG:t beräknades. En blindad, oberoende, legitimerad kardiolog klassificerade sedan rytmremorna som producerats av EKG-appen enligt samma kategorier.

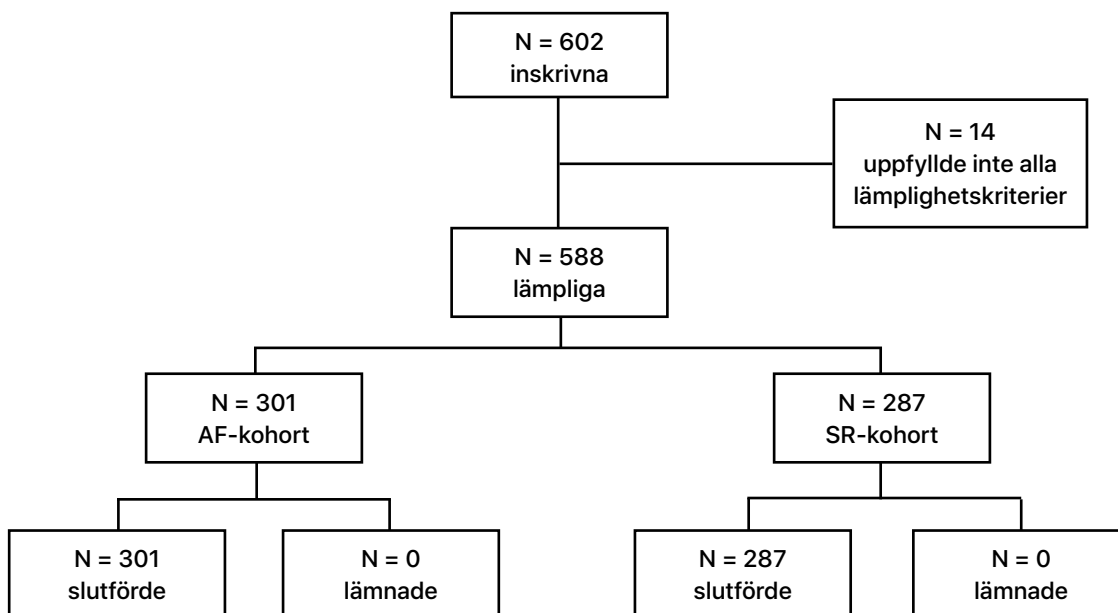
För analys av primärt effektmått beräknades en ensidig exakt undre gräns i ett 97,5 % konfidensintervall separat för sensitivitet och specificitet. Om den undre gränsen för sensitivitet överskred 90 % avvisades nollhypotesen till förmån för sensitiviteten som överskred 90 %. Om den undre gränsen för specificitet överskred 92 % avvisades nollhypotesen till förmån för specificiteten som överskred 92 %.

## Resultat

Studien rekryterade 602 personer, och 588 med lämplighetskriterier. Av dessa 588 tilldelades 301 personer med självrapporterat AF till AF-kohorten, och 287 personer utan självrapporterat AF till SR-kohorten. Dessa kohorttilldelningar användes endast för att säkerställa adekvat rekrytering. Utvärderarna var blindade beträffande kohorterna och förekomsten eller frånvaron av AF baserades enbart på EKG:t som erhöles under testning. De 14 personer som slutförde studien men inte tilldelades till någon rekryteringskohort var olämpliga för studiedeltagande på grund av tidigare paroxysmalt AF utan AF på EKG vid tidpunkten för screening. Alla lämpliga deltagare slutförde studien (bild 2). Inga biverkningar rapporterades under studien.



**Bild 2: Flödesschema över deltagarfördelning**



Tre oberoende legitimerade EKG-tekniker hittade visuell morfologisk likvärdighet mellan EKG- appens vågform och avledning I-EKG:t som användes som referens och som tagits med den vanliga kliniska apparaten för 98,4 % av analyserade remsor i AF-kohorten och 100 % i SR-kohorten (tabell 2). Den totala andelen av personer med beteckningen "godkänd" var 99,2 % (undre gräns i 97,5 % konfidensintervall = 95,7 %). Remsor uteslöts om sex på varandra följande slag (PQRST-komplex) utan artefakt inte kunde identifieras i någon av remsorna (EKG-app eller referens).

**Tabell 2. Jämförelse av vågform**

Egenskaper	Personer med AF (N = 61)	Personer med SR (N = 65)	Totalt (N = 126)	Undre gräns i konfidensintervall*	P-värde**
Antal parade remsor (EKG-app och referensremsor) med beteckningen "godkänd"	60	65	125		
Antal läsbara parade remsor (EKG-app och referensremsor)	61	65	126		
Andel av remsor med beteckningen "godkänd"	60/61 (98,4 %)	65/65 (100 %)	125/126 (99,2 %)	95,7 %	< 0,0001
Antal uteslutna parade remsor	8	5	13		

\*Undre exakt binomial gräns i ensidigt 97,5 % konfidensintervall för total

\*\*Test av hypotes för deltagarframgång > 0,8

Förkortningar: AF = förmaksflimmer, SR = sinusrytm

För ytterligare bekräftelse på att vågformerna som producerades av EKG-appen och referensapparaten var liknande jämfördes klassificeringen som utförts av en blindad kardiolog av EKG-appens remsor med

kardiologklassificering av referensremorna (tabell 3). Andelen överensstämmelse mellan klassificeringen av enhetsremsan och AF- och SR-referensresultaten var 100 % respektive 99,1 %. Oläsbara remсор inkluderades inte i denna analys.

**Tabell 3. Klassificeringar mellan EKG-appens remсор och referensremсор**

Egenskaper	Totalt (N = 522)
Slutligt EKG-referensresultat = AF	263
Klassificering av EKG-apprensa = AF	239/263 (90,9 %)
Klassificering av EKG-apprensa = SR	0/263 (0,0 %)
Klassificering av EKG-apprensa = annan	0/263 (0,0 %)
Klassificering av EKG-apprensa = oläsbar	24/263 (9,1 %)
<b>% överensstämmelse med AF-referensresultat*</b>	<b>239/239 (100,0 %)</b>
Slutligt EKG-referensresultat = SR	244
Klassificering av EKG-apprensa = AF	0/244 (0,0 %)
Klassificering av EKG-apprensa = SR	232/244 (95,1 %)
Klassificering av EKG-apprensa = annan	2/244 (0,8 %)
Klassificering av EKG-apprensa = oläsbar	10/244 (4,1 %)
<b>% överensstämmelse med SR-referensresultat*</b>	<b>232/234 (99,1 %)</b>
Slutligt EKG-referensresultat = annat	15
Klassificering av EKG-apprensa = AF	0/15 (0,0 %)
Klassificering av EKG-apprensa = SR	3/15 (20,0 %)
Klassificering av EKG-apprensa = annan	12/15 (80,0 %)
Klassificering av EKG-apprensa = oläsbar	0/15 (0,0 %)
<b>% överensstämmelse med annat referensresultat*</b>	<b>12/15 (80,0 %)</b>

\*Oläsbara remсор uteslöts

Förkortningar: AF = förmaksflimmer, SR = sinusrytm

Totalt 485 av 602 parade rytmremсор från EKG-appen och referensapparaten bedömdes vara klassificerbara. De återstående paren hade EKG-appremсор eller referensremсор som bedömdes vara oläsbara eller ej klassificerbara. I tabell 4 visas indelningen i AF- och SR-kohorterna.

**Tabell 4. Slutliga resultat av klassificering med EKG-appens algoritm och referensremсор**

Klassificering med EKG-appens algoritm	Klassificering med referensremсор				
	SR	AF	Annat	Oläsbar	Totalt
SR	238	4	4	1	247
AF	1	236	2	2	241
Ej klassificerbar	6	7	6	0	19
Oläsbar	18	30	1	0	49
Inget rapporterat enhetsresultat*	32	13	1	0	46
<b>Totalt</b>	<b>295</b>	<b>290</b>	<b>14</b>	<b>3</b>	<b>602</b>

\*Resultat inte rapporterade baserat på angivna kriterier (till exempel ingen synkronisering detekterad) för alla utom en person  
Förkortningar: AF = förmaksflimmer, SR = sinusrytm

Klassificeringen med EKG-appens algoritm uppnådde 98,3 % sensitivitet och 99,6 % specificitet (tabell 5). När analysen utökas att inkludera de 2,4 % (7 av 290) och 2 % (6 av 295) av remsorna som bedömdes vara ej klassificerbara av enheten i klassificeringskategorierna AF- respektive SR-referensremсор var sensitiviteten 95,5 % (95 % KI: 92,2 %, 97,8 %) och specificiteten 97,1 % (95 % KI: 94,2 %, 98,8 %). Dessa resultat uppfyllde de primära effektmåtten som angetts i studiens upplägg. Dessutom hade 12,2 % (68 av 556) av registreringarna ett osäkert resultat, det vill säga var oläsbara eller ej klassificerbara, och kunde varken klassificeras som SR eller AF. När registreringar med osäkert resultat inkluderades i analysen klassificerade EKG-appen SR korrekt hos 90,5 % (238 av 263) av personer med SR, och AF korrekt hos 85,2 % (236 av 277) av personer med AF. De kliniska verifieringsresultaten återspeglar användning i en kontrollerad miljö. Faktisk användning av EKG-appen kan leda till att ett större antal remsor bedöms ge osäkra resultat och ej klassificerbara.

**Tabell 5. Analys av sensitivitet och specificitet – klassificerbara remsor**

Parameter	Värde	Undre gräns i konfidensintervall*	P-värde**
Slutligt EKG-referensresultat = AF (n)	240		
EKG-appenhetens resultat = AF	236/240 (98,3 %)		
EKG-appenhetens resultat = SR	4/240 (1,7 %)		
<b>Sensitivitet</b>	236/240 (98,3 %)	95,8%	< 0,0001
Slutligt EKG-referensresultat = SR (n)	239		
EKG-appenhetens resultat = AF	1/239 (0,4 %)		
EKG-appenhetens resultat = SR	238/239 (99,6 %)		
<b>Specificitet</b>	238/239 (99,6 %)	97,7%	< 0,0001

\*Undre exakt binomial gräns i ensidigt konfidensintervall

\*\*Test av hypotes för sensitivitet > 0,9 och specificitet > 0,92

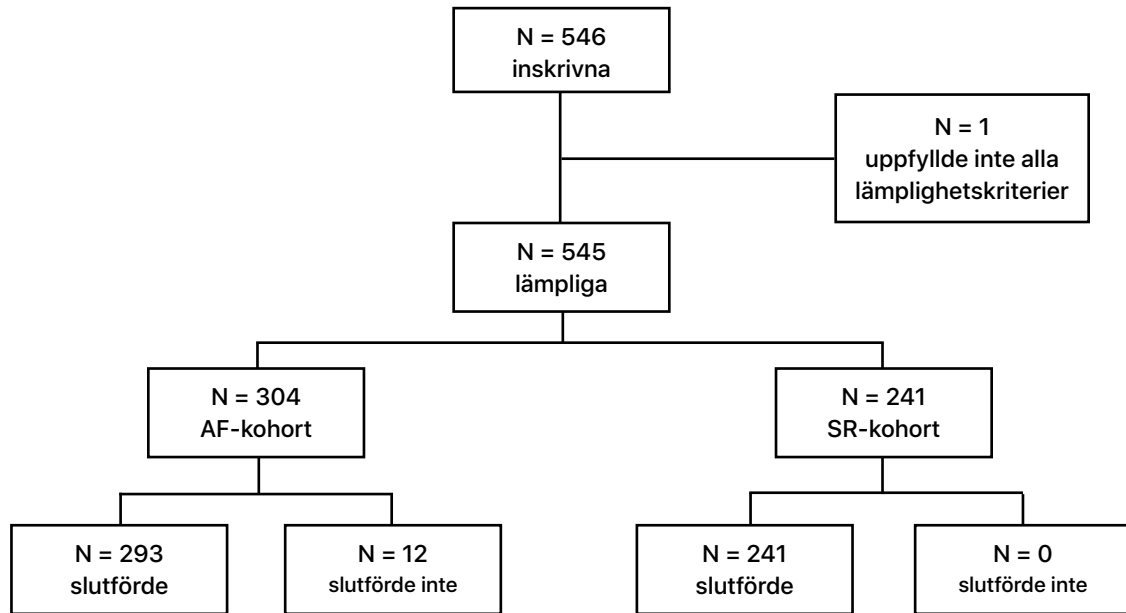
Förkortningar: AF = förmaksflimmer, SR = sinusrytm

### Klinisk verifieringsstudie av EKG 2.0

En andra studie genomfördes för att stödja och verifiera algoritmerna som undersöks i EKG-app 2.0 (testenhet), som utökar det klassificerbara HR-intervallet (50–150 bpm) och inför nya klassificeringsresultat (SR, SR med hög HR, AF, AF med hög HR, osäkert resultat samt dålig registrering). Syftet med denna studie var att utvärdera testenhetens prestation. De primära effektmått som utvärderades var specificitet och sensitivitet. Sekundära effektmått innefattade korrekt klassificering av följande kategorier av personer på specifika läsbara och klassificerbara EKG-testremsor: NSR (normal sinusrytm) (HR 50–150), SR på samtidigt 12-avlednings-EKG som SR, AF (HR 50–99), AF på samtidigt 12-avlednings-EKG som AF, sinustakykardi (HR 100–150), SR på samtidigt 12-avlednings-EKG som hög puls, och AF med hög puls (HR 100–150), AF på samtidigt 12-avlednings-EKG som AF. Dessutom utvärderades även likvärdigheten för EKG-appens vågform och avledning I från 12-avlednings-EKG – uppmätt med godtagbar överenskommelse i morfologin för PQRST-komplex och R-vågsamplitud – som liknade vågformsutvärderingen som utfördes i den kliniska verifieringsstudien av EKG 1.0.

Genomförandet av denna studie godkändes av tillämplig institutionell granskningsnämnd för respektive studieklinik i den prospektiva multicenterstudien som genomfördes i USA. Prövarna erhöll IRB-godkännande, informerat samtycke och eventuellt material till deltagarna på varje studieklinik före deltagande i studien.

**Bild 3: Flödesschema över deltagarfördelning**



Studiedeltagarna innefattade personer med normal SR vid tidpunkten för screening – utan tidigare känt AF – och personer med känt ihållande, permanent eller kroniskt AF som upplevde AF under screening. Alla personer fick anvisningen att samtidigt ta ett 12-avlednings-EKG och ett separat fullständigt 1-avlednings-EKG med Apple Watch. Tre försök genomfördes för varje deltagare. Samma procedur följdes för motionssessionerna. Personer som ansågs lämpliga ombads att cykla i fem minuter på en motionscykel för att uppnå målpulsen. De första försöken av både vilo- och motionssessionerna utvärderades och analyserades.

Två blindade, oberoende legitimerade kardiologer granskade 12-avlednings-EKG:t för att ställa en HR- och rytm diagnos. Om deras utvärdering skiljde sig åt genomförde en tredje granskare en granskning. HR beräknades för varje 12-avlednings-EKG. HR registrerades och HR-diagnoskoden som motsvarar den HR som observerades på referens-EKG:t valdes. Följande hjärtrytm diagnoser utvärderades på 12-avlednings-EKG: SR, AF, supraventrikulär takykardi (SVT), någon annan avvikande rytm (frekventa prematura förmakskontraktioner (PAC), frekventa prematura kammarkontraktioner (PVC), förmaksfladder, kammartakykardi, kammarflimmer, AV-block andra graden typ I, AV block andra graden typ II, AV block tredje graden, och annan) samt omöjlig att tolka. Tre blindade EKG-tekniker eller kardiologer granskade vågformen under utvärderingen av EKG-data från de parade EKG-remsorna. De första sex på varandra följande distinkt läsbara PQRST-komplexen utan artefakt med en match mellan personens enhetsremsa och referensremsan för utvärdering identifierades av en granskare och användes av de två andra granskarna. Remsorna uteslöts om sex på varandra följande slag inte kunde hittas.

För analys av primärt effektmått användes en bootstrap-metod för att erhålla dubbelsidiga 95 % konfidensintervall för sensitivitet och specificitet eftersom data samlades in från samma personer i vila och efter motion. Personer med minst ett utvärderat resultat med AF (för sensitivitet) eller SR (för specificitet) och med ett klassificerbart algoritmrresultat (SR eller AF) valdes ut slumpmässigt med ersättning. 2,5:e och 97,5:e percentilen hos fördelningen av bootstrap-uppskattningar representerade de dubbelsidiga 95 % konfidensintervallen. Om de undre gränserna i konfidensintervallen för både sensitivitet och specificitet överskred det angivna prestationsmålet förknippat med dessa prestationsmätningar (90 % för sensitivitet och 92 % för specificitet) avvisades nollhypotesen till förmån för den alternativa hypotesen. För jämförelsen av vågform användes data från den första försöksrundan med de slumpmässigt utvalda personerna – med deltagarna i vila och efter motion – för att testa hypotesen för vågformsutvärderingen med en bootstrap-samlingsmetod. Om den 2,5:e percentilen i bootstrap-fördelningen för andelen med "godkänd" morfologi överskred 80 % avvisades nollhypotesen (andelen med "godkänd" morfologi var lägre än 80 %). Samma kriterium tillämpades på utvärderingen av andelen överensstämmande R-vågsamplitud.

## Resultat

Totalt 546 personer skrevs in i studien. 305 skrevs in i AF-kohorten och 241 skrevs in i NSR-kohorten (normal sinusrytm). En person i AF-kohorten uteslöts av lämplighetsskäl. Av de 546 personerna slutförde 534 (293 i AF-kohorten och 241 i NSR-kohorten) studien. 12 personer i AF-kohorten slutförde inte studien.

Tre olika EKG-tekniker granskade analysuppsättningen för vågformsutvärderingen för att säkerställa överensstämmelse. Parade Watch- och referensremsor användes för utvärderingen. Totalt 91 personer (100 %) hade en läsbar ihopparring av remsor "i vila", 87 personer (100 %) hade en läsbar ihopparring av remsor "efter motion" och 93 personer (100 %) hade minst en ihopparring av remsor från "i vila" eller "efter motion".

Den klassificerbara analysuppsättningen bestod av alla personer som hade läsbara parade utvärderade resultat från EKG-app 2.0 och 12-avlednings-EKG. Totalt 512 personer (n=279 i AF-kohorten och n=233 i NSR-kohorten) bedömdes vara klassificerbara. Analysuppsättningen för vågformsutvärdering – definierad som slumpmässigt utvalda personer med läsbara parade remsor från EKG-appen och avledning 1 från 12-avlednings-EKG:t – bestod av totalt 93 personer (n=48 i AF-kohorten och n=45 i NSR-kohorten).

I tabell 6 anges korsklassificerade data för EKG-app 2.0 och de utvärderade 12-avlednings-EKG-resultaten för kombinerade data i vila och efter motion. En hög grad av överensstämmelse mellan de utvärderade 12-avlednings-EKG-resultaten och EKG-appenhetens utdata uppnåddes för SR- och AF-kategorierna. För 12-avlednings-EKG-remsorna som utvärderats som SR (n=470) klassificerades också totalt 436 resultat som SR med EKG-appenhetens utdata. För 12-avlednings-EKG-remsorna som utvärderats som AF (n=521) klassificerades totalt 474 resultat som AF med EKG-appenhetens utdata. EKG-appenheten klassificerade 9 resultat som SR, 37 resultat som AF, 13 resultat som osäkert resultat och 9 resultat som dålig registrering som utvärderades som annat (SVT eller annat) på 12-avlednings-EKG:t.

**Tabell 6. Korsklassificeringstabell av resultat från EKG-appen och 12-avlednings-EKG (i vila och efter motion) – klassificerbar analysuppsättning**

Direkta observationer med utvärderade resultat från 12-avlednings-EKG						
EKG-appenhetens utdata	SR (50≤HR≤150)	AF (50≤HR≤150)	Annat (SVT eller annat med HR 50–150, HR<50, HR>50)	Kunde ej tolkas	Inte tillgänglig	Totalt
SR [SR (50–99), NSR sinustakykardi (100–150)]	436	7	9	0	2	454
AF (50≤HR≤150)	3	474	37	0	6	520
Osäkert resultat [inklusive HR<50, HR>150]	10	13	13	0	1	37
Dålig registrering	14	23	9	0	3	49
Inte tillgänglig	7	4	1	1	0	13
<b>Totalt</b>	<b>470</b>	<b>521</b>	<b>69</b>	<b>1</b>	<b>12</b>	<b>1073</b>

Förkortningar: AF=förmaksflimmer, SR=sinusrytm, HR=puls, SVT=supraventrikulär takykardi

Specificiteten för EKG-app 2.0 uppskattades vara 99,3 % och sensitiviteten uppskattades vara 98,5 % (tabell 7). Den undre gränsen i konfidensintervallet för specificitetsuppskattningen (98,4 %) och sensitivitetens uppskattningen (97,3 %) överskred de angivna prestationsmålen för dessa prestationsmätningar. Därför avvisades nollhypoteserna för specificitet och sensitivitet till förmån för den alternativa hypoteserna, och uppnåddes de primära effektmåten för specificitet och sensitivitet. Specificiteten i vila uppskattades till 99,1 % och sensitiviteten i vila uppskattades till 98,5 %. Specificiteten efter motion uppskattades till 99,5 % och sensitiviteten efter motion uppskattades till 98,6 %. En ytterligare analys som tog med algoritmklassificeringen "osäkert resultat" i beräkningen utfördes för kombinerade data för "i vila" och "efter motion". Den resulterande specificiteten uppskattades till 97,1 % och sensitiviteten uppskattades till 96,0 %.

En hög grad av överensstämmelse observerades mellan utvärderade rytmdata från 12-avlednings-EKG:t och EKG-appenhetens utdata. Graden av felaktiga klassificeringar var låg vid jämförelse med referensremсор från 12-avlednings-EKG. I tabell 8 anges en sammanfattning av resultaten som bekräftar att de sekundära effektmåten för korrekt klassificering av patienter med NSR (HR 50–99, SR på samtidigt 12-avlednings-EKG) som SR, AF (HR 50–99, AF på samtidigt 12-avlednings-EKG) som AF, sinustakykardi (HR 100–150, SR på samtidigt 12-avlednings-EKG) som hög puls, och AF med hög puls (HR 100–150, AF på samtidigt 12-avlednings-EKG) som AF (hög puls) på en läsbar och klassificerbar EKG-apprens. Andelen korrekt klassificerad SR (HR 50–99) var 98,3 %, den undre gränsen i konfidensintervallet på 96,6 % överskred det angivna prestationsmålet på 90 % för detta effektmått. Andelen korrekt klassificerat AF (HR 50–99) var 98,9 %, den undre gränsen i konfidensintervallet på 97,5 % överskred det angivna prestationsmålet på 81 %. Andelen korrekt klassificerat SR (HR 100–150) var 90,7 %, den undre gränsen i konfidensintervallet på 86,7 % överskred det angivna prestationsmålet på 81 %. Andelen korrekt klassificerat AF (HR 100–150) var 83,0 %, den undre gränsen i konfidensintervallet på 77,8 % överskred det angivna prestationsmålet på 75,5 %.

<b>Tabell 7. Sensitivitets- och specificitetsresultat (resultat primära effektmått) – klassificerbar analysuppsättning</b>		
Parameter	Värde	Dubbelsidigt 95 % bootstrap-konfidensintervall*
<b>Kombinerat (i vila och efter motion)</b>		
Referensresultat = AF (50–150 HR)	481	
EKG-app = AF (50–150 HR)	474/481 (98,5 %)	
EKG-app = SR (50–150 HR)	7/481 (1,5 %)	
<b>Sensitivitet</b>	<b>474/481 (98,5 %)</b>	<b>(97,3 %, 99,6 %)</b>
Referensresultat = SR (50–150 HR)	439	
EKG-app = AF (50–150 HR)	3/439 (0,7 %)	
EKG-app = SR (50–150 HR)	436/439 (99,3 %)	
<b>Specificitet</b>	<b>436/439 (99,3 %)</b>	<b>(98,4 %, 100,0 %)</b>
*Resultat erhållna från bootstrap-omsampling Förkortningar: AF=förmaksflimmer, SR=sinusrytm, HR=puls		

<b>Tabell 8. Andel korrekt av klassificerbara AF- och SR-rytmer</b>		
Parameter	Värde	Bootstrap-konfidensintervall**
<b>Kombinerat (i vila och efter motion)*</b>		
Andel korrekt för SR- referens (HR 50–99)	230/234 (98,3 %)	(96,6 %, 99,6 %)
Andel korrekt för AF- referens (HR 50–99)	272/275 (98,9 %)	(97,5 %, 100 %)
Andel korrekt för SR- referens (HR 100-150)	186/205 (90,7 %)	(86,7 %, 94,6 %)
Andel korrekt för AF- referens (HR 100-150)	171/206 (83,0 %)	(77,8 %, 88,0 %)
*Beräkning av andel korrekt inkluderar klassificerbara AF- och SR-rytmer med EKG-app 2.0 **Konfidensintervall erhållet från bootstrap-omsampling Förkortningar: AF=förmaksflimmer, HR=puls, SR=sinusrytm		

Dessutom visade vågformsjämförelsen mellan algoritmen i EKG-app 2.0 och avledning I från ett 12-avlednings-EKG att den totala andelen med beteckningen "godkänd" för kombinerade morfologireultat för "i vila" och "efter motion" var 100 %, med totalt fem uteslutna resultat. Den undre gränsen i konfidensintervallet var 97,9 %, vilket överskred den angivna andelen för morfologibeteckningen "godkänd" på 80 %. Den totala andelen med beteckningen "godkänd" för kombinerade resultat av R-vågsutvärdering "i vila" och "efter motion" var 97,2 %, den undre gränsen för konfidensintervallet var 93,6 %, vilket överskred den angivna andelen för överensstämmelse av R-vågsamplitud på 80 %. "I vila"- och "efter motion"-grupper visade liknande R-våg.



# Slutsatser

Apple Watch-kunder har nu tillgång till två valfria funktioner som aktiverar identifiering av oregelbunden hjärtrytm: funktionen för notiser om oregelbunden rytm (tillgänglig på Apple Watch Series 1 och senare) och EKG-appen (tillgänglig på Apple Watch Series 4, Series 5 och Series 6).

Funktionen för notiser om oregelbunden rytm är ett program som klassificerar opportunistiskt insamlade takogram i bakgrunden, och meddelar kunden som aktiverat funktionen om en oregelbunden hjärtrytm förekommer. Apples hjärtstudie visade att bland deltagarna som fick en notis samtidigt som de bar en Apple Watch och ett EKG-plåster visade 78,9 % överensstämmande AF på EKG-plåstret och 98,2 % AF och andra kliniskt relevanta arytmier. Dessa resultat antyder att notisen i majoriteten av fallen korrekt anger förekomsten av AF. I vissa fall kan dock en notis ange förekomsten av en annan arytmier än AF. I Apple Watch Series 4, Series 5 och Series 6 är en elektrisk hjärtsensor tillgänglig som – tillsammans med EKG-appen och algoritmen – producerar en EKG-vågform som liknar ett 1-avlednings-EKG (avledning I) och som kan klassificera vågformen för att hitta förekomsten av AF. Den upphovsrättsskyddade algoritmen som är avsedd att klassificera dessa EKG:n uppvisade sensitivitet på mer än 98 % och specificitet på mer än 99 % beträffande identifiering av AF, jämfört med EKG:n som registrerades med en referensapparat och tolkades av oberoende kliniska experter.

Kunder som vill använda dessa två funktioner måste slutföra en användarupplevelse som ger information om tolkningen och den icke-diagnostiska karaktären av dessa fynd samt algoritmens begränsningar. Båda funktionerna har godkänts av FDA genom De Novo- eller 510K-klassificering.