



Brug af Apple Watch til at registrere arytmi

December 2020

Indhold

Oversigt	3
Introduktion	3
Fotoplethysmografi baseret registrering af arytm	3
Beskrivelse af teknologien og funktionerne	3
Præklinisk udvikling	5
Klinisk validering	5
Apple Heart Study	5
Udformning af AHS-delundersøgelse	5
Resultater.....	6
EKG-baseret registrering	6
Beskrivelse af teknologien og funktionerne	6
Præklinisk udvikling.....	7
Klinisk validering	7
Forsøgets udformning.....	7
Klinisk valideringsundersøgelse af EKG 1.0	8
Resultater.....	8
Klinisk valideringsundersøgelse af EKG 2.0	12
Resultater	14
Konklusioner	17

Oversigt

Apple Watch-brugere har adgang til to funktioner i software, der kan opdage hjertearytmier som atrieflimren (AFLI): funktion til notifikation om uregelmæssig puls (IRNF) og EKG-appen.

På Apple Watch Series 1 eller senere analyserer IRNF pulsdatabaser, der er indsamlet af en fotoplethysmografisk (PPG) sensor, for at identificere perioder med uregelmæssig puls, der kan indikere AFLI, og den sender en notifikation til brugeren, når der registreres en sådan periode. På Apple Watch Series 4, Series 5 og Series 6 kan EKG-appen generere et elektrokardiogram (EKG) i bølgeform, der svarer til en enkeltpunkts-EKG-måling, og klassificerer derefter den pågældende bølgeform som sinusrytme (SR), atrieflimren (AFLI), høj eller lav puls eller ikke entydigt; med EKG 2.0 findes der yderligere AFLI-klassificeringer med høj puls eller dårlig måling.

Formålet med dette dokument er at give en mere detaljeret forståelse af mulighederne i disse funktioner, herunder hvordan de er blevet testet og valideret.

Introduktion

Atrieflimren (AFLI), som er en slags uregelmæssig hjerterytme, hvor hjertets atrier slår uregelmæssigt og nogle gange hurtigt, er en af de mest almindelige årsager til slagtilfælde. AFLI er imidlertid ofte symptomfri, hvilket betyder, at mange personer med AFLI ikke er klar over, at de har denne sygdom. Der er i dag en øget interesse for tidlig identifikation af AFLI uden for de kliniske rammer; dette skyldes risikoen for slagtilfælde, de manglende symptomer, muligheden for effektive farmakologiske behandlinger, der minimerer risikoen for slagtilfælde, og en øget udbredelse af forbrugerudstyr, der kan opdage AFLI.

Med watchOS 5.1.2 eller senere kan Apple Watch Series 1 og senere bruge fotoplethysmografiske signaler med en algoritme til at identificere perioder med uregelmæssig puls, der kan indikere AFLI. Ud over denne fotoplethysmografiske identifikationsalgoritme er Apple Watch Series 4, Series 5 og Series 6 også udstyret med en elektrisk hjertesensor, der sammen med EKG-appen, kan generere og analysere en EKG-måling svarende til en enkeltpunkts-EKG-måling.

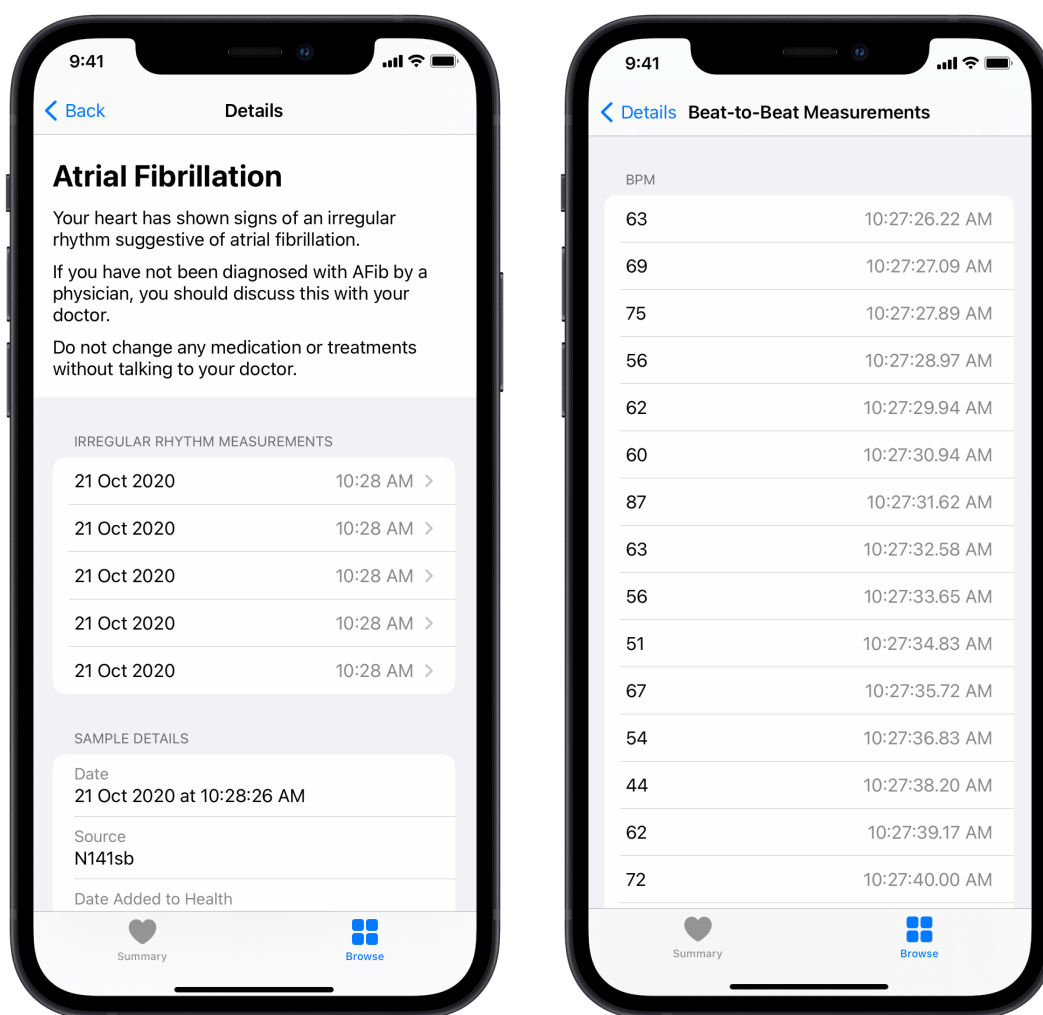
Fotoplethysmografi baseret registrering af arytmi

Beskrivelse af teknologien og funktionerne

Apple Watch har en optisk hjertesensor, der ved hjælp af fotoplethysmografi bruger grønne LED-lys med lysfølsomme fotodioder til at registrere mængden af blod, der strømmer gennem brugerens håndled. Disse sensorer og underliggende algoritmer er grunden til, at Apple Watch Series 1 eller nyere kan registrere puls og hjertefrekvensvariabilitet (HRV). For at fastslå HRV laver Apple Watch en intervalmåling – dvs. en fremstilling af tidsintervallet mellem hvert hjerteslag – hver anden til fjerde time. Fra watchOS 5.1.2 og senere kan brugeren også vælge at slå en funktion til, der bruger disse intervalmålinger til at opdage arytmier. For at bruge funktionen Notifikation om uregelmæssig puls i Apple Watch skal brugeren først registrere sig i appen Sundhed på sin pardannede iPhone for at lære at bruge funktionen og få viden om AFLI. Der findes flere oplysninger om brugeroplevelsen på support.apple.com/da-dk/HT208931.

Hvis den fotoplethysmografibaserede funktion til registrering af arytmier er slået til, klassificeres hver intervalmåling ved hjælp af en algoritme, som fastslår, om der er tale om uregelmæssig hjerterytme. En intervalmåling, der viser uregelmæssig rytme, starter en række hyppigere intervalmålinger (så ofte som muligt, med minimum 15 minutters mellemrum) og analyserer. Apple Watch indsamler og analyserer kun intervalmålingerne, hvis brugeren forholder sig tilstrækkeligt i ro til, at der kan foretages en måling. Af denne årsag overvåges brugeren ikke altid af algoritmen, men gør det forholdsvist tilfældigt, når der er tilstrækkeligt signal til indsamling og analyse. Hvis fem ud af seks fortløbende følgende intervalmålinger – inklusive den første – klassificeres som uregelmæssige inden for en periode på 48 timer, modtager brugeren en notifikation om risikoen for arytmier. Ud over at modtage en notifikation får brugeren adgang til yderligere oplysninger om de uregelmæssige intervalmålinger i appen Sundhed (figur 1). Hvis to intervalmålinger klassificeres som ikke uregelmæssige, inden tærskelværdien er nået, nulstilles cyklusen, og indsamlingen af intervalmålinger skifter til sin normale frekvens på hver anden time.

Figur 1: Målinger af uregelmæssig puls i appen Sundhed



I appen Sundhed kan brugerne se de gange, algoritmen har registreret et uregelmæssigt interval, som har udløst en notifikation (venstre). Ved at trykke på dato og klokkeslæt kan brugeren visualisere de hjerteslagsmålinger, der beregnes i hver intervalmåling.

Præklinisk udvikling

Forud for de kliniske forsøg blev der foretaget undersøgelse for at udvikle den fotoplethysmografibaserede algoritme og for at evaluere algoritmens præstation under forskellige forhold og ved forskellig brugeradfærd. Disse omfattede bl.a. dybe vejrtrækninger, bilkørsel, rystelser og bevægelser i hænderne, nedsat perfusion i hånd/håndled, brug natten over, hurtig ventrikulær respons hos personer med AFLI og andre arytmier. Undersøgelserne blev udført på 2300 kontrolpersoner og mere end 500 personer med AFLI.

Fotoplethysmografi er afhængig af lysabsorption, og derfor blev algoritmen til registrering af arytmier testet på mange forskellige hudtyper og -farver for at sikre, at sensorplatformens justering for hudfarve var tilstrækkelig i forhold til de algoritmer, der bruges til at registrere arytmier. Melanin har høj absorptionsevne ved den bølgelængde, der anvendes af de grønne LED-lys på Apple Watch, hvilket potentielt kan gøre fotoplethysmografibaserede målinger af hjertefrekvens vanskeligere på mørke hudfarver. For at tage højde for dette og sikre passende signalamplitude for alle hudfarver, justerer sensorplatformen på Apple Watch LED-tilstrømningen (og dermed lysudgangen), fotodiodeforøgelsen (lysfølsomheden) og prøvetagningssekvensen.

Af valideringshensyn blev der foretaget analyse af 1,3 millioner intervallmålinger fra 1124 forsøgspersoner (51 procent kvinder) med forskellig hudtype og -farve (Fitzpatrick-hudtype og hudtone målt med spektrofotometer på håndleddet). En af de primære tekniske udfordringer var signalamplitude hos personer med mørk hud, og derfor havde næsten 5 procent af forsøgspersonerne Fitzpatrick-hudtype VI. Gennem valideringen blev der ikke påvist nogen signifikant forskel i algoritmens følsomhed eller specificitet på tværs af hudtyper eller -farver.

Klinisk validering

Apple Heart Study

Apple Heart Study (AHS) er en prospektiv og pragmatisk enkeltarmet undersøgelse, der blev gennemført virtuelt, for at evaluere i hvor høj grad den Apple Watch-baserede algoritme til notifikation om uregelmæssig puls, kan identificere arytmier, som kan indikere AFLI. Hvis en bruger opfyldte fem ud af seks tærskelværdier, fik brugeren en iPhone- og Apple Watch-notifikation og mulighed for at kontakte en af undersøgelsens læger for at få tilsendt et ambulatorisk EKG-plaster (ePatch, BioTelemetry, Inc. Conshohocken, Pennsylvania). Deltagerne fik besked på at bære ePatch i op til syv dage, hvor indsamlede data fra en deltager blev betragtet som tilstrækkelige efter mindst en times analysetid.

De detaljerede resultater af AHS blev publiceret 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).

Udformning af AHS-delundersøgelse

Der blev gennemført en delundersøgelse af de data, som blev indsamlet i AHS, for at afgøre, om algoritmens klassificering af intervallmålingerne (individuelle intervallmålinger eller stikprøver) og algoritmen for bekræftelsescyklus (advarselsniveau, fem ud af seks målinger) havde en tilstrækkelig positiv prædiktiv værdi (PPV) sammenlignet med ePatch-overvågningen i forhold til at detektere uregelmæssige hjerterytmier, der angav AFLI, hos en delgruppe af AHS-deltagerne. AHS-forskere kendte til delundersøgelsen, de efterfølgende analyser, og at dataene blev sendt ind til FDA (United States Food and Drug Administration), men de fik ikke adgang til resultaterne fra delundersøgelsen under den igangværende AHS-undersøgelse. Institutional Review Board (IRB), som godkendte AHS, fastslog, at delundersøgelsen

var fritaget fra IRB-tilsyn. Alle AHS-deltagere gav informeret samtykke, hvilket omfattede samtykke til at bruge deres undersøgelsesdata i delundersøgelsen.

Data til delundersøgelsen blev indsamlet fra AHS-deltagere, der deltog i perioden 30. november 2017 til 22. juni 2018. Forsøgspersonerne i delundersøgelsen modtog en notifikation om uregelmæssig puls fra AHS-appen, hvorefter de modtog og påførte sig et ambulatorisk EKG-plaster (ePatch) til aflæsning af de ambulatoriske EKG-resultater. De første uregelmæssige intervalmålinger, der førte til den første notifikation og eventuelt forårsagede den første telekonsultation, blev ikke analyseret som en del af denne delundersøgelse; der blev kun analyseret uregelmæssige intervalmålinger og notifikationer, der forekom, mens brugeren havde det til undersøgelsen godkendte ePatch-plaster.

To uafhængige EKG-eksperter med U.S. board-certificering i kardiologi eller elektrofysiologi gennemgik og bedømte EKG-strimlerne og klassificerede dem som SR, AFLI, anden uregelmæssig puls eller ulæselig. Hvis eksperterne var uenige, ville en tredje ekspert med samme kvalifikationer vurdere strimlen. Eksperterne kendte ikke til klassificeringen af intervalmålingerne. Eksperternes EKG-klassificering og algoritmens klassificering af intervalmålingerne blev på sikker vis sendt til dataanalyse hos undersøgelsens statistiker.

Resultater

ePatch registrerede AFLI hos 41,6 procent (94 deltagere) af de 226 deltagere i delundersøgelsen, som modtog en første notifikation om arytmier og bar ePatch i ca. en uge. Ved samtidig brug af Apple Watch og et ePatch-plaster modtog 57 ud af 226 deltagere en AFLI-notifikation, dvs. at de havde fem ud af seks fortløbende intervalmålinger, der blev klassificeret som uregelmæssige). Af disse 57 deltagere viste 78,9 procent (45 deltagere) tilsvarende AFLI på deres ePatch-data, og 98,2 procent (56 deltagere) viste AFLI eller anden klinisk relevante arytmier. Resultaterne viser, at selvom en notifikation i de fleste tilfælde indikerer korrekt tilstedeværelse af AFLI, kan en notifikation i andre tilfælde indikere tilstedeværelse af en anden slags arytmier end AFLI. Der blev ikke registreret alvorlige bivirkninger i forbindelse med brug af udstyret.

EKG-baseret registrering

Beskrivelse af teknologien og funktionerne

Apple Watch Series 4, Series 5 og Series 6 er udstyret med en elektrode af titanium inden i Digital Crown og et ultratyndt lag af krom, silikone og karbonnitrid på safirkristallen på bagsiden af Apple Watch. EKG-appen læser og registrerer de elektriske impulser, der kontrollerer hjertet, fra brugerens fingerspids (med Digital Crown) og håndleddet (via bagsiden af Apple Watch), hvilket skaber et lukket kredsløb. For at bruge EKG-appen på Apple Watch skal brugeren først registrere sig i appen Sundhed på sin pardannede iPhone for at lære at bruge funktionen og få viden om AFLI. Brugere udfører en EKG-måling ved at åbne EKG-appen, der er installeret på Apple Watch, og derefter placere en finger fra den hånd, der ikke bærer Apple Watch, på Digital Crown i 30 sekunder. Polariteten bestemmes af, hvilket håndled brugeren har placeret Apple Watch på, hvilket angives i Indstillinger.

Efter at have udført EKG-målingen bruges en proprietær algoritme til at klassificere målingen som SR, AFLI eller ikke entydig i EKG 1.0. Hvor EKG 2.0 er tilgængelig findes der desuden yderligere klassificeringer, f.eks. AFLI med høj puls (puls 100-150), og yderligere differentiering mellem dårlig måling og ikke entydige målinger. Rytmeklassificeringen – den gennemsnitlige puls, brugerrapporterede symptomer og bølgeform – føjes til appen Sundhed, hvor de gemmes og kan deles af brugeren som PDF-arkiv fra appen på brugerens pardannede iPhone. Der findes flere oplysninger om brugeroplevelsen på support.apple.com/da-dk/HT208955.

Præklinisk udvikling

Algoritmen til registrering af EKG-signal og klassificeringsalgoritmen blev også testet i flere undersøgelser forud for den kliniske validering. Sensor- og -klassificeringsalgoritmen blev testet på tværs af forskellige etniske grupper, håndledsomkredse, BMI (kropsmasseindeks), aldre, arytmier, som ikke var AFLI, hvor stramt remmen sad, stillinger, aktivitetsniveauer og sved. Ca. 2500 forsøgspersoner deltog i disse tests, hvoraf ca. 25 procent tidligere var blevet diagnosticeret med AFLI eller anden form for uregelmæssig puls.

Øget frekvens af "ulæselige" EKG-målinger var den primære variation i algoritmens præstation. Faktorer, der førte til denne variation, var lav signalamplitude (som følge af højre akse-afvigelse – især bemærket hos deltagere med lavt BMI – eller som svedte under tests efter træning) og bevægelsesartefakter som følge af brugeradfærd. Apple Watch bruger tørre elektroder, der er designet til at være mekanisk stærkere og korrosionsbestandigt, som passende for en bærbar enhed. Tørre elektroder – især dem, der placeres på arme og ben – har dog en tendens til at producere støj som den, der beskrives ovenfor, sammenlignet med de midlertidige gelelektroder, der bruges i medicinske enheder.

Ud over de ovennævnte faktorer ændrede tilstedeværelsen af visse arytmier, der ikke var AFLI, i betydelig grad algoritmens præstation sammenlignet med forsøgspersoner med SR. Disse tilstande og resultater beskrives i tabel 1 nedenfor.

Tabel 1. Arytmier, der ikke er AFLI, som påvirker algoritmens præstation (EKG 2.0)

Arytmi	Variation
Venstre eller højre grenblok	7,9 % af forsøgene klassificeret som AFLI
AV-blok, grad 1	10,2 % af forsøgene uklassificerede; af de klassificerede er 2,4 % klassificeret som AFLI
Bigemini	92,5 % uklassificerede
Hyppe PVC-hjerteslag	48,8 % af forsøgene uklassificerede; af de klassificerede er 24,1 % klassificeret som AFLI
Hyppe PAC-hjerteslag	19,8 % af forsøgene uklassificerede; af de klassificerede er 23,5 % klassificeret som AFLI
Atrial takykardi	7,7 % af forsøgene klassificeret som AFLI
Blandet rytme	46,5 % af forsøgene uklassificerede; af de klassificerede er 29,7 % klassificeret som AFLI
Høj/lav puls (uden for 50-150 slag i minuttet)	94,8 % uklassificerede

Klinisk validering

Forsøgets udformning

Apple sponsorerede to undersøgelser udført på flere forskellige klinikker med henblik på at validere EKG-appens evne til at: (1) generere en EKG-bølgeform svarende til en enkeltpunkts-EKG-måling i en standard EKG-måling med 12 elektroder og (2) anvende en rytmeklassificeringsalgoritme til at klassificere denne enkeltpunkts-EKG-måling som enten SR eller AFLI.

Klinisk valideringsundersøgelse af EKG 1.0

Der blev foretaget en klinisk valideringsundersøgelse for at validere præstationen af EKG-appen 1.0. Hovedformålet med undersøgelsen var at undersøge, hvor følsom rytmeklassificeringsalgoritmen var i forhold til registrering af AFLI, og hvor specifik den var i forhold til registrering af SR. En ekstern enhed Institutional Review Board (IRB) godkendte protokollen, samtykkeformularen og alt andet relevant materiale, før de første forsøgspersoner blev tilmeldt, og alle forsøgspersoner gav skriftligt samtykke inden tilmeldingen.

Deltagere med diagnosticeret AFLI og andre uden kendte hjerterytmeforstyrrelser blev tilmeldt. De blev bedt om at registrere tre EKG-målinger med én elektrode ved hjælp af EKG-appen, mens undersøgelsens medarbejdere samtidigt registrerede tre 12-punkts EKG-målinger ved hjælp af en FDA-godkendt klinisk enhed (EKG-enhed fra GE Healthcare CardioSoft). De første prøver blev overvejet til bedømmelse og analyse. Deltagerne fik hjælp til placering af Apple Watch og blev bedt om at holde armen stille – evt. hvile armen på et bord eller deres ben – og de fik lov til at øve sig på målingerne inden testene.

For at teste punkt (1) tog tre uafhængige, certificerede hjerteteknikere de genererede strimler fra 140 tilfældigt udvalgte forsøgspersoner (70 med AFLI og 70 med SR) og lagde dem over på den tilsvarende strimmel fra enkeltpunktsmålingen fra den kliniske enhed for at foretage en visuel sammenligning af morfologien af seks fortløbende PQRST-komplekser. Hjerteteknikerne registrerede hver strimmel som bestået eller dumpet på baggrund af ligheder i den visuelt vurderede morfologi. De blev også bedt om at måle R-amplituden fra den isoelektriske basislinje til den nærmeste millimeter for de første to QRS-komplekser på både referencestrimlerne og EKG-appens genererede strimler, hvorefter overensstemmelsen mellem de to blev vurderet.

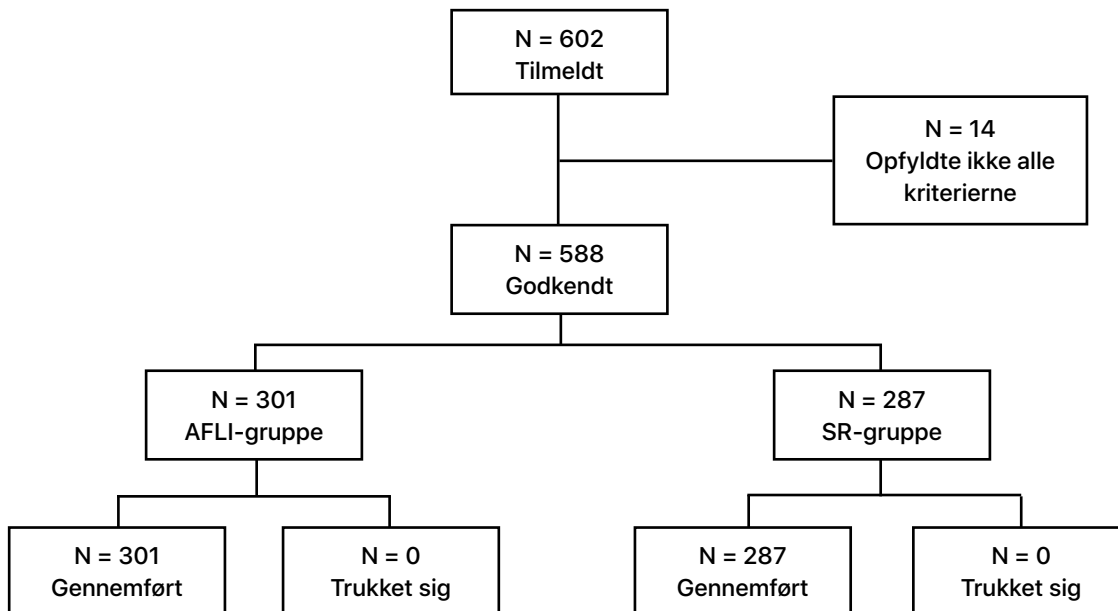
For at teste punkt (2) blev hver EKG-referencestrimmel fra 12-punktsmålingerne gennemgået af tre uafhængige certificerede kardiologer (U.S. board certification), som ikke kendte til undersøgelsen, der skulle klassificere rytmerne som SR, AFLI, andet (alt, der ikke var SR eller AFLI inden for parametrene for hjerterytme) eller ulæselig (en diagnose kunne ikke stilles, da strimlen var utilstrækkelig til aflæsning). EKG-appens algoritme klassificerede det i EKG-appen genererede EKG som SR, AFLI, uklassificerbar eller ulæselig. Derefter blev følsomheden og specificiteten af EKG-appens klassificering af SR og AFLI (for klassificerbare EKG-målinger) beregnet i sammenligning med kardiologernes tolkning af 12-punkts EKG-målinger. En uafhængig, certificeret kardiolog (U.S. board-certificeret), der ikke kendte til undersøgelsen, blev derefter bedt om at klassificere EKG-appens genererede strimler ud fra samme kategorier.

For at analysere hovedformålene blev der separat beregnet en præcis ensidig nedre konfidensgrænse på 97,5 procent for følsomhed og specificitet. Hvis den nedre grænse for følsomhed oversteg 90 procent, blev nulhypotesen afvist til fordel for følsomheden, der oversteg 90 procent. Hvis den nedre grænse for specificitet oversteg 92 procent, blev nulhypotesen afvist til fordel for specificiteten, der oversteg 92 procent.

Resultater

Undersøgelsen havde 602 deltagere, hvoraf 588 forsøgspersoner opfyldte alle krav. Ud af de 588 godkendte forsøgspersoner blev 301 deltagere med egenrapporteret AFLI placeret i en AFLI-gruppe, og 287 deltagere, der ikke havde rapporteret om AFLI, blev placeret i en SR-gruppe. Placeringerne i grupper blev kun brugt til at sikre korrekt tilmelding – personerne, der foretog vurderingerne, kendte intet til grupperne, og tilstedeværelse eller fravær af AFLI var kun baseret på EKG-målinger, der blev foretaget under testene. De 14 forsøgspersoner, der gennemførte undersøgelsen, men som ikke blev placeret i en gruppe, kunne ikke deltage i undersøgelsen, fordi de tidligere havde haft paroxysmal AFLI, men uden at kunne påvise AFLI under EKG-målingerne. Alle godkendte forsøgspersoner gennemførte undersøgelsen (figur 2). Der blev ikke rapporteret om bivirkninger under undersøgelsen.

Figur 2: Diagram over forsøgspersoners disposition



Tre uafhængige, certificerede hjerteteknikere fandt visuel morfologisk lighed mellem bølgeformen i strimlerne fra EKG-appen og enkeltpunkts-EKG-målingerne, der blev udført af en standard klinisk enhed for 98,4 procent af de analyserede strimler i AFLI-gruppen og 100 procent i SR-gruppen (tabel 2). Andelen af de samlede forsøgspersoner med godkendte målinger var 99,2 procent (97,5 procent nedre konfidensgrænse, 95,7 procent). Strimlerne blev udelukket, hvis det ikke var muligt at identificere seks fortløbende slag (PQRST-komplekser) uden rystelser på nogen af de to typer strimler (fra EKG-appen eller referenceenheden).

Tabel 2. Sammenligning af bølgeform

Karakteristika	AFLI-forsøgspersoner (N = 61)	SR-forsøgspersoner (N = 65)	I alt (N = 126)	Nedre konfidensgrænse*	P-værdi**
Antal godkendte sammenlignede strimler fra forsøgspersoner (fra EKG-appen og referenceenheden)	60	65	125		
Antal læsbare, sammenlignede strimler fra forsøgspersoner (fra EKG-appen og referenceenheden)	61	65	126		
Andel af godkendte strimler fra forsøgspersoner	60/61 (98,4 %)	65/65 (100 %)	125/126 (99,2 %)	0%	< 0,0001
Antal udelukkede sammenlignede strimler fra forsøgspersoner	8	5	13		

* Lavere præcis binomial ensidig konfidensgrænse på 97,5 % i alt ** Hypotesetest for vellykkede forsøg > 0,8 Forkortelser: AFLI = atrieflimren, SR = sinusrytme

For yderligere at bekræfte, at EKG-appens genererede bølgeforme og referenceenheden var sammenlignelige, blev EKG-strimlerne fra appen, som var blevet klassificeret af kardiologerne uden at kende til undersøgelsen, sammenlignet med referencestrimlerne, der blev klassificeret af kardiologerne (tabel 3). Overensstemmelsen mellem klassificeringen af enhedsstrimlerne og AFLI- og SR-referenceresultaterne var på 100 procent og 99,1 procent. Ulæselige strimler blev ikke medtaget i denne analyse.

Tabel 3. Klassificeringer mellem EKG-appen og referenceenheden

Karakteristika	I alt (N = 522)
Endeligt EKG-referenceresultat = AFLI	263
Klassificering af EKG-strimler fra appen = AFLI	239/263 (90,9 %)
Klassificering af EKG-strimler fra appen = SR	0/263 (0,0 %)
Klassificering af EKG-strimler fra appen = Andet	0/263 (0,0 %)
Klassificering af EKG-strimler fra appen = Ulæselig	24/263 (9,1 %)
% overensstemmelse med referenceresultatet AFLI*	239/239 (100,0 %)
Endeligt EKG-referenceresultat = SR	244
Klassificering af EKG-strimler fra appen = AFLI	0/244 (0,0 %)
Klassificering af EKG-strimler fra appen = SR	232/244 (95,1 %)
Klassificering af EKG-strimler fra appen = Andet	2/244 (0,8 %)
Klassificering af EKG-strimler fra appen = Ulæselig	10/244 (4,1 %)
% overensstemmelse med referenceresultatet SR*	232/234 (99,1 %)
Endeligt EKG-referenceresultat = Andet	15
Klassificering af EKG-strimler fra appen = AFLI	0/15 (0,0 %)
Klassificering af EKG-strimler fra appen = SR	3/15 (20,0 %)
Klassificering af EKG-strimler fra appen = Andet	12/15 (80,0 %)
Klassificering af EKG-strimler fra appen = Ulæselig	0/15 (0,0 %)
% overensstemmelse med referenceresultatet Andet*	12/15 (80,0 %)

*Ulæselige strimler blev ekskluderet
 Forkortelser: AFLI = atrieflimren, SR = sinusrytme

I alt blev 485 af de 602 sammenlignede strimler fra EKG-appen og referenceenheden anset for at være klassificerbare. De resterende indeholdt strimler fra EKG-appen og/eller referenceenheden, der var ulæselige eller uklassificerbare. Tabel 4 viser fordelingen blandt AFLI- og SR-grupperne.

Tabel 4. Endeligt klassificeringsresultat for EKG-appens algoritme og referencestrimler

Klassificering via EKG-appens algoritme	Klassificering af referencestrimler				
	SR	AFLI	Andet	Ulæselig	I alt
SR	238	4	4	1	247
AFLI	1	236	2	2	241
Uklassificerbar	6	7	6	0	19
Ulæselig	18	30	1	0	49
Enhedsresultater ikke rapporteret*	32	13	1	0	46
I alt	295	290	14	3	602

*Resultater, som ikke blev rapporteret på baggrund af forudbestemte kriterier (f.eks. manglende synkronisering) for alle forsøgspersoner bortset fra én.
 Forkortelser: AFLI = atrieflimren, SR = sinusrytme

Klassificeringen via EKG-algoritmen opnåede en følsomhed på 98,3 procent og en specificitet på 99,6 procent (tabel 5). Hvis analysen udvides til at inkludere 2,4 procent (7 ud af 290) og 2,0 procent (6 ud af 295) strimler, der blev kategoriseret som uklassificerbare af enheden i AFLI- og SR-referencestrimlernes klassificeringskategorier, ville følsomheden have ligget på 95,5 procent (95 procent CI: 92,2 procent, 97,8 procent), og specificiteten ville være 97,1 procent (95 procent CI: 94,2 procent, 98,8 procent). Resultaterne opfyldte de hovedformål, der blev angivet forud for undersøgelsen. Derudover var 12,2 procent (68 ud af 556) af målingerne ikke entydige – enten ulæselige eller uklassificerbare – og kunne ikke klassificeres som hverken SR eller AF. Når ikke entydige målinger blev inkluderet i analysen, klassificerede EKG-appen korrekt SR i 90,5 procent (238 ud af 263) af forsøgspersoner med SR, og AFLI i 85,2 procent (236 ud af 277) af forsøgspersoner med AFLI. De kliniske valideringsresultater stammer fra tests i et kontrolleret miljø. Brug af EKG-appen hos brugeren kan resultere i et større antal ikke entydige strimler, der ikke kan klassificeres.

Tabel 5. Analyse af følsomhed og specificitet – klassificerbare strimler

Parameter	Værdi	Nedre konfidensgrænse*	P-værdi**
Endeligt EKG-referencerresultat = AFLI (n)	240		
Enhedsresultat for EKG-app = AFLI	236/240 (98,3 %)		
Enhedsresultat for EKG-app = SR	4/240 (1,7 %)		
Følsomhed	236/240 (98,3 %)	0%	< 0,0001
Endeligt EKG-referencerresultat = SR (n)	239		
Enhedsresultat for EKG-app = AFLI	1/239 (0,4 %)		
Enhedsresultat for EKG-app = SR	238/239 (99,6 %)		
Specificitet	238/239 (99,6 %)	0%	< 0,0001

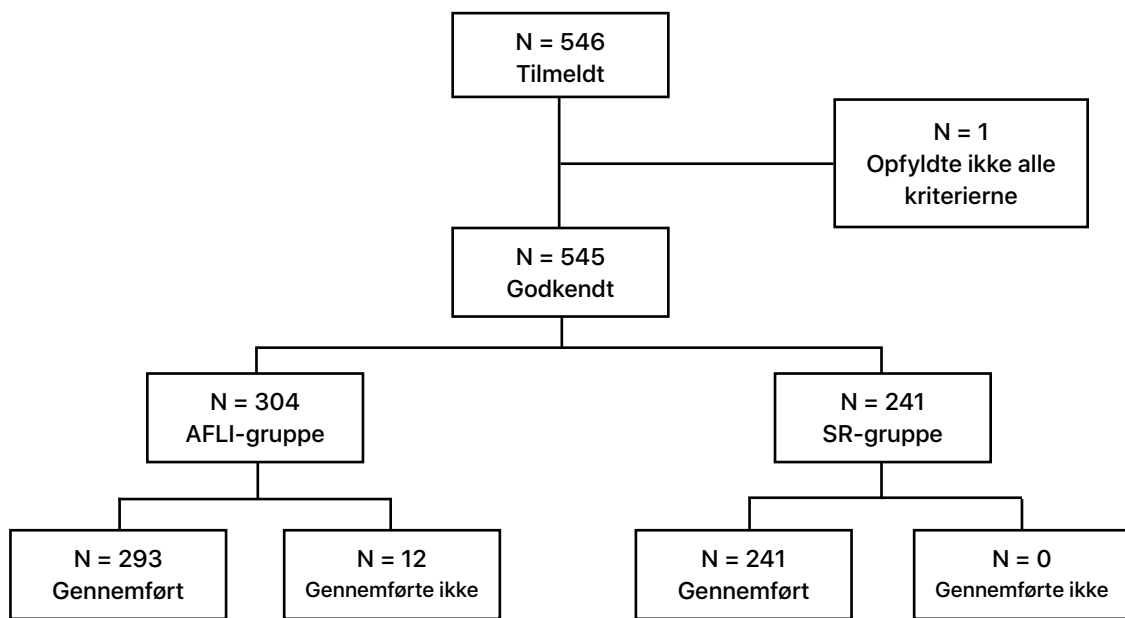
*Nedre nøjagtig binomial ensidig konfidensgrænse **Hypotesetest for følsomhed > 0,9 og specificitet > 0,92. Forkortelser: AFLI = atrieflimren, SR = sinusrytme

Klinisk valideringsundersøgelse af EKG 2.0

En anden undersøgelse blev udført for at understøtte og validere undersøgelse af EKG 2.0-appens algoritmer (testenhed), der udvider pulsens klassificerbare område (50-150 slag i minuttet) og introducerer nye klassificeringsresultater (SR, SR med høj puls, AFLI, AFLI med høj puls, ikke entydig og dårlig måling). Formålet med undersøgelsen var at evaluere testenhedens præstation. Hovedformålet var vurdering af specificitet og følsomhed. Det sekundære formål omfattede korrekt klassificering af følgende emnekategorier på bestemte læselige og klassificerbare EKG-teststrimler: NSR (puls 50-150), SR på samtidig 12-punkts EKG-måling som SR; AFLI (puls 50-99), AFLI på samtidig 12-punkts EKG-måling som AFLI, sinustakykardi (puls 100-150), SR på samtidig 12-punkts EKG-måling som høj puls og AFLI med høj puls (puls 100-150), AFLI på samtidig 12-punkts EKG-måling som AFLI. Derudover vurderedes også lighed mellem EKG-appens bølgeform svarende til en enkeltpunkts-EKG-måling i en 12-punkts EKG-måling – som målt med acceptabel morfologi for PQRST-komplekser og R-amplitudeoverensstemmelse – svarende til vurderingen af bølgeformen i den klinisk valideringsundersøgelse af EKG 1.0.

Denne undersøgelse blev godkendt af Institutional Review Board (IRB) for de respektive undersøgelsessteder for den prospektive, USA-baserede undersøgelse på flere forskellige klinikker. Før deltagelse i undersøgelsen indhentede stedets forskere IRB-godkendelse, samtykkeformularer og andet relevant deltagermaterialer på hvert undersøgelsessted.

Figur 3: Diagram over forsøgspersoners disposition



Undersøgelsens deltagere omfattede personer med normal SR på screeningtidspunktet – uden kendt AFLI – og deltagere med kendt vedvarende, permanent eller kronisk AFLI, der blev behandlet for AFLI under screeningen. Alle deltagere blev instrueret i at udføre en 12-punkts EKG-måling samtidig med, at der blev udført en særskilt, komplet enkeltpunkts-EKG-måling med Apple Watch; hver deltager udførte tre prøver. Samme fremgangsmåde blev fulgt for træningssessionerne, hvor forsøgspersoner, der blev vurderet som egnede, blev bedt om at cykle i fem minutter på en motionscykel for at nå en ønsket puls. For både hvile- og træningssessioner gjaldt, at de første prøver blev overvejet til bedømmelse og analyse.

To uafhængige U.S. board-certificerede kardiologer, der ikke kendte til undersøgelsen, gennemgik de 12-punkts EKG-målinger for diagnosticering af puls og rytme. I tilfælde af en uoverensstemmelse blev målingen gennemgået af en tredje ekspert. Pulsen blev beregnet for hver 12-punkts EKG-måling. Pulsen blev registreret, og der blev valgt en pulsdagnosticeringskode, der svarede til den puls, som blev observeret på reference-EKG-målingen. Følgende pulsdagnosticeringer blev bedømt af eksperter og holdt sammen med data fra 12-punkts EKG-målingen: SR, AFLI, supraventrikulær takykardi (SVT), anden uregelmæssig rytme (hyppige for tidlige atriske sammentrækninger, hyppige for tidlige ventrikulære sammentrækninger, atrieflagren, ventrikulær takykardi, ventrikulær fibrillation, andengrads AV-blok type I, andengrads AV-blok type II, tredje grads AV-blok og andet) og kunne ikke tolkes. Tre hjerteteknikere, der ikke kendte til undersøgelsen, gennemgik bølgeformen under vurdering og bedømmelse af EKG-data fra de sammenlignede EKG-strimler. De første seks fortløbende, forskellige og læselige PQRST-komplekser uden artefakter, der matchede mellem forsøgspersonens enhedsstrimmel og referenceenhedens strimmel til evaluering, blev identificeret af én reviewer og brugt af de to andre reviewers. Strimlerne blev ekskluderet, hvis der ikke kunne findes seks fortløbende slag.

Til analyse af hovedformålene blev der implementeret en bootstrap-tilgang for at opnå tosidede 95 procent konfidensintervaller for følsomhed og specificitet, da dataene blev indsamlet fra samme forsøgspersoner i hvile og efter træning. Forsøgspersoner med mindst et bedømt AFLI-resultat (for følsomhed) eller SR (for specificitet) og med et klassificerbart algoritmeresultat (SR eller AFLI) blev valgt tilfældigt med erstatning. 2,5- og 97,5-percentilen af fordelingen af bootstrap-estimer repræsenterede de tosidede 95 procent konfidensgrænser. Hvis de nedre konfidensgrænser for både følsomhed og specificitet overskred de forudangivne præstationsmål i forbindelse med disse præstationsmålinger (90 procent for følsomhed og 92 procent for specificitet), blev nulhypotesen afvist til fordel for den alternative hypotese. Til sammenligning af bølgeform blev der brugt data fra den første prøvekørsel af tilfældigt valgte forsøgspersoner – hvor personerne var i hvile eller efter træning – til at teste bølgeformens vurderingshypotese ved hjælp af en bootstrap-samlingtilgang. Hvis 2,5-percentilen af bootstrap-fordelingen for den godkendte andel af morfologi, oversteg 80 procent, blev nulhypotesen (andelen af godkendt morfologi var mindre end 80 procent) afvist. Det samme kriterium blev anvendt til evaluering af R-amplitudeoverensstemmelse.

Resultater

I alt 546 forsøgspersoner blev tilmeldt undersøgelsen; 305 blev tilmeldt i AFLI-gruppen, og 241 blev tilmeldt i gruppen med normal sinusrytme (NSR). En forsøgsperson fra AFLI-gruppen blev ekskluderet, da vedkommende ikke var kvalificeret. Af de 546 forsøgspersoner gennemførte de 534 personer undersøgelsen (293 i AFLI-gruppen og 241 i NSR-gruppen); 12 forsøgspersoner i AFLI-gruppen gennemførte ikke undersøgelsen.

Tre separate hjerteteknikere gennemgik sættet med vurderingsanalyse af bølgeformen med henblik på ansvar. Sammenlignede strimler fra ur og reference blev brugt til vurderingen. I alt 91 forsøgspersoner (100 procent) havde en læsbar "under hvile"-strimmelsammenligning, 87 forsøgspersoner (100 procent) havde en læsbar "efter træning"-strimmelsammenligning, og 93 forsøgspersoner (100 procent) havde mindst én strimmelsammenligning fra "under hvile" eller "efter træning".

Det klassificerbare analysesæt bestod af alle forsøgspersoner, der havde læsbare, sammenlignede EKG-app 2.0- og 12-punkts EKG-bedømte resultater. I alt 512 forsøgspersoner (n = 279 i AFLI-gruppen og n = 233 i NSR-gruppen) blev vurderet som klassificerbare. Sættet med vurderingsanalyse af bølgeformen – defineret som tilfældigt valgte forsøgspersoner med læsbare, sammenlignede strimler fra EKG-appen og en enkeltpunkts-EKG-måling i en standard EKG-måling med 12 elektroder – bestod af i alt 93 forsøgspersoner (n = 48 i AFLI-gruppen og n = 45 i NSR-gruppen).

Tabel 6 viser de krydsklassificerede data for EKG-app 2.0 og de bedømte 12-punkts EKG-resultater for kombinerede data under hvile og efter træning. En høj grad af lighed mellem de bedømte 12-punkts EKG-resultater og EKG-appens output blev opnået for SR- og AFLI-kategorier. For 12-punkts EKG-strimler, der blev bedømt som SR (n = 470), blev i alt 436 resultater også klassificeret som SR af EKG-appens output. For 12-punkts EKG-strimler, der blev bedømt som AFLI (n = 521), blev i alt 474 resultater klassificeret som AFLI af EKG-appens output. EKG-appen klassificerede 9 resultater som SR, 37 resultater som AFLI, 13 resultater som ikke entydige og 9 resultater som dårlige målinger, der blev bedømt som andet (SVT eller andet) på 12-punkts EKG-målingen.

Tabel 6. Krydsklassificeringstabel for EKG-app og bedømt 12-punkts EKG-resultater (under hvile og efter træning) – klassificerbart analysesæt

Ground Truth med bedømte 12-punkts EKG-resultater						
EKG-appenhedens output	SR (50≤puls≤150)	AFLI (50≤puls≤150)	Andet (SVT eller andet med puls 50-150; puls<50; puls>150)	Kunne ikke tolkes	Ikke tilgængelig	I alt
SR [SR (50-99); NSR sinustakykardi (100-150)]	436	7	9	0	2	454
AFLI (50≤puls≤150)	3	474	37	0	6	520
Ikke entydig [herunder puls<50; puls>150]	10	13	13	0	1	37
Dårlig måling	14	23	9	0	3	49
Ikke tilgængelig	7	4	1	1	0	13
I alt	470	521	69	1	12	1073

Forkortelser: AFLI = atrieflimren, SR = sinusrytme, puls = hjerteslag, SVT = supraventrikulær takykardi

Specificiteten af EKG-app 2.0 blev anslået til at være 99,3 procent, og følsomheden blev anslået til at være 98,5 procent (tabel 7). De nedre konfidensgrænser for specificitetsestimater (98,4 procent) og følsomhedsestimater (97,3 procent) overskred de forudbestemte præstationsmål i forbindelse med disse præstationsmålinger, og derfor blev nulhypotesen for specificitet og følsomhed afvist til fordel for den alternative hypotese, og hovedformålet for specificitet og følsomhed blev opfyldt. Specificiteten under hvile blev anslået til 99,1 procent, og følsomheden under hvile blev anslået til 98,5 procent. Specificiteten efter træning blev anslået til 99,5 procent, og følsomheden efter træning blev anslået til 98,6 procent. En yderligere analyse, der integrerer algoritmeklassificeringskategorien "ikke entydig" i beregningen, blev udført for de kombinerede under hvile- og efter træning-data. Den resulterende specificitet under hvile blev anslået til 97,1 procent, hvor følsomheden blev anslået til 96,0 procent.

En høj grad af lighed blev observeret mellem de bedømte 12-punkts EKG-resultater, pulsresultater og EKG-appenhedens output. Niveaulet af fejlklassificering var lavt, når sammenlignet med referenceresultaterne fra den 12-punkts EKG-måling. Tabel 8 viser de opsummerede resultater, der bekræfter det sekundære formål for korrekt klassificering af patienter med NSR (puls 50-99, SR på samtidig 12-punkts EKG) som SR, AFLI (puls 50-99, AFLI på samtidig 12-punkts EKG) som AFLI, sinustakykardi (puls 100-150, SR på samtidig 12-punkts EKG) som høj puls og AFLI med høj puls (puls 100-150, AFLI på samtidig 12-punkts EKG) som AFLI (høj puls) på en læsbar og klassificerbart EKG-appstrimmel. Procentdelen af SR (puls 50-99), der blev korrekt klassificeret, var 98,3 procent; den nedre konfidensgrænse på 96,6 procent overskred det forudbestemte præstationsmål på 90 procent for dette formål. Procentdelen af AFLI (puls 50-99), der blev korrekt klassificeret, var 98,9 procent; den nedre konfidensgrænse på 97,5 procent overskred det forudangivne præstationsmål på 81 procent. Procentdelen af SR (puls 100-150), der blev korrekt klassificeret, var 90,7 procent; den nedre konfidensgrænse på 86,7 procent overskred det forudangivne præstationsmål på 81 procent for dette formål. Procentdelen af AFLI (puls 100-150), der blev korrekt

klassificeret, var 83,0 procent; den nedre konfidensgrænse på 77,8 procent overskred det forudangivne præstationsmål på 75,5 procent.

Tabel 7. Resultater for følsomhed og specificitet (hovedformål) – klassificerbart analysesæt		
Parameter	Værdi	Tosidet 95 % bootstrap-konfidensinterval*
Kombineret (under hvile og efter træning)		
Referenceresultat = AFLI (puls 50-150)	481	
EKG-app = AFLI (puls 50-150)	474/481 (98,5 %)	
EKG-app = SR (puls 50-150)	7/481 (1,5 %)	
Følsomhed	474/481 (98,5 %)	97,3 %, 99,6 %).
Referenceresultat = SR (puls 50-150)	439	
EKG-app = AFLI (puls 50-150)	3/439 (0,7 %)	
EKG-app = SR (puls 50-150)	436/439 (99,3 %)	
Specificitet	436/439 (99,3 %)	(98,4 %, 100,0 %).
*Resultater hentet fra bootstrap-resampling Forkortelser: AFLI = atrieflimren, SR = sinusrytme, puls = hjerteslag		

Tabel 8. Procent korrekt for klassificerbare rytmer for AFLI og SR		
Parameter	Værdi	Bootstrap-konfidensinterval**
Kombineret (under hvile og efter træning)*		
Procent korrekt for SR-reference (puls 50-99)	230/234 (98,3 %)	(96,6 %, 99,6 %).
Procent korrekt for AFLI-reference (puls 50-99)	272/275 (98,9 %)	(97,5 %, 100 %).
Procent korrekt for SR-reference (puls 100-150)	186/205 (90,7 %)	(86,7%, 94,6 %).
Procent korrekt for AFLI-reference (puls 100-150)	171/206 (83,0 %)	(77,8 %, 88,0 %).
*Procent korrekt beregning inkluderer EKG-app 2.0-klassificerbare rytmer for AFLI og SR **Konfidensinterval hentet fra bootstrap-resampling. Forkortelser: AFLI = atrieflimren, puls = hjerteslag, SR = sinusrytme		

Sammenligning af bølgeform mellem EKG-app algoritme 2.0 og en enkeltpunkts-EKG fra en 12-punkts EKG-måling viste, at den samlede godkendelsesprocent for kombinerede morfologiske resultater under hvile- og efter træning var 100 procent med i alt fem ekskluderede resultater. Den nedre konfidensgrænse på 97,9 procent overskred den forudbestemte, godkendte morfologi, på 80 procent.

Den samlede godkendelsesprocent for kombinerede R-vurderingsresultater under hvile- og efter træning var 97,2 procent; den nedre konfidensgrænse på 93,6 procent overskred den forudbestemte R-amplitudeoverensstemmelse på 80 procent. Under hvile- og efter træning-grupper viste tilsvarende R-bølge.

Konklusioner

Apple Watch-kunder har nu adgang til to funktioner, der kan registrere uregelmæssig puls: Notifikation om uregelmæssig puls (Apple Watch Series 1 og senere) og EKG-appen (Apple Watch Series 4, Series 5 og Series 6).

Funktionen Notifikation om uregelmæssig puls er software, der klassificerer tilfældigt indsamlede intervalmålinger i baggrunden, og gør brugere, der har slået funktionen til, opmærksom på uregelmæssig puls. Apple Heart Study viste, at ud af de deltagere, der fik en notifikation ved samtidig brug af Apple Watch og et EKG-plaster, viste 78,9 procent tilsvarende AFLI på EKG-plasteret, og 98,2 procent viste AFLI og andre klinisk relevante arytmier. Resultaterne viser, at selvom en notifikation i de fleste tilfælde indikerer korrekt tilstedeværelse af AFLI, kan en notifikation i andre tilfælde indikere tilstedeværelse af en anden slags arythmi end AFLI.

Apple Watch Series 4, Series 5 og Series 6 har en elektrisk hjertesensor, der i samarbejde med EKG-appen og algoritmen, genererer en EKG-måling svarende til en enkeltpunkts-EKG-måling, der kan klassificere bølgeformen til at søge efter forekomst af AFLI. Den proprietære algoritme, der er designet til at klassificere disse EKG-målinger viste en følsomhed over 98 procent og en specificitet over 99 procent i identificering af AFLI sammenlignet med EKG-målinger, der blev udført med en referenceenhed og tolket af uafhængige, kliniske eksperter.

Forbrugere, der ønsker at bruge disse to funktioner, skal gennemføre en brugerorienteret træning, hvor de lærer, hvordan resultaterne skal tolkes, lærer om algoritmernes begrænsninger og får at vide, at resultaterne ikke skal opfattes som diagnostiske. Begge funktioner er FDA-godkendt af De Novo eller 510K-klassificeret.