

# Apple Watch bijt weer in het zand als medisch diagnosticum



Twee keer, in [2017](#) en [2018](#), publiceerde ik op deze website artikelen over het tevergeefs ziekten proberen vast te stellen met behulp van smartwatches, in het bijzonder de Apple watch. Uit [een zeer recente publicatie op 21 maart 2018 in de Journal of the American Medical Association \(JAMA\)](#) blijkt dat de genoemde smartwatch in de verste verte niet als betrouwbaar diagnostisch hulpmiddel voor het vaststellen van atriumfibrilleren, een bepaalde vorm van een onregelmatige hartslag, te gebruiken is. De publicatie is weer afkomstig van een onderzoeksgroep aan de University of California, San Francisco (UCSF) met Brandon Ballinger als één van de auteurs. Het falen van de Apple smartwatch voor dit doel is niet direct af te leiden uit de kop en oppervlakkige lezing van de samenvatting van het artikel (“Passive Detection of Atrial Fibrillation Using a Commercially Available Smartwatch”). Pas als men de inhoud tot zich door laat dringen wordt het duidelijk dat de smartwatch jammerlijk faalde om betrouwbaar atriumfibrilleren vast te stellen en dus ten enenmale ongeschikt is om ook deze ziekte toestand vast te stellen.

## Cardiogram

Het onderzoek dat nu naar buiten komt is net als de vorige twee waarover ik schreef (zie hierboven) gedaan met behulp van

de Apple smartwatch in samenhang met de app Cardiogram, gemaakt door het gelijknamige bedrijf. De baas daarvan is Brandon Ballinger, die net als bij de eerdere publicaties mede-auteur is. De technologie die gebruikt is betreft [de photoplethysmografie](#). Hierbij wordt de doorbloeding van de huid gemeten door middel van uitgezonden en weerkaatst rood licht afkomstig van de smartphone. Per definitie is deze methode al inferieur aan het vastleggen van eventueel atriumfibrilleren met een elektrocardiogram. Die meet de elektrische activiteit die het gevolg is van een hartslag en is éénduidig vast te leggen met elektroden. De methode met licht die de doorbloeding van de huid meet kent evidente foutenbronnen als omgevingslicht, huidskleur, lokale pigmentatie etc en zal dus per definitie onbetrouwbaarder zijn.

### **Resultaat(1)**

Uiteraard bleek dat ook uit het eerste dele van het onderzoek dat in de JAMA gepubliceerd is. In een door hen zo genoemd "extern validatie cohort" werd bij een groep van 51 in bed liggende personen, die in ziekenhuizen opgenomen waren voor cardioversie bekeken hoe nauwkeurig de Apple smartwatch met de Cardiogram-app werkte. Cardioversie is het geven van een elektrische schok om het atriumfibrilleren te behandelen en daarmee een regelmatige hartslag te bereiken. Men gebruikte om de data te analyseren kunstmatige intelligentie in de vorm van een deep neural network. De sensitiviteit van de test met de smartwatch was 98 procent en de specificiteit 90,2 procent. Dat **lijkt** zeer hoog en zeer fraai, maar omdat het geen 100 % is betekent het wel dat bij mensen waarvan al bekend is dat ze de ziekteconditie hadden door het maken van een elektrocardiogram(ECG), de meting met het ECG niet volledig gerepliceerd kon worden. De smartwatch was dus inferieur is aan het ECG.

### **Resultaat(2)**

Daarnaast heeft men gekeken wat de nauwkeurigheid was als men bij een niet op atriumfibrilleren geselecteerde groep van 1600 ambulante mensen die mee deden aan het onderzoek en de smartwatch dagelijks om hadden, ging kijken. Dan bleek de nauwkeurigheid van de voorspelling dramatisch te dalen tot slechts 8 procent. Het ging dan om mensen die zelf meldden dat ze atriumfibrilleren hadden en het neurale netwerk met de smartwatch-data het atriumfibrilleren "vaststelde". De sensitiviteit en specificiteit waren dan resp. 67,7 en 67,6 procent en dat is zeer laag. Het betekent dat de methodiek zeer veel fout positieve en fout negatieve uitslagen kent met enerzijds dus ongerust maken van mensen en anderzijds onderdiagnostiek. Eerdere publicaties van de UCSF in samenwerking met Cardiogram lieten hetzelfde beeld zien bij het onderzoek om met smartwatches hypertensie, slaap-apneu en diabetes mellitus vast te stellen. De methode met de smartwatches is gewoonweg niet geschikt om deze ziektecondities betrouwbaar vast te stellen.

### **Proof of Concept???**

Om er toch nog een positief bericht uit persen noemt men het onderzoek een "Proof Of Concept" om daarnaast aan te kondigen dat men de systematiek verder wil verbeteren. Deze opstelling is te karakteriseren als het tegen beter weten in proberen het gebruik van smartwatches voor het opsporen van ziektecondities te blijven promoten. Het is trouwens zeer de vraag of de Amerikaanse Food and Drug Administration ooit haar fiat zou geven voor propageren van de smartwatch als diagnosticum.

### **Triviaal?**

Het lijkt triviaal om je druk te maken over het op slechte wijze propageren van het gebruik van smartwatches voor medisch-diagnostische doeleinden. Men moet echter niet vergeten dat het om zeer grote belangen gaat van producenten van smartwatches, zoals Apple en tevens van bedrijven zoals Apple, Cardiogram en vele anderen. Zij willen grote aantallen

data verkrijgen van personen om daarmee hun eigen verdienmodel te ontwikkelen c.q. bevorderen. We staan hiermee ook weer te kijken naar een vorm van data-grabbing/resourcegrabbing. Positieve berichtgeving, hoe geforceerd ook, dient ertoe om de verkoop en het gebruik van smartwatches verder op te stuwen.

## **Kritisch**

Opvallend is hoe verschillend in diverse media om wordt gegaan met de berichtgeving over het laatste onderzoek. Meerdere media in de V.S([A](#),[B](#),) sluiten geheel of gedeeltelijk hun ogen voor de slechte voorspelbaarheid in een min of meer open onderzoeksgroep, terwijl bijvoorbeeld [The Washington Post](#) een behoorlijk kritisch artikel erover publiceerde. Ook in Nederland was het doorgaans wat minder kritische [online magazine ICT&Health](#) behoorlijk scherp over het beschreven onderzoek.

De gezondheidszorg is ook niet gebaat het pushen van smartwatches als diagnosticum terwijl er veel betere en betrouwbaardere methoden zijn.

W.J. Jongejan