

How to produce bullshit about diabetes detection with heartrate sensors on smartwatches



De afgelopen paar dagen is er in de Amerikaanse media, die publiceren over de zegeningen van de “wearables”, zoals smartwatches en fitnesstrackers a la Fitbit, fors uitgekapt over het kunnen detecteren van diabetes met smartwatches. Ook Nederlandse media doken erop, zoals onder andere [de website Smarthealth](#) en [de iCulture-website](#). De aanleiding is de presentatie van [een onderzoeksgroep van de universiteit van San Francisco samen met de mensen achter de medische startup Cardiogram](#), (klik hier op de link [Recent Claim](#)) die de gelijknamige app maakt, op een congres van de Association for the Advancement of Artificial Intelligence (AAAI) in New Orleans van 2 tot 7 februari. Het onderzoek dat nog niet in vakliteratuur gepubliceerd is pretendeert met behulp van gebruik van hartfrequentie-data van smartwatches (Apple), geanalyseerd met een diep neurale netwerk genaamd DeepHeart, diabetes mellitus met een accuratesse van 85 procent vast te stellen. Eerder publiceerden dezelfde mensen resultaten van onderzoek waarbij met dezelfde methodiek slaapapneu en hypertensie vastgesteld zouden kunnen worden. Dat onderzoek kon op zijn zachtst gezegd “sloppy science” genoemd worden, omdat de beschreven nauwkeurigheid op de keper beschouwd nogal fors tegenvalt en tot forse aantallen fout-positieven en fout-

negatieven leidt. [Ik publiceerde daar op 24 november 2017 over.](#)

Achtergrond

Nog steeds zijn wearables in de ogen van techno-optimisten de heilige graal waarmee men vele ziekten denkt te kunnen voorspellen. Het probleem is dat sensors in wearables een aantal zeer specifieke lichaamsfuncties kunnen volgen, zoals hartslag, temperatuur, bewegingen. Wat men met gebruikmaking van kunstmatige intelligentie probeert is om op indirecte wijze uit de meetwaarden van die sensors conclusies te trekken ten aanzien van ziekten waarvan de sensors niet direct de kenmerken meten. Zo is de achterliggende gedachte bij het proberen diabetes te vast te stellen met gebruikmaking van hartslagmetingen dat er bij diabetes vaak sprake is van bepaalde patronen (andere hartfrequentie in rust en andere variabiliteit in het hartritme) die anders zijn dan bij niet-diabeten.

Wat deed men?

Bij de melding van een accuratesse van 85 moet men zich altijd afvragen wat daarmee nu precies bedoeld wordt. Bij diagnostische apparatuur dient men ook altijd [de sensitiviteit en de specificiteit](#) van de metingen te weten om een enigszins gefundeerd oordeel te kunnen vellen. In dit geval voedde men het neurale netwerk DeepHeart 70 procent van de metingen van in totaal 14000 Apple Watch-gebruikers en ontwikkelde een algoritme dat bij de overige 30 procent uitgetest werd. Daarbij slaagde het algoritme erin om bij de mensen waarvan eerder diabetes was vastgesteld 85 procent te kunnen identificeren. Dat klinkt allemaal heel aardig, maar men miste toch blijkbaar 15 procent van de anderszins bewezen diabeten. Men mat dus uitsluitend de sensibiliteit, het percentage terecht positieve uitslagen onder de zieke personen. Met 15 procent niet gediagnostiseerd en het volledig ontbreken van gegevens over de specificiteit (het percentage terecht

negatieve testuitslagen onder de niet-zieke personen) zegt deze methode om diabetes vast te stellen niets, omdat die in wezen te onnauwkeurig is.

Geen FDA-goedkeuring aangevraagd

Om in de Verenigde Staten eendiagnostische test te mogen verkopen, heeft men goedkeuring van de Food and Drug Administration nodig, zulks om de betrouwbaarheid van die tests te borgen. De oprichter van Cardiogram en drijvende kracht achter alle publiciteit, [Brendan Ballinger](#)([geen arts, maar ICT-er](#)) heeft echter geenszins de bedoeling om voor zijn methodiek een FDA-goedkeuring aan te vragen. Hij weet ook dat de Cardiogram-app nooit als een standalone diagnostische test gebruikt kan worden, maar als een soort vriendelijk advies, en [zegt](#):

“To stay on the right side of the US Food and Drug Administration, the app can’t function as a standalone diagnostic, more like some friendly advice. But the kind of advice an insurer might cover if they thought it would get people into treatment earlier and save healthcare costs .”

Wonderlijk

Hoewel de Cardiogram-app gratis is, ziet Ballinger wel een businessmodel in zijn black-box-intelligentie, dus nadere analyse tegen betaling. Maar als er dan een afwijkende uitslag komt zoals diabetes is zijn advies:

“Right now, we’d always use an existing FDA-cleared (or CLIA-waved) test to confirm the diagnosis.”

Dus, even diep ademhalen, hij adviseert zijn uitslag te laten bevestigen door een test te doen met het gebruikelijk onderzoek van diabetes in bloed met gevalideerde apparatuur en zeer betrouwbare media. Dan is het toch veel simpeler voor een persoon die zich afvraagt of hij/zij diabetes heeft, om de te betalen service van Cardiogram over te slaan en meteen een

gevalideerde bloedtest te laten doen bij de eigen arts.

Ongerustheid

In feite is men met dit soort zogenaamde doorbraken alleen maar bezig de ongerustheid bij mensen over hun gezondheid aan te wakkeren en te exploiteren. Het lijkt allemaal heel mooi en voedt de gedachte dat met hightech-toepassingen van alles gemeten kan worden, maar niets is minder waar. Dit soort propaganda van hightech wordt klakkeloos in de media overgenomen in de vorm van copy-paste-journalistiek zonder het afvragen wat de basale vereisten zijn van betrouwbaar te achten diagnostisch onderzoek. Het nu gepubliceerde verhaal van Cardiogram kan ik slechts als "bullshit" karakteriseren.

W.J. Jongejan