

## **Stopzetting van preventieve cardiovasculaire medicatie bij patiënten in de eerstelijnszorg: het in vrije tekst door middel van natuurlijke taalverwerking openbaar maken van medische gegevens voor epidemiologisch onderzoek**

Meer dan 75% van de Nederlanders van 70 jaar of ouder gebruikt preventieve cardiovasculaire medicatie (PCVM), hoewel de voor- en nadelenbalans naarmate de leeftijd vordert minder gunstig wordt. De nationale richtlijn beschouwt neveneffecten en een beperkte gezondheidsverwachting als gegronde redenen om PCVM stop te zetten. Toch is de bestaande praktijk wat betreft stopzetting van PCVM in de eerstelijnszorg soms anders. Wij willen aan de hand van stelselmatig verzamelde elektronische patiëntgegevens van meer dan 30 huisartspraktijken onderzoeken welke oudere patiënten met PCVM stoppen, en waarom. Gecodeerde informatie over diagnoses, medicatie en laboratoriumtests is eenvoudig toegankelijk en wordt al gebruikt in epidemiologisch onderzoek. De informatie in de beschrijvende verslagen van huisartsen is echter niet eenvoudig te achterhalen. Hierin staan persoonlijke en contextgebonden redenen om te stoppen met PCVM en relevante symptomen en diagnoses die niet gecodeerd zijn. Deze informatie zou de gestructureerde data aanvullen en zou de statistische analyses nauwkeuriger maken. Dit project heeft tot doel de ongestructureerde data beschikbaar te stellen voor kwantitatief onderzoek met natuurlijke taalverwerking. Een data scientist zal de data opschonen, woordinbeddingen maken en de redenen om met PCVM te stoppen, sorteren. De interne validiteit van de opgehaalde data zal worden beoordeeld volgens handmatige extractie overeenkomstig de gouden standaard, en de externe validiteit met data die tot dat moment ongezien zijn. Het script en de expertise die dit project zal opleveren, zullen ook nuttig zijn voor ander onderzoek binnen onze afdeling. In de toekomst wil de afdeling onder andere met behulp van zelflerende systemen vooralsnog onbekende factoren in de (on)gestructureerde data opsporen die relevant zijn voor een vroege opsporing van ziekte, het verloop van ziekten en therapietrouw.