



# Bij welke sociaal-demografische groepen komt suïcide het meest voor?

## Inzicht in interacties van risicofactoren met behulp van machine learning

Guus Berkelmans · Renske Gilissen · Sandjai Bhulai · Rob van der Mei · Lizanne Schweren

Geaccepteerd op: 14 april 2023  
© The Author(s) 2023

### Samenvatting

**Inleiding** Voor suïcidepreventie is het van belang om groepen met een verhoogd risico op suïcide zo goed mogelijk te identificeren. Tot nog toe is er weinig bekend over interacties van meerdere risicofactoren. Machine learning-methoden bieden nieuwe mogelijkheden voor flexibel, datagedreven, hypothesevrij en robuust onderzoek naar de interacties van risicofactoren voor suïcide.

**Methode** In dit onderzoek is gebruikgemaakt van microdata van het Centraal Bureau voor de Statistiek met sociaal-demografische gegevens van alle Nederlandse inwoners. Er is een machine learning-model ontwikkeld en uitgevoerd om interacties van risicofactoren voor suïcide te vinden.

**Resultaten** De resultaten laten disproportioneel hoge suïcide-incidenties zien onder 1) mensen met een arbeidsongeschiktheidsuitkering die nooit getrouwd zijn geweest, 2) arbeidsongeschikte mannen, 3) mensen tussen de 55 en 69 jaar oud, die alleen wonen, nooit getrouwd zijn en wiens huishoudinkomen laag is. Daarnaast hebben we hoge suïcidecijfers gevonden onder mensen tussen de 25 en 39 jaar met een laag opleidingsniveau en onder verweduwd mannen.

**Conclusie** De bevindingen helpen om gericht interventies in te kunnen zetten om het aantal suïcides te verminderen.

**Trefwoorden** zelfmoordpreventie · machine learning · interacties van risicofactoren

**Among which socio-demographic groups is suicide most common? Understanding interactions of risk factors using machine learning**

### Abstract

**Introduction** For suicide prevention it is important to identify high-risk groups as best we can. So far not much is known about the interactions between multiple risk factors for suicide. Machine learning methods offer new possibilities for flexible, data driven, hypothesis-free and robust research into the accumulation of risk factors for suicide.

**Methods** In this study microdata from Statistics Netherlands was used. This encompasses socio-demographic data from all Dutch inhabitants. A machine learning model was developed and estimated to find interactions of risk factors for suicide.

**Results** The results show disproportionately high suicide rates among 1) people who were never married and are unfit for work, 2) unfit for work males, 3) people between the ages of 55 and 69, who live alone, were never married, and have a low household income. Additionally, we found high suicide rates among people between the ages of 25 and 39 with a low level of education, and among widowed males.

**Conclusion** These findings will be helpful in implementing targeted interventions to reduce the amount of suicides.

**Keywords** Suicide prevention · Machine learning · Interactions of risk factors

G. Berkelmans · R. Gilissen (✉) · L. Schweren  
113 Zelfmoordpreventie, Amsterdam, Nederland  
r.gilissen@113.nl

S. Bhulai  
VU Business Analytics, Amsterdam, Nederland

G. Berkelmans · R. van der Mei  
Centrum Wiskunde & Informatica, Amsterdam, Nederland



## Inleiding

In Nederland sterven gemiddeld vijf mensen per dag aan suïcide. Elk van deze gevallen is een tragedie, voor zowel het slachtoffer als de nabestaanden. De oorzaak van en aanleiding tot suïcide is complex en daarmee het terugdringen van suïcides ook. Net als bij hiv, verkeersongevallen, roken, enzovoort, waarbij het is gelukt om het aantal slachtoffers terug te dringen, vereist suïcidepreventie maatregelen op meerdere niveaus [1]: 1) universele preventiestrategieën die zich op de gehele populatie richten, zoals verantwoorde berichtgeving in de media, publiekscampagnes en de beperking van de toegang tot middelen om suïcide te plegen; 2) selectieve preventiestrategieën die zich op risicogroepen richten: doelgroepen binnen een populatie waarvan we weten dat ze meer risico hebben op suïcide. Voorbeelden van interventies op dit niveau zijn gatekeepertraining (waar professionals getraind worden om suïcidaliteit te signaleren en bespreekbaar te maken) op plekken waar relatief veel mensen met meer risico komen en preventieprogramma's op scholen om mentale problematiek in een vroeg stadium aan te pakken [2, 3]. En 3) geïndiceerde preventiestrategieën die gericht zijn op de specifiek kwetsbare personen binnen de bevolking: mensen die suïcidegedachten hebben of een suïcidepoging hebben gedaan [1]. Voorbeelden van deze interventies zijn cognitieve gedragstherapie voor mensen met depressieve klachten en nazorg voor mensen die na een suïcidepoging op de spoedeisende hulp terechtkomen.

Voor interventies op het tweede niveau (selectieve preventie) is het van belang om groepen met een verhoogd risico op suïcide te identificeren en te detecteren. Er zijn al meerdere onderzoeken naar risicofactoren voor zelfdoding gedaan [4–9]. Niet geheel onverwacht zijn de belangrijkste voorspellers voor overlijden door suïcide een eerdere niet-fatale suïcidepoging of een eerdere psychiatrische opname [7]. Traumatische levensgebeurtenissen en het hebben van mentale problemen, zoals depressie of verslaving, verhogen het risico op suïcidale gedachten en pogingen substantieel, wat vervolgens het risico op overlijden door suïcide verhoogt [7]. Daarnaast is er een aantal sociaal-demografische groepen met een verhoogd risico, zoals mannen, mensen van middelbare leeftijd en alleenwonenden [7, 9].

Bij complexe en multifactoriële gebeurtenissen zoals suïcide is bekend dat interacties van risicofactoren een substantiële rol spelen. Met andere woorden, een risicofactor 'A' die het risico groter maakt onder mensen die al risicofactor 'B' hebben dan onder mensen die risicofactor 'B' niet hebben. Het is bijvoorbeeld bekend dat stressvolle levensgebeurtenissen sneller tot depressie leiden bij mensen met een genetische kwetsbaarheid voor depressie [10]. Over interacties van sociaal-demografische risicofactoren voor suïcide is echter nog weinig bekend. In een hypothetisch voorbeeld: het is aannemelijk dat werkloosheid een

grotere risicofactor is voor mannen die alleen wonen dan voor de rest van de bevolking. De detectie van relevante interacties van risicofactoren maakt het makkelijker om de hoogrisicogroepen op te sporen, wat de effectiviteit van selectieve preventiemaatregelen ten goede kan komen en zo kan bijdragen aan minder suïcides.

Machine learning-methoden bieden nieuwe mogelijkheden voor flexibel, datagedreven, hypothesevrij en robuust onderzoek naar de interacties van risicofactoren voor suïcide. In dit artikel presenteren we een nieuw machine learning-model dat toestaat om complexe interacties van sociaal-demografische risicofactoren te onderzoeken. Ons model resulteert in gedetailleerde en interpreteerbare resultaten die het mogelijk maken groepen met een hoog risico op suïcide te identificeren, waardoor het model kan helpen om gerichte interventies in te zetten. Een Engelstalig artikel over dit onderzoek is onlangs verschenen in een internationaal wetenschappelijk peer-reviewed tijdschrift [11].

## Methode

### *De data*

Voor dit onderzoek is gebruikgemaakt van microdata van het Centraal Bureau voor de Statistiek (CBS). Het CBS onderhoudt een database van hoge kwaliteit, die onder andere sociaal-demografische data van iedere Nederlandse ingezetene bevat. Analyses op de data van het CBS worden uitgevoerd via een remote access-verbinding naar hun servers. Alle resultaten worden gecontroleerd voordat ze vrijgegeven worden om te zorgen dat aan de privacyregelgeving wordt voldaan.

Er zijn data gebruikt van alle Nederlandse ingezetenen die ouder zijn dan tien jaar (onder de tien jaar komt zelfdoding niet tot zeer zelden voor), op 31 december van negen opeenvolgende jaren (2011 tot en met 2019), voor een totaal van 137.666.515 persoonsjaren. Een persoonsjaar omvat één persoon gedurende één jaar, dus als iemand van 2011 tot en met 2019 elk jaar in de data voorkomt, telt diegene voor negen persoonsjaren. Is deze persoon bijvoorbeeld geëmigreerd en komt deze voor het laatst in 2014 in de data voor, dan telt deze persoon voor vier persoonsjaren mee. In 16.417 persoonsjaren was er sprake van suïcide.

De volgende sociaal-demografische factoren zijn in dit onderzoek meegenomen: geslacht, leeftijd, migratieachtergrond, persoonlijk inkomen, huishoudinkomen, netto vermogen van het huishouden, opleidingsniveau, kosten voor somatische zorg, rol in het huishouden (thuiswonend kind, alleenwonend, partner van koppel met thuiswonende kind(eren), partner van koppel zonder thuiswonende kind(eren)), burgerlijke stand (nooit getrouwd of geregistreerd partnerschap, getrouwd of geregistreerd partnerschap, verwe-

duwd, gescheiden), WW-uitkering, bijstandsuitkering en arbeidsongeschiktheidsuitkering.

### Het model

Er is een model ontwikkeld om interacties te vinden die een toegevoegd risico (of een verminderd risico) opleveren. De verkregen interacties zijn geprioriteerd op basis van zowel statistische significantie als verbetering van het model. Het algoritme bevat vier stappen: 1) splitsing van de data, 2) opsporen van interacties, 3) schatten van het model en 4) meten van de performance. Zie [11] voor details van de methode.

1. Splitsing van de data: de data zijn gesplitst in drie niet-overlappende datasets: een trainingsset, een validatieset en een testset. De trainingsset bevat 50% van de persoonsjaren waarin sprake was van een suïcide ( $n=8.214$ ) en 1% van de referentiepopulatie ( $n=1.377.055$ ) en is gebruikt om interacties op te sporen. De validatieset bevat 40% van de persoonsjaren die resulteerden in een suïcide ( $n=6.512$ ) en 1% van de referentiepopulatie ( $n=1.377.870$ ) en is gebruikt om het uiteindelijke logistisch regressiemodel te schatten. De testset bestaat uit de overige 10% van de persoonsjaren die resulteerden in een suïcide ( $n=1.691$ ) en 1% van de referentiepopulatie ( $n=1.375.966$ ) en is gebruikt om de prestatie van het uiteindelijke model te meten.
2. Opsporen van de interacties: het model detecteert binnen de trainingsset relevante interacties van risicofactoren. Dit doet het door steeds een interactie van de vorm 'A en B' toe te voegen (dus bijvoorbeeld 'man en nooit getrouwd' of '55–69 jaar oud en alleenwonend'). 'A' kan hiervoor zelf ook een interactieterm zijn en 'B' houden we altijd bij een van onze originele variabelen. Hierdoor evalueert het model dus ook interacties met drie of meer variabelen, zoals 'man en nooit getrouwd en 55–69 jaar oud'. We bepalen welke interacties worden toegevoegd aan de hand van een statistiek die zowel de mate meeneemt waarin het model er beter van wordt, als de mate meeneemt waarin het waarschijnlijk is dat het model statistisch significant bijdraagt.
3. Schatten van het model: een logistisch regressiemodel wordt geschat binnen de validatieset die alle interacties die in stap 2 zijn verkregen samen met de originele variabelen meeneemt. Omdat de data in de validatieset in stap 2 niet zijn gebruikt, is er geen sprake van overfitting en kunnen de reguliere teststatistieken op de normale manier geïnterpreteerd worden.
4. Meten van de performance: op de testset zijn vervolgens performance statistics berekend: log-likelihood als indicator van model fit (hoe goed de voorspellingen en de observaties matchen) en area under the receiver operating characteristics (ROC) curve (AUC) als indicator van hoe goed het model kan onderscheiden wie er aan suïcide overleed en wie niet.

Voor elke risicofactor en interactie van risicofactoren rapporteren we de  $\beta$ -parameters van het logistische regressiemodel, de oddsratio's (OR) en de corresponderende betrouwbaarheidsintervallen (BI). Voor de interactietermen rapporteren we ook de compound oddsratio's (COR) en de corresponderende betrouwbaarheidsintervallen. De COR's zijn het totale effect van de risicofactoren wanneer ze gecombineerd zijn (bijvoorbeeld voor verweeduwde mannen:  $\exp(\beta_{\text{man}} + \beta_{\text{verweuwd}} + \beta_{\text{man en verweuwd}})$ ). Ook wordt het aantal suïcides van de corresponderende deelbevolkingen in de validatieset gerapporteerd, zowel in absolute zin als in relatieve zin per 100.000 per jaar.

## Resultaten

### Individuele risicofactoren

De belangrijkste individuele risicofactoren voor suïcide zijn man zijn, een uitkering ontvangen, middelbare leeftijd, alleen wonen, hoge zorgkosten hebben en gescheiden zijn (respectievelijk  $OR_{\text{man vs vrouw}} = 2,29$ ; 95%-BI = 2,14–2,46;  $OR_{\text{arbeidsongeschikt vs niet}} = 3,67$ ; 95%-BI = 3,18–4,23;  $OR_{\text{bijstand vs niet}} = 1,72$ ; 95%-BI = 1,52–1,95;  $OR_{\text{WW vs niet}} = 1,21$ ; 95%-BI = 1,09–1,35;  $OR_{40-54 \text{ vs } 25-39} = 1,62$ ; 95%-BI = 1,48–1,76;  $OR_{55-69 \text{ vs } 25-39} = 1,45$ ; 95%-BI = 1,24–1,68;  $OR_{\text{alleenwonend vs deel stel zonder kinderen}} = 2,41$ ; 95%-BI = 2,17–2,66;  $OR_{5-10 \text{ k zorgkosten/jaar vs geen}} = 2,39$ ; 95%-BI = 1,89–3,02;  $OR_{>10 \text{ k/jaar vs geen}} = 4,62$ ; 95%-BI = 3,54–6,05;  $OR_{\text{gescheiden vs nooit getrouwd}} = 1,67$ ; 95%-BI = 1,48–1,86).

De belangrijkste beschermende factoren voor suïcide zijn vrouw zijn, jongere leeftijd, het hebben van een niet-westerse migratieachtergrond en het hebben van een hoger inkomen. ( $OR_{\text{vrouw vs man}} = 0,44$ ; 95%-BI = 0,41–0,47;  $OR_{10-24 \text{ vs } 25-39} = 0,43$ ; 95%-BI = 0,37–0,49;  $OR_{1\text{e: generatie niet-westers vs geen migratieachtergrond}} = 0,36$ ; 95%-BI = 0,32–0,41;  $OR_{2\text{e: generatie niet-westers vs geen migratieachtergrond}} = 0,59$ ; 95%-BI = 0,50–0,70;  $OR_{\text{persoonlijk inkomen in het 4e: kwartiel vs het 1e: kwartiel}} = 0,54$ ; 95%-BI = 0,48–0,61).

In de hele bevolking van tien jaar of ouder zijn er 11,8 suïcides per 100.000 inwoners gedurende de periode van dit onderzoek. Het hoogste relatieve aantal komt voor onder mensen met een arbeidsongeschiktheidsuitkering (47,0 suïcides per 100.000), gevolgd door mensen met een bijstandsuitkering (32,1 suïcides per 100.000). De overige individuele risicofactoren zitten onder de 30,0 per 100.000.

### Interacties van risicofactoren

In fig. 1 staan de twintig interacties die gevonden zijn en in het uiteindelijke logistische regressiemodel zijn meegenomen. Van deze twintig waren er zeventien significant op de validatieset ( $p < 0,05$ ). Onder de interacties waren twaalf groepen met relatieve suïcidecijfers boven de dertig per 100.000, waarvan vier met relatieve suïcidecijfers boven de vijftig per 100.000.

**Figuur 1** Interactietermen gevonden door het algoritme en getest op de validatieset.  $\beta$ -parameters, oddsratio's, compound-oddsratio's en absolute en relatieve aantallen suïcides binnen de validatieset zijn aangegeven. Interacties in het rood hebben een relatief aantal suïcides boven de 50/100.000 en die in het oranje tussen de 30 en 50/100.000. Niet-significante interacties zijn *cursief*

interactie	$\beta$ (95%-BI)	oddsratio (95%-BI)	compound oddsratio (95%-BI)	aantal suïcides	Relatief aantal suïcides
25-39 jaar oud en laagopgeleid	0,46 (0,30-0,62)	1,58 (1,35 - 1,86)	1,63 (1,38 - 1,93)	259	20,07
40-54 jaar oud en bijstand	-0,22 (-0,41 - -0,04)	0,80 (0,67 - 0,96)	2,23 (1,90 - 2,61)	234	35,58
55-69 jaar oud en alleenwonend	-0,42 (-0,67 - -0,17)	0,66 (0,51 - 0,84)	2,27 (1,78 - 2,90)	833	35,54
55-69 jaar oud en alleenwonend en Nederlandse migratieachtergrond	0,18 (-0,04 - 0,39)	1,20 (0,96 - 1,48)	2,71 (2,30 - 3,19)	728	39,37
55-69 jaar oud en alleenwonend en huishoudinkomen in het eerste kwartiel en nooit getrouwd	-0,21 (-0,43 - 0,01)	0,81 (0,65 - 1,01)	3,44 (2,60 - 4,55)	229	57,22
55-69 jaar oud en nooit getrouwd	0,32 (0,15 - 0,50)	1,38 (1,16 - 1,65)	2,00 (1,64 - 2,44)	427	34,81
55-69 jaar oud en deel koppel zonder thuiswonend kind	-0,46 (-0,63 - -0,29)	0,63 (0,53 - 0,75)	0,91 (0,79 - 1,05)	622	9,38
55-69 jaar oud en zorgkosten van meer dan € 10.000	-0,44 (-0,63 - -0,25)	0,64 (0,53 - 0,78)	4,30 (3,16 - 5,86)	238	30,70
70 of ouder en zorgkosten van meer dan € 10.000	-0,66 (-0,88 - -0,44)	0,52 (0,41 - 0,64)	2,14 (1,58 - 2,90)	175	15,59
Man en arbeidsongeschikt	-0,39 (-0,54 - -0,24)	0,68 (0,59 - 0,78)	2,48 (2,21 - 2,79)	642	58,56
Man en deel van koppel met thuiswonend(e) kind(eren)	0,64 (0,48 - 0,80)	1,90 (1,61 - 2,22)	0,82 (0,73 - 0,92)	801	10,94
Man en verweduwd	0,54 (0,33 - 0,74)	1,72 (1,40 - 2,09)	1,56 (1,31 - 1,86)	218	31,31
Man en zorgkosten van meer dan € 10.000	-0,30 (-0,46 - -0,14)	0,74 (0,63 - 0,87)	3,42 (2,64 - 4,43)	456	27,48
Nooit getrouwd en arbeidsongeschikt	-0,03 (-0,26 - 0,19)	0,97 (0,77 - 1,21)	3,54 (2,77 - 4,53)	441	88,48
Arbeidsongeschikt en persoonlijk inkomen in het tweede kwartiel	-0,38 (-0,53 - -0,23)	0,68 (0,59 - 0,80)	1,98 (1,65 - 2,38)	234	35,58
Opleidingsniveau onbekend en zorgkosten tussen € 1 en € 5000	0,28 (0,16 - 0,41)	1,32 (1,17 - 1,51)	1,21 (0,95 - 1,54)	833	35,53

Vergeleken met de algemene Nederlandse bevolking ouder dan tien jaar (11,8 suïcides per 100.000 personen per jaar) hebben mensen die nooit getrouwd zijn geweest en arbeidsongeschikt waren (en onder hen degenen met lage zorgkosten) tot wel 7,5 keer zo'n groot risico om te overlijden door

suïcide (88 suïcides per 100.000 personen per jaar). Hoewel deze groep relatief klein is, vonden er over de periode 2012–2020 gemiddeld per jaar meer dan 100 suïcides plaats in deze groep (7% van alle suïcides over deze periode). De derde hoogrisicogroep zijn de arbeidsongeschikte mannen, met 59 suïcides



per 100.000 personen per jaar, en de vierde groep zijn mensen van 55–69 jaar oud, die nooit getrouwd waren, alleen wonen en daarnaast in de laagste 25% huishoudinkomens zitten (57 suïcides per 100.000 personen per jaar).

Hiernaast vonden we twee interacties die enkel samen met andere factoren een verhoging opleveren. Ten eerste: hoewel noch jongvolwassenen (25–39), noch laagopgeleiden individueel hogererisicogroepen zijn, zien we bij de interactie (laag opgeleiden van 25–39) wel een verhoogd risico. Onder jongvolwassenen kwam suïcide meer dan twee keer zo vaak voor onder mensen met een laag opleidingsniveau dan onder andere jongvolwassenen (20,1 versus 8,8 suïcides per 100.000 per jaar). Ten tweede vonden we dat verweerd zijn voor vrouwen geen risicofactor is (OR=0,91; 95%-BI van OR=0,76–1,10), maar wel voor mannen (OR=1,72; 95%-BI van OR=1,40–2,09; 95%-BI van COR=1,31–1,86). Het relatieve aantal suïcides onder verweerde mannen is 31 suïcides per 100.000 personen per jaar (zie fig. 1).

## Beschouwing

In dit onderzoek hebben we een nieuw model ontwikkeld om interacties tussen hoogrisicofactoren te vinden. We hebben vier groepen gevonden waarin meer dan vijftig per 100.000 personen per jaar overlijden aan suïcide. Gegeven dat in de gehele populatie gemiddeld 11,8 personen per 100.000 overlijden aan suïcide, vormen deze groepen zeerhoogrisicogroepen. Daarnaast hebben we factoren geïdentificeerd die samen het risico op zelfdoding verhogen, terwijl ze dat in isolatie niet noodzakelijkerwijs doen. Deze interacties hadden met standaardvoorspellingsmodellen niet gedetecteerd kunnen worden.

We hebben vier hoogrisicogroepen geïdentificeerd die gerelateerd lijken aan sociaal-economische problematiek: mensen die nooit getrouwd zijn geweest en een arbeidsongeschiktheidsuitkering ontvangen (88 suïcides per 100.000 personen per jaar); mensen die nooit getrouwd zijn geweest, een arbeidsongeschiktheidsuitkering ontvangen en lage zorgkosten hadden (83 suïcides per 100.000 per jaar); mannen met een arbeidsongeschiktheidsuitkering (59 suïcides per 100.000 personen per jaar); en mensen van 55–69 jaar oud die nooit getrouwd waren, alleen wonen en daarnaast in de laagste 25% huishoudinkomens zitten (57 suïcides per 100.000 personen per jaar).

Op basis van deze bevindingen raden wij suïcidepreventiemaatregelen aan bij instanties en professionals die vaak met deze mensen werken, zoals mensen die een arbeidsongeschiktheidsuitkering ontvangen (zorgmedewerkers, arbo-artsen, arbeidsdeskundigen en maatschappelijke dienstverleners). Een van de mogelijkheden is een suïcidepreventietraining waarbij je leert hoe je signalen herkent en suïcidaliteit bespreekbaar kan maken [2, 3]. Om er achter te komen

wat er exact aan de hand is bij deze combinaties van risicofactoren dat maakt dat er zo'n hoog risico is op suïcide is meer (kwalitatief) onderzoek nodig.

Naast de hoogrisicogroepen van meer dan vijftig suïcides per 100.000 personen per jaar hebben we ook risicofactoren gevonden die enkel in het bijzijn van andere factoren een verhoogd risico opleveren. Ten eerste, hoewel noch jongvolwassenen (25–39 jaar), noch laagopgeleiden afzonderlijk gerelateerd zijn aan een hoger risico op suïcide, zien we dat de overlappende groep (laag opgeleiden van 25–39 jaar) wel een verhoogd risico heeft om door suïcide te overlijden. Om te onderzoeken waarom ze meer risico lopen, raden wij aan om nader onderzoek te doen naar deze groep, waarin onder andere sociaal-economische factoren, externe stressoren, psychosociale omstandigheden en psychische kwetsbaarheden worden meegenomen.

Ten tweede verhoogt het verweerd zijn het risico op suïcide niet in de algemene bevolking, maar wel sterk onder mannen. Onder verweerde mannen is het suïcidecijfer twee keer zo hoog als onder mannen in het algemeen. In eerdere onderzoeken werd een relatief hoog risico op suïcide gevonden onder verweerden [12–14], maar voor zover wij weten is dit het eerste onderzoek dat aantoonde dat in Nederland dit hogere risico onder verweerden hoofdzakelijk voorkomt onder mannen. Een eerder onderzoek heeft aangetoond dat verweerde mannen vergeleken met verweerde vrouwen minder inkomensverlies hebben, maar vaker negatieve gevolgen zoals eenzaamheid en depressie ervaren [15]. Onze bevindingen onderstrepen de noodzaak van sociale steun aan mannen die hun partner hebben verloren, en sporen aan tot het trainen van gatekeepers onder professionals die in contact komen met deze groep.

Ten slotte wijzen we op twee risicofactoren die in een groot aantal interacties voorkomen: 1) tussen de 55–69 jaar oud zijn en 2) nooit getrouwd zijn. Het grote aantal significante interacties met deze risicofactoren suggereert dat de risicoprofielen binnen deze groepen anders zijn dan in de algemene bevolking.

Ons onderzoek kent enkele beperkingen die voor de interpretatie van de resultaten van belang zijn. We hebben alle suïcides van Nederlandse ingezetenen tussen 2012 en 2020 meegenomen. We kunnen niet zien of de resultaten stabiel zijn over de tijd. Hiernaast is het model bottom-up opgebouwd door significante interactie-effecten toe te voegen. Met een top-down-aanpak, waar je met alle mogelijke interacties begint, kunnen mogelijk meer hoogrisico-interacties gevonden worden. Zo'n aanpak zal echter meer foutpositieven genereren en vereist disproportioneel veel rekenkracht. Ten derde willen we benadrukken dat dit model geschikt is voor het identificeren van interacties van risicofactoren en niet gebruikt kan worden als taxatie- of screeningsinstrument om het risico in te schatten dat iemand zal overlijden door suïcide. Dit onderzoek draagt bij aan het opsporen van hoog-



risicogroepen, zodat we beter weten waar selectieve preventiemaatregelen het meest effect kunnen hebben om suïcides te verminderen. Een ander voordeel van ons onderzoek is dat ons model hypothesevrij is, waardoor niet eerder bekende of veronderstelde interacties van risicogroepen gevonden konden worden. Tot slot genereert het model, zeker in vergelijking met andere machine learning-methoden, resultaten die goed interpreteerbaar zijn.

Samenvattend hebben we een machine learning-model ontwikkeld en uitgevoerd om interacties van risicofactoren voor suïcide te vinden. We hebben disproportioneel hoge suïcide-incidenties gevonden onder 1) arbeidsongeschikten die nooit getrouwd waren, 2) arbeidsongeschikte mannen, 3) mensen tussen de 55 en 69 jaar die alleen wonen, nooit getrouwd zijn en wier huishoudinkomen laag was. Daarnaast hebben we hoge suïcideaantallen gevonden onder mensen tussen de 25–39 jaar met een laag opleidingsniveau en onder verweduwde mannen. Het is wenselijk deze bevindingen in toekomstig suïcidepreventiebeleid op te nemen.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## Literatuur

- World Health Organization. Preventing suicide: a global imperative. 2014. <https://apps.who.int/iris/handle/10665/131056>. Geraadpleegd op: 24 mei 2022.
- 113 Zelfmoordpreventie. Suïcidepreventietraining gatekeeper. <https://www.113.nl/voor-professionals/113-academy-voorlichting-workshops-en-trainingen>. Geraadpleegd op: 12 apr 2023.
- Terpstra S, Beekman A, Abbing J, Jaken S, Steendam M, Gilissen R. Suicide prevention gatekeeper training in the Netherlands improves gatekeepers' knowledge of suicide prevention and their confidence to discuss suicidality, an observational study. *BMC Public Health*. 2018;18(1):637.
- Ayhan G, Arnal R, Basurko C, About V, Pastre A, Pinganaud E, et al. Suicide risk among prisoners in French Guiana: prevalence and predictive factors. *BMC Psychiatry*. 2017;17(1):156.
- Bhatt M, Perera S, Zielinski L, Eisen RB, Yeung S, El-Sheikh W, et al. Profile of suicide attempts and risk factors among psychiatric patients: a case-control study. *PLoS ONE*. 2018;13(2):e192998.
- Choi SB, Lee W, Yoon JH, Won JU, Kim DW. Risk factors of suicide attempt among people with suicidal ideation in South Korea: a cross-sectional study. *BMC Public Health*. 2017;17(1):579.
- Franklin JC, Ribeiro JD, Fox KR, Bentley KH, Kleiman EM, Huang X, et al. Risk factors for suicidal thoughts and behaviors: a meta-analysis of 50 years of research. *Psychol Bull*. 2017;143(2):187–232.
- Parra-Urbe I, Blasco-Fontecilla H, Garcia-Parés G, Martínez-Naval L, Valero-Coppin O, Cebrià-Meca A, et al. Risk of re-attempts and suicide death after a suicide attempt: a survival analysis. *BMC Psychiatry*. 2017;17(1):163.
- Berkelmans G, Mei R van der, Bhulai S, Gilissen R. Identifying socio-demographic risk factors for suicide using data on an individual level. *BMC Public Health*. 2021;21(1):1702.
- Uher R. Gene-environment interactions in severe mental illness. *Front Psychiatry*. 2014;5:48.
- Berkelmans G, Schwenen L, Bhulai S, Van der Mei R, Gilissen R. Identifying populations at ultra-high risk of suicide using a novel machine learning method. *Compr Psychiatry*. 2023;123:152380.
- Bower KL, Emerson KG. Exploring contextual factors associated with suicide among older male farmers: results from the CDC NVDRS dataset. *Clin Gerontol*. 2021;44(5):528–35.
- Yang J, He G, Chen S, Pan Z, Zhang J, Li Y, et al. Incidence and risk factors for suicide death in male patients with genital-system cancer in the United States. *Eur J Surg Oncol*. 2019;45(10):1969–76.
- Richardson C, Robb KA, O'Connor RC. A systematic review of suicidal behaviour in men: a narrative synthesis of risk factors. *Soc Sci Med*. 2021;276:113831.
- Streeter JL. Gender differences in widowhood in the short run and long run: financial and emotional well-being. *Innov Aging*. 2019;3(Supplement\_1):S736–S736.