

Kennisontwikkeling in de klinische psychiatrie: leren van elektronische patiëntendossiers

V.J. Menger, M.R. Spruit, F.E. Scheepers

- Achtergrond** Vandaag de dag leggen vrijwel alle psychiatrische zorginstellingen informatie rondom de zorg die zij leveren vast in een elektronisch patiëntendossier (EPD). Door deze zorgdata met innovatieve en geavanceerde technieken te analyseren kunnen ze in de nabije toekomst een belangrijke bron van nieuwe kennis zijn, en daarmee bijdragen aan verbetering van psychiatrische zorg.
- Doel** Onderzoeken op welke manier data uit EPD's relevante kennis en inzichten kunnen opleveren voor de psychiatrische zorg.
- Methode** We beschrijven oplossingen die wij ontwierpen voor enkele technische, organisatorische en ethische vraagstukken rondom het ontsluiten van zorgdata. Daarna analyseren we de verkregen zorgdata uit EPD's met technieken uit de knowledge discovery, het proces waarbij nieuwe en bruikbare informatie uit data wordt geëxtraheerd. We gebruiken onder meer datavisualisatie, machine learning en natural language processing om te demonstreren wat voor soort resultaten te behalen zijn.
- Resultaten** Onze aanpak toont aan dat het mogelijk is om in samenwerking met behandelaars en patiënten nieuwe en interessante inzichten te genereren die op geaggregeerd niveau verscholen zitten in EPD's. In het bijzonder laten we zien hoe het risico op agressief gedrag van patiënten tijdens opname effectief en nauwkeurig ingeschat kan worden aan de hand van klinische tekst in het EPD.
- Conclusie** Als men aandacht besteedt aan enkele belangrijke vraagstukken rondom het analyseren van EPD's, is leren van data uit EPD's een nieuwe en interessante benadering met duidelijke potentie voor het verbeteren van de psychiatrische zorg.

In de gezondheidszorg van vandaag de dag is al veel vooruitgang geboekt in het verbeteren van de gezondheid en het welzijn van de algehele populatie. In vrijwel alle medische specialismen bestaan er echter nog altijd vele open vragen rondom aspecten van diagnose en behandeling, en de psychiatrie is daarop geen uitzondering (Krumholz 2014). Verdere uitbreiding van onze kennis over psychiatrische aandoeningen is dringend nodig om de kwaliteit van zorg te kunnen blijven verbeteren. Om tegelijkertijd ook de stijgende kosten van de geestelijke gezondheidszorg in de hand te houden zijn nieuwe en innovatieve benaderingen van groot belang (Appleby 2012).

Data-analyse

Het analyseren van data uit elektronische patiëntendossiers (EPD's) kan één zo'n innovatieve benadering zijn (Raghupathi & Raghupathi 2014). Hoewel de gege-

vens in het EPD primair worden vastgelegd om goede zorg te kunnen leveren, ligt er op geaggregeerd niveau ook allerlei kennis in verscholen. Deze kennis is met *knowledge discovery*-technieken expliciet te maken, een term uit de informatica die Frawley e.a. (1992) definiëerden als '*the non-trivial extraction of implicit, previously unknown, and potentially useful information from data.*' Daarbij maakt men gebruik van nieuwe technieken uit verschillende vakgebieden binnen en buiten de informatica, zoals statistiek, *machine learning* en datavisualisatie.

Deze technieken zijn ook zeer bruikbaar om grote datasets uit EPD's te integreren en te analyseren. In veel domeinen buiten de gezondheidszorg worden routinematig verzamelde data al op die manier geanalyseerd, en ook binnen de gezondheidszorg neemt het aantal wetenschappelijke studies met gebruikmaking van data uit EPD's toe (Zhu & Zheng 2018). Tot dusver gebeurt dit

AUTEURS

Vincent Menger, postdoctoraal onderzoeker, afd. Psychiatrie, UMC Utrecht.

Marco Spruit, hoogleraar Advanced data science in population health, afd. Public Health & Eerstelijns Geneeskunde, LUMC, Leiden.

Floortje Scheepers, hoogleraar Innovatie in de GGZ, psychiater, medisch afdelingshoofd, afd. Psychiatrie, UMC Utrecht.

Correspondentieadres

Prof. dr. F.E. Scheepers, afd. Psychiatrie, divisie Hersenen, UMC Utrecht, Postbus 85.500, 3508 GA Utrecht.
E-mail: f.e.scheepers-2@umcutrecht.nl

Geen strijdige belangen meegedeeld.

Het artikel werd voor publicatie geaccepteerd op 3-11-2020.

Citeren

Tijdschr Psychiatr. 2021;63(4):294-300

in de psychiatrie echter slechts sporadisch en op kleine schaal.

Toepassing in psychiatrie

In het bijzonder in de psychiatrie liggen er diverse kansen voor dit soort analyses (Torous & Baker 2015; Scheepers e.a. 2018). De huidige evidence-based richtlijnen op groepsniveau kunnen verder gepersonaliseerd worden met flexibele *beslissingsondersteunende systemen*, die meer rekening houden met kenmerken van de individuele patiënt (Fernandes e.a. 2017).

Ook voor het meetbaar maken en begrijpen van gedrag, symptomen en uitkomst kunnen we nieuwe analyse-technieken gebruiken, bijvoorbeeld door relevante concepten automatisch en nauwkeurig te extraheren uit de vrije tekst uit het dossier (Insel 2017). Dit gebeurt tot dusver vrijwel uitsluitend handmatig en daarmee op kleine schaal.

Tot slot kunnen we zorgdata inzetten om van de heterogene diagnostische categorieën uit de DSM te bewegen naar *risicoprofielen* van kwetsbaarheid en weerbaarheid, die meer recht doen aan de context en variabiliteit van patiënten (Dean 2017).

Het gebruik van grote zorgdatasets uit het EPD maakt daarnaast onderzoek mogelijk dat met een traditionele studieopzet niet altijd haalbaar is. Ten eerste kan men met een grotere *steekproefgrootte* werken, omdat in potentie alle behandelde patiënten geïnccludeerd kunnen worden. Ten tweede kan men *selectiebias* tegengaan, doordat in EPD's ook informatie wordt vastgelegd van patiënten die vaak uitgesloten zijn van deelname aan trials en cohorten (Pannucci & Wilkins 2010). Tot slot wordt de drempel tot *exploratief* onderzoek verlaagd door de hoeveelheid beschikbare data, waarmee hypothesen en nieuwe ideeën kunnen worden gevonden. Verdere validatie kan vervolgens plaatsvinden in andere studies (Angus 2015).

Natuurlijk is het analyseren van zorgdata uit EPD's niet zonder verdere uitdagingen en barrières. De data zelf hebben meerdere modaliteiten, zijn onregelmatig, bevatten missende waarden, en kunnen diverse vormen van *bias* vertonen. Op technisch gebied moet men de juiste manier om zulke data te verwerken en te analyseren vooraf uitdenken.

Organisatorisch gezien is het belangrijk aandacht te besteden aan de samenwerking tussen behandelaars en data-experts, die elkaar nodig hebben om tot een succesvol resultaat te komen. Ook op ethisch vlak kan het analyseren van EPD-data diverse vragen oproepen op het gebied van privacy van patiënten, wetgeving, en potentiële inzet in de spreekkamer. Door het ontwikkelen en toepassen van de juiste methoden, tools en algoritmen, en daarbij verbinding te zoeken tussen het medische vakgebied enerzijds, en informatica, kunstmatige intelligentie en *datascience* anderzijds, is het echter goed mogelijk om oplossingen te vinden voor dit soort uitdagingen.

Op de afdeling Psychiatrie van het Universitair Medisch Centrum Utrecht begonnen we in 2015 met het analyseren van geanonimiseerde datasets uit het EPD, met zorgverbetering als uiteindelijk doel. In dit proefschriftartikel gaan we in op enkele bevindingen en resultaten die we daarbij hebben behaald. Centraal staat de onderzoeksvraag: 'Hoe kunnen relevante inzichten voor de psychiatrische zorg worden verkregen uit EPD's?'. We beschrijven eerst enkele concrete oplossingen voor vraagstukken rondom het werken met datasets uit EPD's, zoals in de voorgaande alinea genoemd. Vervolgens analyseren we de hiermee verkregen data. Daarbij kijken we specifiek naar taxatie van agressierisico, en passen daarvoor geavanceerde technieken toe op geanonimiseerde klinische teksten uit het EPD.

Achtergrond

In dit artikel benadrukken we steeds het belang van samenwerking tussen behandelaars, die veel kennis hebben over de zorgpraktijk, en data-experts, die de technische aspecten van analyse beheersen. We gaan daarom kort in op *machine learning* en *natural language processing*, twee veelgehoorde termen bij geavanceerde analyses. Dit om de leesbaarheid van dit artikel te bevorderen, en een kleine bijdrage te leveren aan het onderlinge begrip.

Machine learning

Machine learning houdt een aantal statistische technieken in die een computer in staat stellen om taken uit te voeren door te leren van (patiënten)data. Een belangrijk onderscheid is dat tussen *supervised* en *unsupervised learning*. In het eerste geval is de uitkomst voor een specifieke patiënt bekend, bijvoorbeeld een gestelde diagnose, het aantal agressie-incidenten dat plaatsvond tijdens opname, of een langetermijnuitkomst. Het doel is om de relatie tussen bekende (patiënt)kenmerken en deze uitkomst te leren, en daarmee de uitkomst voor nieuwe gevallen te weten te komen. In het geval van *unsupervised learning* zet men de computer in om te zoeken naar onbekende en/of onverwachte patronen en verbanden tussen variabelen uit het dossier, bijvoorbeeld medicatiegebruik, woorden in de vrije tekst, diagnose, en/of symptomen.

Natural language processing

Vrije tekst uit het EPD is niet zo eenvoudig te analyseren als gestructureerde data, maar door *natural language processing* toe te passen kunnen we toch gebruikmaken van de genuanceerde en flexibele inhoud (Abbe e.a. 2016). Zo kan men geautomatiseerd op zoek gaan naar thema's in opnamegesprekken, bijwerkingen van medicatie extraheren uit de verpleegkundige rapportage, of de voorgeschiedenis uit het behandelplan onder de loep nemen.

Extra uitdagend in het klinische domein is het omgaan met de vele afkortingen en specialismespecifieke termen. Ook worden de grammaticale regels van de taal onder tijdsdruk nog wel eens uit het oog verloren. Steeds beter kan men wel de context en samenhang in een tekst meenemen in analyses, zoals bij negaties (bijv. 'depressieve kenmerken afwezig') of historische en familiefactoren (bijv. 'depressie in familiegeschiedenis').

METHODEN

Bij het analyseren van EPD's zijn drie aspecten van belang:

- interactieve data-analyse om samenwerking tussen data-experts en behandelaars te verbeteren;
- het ontwerpen van een passende data-infrastructuur die analyse mogelijk maakt;
- het de-identificeren van privacygevoelige gegevens.

We gaan in op deze drie aspecten en beschrijven vervolgens een toegepaste analyse over het inschatten van agressierisico op basis van klinische tekst, en de potentie daarvan voor de psychiatrische zorg. Met oog op de overzichtelijkheid zullen we de methoden en (tussen) resultaten van deze deelprojecten gekoppeld aan elkaar bespreken.

Interactieve data-analyse

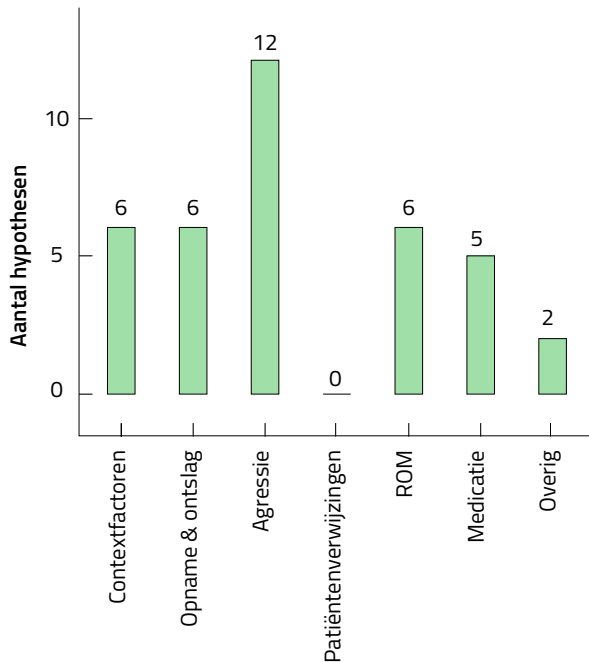
Om te onderzoeken hoe data-experts en behandelaars kunnen samenwerken, bekeken we of zij door datavisualisatie op zoek kunnen gaan naar nieuwe patronen, verbanden en hypothesen in zorgdata (Menger e.a. 2016). Bij het analyseren van EPD's is de samenwerking tussen behandelaars en data-experts essentieel (Brennan & Bakken 2015). Enerzijds kunnen data-experts zonder inhoudelijke zorgkennis geen accurate analyses uitvoeren, omdat zij doorgaans geen kennis hebben van de zorgpraktijk. Anderzijds zullen behandelaars voldoende begrip moeten hebben van de toegepaste analyse om deze uiteindelijk op correcte wijze in te kunnen zetten in de praktijk.

We organiseerden interactieve sessies waarin behandelaars met interactieve visualisatiesoftware op zoek gingen naar verbanden in data. Dit deden we aan de hand van een aantal thema's die behandelaars aangedragen hadden (zie **figuur 1**), zoals opname, contextfactoren en medicatie. Tijdens de sessies bedienden de data-experts de software, en hielpen zij de uitkomsten te duiden. Om de sessies te faciliteren verzorgden we enkele aanpassingen aan het Cross Industry Standard Process for Data Mining (CRISP-DM), om uit te komen op de CRISP-IDM, waar de I staat voor interactief (Chapman e.a. 2000). Een belangrijke wijziging is weergegeven in **figuur 2**: het modelleren en het evalueren aan de hand van de visualisatiesoftware gebeuren tegelijkertijd, zodat behandelaars en data-experts onmiddellijk gebruik kunnen maken van elkaars kennis.

In totaal lieten we in 19 sessies telkens 1-3 behandelaars het zoekproces leiden door datavisualisatie. We brachten diverse gegevens uit patiëntendossiers met elkaar in verband, zoals diagnose, medicatievoorschriften, verpleegkundige rapportage, scores op diverse vragenlijsten, en het behandelplan. Tijdens de sessies vonden we in totaal 29 interessante resultaten in de visualisaties die relevant werden geacht door behandelaars.

Zo vonden we een mogelijk verband tussen agressie-incidenten en het afnemen van de werking van medicatie bij patiënten met ADHD, en tussen de lengte van opname en socio-economische status in de woonomgeving. In beide gevallen waren het de behandelaars die de onmisbare kennis van de praktijk inbrachten – zo misten data-experts bijvoorbeeld kennis over het afnemen van de werking van ADHD-medicatie. Tegelijkertijd brachten de data-experts de onmisbare technische kennis in – bijvoorbeeld door de sterkte van de verbanden statistisch te interpreteren.

Figuur 1. Aantal geformuleerde hypothesen voor verder onderzoek per thema. In totaal werden 29 hypothesen gegenereerd, daarbij kan een hypothese tot meerdere thema's behoren



Van de 29 hypothesen voor verder onderzoek die naar voren kwamen, waren er slechts 5 vooraf in de interviews direct of indirect benoemd. Hiermee laten we zien dat het daadwerkelijk mogelijk is om in EPD's op exploratieve wijze op zoek te gaan naar tot dusver niet bedachte kennis en hypothesen. Bij deze aanpak moeten we wel opmerken dat verdere validatie in externe datasets noodzakelijk is om de gegenereerde hypothesen te toetsen. De kennis die we opdeden, was vaak echter essentieel voor het doen van verdere analyses.

Vorbewerking van data

Het hergebruiken van zorgdata voor analyse stelt ons op het terrein van datamanagement voor grote uitdagingen (Hersh e.a. 2014; Meystre e.a. 2017). Zorgdata worden bijvoorbeeld vaak in diverse formats en uit meerdere interne bronnen aangeleverd. Waar in traditionele onderzoekdatasets alle metingen zijn gedocumenteerd in een protocol, zijn zorgdata amper gedocumenteerd, en zijn registratieprotocollen in de praktijk veranderlijk. Daarnaast worden dagelijks nieuwe patiënten behandeld, waardoor de dataset constant aangevuld wordt. Dit vraagt om een passende data-infrastructuur, waarbij men de juiste hardware en software gebruikt om de data-analyse van begin tot eind in goede banen te leiden (Safran 2014).

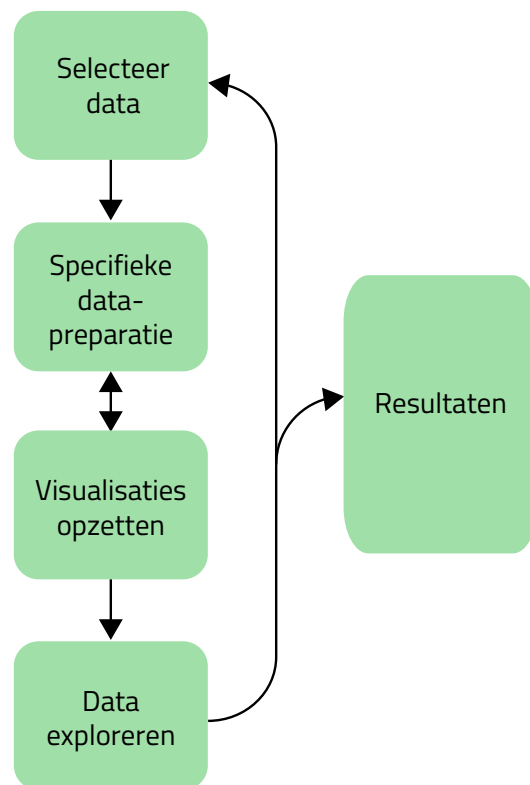
Samen met relevante experts formuleerden we negen vereisten waaraan een infrastructuur voor het analy-

seren van zorgdata moet voldoen (Menger e.a. 2019a). Naast een aantal technische benodigheden, zoals het integreren, voorbereiden en opslaan van data, gaat het voornamelijk om organisatorische voorwaarden. Denk bijvoorbeeld aan ondersteuning voor samenwerking en documentatie, zodat onderzoekers in staat worden gesteld hun kennis over data, processen en algoritmen uitgebreid te documenteren, en daarmee te profiteren van elkaars inspanningen. Ook het mogelijk maken van herhaalbare analyses, en het waarborgen van privacy en veiligheid van data behoren tot belangrijke vereisten. De negen vereisten organiseerden we vervolgens in het Capable Reuse of Electronic Health Record Data(CARED)-framework, dat houvast kan bieden voor andere zorginstellingen die graag hun EPD's willen analyseren. Dat de implementatie van een technische infrastructuur voor analyse van zorgdata haalbaar is, toonden we aan op onze afdeling psychiatrie. Onderzoekers met diverse achtergronden maken dagelijks gebruik van deze infrastructuur, die analyse niet alleen mogelijk maakt, maar voornamelijk de kwaliteit en betrouwbaarheid ervan sterk verbetert.

De-identificatie van tekst

Om de privacy van patiënten te waarborgen is het noodzakelijk de gebruikte medische gegevens te de-identifi-

Figuur 2. Het modelleren en evalueren in het CRISP-IDM-model, waarin behandelaars en data-experts iteratief op interactieve wijze samenwerken bij het exploreren van data



ceren. Identificerende variabelen zoals woonplaats of BSN-nummer hebben vaak geen toegevoegde waarde, en worden daarom simpelweg weggelaten (El Emam e.a. 2006). In het geval van vrije tekst is dit een ander verhaal, omdat er regelmatig identificerende gegevens zoals namen, adresgegevens, telefoonnummers en datums in vóórkomen.

Het handmatig de-identificeren van teksten is foutgevoelig, en bovendien onhaalbaar op grote schaal (Deleger e.a. 2013). Daarom ontwikkelden we een methode voor het automatisch de-identificeren van medische tekst: DEDUCE (Menger e.a. 2017). In diverse talen, zoals het Engels, Zweeds en Koreaans, was deze toepassing al beschikbaar, maar in het Nederlands nog niet.

In de Nederlandse of Europese wetgeving is niet vastgelegd welke specifieke kenmerken als identificerend moeten worden beschouwd. Op basis van Amerikaanse wetgeving en beschikbare literatuur kwamen wij tot de volgende categorieën:

- namen en initialen van personen, inclusief patiënten, familie en behandelaars;
- locaties, kleiner dan een land;
- namen van zorginstellingen;
- datums;
- leeftijden;
- patiëntnummers;
- telefoonnummers;
- URL's en e-mailadressen.

We ontwierpen een op regels gebaseerde methode, in tegenstelling tot een statistische benadering. Het grote voordeel hiervan is dat we onze tool openbaar kunnen publiceren. Daarbij annoteren we onder andere op basis van lijsten van voor- en achternamen, instellingen en plaatsnamen. Deze informatie wordt gecombineerd met contextinformatie uit de zin – als een woord bijvoorbeeld voorafgegaan wordt door 'dr.' of 'mw.' verhoogt dat de kans dat er sprake is van een persoonsnaam, terwijl een achternaam gevolgd door een koppelteken (-) op een dubbele achternaam kan wijzen.

Op basis van een evaluatie op 200 verpleegkundige rapportages en 200 behandelplannen, toonden we aan hoe nauwkeurig onze toepassing is. Zo werden er geen namen van patiënten gemist, werd 96% van overige namen gevonden, en werd 92% van identificerende gegevens correct verwijderd. De implementatie van DEDUCE in de programmeertaal Python is vrij beschikbaar (via <https://github.com/vmenger/deduce/>), zodat ook andere onderzoeks- of zorginstellingen het risico op identificatie in tekstuele datasets sterk kunnen beperken op een schaalbare en efficiënte manier.

RESULTATEN

Huidige risicotaxatie agressie

We kozen ervoor om met een toegepaste analyse dieper in te gaan op het inschatten van het risico op agressie, een belangrijk onderdeel van agressie-management op een afdeling. Dit gebeurt doorgaans aan de hand van

het ongestructureerde klinische oordeel van de behandelaar, of met een gestructureerd taxatie-instrument (bijv. de *Historical Clinical Risk Management-20*). Taxatie-instrumenten leiden tot een objectiever oordeel over risico en onderzoek wijst uit dat ze vaak accurater zijn dan het ongestructureerde klinische oordeel (Singh e.a. 2014). De kwaliteit van taxatie op basis van een standaardinstrument, vaak ontwikkeld en gevalideerd met een bepaalde patiëntengroep, verschilt echter behoorlijk voor verschillende populaties, en laat ook nog ruimte voor verbetering qua nauwkeurigheid (Singh e.a. 2011).

Voorspellen met data-analyse

We onderzochten of we het risico op agressie tijdens opname konden inschatten aan de hand van de al beschikbare verpleegkundige rapportage en psychiatrische decursus (Menger e.a. 2019b). We gebruikten hiervoor alle teksten geschreven in de eerste 24 uur van opname, en bepaalden het risico op agressie binnen de eerste 4 weken van opname aan de hand van verplicht gemelde agressie-incidenten. Agressie definieerden we als alle vormen van dreigend of agressief gedrag, zowel verbaal als fysiek, richting een ander persoon. Zelfbeschadiging of ongepast gedrag (inclusief midde-lenmisbruik) vielen hier niet onder.

Als cohort namen we hiervoor alle opnames (n = 3189) van patiënten die op een van de afdelingen voor volwassenpsychiatrie van het UMCU behandeld werden tussen 2013 en 2018. In eerste instantie onderzochten we of klinische tekst enige voorspellende waarde heeft, door de associatie tussen individuele woorden en de uitkomst van agressie tijdens opname te exploreren. In **tabel 1** is te zien dat diverse woorden significant geassocieerd kunnen worden met de uitkomst, zoals 'agressief', 'boos', 'ibs', en 'noodmedicatie'.

Omdat de correlaties tussen individuele woorden en de uitkomst echter zwak waren, kozen we ervoor om alle tekst tot de eerste 24 uur van opname integraal te

Tabel 1. Exploratieve analyse van associatie tussen individuele termen in klinische tekst, en agressieve uitkomst, met Matthews' correlatiecoëfficiënt (MCC) en de volgens Holm-Bonferroni gecorrigeerde p-waarde

Rang	Term	MCC	Pcorr (χ^2)
1	agressief	0,17	< 0,001**
2	reageert	0,15	< 0,001**
3	aangeboden	0,14	< 0,001**
4	boos	0,16	< 0,001**
5	deur	0,14	< 0,001**
6	loopt	0,15	< 0,001**
7	ibs	0,14	< 0,001**
8	aanbieden	0,12	< 0,001**
9	noodmedicatie	0,14	< 0,001**
10	liep	0,12	< 0,001**

Tabel 2. Voorspellende waarde van getrainde modellen op de twee locaties, zowel intern gevalideerd met cross validation (CV) als gevalideerd in de andere locatie

Evaluatie	Intern (CV), locatie 1	Intern (CV), locatie 2	Extern model, locatie 1	Extern model, locatie 2
Model geëvalueerd op locatie	1	2	1	2
Model getraind op locatie	1	2	2	1
AUC (95%-BI)	0,797 (0,771-0,822)	0,764 (0,732-0,797)	0,722 (0,690-0,753)	0,643 (0,610-0,675)
N	3189	3253	3189	3253
Specificiteit (95%-BI)	0,935 (0,930-0,940)	0,947 (0,943-0,951)	0,925 (0,921-0,930)	0,929 (0,926-0,933)
Sensitiviteit (95%-BI)	0,334 (0,287-0,383)	0,336 (0,285-0,389)	0,248 (0,205-0,296)	0,134 (0,097-0,179)
Relatief risico (95%-BI)	5,121 (4,109-6,330)	6,297 (4,956-7,922)	3,314 (2,581-4,214)	1,885 (1,305-2,673)

gebruiken. Hiervoor gebruikten we het *paragraph2vec*-algoritme, dat op basis van een groot klinisch corpus een interpretatie voor lange teksten leert, in combinatie met een *support vector machine*, een machine-learningalgoritme dat bij uitstek geschikt is voor classificatie. We valideerden deze modellen bovendien op een tweede, vergelijkbaar maar onafhankelijk cohort van patiënten die onder behandeling stonden bij Antes GGZ in de regio Rotterdam.

In **tabel 2** is te zien dat in beide cohorten de getrainde modellen op basis van interne *cross validation* een zeer behoorlijke nauwkeurigheid bereiken, met een respectievelijke *area under the curve* (AUC) van 0,80 en 0,76. Deze resultaten kunnen beslist als bruikbaar voor in de praktijk worden gezien. Zo vonden Teo e.a. (2013) een AUC van 0,70 voor het ongestructureerde klinische oordeel en Fazel e.a. (2012) vonden in een meta-analyse een gemiddelde AUC van 0,72 voor gestructureerde risicotaxatie-instrumenten. Ook de in Nederland veelgebruikte *Short-Term Assessment of Risk and Treatability* (START) toont een vergelijkbare validiteit (Braitwaite 2011). Het uitwisselen van modellen onder de twee cohorten liet echter significant minder nauwkeurige modellen zien (AUC: 0,72 en 0,64 respectievelijk). Dit geeft aan dat ook machine-learningmodellen niet per definitie robuust zijn tegen variatie in populatie en taalgebruik. De kracht van de machine-learningbenadering schuilt echter juist in de mogelijkheid om de risico-inschatting op maat te maken, en daarmee tegelijkertijd de nauwkeurigheid te kwantificeren.

Inzetbaarheid

Een voorspelmodel voor agressie heeft potentie voor inzetbaarheid in de klinische praktijk, op vergelijkbare manier als de nu bestaande taxatie-instrumenten worden gebruikt. De automatisch berekende uitkomst van het model zou in het bijzonder kunnen dienen als input voor gerichte preventie in een agressiepreventie- en/of -werkplan. Door het risico in te schatten op basis van toch al geregistreerde vrije tekst in het EPD kunnen de handmatig ingevulde taxatie-instrumenten in theorie daarnaast achterwege blijven, en beperken we dus de

administratieve last van professionals zo veel mogelijk. In de toekomst onderzoeken we of we voorspellingen nog accutater kunnen maken door het toevoegen van gestructureerde data, zoals medicatiegebruik en juridische status.

Uitdagingen

Ondanks het veelbelovende resultaat heeft ook het in de praktijk brengen van een voorspelmodel diverse uitdagende aspecten. Niet alle machine-learningmodellen kunnen de inschatting die ze geven, ook onderbouwen, zodat de uitkomst in de spreekkamer inhoudelijk besproken kan worden. Daarbij gaat transparantie doorgaans ten koste van nauwkeurigheid. Ook zijn er duidelijkere ethische kaders nodig, bijvoorbeeld in het kader van autonomie van de behandelaar, en mogelijke bias tegen bepaalde (kwetsbare) groepen. Toch is het goed mogelijk dat in de toekomst algoritmen zullen bijdragen aan agressiepreventie, en daarmee aan een veiliger praktijk voor patiënt en medewerker.

DISCUSSIE

In de nabije toekomst kan datagedreven onderzoek nieuwe doorbraken in de psychiatrie mogelijk maken. In dit artikel beschreven we eerst hoe het doen van kwalitatief goede analyses op zorgdata mogelijk wordt. We beschreven hoe behandelaars en data-analisten data exploreerden met datavisualisaties, om daarmee de samenwerking te versterken. Daarna omschreven we waar een data-infrastructuur aan moet voldoen die de kwaliteit van databewerking en -analyse kan vergroten. Tot slot beschreven we hoe zorgdata automatisch gede-identificeerd kunnen worden, om daarmee privacy van patiënten te waarborgen. We lieten ook zien wat de toegevoegde waarde is van klinische teksten bij het inschatten van agressierisico, een resultaat dat potentie heeft voor implementatie in de psychiatrische praktijk. Om in de toekomst op zoek te gaan naar nieuwe kennis met nog grotere zekerheid is vooral samenwerking tussen instellingen noodzakelijk. Hiermee wordt de invloed van ruis en bias in lokale datasets verder verkleind, en

de externe validiteit van analyses vergroot. Het delen van zorgdata is vaak niet mogelijk in het kader van de privacy van patiënten en wetgeving op dit terrein. Momenteel onderzoeken we hoe samenwerking door uitwisseling van kennis en algoritmen en het repliceren van elkaars bevindingen mogelijk is in ons Compute Visits Data(CoViDa)-model. Ook de mogelijkheid van *federated learning*, waarbij kennis van meerdere datasets kan worden afgeleid zonder deze samen te voegen, is een veelbelovende optie. Daarnaast ontwikkelen we specifieke tools om klinische tekst in de psychiatrie nog beter te begrijpen en te analyseren, en onderzoeken we wat de beste manier is om opgedane kennis daadwerkelijk in de praktijk in te zetten.

CONCLUSIE

Er zijn nog genoeg uitdagingen te overwinnen om de volledige potentie van analyses op zorgdata waar te maken. Daarbij zullen informatici en klinici zich moeten blijven inspannen om methoden, technieken en tools te verbeteren. Onze bevindingen tonen aan dat het op lokaal niveau al haalbaar is om zorgdata uit EPD's te analyseren met toegevoegde waarde voor de praktijk. Door nu de schaal en kwaliteit van analyses verder te vergroten zullen we zorgdata nog beter kunnen inzetten om de globale kennis van psychiatrische aandoeningen verder uit te breiden, en daarmee kwaliteit van zorg te verbeteren.

NOOT

1 Dit artikel is gebaseerd op het proefschrift van de eerste auteur, 'Knowledge discovery in clinical psychiatry: Learning from electronic health records' (Universiteit Utrecht, <http://dspace.library.uu.nl/handle/1874/385129>).

LITERATUUR

Abbe A, Grouin C, Zweigenbaum P, Falissard B. Text mining applications in psychiatry: A systematic literature review. *Int J Methods Psychiatr Res* 2016; 25: 86-100.

Angus DC. Fusing randomized trials with big data: the key to self-learning health care systems? *JAMA* 2015; 314: 767-8.

Appleby J. Rises in healthcare spending: Where will it end? *BMJ* 2012; 345: e7127.

Braithwaite E, Charette Y. The predictive validity of clinical ratings of the Short-Term Assessment of Risk and Treatability (START). *Int J Forensic Mental Health* 2011; 9:4

Brennan PF, Bakken S. Nursing needs big data and big data needs nursing. *J Nurs Scholarsh* 2015; 47: 477-84.

Chapman P, Clinton J, Kerber R, Khabaza T, Reinartz T, Shearer C, e.a. CRISP-DM 1.0 Step-by-step data mining guide. 2000.

Dean CE. Social inequality, scientific inequality, and the future of mental illness. *Philos Ethics Humanit Med* 2017; 12: 10.

Deleger L, Molnar K, Savova G, Xia F, Lingren T, Li Q, e.a. Large-scale evaluation of automated clinical note de-identification and its impact on information extraction. *J Am Med Inform Assoc* 2013; 20: 84-94.

El Emam K, Jabbouri S, Sams S, Drouet Y, Power M. Evaluating common de-identification heuristics for personal health information. *J Med Internet Res* 2006; 8: e28.

Fazel S, Singh JP, Doll H, Grann M. Use of risk assessment instruments to predict violence and antisocial behaviour in 73 samples involving 24 827 people: Systematic review and meta-analysis. *BMJ* 2012; 345: e4692.

Fernandes BS, Williams LM, Steiner J, Leboyer M, Carvalho AF, Berk M. The new field of 'precision psychiatry'. *BMC Med* 2017; 15: 80.

Frawley WJ, Piatetsky-Shapiro G, Smyth P. Knowledge discovery in databases: an overview. *AI Mag* 1992; 13: 57-70.

Hersh WR, Weiner MG, Embi PJ, Logan JR, Payne PR, Bernstam E V, e.a. Caveats for the use of operational electronic health record data in comparative effectiveness research. *Med Care* 2014; 51: S30-S37.

Insel TR. Digital phenotyping: Technology for a new science of behavior. *JAMA* 2017; 218: 1215-6.

Krumholz HM. Big data and new knowledge in medicine: The thinking, training, and tools needed for a learning health system. *Health Aff* 2014; 33: 1163-70.

Menger V, Scheepers F, van Wijk LM, Spruit M. DEDUCE: A pattern matching method for automatic de-identification of Dutch medical text. *Telemat Informatics* 2017; 35: 727-36.

Menger V, Spruit M, de Bruin J, Kelder T, Scheepers F. Supporting reuse of EHR data in healthcare organizations: the CARED research infrastructure framework. In: Proceedings of the 12th conference on Health Informatics. 2019; 41-50.

Menger V, Spruit M, Van Est R, Nap E, Scheepers F. Machine learning approach to inpatient violence risk assessment using routinely collected clinical notes in electronic health records. *JAMA Netw Open* 2019; 2: e196709.

Menger V, Spruit M, Hagoort K, Scheepers F. Transitioning to a data driven mental health practice: collaborative expert sessions for knowledge and hypothesis finding. *Comput Math Methods Med* 2016; 2016.

Meystre SM, Lovis C, Bürkle T, Tognola G, Budrionis A, Lehmann CU. Clinical data reuse or secondary use: current status and potential future progress. *Yearb Med Inform* 2017; 26: 38-52.

Pannucci CJ, Wilkins EG. Identifying and avoiding bias in research. *Plast Reconstr Surg* 2010; 126: 619-25.

Raghupathi W, Raghupathi V. Big data analytics in healthcare: promise and potential. *Heal Inf Sci Syst* 2014; 2: 3.

Safran C. Reuse of clinical data. *IMIA Yearb* 2014; 9: 52-54.

Scheepers FE, Menger V, Hagoort K. Datascience in de psychiatrie. *Tijdschr Psychiatr* 2018; 60: 205-9.

Singh JP, Desmarais SL, Hurducas C, Arbach-Lucioni K, Condemarin C, Dean K, e.a. International perspectives on the practical application of violence risk assessment: a global survey of 44 countries. *Int J Forensic Ment Health* 2014; 13: 193-206.

Singh JP, Grann M, Fazel S. A comparative study of violence risk assessment tools: A systematic review and meta-regression analysis of 68 studies involving 25,980 participants. *Clin Psychol Rev* 2011; 31: 499-513.

Teo AR, Holley SR, Leary M, McNiel DE. The relationship between level of training and accuracy of violence risk assessment. *Psychiatr Serv* 2012; 63: 1089-94.

Torous J, Baker JT. Why psychiatry needs data science and data science needs psychiatry: connecting with technology. *JAMA Psychiatry* 2015; 73: 1-2.

Zhu L, Zheng WJ. Informatics, data science, and artificial intelligence. *JAMA* 2018; 320: 1103-4.

SUMMARY

Knowledge discovery in clinical psychiatry: learning from electronic health records

V.J. Menger, M.R. Spruit, F.E. Scheepers

Background Today, almost every psychiatric care institution registers information concerning the care they provide in an electronic health record (EHR). By analyzing these health care data with innovative and advanced techniques, they can be an important source of new knowledge in the near future, and thereby contribute to improving psychiatric care.

Aim To investigate how data from EHRs can provide relevant knowledge and insights for psychiatric care.

Method We designed and discussed solutions for some technical, organizational and ethical barriers surrounding unlocking health care data, in order to make analysis possible. We then analyzed the obtained health care data using techniques from knowledge discovery, the process in which new and useful information is extracted from data. We used techniques from data visualization, machine learning and natural language processing, among others, to demonstrate which types of results can be achieved.

Results Our approach showed that it is possible to find new and interesting insights that are hidden in EHRs on an aggregated level, in collaboration with healthcare professionals and patients. In particular we showed how the risk of violent behavior can effectively and accurately be assessed based on clinical text in the EHR.

Conclusion After addressing some of the important challenges surrounding analyzing EHR data, learning from data from EHRs is a new and interesting approach with clear potential for improving psychiatric care.