

Ethiek van (big)datatoepassingen in de psychiatrische praktijk

F.E. Scheepers, M. Mostert, M.M. Milota, G.J.M. van Thiel

Achtergrond	Door snelle digitalisering zijn er steeds meer data beschikbaar en ook in de zorg doen big data en artificiële intelligentie (AI) hun intrede.
Doel	Inzicht geven in de verschillende ethische dilemma's waar rekening mee gehouden moet worden bij (big-) datatoepassingen in de psychiatrische praktijk.
Methode	Beschrijven en analyseren van de ethische aspecten die samenhangen met gebruik van klinische data in de context van psychiatrische zorg.
Resultaten	Verschiedende ethische aspecten spelen een rol in vier fasen; dataverzameling, analyse, disseminatie en toepassing van resultaten. Om op een verantwoorde manier het gebruik van klinische data in te zetten moet rekening gehouden worden met deze aspecten.
Conclusie	Het gebruik van big data en AI in de zorg zou als doel moeten hebben het leren en verbeteren van de zorg samen met patiënten en professionals te stimuleren. Big data en AI moeten niet gezien worden als de heilige graal maar als ondersteunend aan de zorg waarin veel aspecten een rol spelen die niet in meetbare data zijn om te zetten.

De manier waarop *high-performance*computers (HPC's) data kunnen verwerken wordt steeds geavanceerder. Artificiële intelligentie (AI) verwijst naar computers die data kunnen 'begrijpen' door verbanden te leggen en te 'leren' van data waardoor uitkomsten toegepast en steeds weer aangepast kunnen worden. Bij AI wordt vaak gebruikgemaakt van complexe analysemethoden (zoals *deep learning* en *mining*) waardoor soms nauwelijks nog te begrijpen is hoe een uitkomst tot stand is gekomen. Zelflerende algoritmes zijn de rekenformules waar AI door tot stand komt.

Ook in de zorg wordt steeds meer gebruikgemaakt van klinische data en AI. Thuismonitoring, digitale dossiers en de inzet van robotica bieden nieuwe mogelijkheden op het gebied van diagnostiek en beslisondersteuning gebaseerd op data uit de klinische praktijk, waardoor snelle leer- en verbetercycli gerealiseerd kunnen worden.

Tegenover deze voordelen zijn er ook ethische dilemma's bij deze ontwikkelingen. Om technologie te selecteren die bijdraagt aan de juiste zorgdoelen is aandacht nodig voor de waarden die door technologie bevordert of juist bedreigd kunnen worden. In de geestelijke gezondheidszorg (ggz) is dit bij uitstek van belang. Niet alleen zijn psychische klachten vaak nauw verweven met een gevoel van identiteit, persoonlijke levenssfeer

en sociale relaties. Ook kan iemand met psychische klachten soms minder goed afwegen wat de impact is van bepaalde keuzes die gemaakt moeten worden bij diagnostiek en behandeling.

In dit artikel focussen we op de belangrijkste ethische aspecten die zich voordoen in de vier fasen van dataverwerking bij de toepassing van (big)data-analyses in de klinische ggz-praktijk.

Big data in de psychiatrische praktijk

In de psychiatrie zijn steeds meer voorbeelden te vinden van het gebruik van klinische data in combinatie met AI om de zorg te verbeteren (Scheepers e.a. 2018; Russ e.a. 2019). Dit varieert van clustering van patiënten op basis van een bepaald profiel tot predictiemodellen om individuele behandel-effecten te voorspellen. De context waarin psychische ontregeling plaatsvindt, is heel bepalend. Door data uit de dagelijkse praktijk te gebruiken kan men meer inzicht krijgen in het dynamische, interactieve en adaptieve proces bij ontregeling. Dit draagt bij aan goede toekomstige zorg voor psychisch kwetsbare mensen.

Er is een aantal ethische aspecten van belang bij het geaggregeerd gebruik van data uit de praktijk om zorg te verbeteren. Deze aspecten spelen in vier fasen van het proces van dataverwerking:

AUTEURS

Floortje Scheepers, (kinder- en jeugd)psychiater, medisch afdelingshoofd psychiatrie UMC Utrecht, hoogleraar Innovatie in de psychiatrie.

Menno Mostert, universitair docent Gezondheidsrecht, Julius Centrum, UMC Utrecht.

Megan Milota, universitair docent Medical humanities, Julius Centrum, UMC Utrecht.

Ghislaïne van Thiel, universitair hoofddocent Medische ethiek, Julius Centrum voor gezondheidswetenschappen en eerstelijns geneeskunde, UMC Utrecht.

Correspondentie

Prof. dr. F.E. Scheepers (f.e.scheepers-2@umcutrecht.nl).

Geen strijdige belangen meegedeeld.

Het artikel werd voor publicatie geaccepteerd op 4-5-2021.

Citeren

Tijdschr Psychiatr. 2021;63(10):699-702

- in de fase van dataverzameling zijn geïnformeerde toestemming (informed consent), eigenaarschap en verantwoordelijkheid belangrijke aandachtspunten;
- in de fase van analyseren zijn dat onder andere validiteit of betrouwbaarheid, het vermijden van bias en non-discriminatie;
- in de fase van dissemineren gaat het om belangen, herleidbaarheid en repliceerbaarheid;
- in de fase van toepassing ten slotte komen transparantie of uitlegbaarheid en de impact op de relatie tussen zorgverlener en patiënt aan de orde (Mittelstadt & Floridi 2016; Morley e.a. 2020).

Fase van dataverzameling

Als een patiënt in het ziekenhuis komt, wordt er gehandeld volgens geldende wet- en regelgeving, in het bijzonder betreffende patiëntenrechten, dossiervoering en privacy. Er is toestemming van de patiënt nodig om over te gaan tot behandelen. Onderzoekuitslagen, overwegingen en conclusies worden opgeslagen in het medisch dossier. Zolang deze ingezet worden voor de individuele zorg van de patiënt en zolang professionals zich houden aan het medisch beroepsgeheim en de Algemene Verordening Gegevensbescherming (AVG) is het gebruik van deze zorgdata toegestaan. In juridische zin zijn data in beginsel niet onderworpen aan eigendomsrechten. Nationale wetgeving en de AVG kennen wel zeggenschapsrechten toe aan de patiënt, zoals het opvragen of verzoeken tot vernietiging van gegevens. Diezelfde wetten leggen plichten op aan de hulpverlener en de zorginstelling, zoals het bijhouden en goed beveiligen van het dossier.

Maar de zorginstelling kan ook gebruikmaken van data op geaggregeerd niveau om zorg te verbeteren of wettelijk verplichte verantwoording af te leggen. Data mogen zelfs, onder bepaalde voorwaarden (zoals de privacy die geborgd is of dat onderzoek in het algemeen belang is),

zonder toestemming van de patiënt gebruikt worden voor medisch-wetenschappelijk onderzoek. Dit soort mogelijkheden tot secundair gebruik van klinische data maken het 'eigenaarschap' soms wat diffuus.

In de psychiatrie

In de psychiatrie wordt relatief veel informatie verzameld in de vorm van geschreven tekst. Tekst is ingewikkelder om te ontdoen van (in)direct tot een persoon herleidbare gegevens. Menger e.a. (2018) ontwikkelden een methode om Nederlandse tekst uit het medisch dossier te pseudonimiseren (data zijn dan ontdaan van direct herleidbare gegevens, maar kunnen in combinatie met andere databronnen mogelijk toch herleidbaar zijn). Hierdoor kunnen ook teksten op geaggregeerd niveau omgezet worden in bruikbare data voor onderzoek of zorgverbetering.

In de psychiatrie kan een complicerende factor zijn dat sommige patiënten onvrijwillig worden behandeld (vanuit de Wet verplichte ggz (Nederland) of de Wet bescherming van de persoon van de geesteszieke (België)). Bij onvrijwillige zorg hebben mensen geen toestemming gegeven voor behandeling en zijn ze extra kwetsbaar. Er ontstaat een ethisch spanningsveld als data afkomstig uit de zorg voor deze mensen worden gebruikt: ze dragen bij aan zorgverbetering in het belang van de patiënt, maar aan essentiële waarden van vrijwilligheid en toestemming kan niet worden voldaan. Al is toestemming niet altijd nodig op grond van de wet, dan nog moet men juist bij deze groep patiënten zorgvuldig afwegen of het secundair gebruik van klinische data verantwoord is. In die afweging moet men ook de nadelen van het niet gebruiken van data meenemen. Juist deze doelgroep wordt vrijwel nooit geïncludeerd in wetenschappelijk onderzoek naar de effectiviteit van interventies. Richtlijnen en protocollen die voor een groot deel gebaseerd zijn op wetenschappelijk

onderzoek sluiten in de praktijk daarom bij deze complexere doelgroep het minst goed aan. Als klinische data gebruikt kunnen worden om te onderzoeken wat bij deze groep patiënten wel en niet werkt in de echte praktijk zouden we wellicht veel beter zorg op maat kunnen bieden.

Ingewikkeld wordt het als data niet in het ziekenhuis verzameld zijn, maar bijvoorbeeld met een mobiele telefoon (zoals apps die je gewicht en je bloeddruk bijhouden of hoeveel je beweegt op een dag). Niet alleen is het bij veel applicaties onduidelijk of data door de ontwikkelaar opgeslagen (en gebruikt) worden, ook het anonimiseren van data is meestal niet realistisch. Als veel unieke data bij elkaar komen, bestaat het risico dat deze data (in)direct herleidbaar blijken tot een persoon. In beginsel zal expliciet toestemming gevraagd moeten worden voor het koppelen van persoonsgegevens die niet in het zorgproces verzameld zijn aan klinische data. Voorbeelden van dilemma's in de praktijk hierbij zijn hoe specifiek die toestemming moet zijn en of een eenmalige toestemming voldoende is om langdurig en veelvuldig gebruik te rechtvaardigen.

Fase van analyseren van data

Bij de fase van analyseren speelt de betrouwbaarheid en validiteit een belangrijke rol. Als data voor klinische doeleinden geregistreerd worden (en niet door derden op juistheid gecontroleerd, zoals bij wetenschappelijk onderzoek wel vaak het geval is), zal er vaker een fout in de data sluipen (ruis) die invloed kan hebben op de analyses en dus ook de uitkomsten. 'Garbage in, garbage out' is in dit verband een bekende opmerking. Ook bias of discriminatie kan optreden als algoritmes op een te eenzijdige manier informatie selecteren of analyseren (Vabalas e.a. 2019). Als er aandacht is voor interne en externe kruisvalidatie van algoritmes kunnen deze effecten zoveel mogelijk beperkt worden. Het systeem wordt met zo'n validatie extra getraind om ervoor te zorgen dat het goed presteert op nieuwe data. Omgekeerd zou je klinische data juist als betrouwbaarder (ecologisch meer valide) kunnen beschouwen dan data uit gecontroleerde settings zoals klinische trials. Richtlijnen en protocollen zijn vaak gebaseerd op gerandomiseerde gecontroleerde trials waarin deelnemers aan strenge selectiecriteria moeten voldoen. In de praktijk blijken resultaten van behandelingen vaak tegen te vallen omdat relatief veel patiënten niet aan zulke criteria voldoen (RVS 2018). Klinisch verzamelde data doen in dat opzicht dus meer recht aan de praktijk en zijn daarom misschien geschikter om inzicht te geven in wat bij wie werkt.

In de fase van analyseren is het tevens van belang om te bepalen met welk doel men een analyse uitvoert. Vooral bij hypothesevrij onderzoek en/of *unsupervised learning* (in je analyses neem je geen uitkomstmaat mee, maar laat je de computer zelf tot een classificatie of uitkomst komen) is het goed in de gaten te houden of een algoritme leidt tot zorgverbetering. In de literatuur noemt

men hiervoor als algemene ethische uitgangspunten: AI bevordert rechtvaardigheid en welzijn met specifieke aandacht voor diversiteit en inclusiviteit, staat in dienst van mensen en mag geen schade berokkenen (Florida e.a. 2018). Een algoritme dat zorgdoelstellingen niet op een positieve manier beïnvloedt, maar wel bijvoorbeeld een bepaald medicijn steeds vaker adviseert, zou argwaan moeten oproepen.

In de psychiatrie

Binnen de (ggz-)zorginstelling moet men tot slot een veilige infrastructuur creëren, met beperkte toegang en regelmatige controles voor wat betreft herleidbaarheid van data. Daarnaast moet men nadenken over het beheer van data en AI-initiatieven. Voorbeelden van beheertoepassingen uit de praktijk zijn de inzet van een *trusted third party* bij het koppelen van persoonsgegevens, of het toetsen van toegang tot data door een interdisciplinaire *data access committee* waarin bij voorkeur ook een patiëntvertegenwoordiger zitting heeft.

Fase van dissemineren

In de fase van disseminatie is het van belang transparant te zijn over de totstandkoming en repliceerbaarheid van algoritmes. Een belangrijke voorwaarde is dat data FAIR moeten zijn: *findable*, *accessible*, *interoperable*, en *reusable* (vindbaar, toegankelijk, uitwisselbaar en herbruikbaar) (Boeckhout e.a. 2018). Voor het verbeteren van algoritmes kunnen belangrijke uitgangspunten als maximale privacybescherming een obstakel vormen.

Omdat klinische data zonder toestemming in principe alleen geanonimiseerd en op geaggregeerd niveau gedeeld kunnen worden, is het niet altijd mogelijk algoritmes exact te repliceren in een dataset van een andere organisatie voor externe validatie. Hierbij kan het principe van *federated learning* of decentrale *machine learning* nuttig zijn. Met deze aanpak worden algoritmes gebouwd op een specifieke dataset in een beveiligde omgeving van bijvoorbeeld een ziekenhuis en wordt vervolgens het algoritme opnieuw getraind op een dataset in een andere omgeving zonder data van de verschillende organisaties te koppelen. Deze manier van externe validatie noemt men ook wel horizontale federated learning omdat van verschillende mensen hetzelfde type data in een andere context worden geanalyseerd. Verticale federated learning is de aanpak waarbij van dezelfde patiënten na data-encryptie data uit verschillende echelons (eerste-, tweede- en derdelijnszorg) in een serie analyses gebruikt worden, wederom zonder data te delen of bij elkaar te brengen.

In de psychiatrie

In de psychiatrie is zowel horizontale als verticale federated learning van belang omdat ggz-organisaties van elkaar verschillen en in verschillende echelons (huisartsenpraktijk, basiszorg en specialistische zorg) mentale ondersteuning geboden wordt.

Fase van toepassing

Algoritmes kunnen toepasbaar zijn in de zorg in de vorm van bijvoorbeeld een *decision support system* (DSS) oftewel een beslisondersteuning. Dit kan de zorg verbeteren omdat in een DSS grote hoeveelheden data verwerkt zijn van heel veel professionals en patiënten, waardoor een professional als het ware in één keer al zijn collega's kan raadplegen en niet alleen afhankelijk is van eigen praktijkervaring in combinatie met kennis uit protocollen en richtlijnen.

Echter, er ontstaat een spanningsveld op het moment dat de uitkomst van een algoritme niet meer te begrijpen is voor de gebruiker. Machinelearningtechnieken kunnen zo complex in elkaar zitten dat niet meer te volgen is hoe een bepaalde beslissing tot stand is gekomen (Burrell 2016). Als een behandeling als meest passend wordt geadviseerd door de computer, maar niet duidelijk is waarom, en de professional op basis van richtlijnen tot een ander advies zou komen, is het moeilijk af te wegen welk advies men het beste kan volgen. De transparantie van het algoritme versus de uitlegbaarheid of begrijpbaarheid moet daarom in balans zijn (Char 2018; Holm 2019): *explainable AI*.

Een manier om hiermee om te gaan is algoritmes iteratief te ontwikkelen in cocreatie met patiënten en professionals. Van belang hierin is dat in het proces steeds kleine stapjes gezet worden en professionals en/of patiënten betrokken worden om betekenis te geven aan de data (Menger e.a. 2016). Ook is het belangrijk een DSS te zien als slechts één aspect van de zorg.

In de psychiatrie

Bij het nemen van beslissingen moeten persoonlijke doelen van de patiënt steeds centraal staan. Ook de relatie tussen patiënt en professional bepalen mede het effect van de behandeling, zeker in de psychiatrie. In deze relatie spelen factoren een rol die niet allemaal meetbaar zijn en daarom nooit als digitale data in een dossier terug te vinden, laat staan dat ze meegenomen kunnen worden in een algoritme. Een DSS is dus ondersteunend, zeker niet leidend, en de beslissing welke behandeling men zal inzetten, blijft een uitkomst die in de dialoog tussen professional, patiënt en diens naasten tot stand komt.

Conclusie

Onderzoek naar de inhoudelijke en technische kanten van bigdata-analyses zou hand in hand moeten gaan met onderzoek naar de ethische aspecten van het gebruik van (big) data in de zorg en de psychiatrie in het bijzonder. De mening van patiënten zelf is hierbij altijd een belangrijke factor die we moeten meewegen om echt waarde toe te voegen aan de zorg (zoals ook beschreven door Fiske e.a. 2019 en Richter e.a. 2020).

LITERATUUR

- Boeckhout M, Zielhuis GA, Bredenoord AL. The FAIR guiding principles for data stewardship: fair enough? *Eur J Hum Gen* 2018; 26: 931-6.
- Burrell J. How the machine 'thinks': Understanding opacity in machine learning algorithms. *Big Data & Society* 2016; 3: 1-12.
- Char DS, Shah NH, Magnus D. Implementing machine learning in health care - addressing ethical challenges. *N Engl J Med* 2018; 378: 981-3.
- Fiske A, Henningsen P, Buyx A. Your robot therapist will see you now: ethical implications of embodied artificial intelligence in psychiatry, psychology, and psychotherapy. *J Med Internet Res* 2019; 21: e13216.
- Floridi L, Cows J, Beltrametti M, e.a. AI4People - An ethical framework for a good AI society: opportunities, risks, principles, and recommendations. *Minds Mach (Dordr)* 2018; 28: 689-707.
- Holm EA. In defense of the black box. *Science* 2019; 364: 26-7.
- London AJ. Artificial intelligence and black-box medical decisions: accuracy versus explainability. *Hastings Center Report* 2019; 49: 15-21.
- Menger V, Spruit M, Hagoort K, e.a. Transitioning to a data driven mental health practice: collaborative expert sessions for knowledge and hypothesis finding. *Comput Math Methods Med* 2016; doi: 10.1155/2016/9089321.
- Menger V, Scheepers FE, van Wijk LM, e.a. DEDUCE: A pattern matching method for automatic de-identification of Dutch medical text. *Telematics and Informatics* 2018; 35: 727-36.
- Mittelstadt BD, Floridi L. The ethics of big data: current and foreseeable issues in biomedical contexts. *Sci Eng Ethics* 2016; 22: 303-41.
- Morley J, Machado CCV, Burr C, e.a. The ethics of AI in health care: A mapping review. *Soc Sci Med* 2020; 260: 113172.
- Richter G, Borzikowsky C, Lesch W, e.a. Secondary research use of personal medical data: attitudes from patient and population surveys in The Netherlands and Germany. *Eur J Hum Genet* 2020; doi: 10.1038/s41431-020-00735-3.
- Russ TC, Woelbert E, Davis KAS, e.a. How data science can advance mental health research. *Nat Hum Behav* 2019; 3: 24-32.
- Raad voor Volksgezondheid en Samenleving (RVS). Zonder context geen bewijs. Over de illusie van evidence-based practice in de zorg. Dan Haag: RVS; 2017. <https://www.raadrvs.nl/documenten/publicaties/2017/06/19/zonder-context-geen-bewijs>.
- Scheepers FE, Menger V, Hagoort K. Datascience in de psychiatrie. *Tijdschr Psychiatr* 2018; 60: 205-9.
- Vabalas A, Gowen E, Poliakoff E, e.a. Machine learning algorithm validation with a limited sample size. *PLoS One* 2019; 14: e0224365.

SUMMARY

Ethics of (big) data applications in psychiatric practice

F.E. Scheepers, M. Mostert, M.M. Milota, G.J.M. van Thiel

- Background** Due to rapid digitalization, an increasing amount of data is available in healthcare settings; big data and artificial intelligence (AI) have also made their appearance.
- Aim** To provide insight into various ethical dilemmas that need to be considered when applying big data in clinical practice.
- Method** Description and analyses of the ethical aspects associated with the use of clinical data in the context of psychiatric care.
- Results** Various ethical aspects play a role in four phases; data collection, analysis, dissemination and application of results. In order to use clinical data and AI in a responsible manner, these aspects must be taken into account.
- Conclusion** The use of big data and AI in healthcare should aim to stimulate learning and improving care together with patients and professionals. Big data and AI should not be seen as the holy grail, but as a supporting tool in healthcare - a field in which many of the aspects that play a role in clinical care cannot be converted into measurable data.