

Personlig medicin nødvendiggør paradigmeskifte i evidensbegrebet

Ingolf Nielsen, Bjarne Sigurd & Bjarne Ørskov

STATUSARTIKEL

Ugeskr Læger
2019;181:V07180480

Med personlig medicin stiler man mod at optimere behandlingen af den enkelte patient ved at skræddersy en individuel medicinsk behandling baseret på livsstil, miljø og mønstre af dybe fænotypiske og genomiske data [1]. Behandling med personlig medicin adskiller sig derfor fra den traditionelle evidensbaserede behandling, som er baseret på dobbeltblindede randomiserede undersøgelser (RCT), hvor der ofte forekommer stor spredning i behandlingseffekt, strækkende sig fra helbredende over ingen til skadelig effekt [2-4].

Behandling med personlig medicin kræver imidlertid et paradigmeskifte med revision af de evidensbegreber, som den medicinske prognostiske prædiktion hviler på nu, til fordel for accept af korrelationer og mønstre fremkommet ved brug af kunstig intelligens (AI) som basis for navigation (**Figur 1**).

EVIDENSBASERET FARMAKOLOGISK BEHANDLING

Evidens, der er baseret på dobbeltblindede randomiserede undersøgelser (RCT) kræver en begrænsning i observationsforhold af et omfang, så evidensen sjældent kan ekstrapoleres til virkeligheden. Med RCT, som jo også hviler på statistiske sandsynligheder, har man opnået mange betydningsfulde succeser under standardiserede forhold i de forskellige individuelle forsøg. Når man skal omsætte denne viden til handling i »virkeligheden«, foretages der nødvendigvis en ekstrapolation. Jo mindre lighed mellem forsøgsomstændighederne ved RCT og de »virkelighedspatienter«, vi behandler, des mere usikkert er grundlaget for deduktionen. Det er elementært. Dertil kommer, at man ved en farmakologisk behandling påvirker den totale organisme med interaktive konsekvenser for det, man påvirker, og det

må formodes, at det i et vist omfang også gælder placebo [2]. Disse problemer må antages at være særlig udtalte hos patienter, som er i behandling med flere forskellige farmaka, altså polyfarmacipatienterne. Ved »evidens«, som er indhentet i RCT, agerer man på baggrund af et patientmateriale, som er »renset« for patienter med sygdomme, som man ville forvente kunne interagere med den tilstand, som blev behandlet. Disse forudsætninger er langt fra altid til stede i den kliniske virkelighed uden for studierne. Dertil kommer, at RCT udføres med køns- og aldersgrupper, der kun i begrænset grad svarer til det mønster, der dominerer i praksis. F.eks. angår langvarig/permanent farmakoterapi overvejende ældre (> 65 år) med flere konkurrerende lidelser, og fraktionen af meget gamle (> 85 år) stiger støt i disse år. Da det f.eks. er kendt, at metaboliseringen af individuelle lægemidler falder med alderen, undertiden forskelligt fra lægemiddel til lægemiddel, er der også her et generelt uafklaret problem. En anden væsentlig faktor hos polyfarmacipatienterne er problemet med komplians [3].

Der er naturligvis gjort rede for sådanne problematikker i de RCT, som vi anvender som referencer, men som ofte kun tilnærmelsesvis er sammenlignelige med den kliniske virkelighed uden for studierne. Realiteten er, at virkelighedens behandlingsbeslutninger, ud over de metodeafhængige statistiske sandsynligheder, kommer til at hvile på et usikkert teoretisk grundlag, der som hovedregel ikke er reflekteret over i de tilgrundsiggende RCT-resultater.

Med alle disse og formentlig flere forbehold, herunder overrapportering af positive resultater [4], kan man så spørge, hvorfor evidenstankegangen har holdt sig så længe? Det korte svar er: i mangel af et bedre paradigme.

PERSONLIG MEDICIN FORUDSÆTTER ET PARADIGMESKIFTE OG OPGØR MED EVIDENSBEGREBET INDEN FOR KOMPLEKSE SYSTEMER

Et paradigmeskifte er imidlertid nu ved at forme sig. Det drejer sig om et computerbaseret dynamisk begrebsapparat, som vedvarende udvikles og forbedres, styret af resultaterne (konsekvenser). Det navigerer efter systemer, der tilbyder styrket statistisk sandsynlig-

HOVEDBUDSKABER

- ▶ Personlig medicin baseret på »sundhedsdata«, herunder oprettelse af en central biobank, kræver et paradigmeskifte fra deduktion til induktion af det evidensbegreb, som medicinsk prædiktion hviler på.
- ▶ Skal data udnyttes effektivt i terapi, kræver det maskinintelligens med stor kapacitet for mønsterpåvisning.
- Dette er baseret på *unsupervised learning*, hvor algoritmen under dataindsamlingen stadigt forbedrer sin prædiktion på baggrund af mønsterpåvisning – en induktion.
- ▶ Paradigme, der ikke er baseret på logisk tænkning – en deduktion – kan ikke kontrolleres og »forstås« logisk.

hed, nemlig selvlærende algoritmer/maskinlæring, AI, baseret på Bayes' teorem [5, 6].

AI giver forslag til prognostik, men kan naturligvis aldrig resultere i perfektion. Den findes kun i den deduktive logiks teoriverden. Men det er bedre at navigere efter systemer, der kan tilbyde forbedret statistisk sandsynlighed, end at navigere efter systemer, som er baseret på en tilnærmet virkelighed [7]. AI hviler på store heterogene datamængder i form af »dybe fænotypiske data«, som bl.a. omfatter data fra »omics« (genetics, genomics, transcriptomics, epigenomics, proteomics, and metabolomics), sociale medier, ekkokardiografi, myokardiescintigrafi, CT, MR-skanning, koronararteriografi etc. Disse datamængder har til dato været alt for store og heterogene og er blevet ændret for hurtigt til, at de kan lagres, analyseres og anvendes. AI har potentiale til at kunne udnytte disse data, så de kan anvendes i avanceret diagnosticering og patientbehandling. F.eks. er kardiovaskulær sygdom særdeles kompleks og heterogen med baggrund i multiple miljøfaktorer samt genetiske og adfærdsmæssige faktorer, hvorfor specielt AI her har et stort potentiale [3, 7].

HVORLEDES FUNGERER DE SELVLÆRENDE ALGORITMER, OG HVAD ER MULIGHEDERNE MED KUNSTIG INTELLIGENS I LÆGEVIDENSKABEN?

Maskinlæring

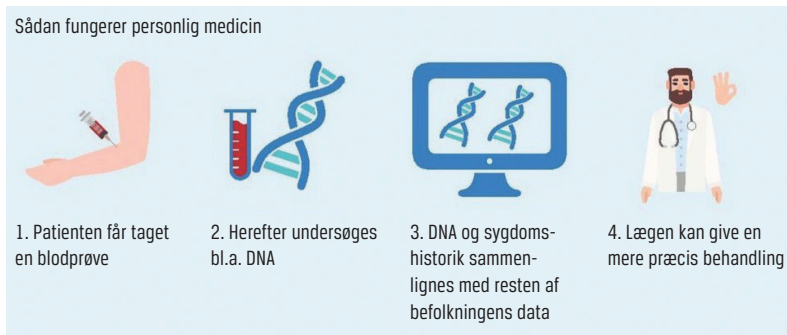
I modsætning til traditionel statistik er man ved maskinlæring i form af *deep learning* fokuseret på at konstruere automatiserede kliniske beslutningssystemer (som f.eks. genindlæggelses- og mortalitetsscorings-systemer), der kan hjælpe behandleren til at foretage mere præcise estimater, end det er muligt med mere simple scoringsystemer.

Deep learning kan opdeles i tre læringsformer: *supervised*, *unsupervised* og *reinforced*. Ved *supervised learning* anvender algoritmer datasæt, der i forvejen er mærket (etiketteret) til forudsigelse af et ønsket og kendt resultat. *Supervised learning* er velegnet til klassificering af regressionsproblemer, dvs. til forudsigelse af værdien af en eller flere kontinuerte variabler, men der kræves store mængder data, og metoden er meget tidskrævende, da data skal mærkes af en operatør. Omvendt kan *unsupervised learning* anvendes til identifikation af nye sygdomsmekanismer, genotyper eller fænotyper fra skjulte mønstre, der kan findes i data. Algoritmen forbedrer løbende sin prædiktions under dataindsamlingen på baggrund af mønsterpåvisning. Med *unsupervised learning* er målet således at finde de skjulte mønstre i data uden feedback fra operatører. F.eks. kan undervisning (etikettering) af en nyuddannet læge, før han/hun tilser patienten, betegnes som *supervised learning*. I modsætning hertil kan f.eks. en nyuddannet læge tilse patienten uden instruktion (ingen etikette). Han/hun kan herved »lære« af sine fejl-

FIGUR 1

Illustration fra annonce for personlig medicin bragt i et annoncetillæg om Folkemødet på Bornholm 2018 under overskriften »Personlig medicin – fremtidens mulighed for kronisk syge patienter« distribueret med Berlingske Tidende juni 2018. Annoncen lover debat i Meningsministeriet fredag den 15. juni kl. 09.00-10.00, hvor »Hjernemadsen tager os med til fremtiden«. Debatten var arrangeret af Giftforeningen, Psoriasisforeningen og Colitis-Crohn Foreningen og sponsoreret af et medicinalfirma. Spørgsmålet er, om danske myndigheder, læger og patienter er parate til at acceptere medicinordinationer baseret på maskinlæring, der ikke kan kontrolleres og forstås logisk.

(Bragt med tilladelse fra PartnerMedier, Henning Andersen).



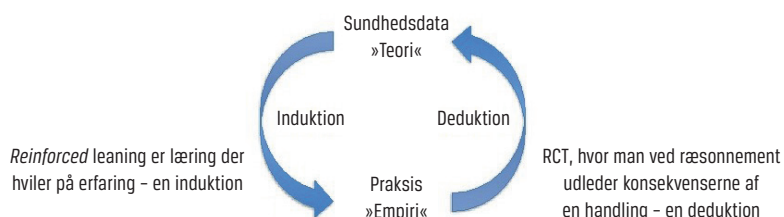
tagelser og fremkomme med egne forslag til diagnose, behandling etc. Dette betegnes *unsupervised learning*. Nogle algoritmer, såsom kunstige neurale netværk, kan trænes vha. *supervised* eller *unsupervised learning* for at optimere nøjagtigheden af automatiseret prædiktions.

Reinforced learning kan ses som en hybrid af *supervised* og *unsupervised learning*. Ideen bag *reinforced learning* er at maksimere nøjagtigheden af algoritmer med løbende justering, så fremtidige handlinger baseres på erindring af de »erfaringer«, man tidligere har gjort i form af forsøg og fejl. *Reinforced learning* er således læring, der hviler på erfaring, en induktion, til forskel for ræsonnementet kendt fra RCT, hvor man ved en logisk proces [3, 7] (Aristoteles' syllogisme) udleder konsekvenserne af en handling, en deduktion [6]. Ved introduktion af personlig medicin anvendes *reinforced learning* med det deraf følgende paradigmeskifte fra deduktion til induktion som grundlag for medicinordinationen (Figur 2).

Reinforced learning har givet positive resultater i andre sammenhænge. DeepMind (en Google-algoritme) udnytter f.eks. i algoritmen AlfaZero et begreb fra erkendelsesteorien kendt som tabula rasa. Med dette princip lagrer man ikke en forhåndsviden om metoder til problemets løsning i det neurale netværk, men lader det fra helt bar bund selv lære at regulere styrken mellem de interne forbindelser, som danner et hjerne-lignende netværk (neuralt netværk), indtil det opnår en form for perfektion inden for en veldefineret opgave som f.eks. at spille skak. Hos Googles DeepMind er træning, træning og atter træning nøglen til succes [8, 9]. Disse algoritmer er således ikke baseret på deduktiv

FIGUR 2

Personlig medicin. Udnyttelse af sundhedsdata i medicinsk praksis kræver maskinintelligens med stor kapacitet for mønsterpåvisning baseret på *unsupervised learning*, hvor algoritmen under dataindsamlingen stadigt forbedrer sin prædiktions på baggrund af mønsterpåvisning – en induktion. Dette bryder med det hidtidige grundlag for medicinsk prognostisering og prædiktions baseret på logisk ræsonnement, hvor man i medicinsk praksis vejledes af et randomiseret kontrollet forsøg (RCT) i valg af medikamentel behandling – en deduktion.



logik, men (induktiv) regelfri læring. De modstår derfor en »logisk« forklaring eller efterregning/kontrol og er blevet kaldt »sorte bokse« [10, 11].

Kunstig intelligens' rolle i medicinsk-farmakologisk behandling

Vha. AI vil der nok langt hurtigere end hidtil kunne gennemføres kliniske forsøg med et givet lægemiddel, ligesom det potentielt vil kunne bringes betydeligt hurtigere på markedet og dermed til anvendelse i den daglige klinik [12]. Mens RCT som hovedregel formentlig fortsat vil være grundlag for registrering, kan man vha. dataanalyse på en anderledes måde matche patienter til relevante kliniske forsøg. Forskellige onlinetjenester gør det muligt at inkludere flere og flere patienter i kliniske forsøg med nye lægemidler eller allerede anvendte lægemidler ved nye diagnoser.

Et andet vigtigt område for AI er design af nye lægemidler. F.eks. bruger firmaet Atomwise supercomputere til design af lægemidler ud fra en database, som indeholder store mængder molekylære strukturer. I 2016 igangsatte firmaet således en virtuel søgning efter sikre, allerede anvendte lægemidler, der kunne redesignes til behandling af ebolavirus. Man fandt her vha. AI to lægemidler, som kunne reducere en ebolainfektion signifikant. Denne analyse, som typisk ville have taget måneder eller år, blev gennemført på mindre end en dag [13]. AI kan således være et supplement til faseundersøgelserne til RCT op til en egentlig registrering. Man kan godt forestille sig, hvordan effektive lægemidler vil kunne designes, hvis sådanne kliniske forsøg ville blive gennemført på »ground zero-niveau« i sundhedsvæsenet.

KONKLUSION

Personlig medicin er en ny tilgang til sekundær sygdomsforebyggelse og behandling, hvor man tager hensyn til individuel variabilitet i gener, miljø og livsstil for

den enkelte patient. Dette kræver betydelig computerkraft og AI-selvlærende algoritmer (*deep learning*) i form af *reinforced learning*. *Deep learning*-algoritmer har vist sig at kunne diagnosticere mindst lige så præcist som speciallæger i kardiologi, dermatologi og onkologi [7]. Det tidligere knæsatte evidensbegreb påvist gennem RCT bliver hermed afløst af det, der kommer ud af den såkaldt sorte boks som et resultat af, at maskinerne lærer, efterhånden som de fodres med data, hvilket medfører et paradigmeskifte i grundlaget for medicinordination, nemlig fra deduktion via RCT til induktion via AI. AI fungerer i forbindelse med sundhedsydelse, f.eks. identifikation og diagnosticering af sygdom, ved udvikling af og produktion af lægemidler og ved opdateringer af patientjournaler. Med AI kan man imidlertid ikke kontrollere »udregningen« på en forståelig måde. Mellem datainput og dataoutput er der ganske mange lag og variabler, som afvejes og kombineres på forskellige måder. Dette skaber et dybt komplekst net, et virvar, hvor det er vanskeligt at finde en forståelig sammenhæng. Et sådant paradigme, der er baseret på maskinlæring, kan selvsagt ikke kontrolleres og forstås logisk. Dets karakter af »sort boks« har medført, at man arbejder intenst på at udvikle metoder til at forstå beslutningsprocesserne i netværkene, og dermed kunne kontrollere resultaterne, som i sidste ende kan være en individuel medicinordination i form af personlig medicin [13, 14]. Indtil dette problem er løst, vil lægens endelige vurdering og beslutning om medicinsk behandling beholde sin essentielle position.

SUMMARY

Ingolf Nielsen, Bjarne Sigurd & Bjarne Ørskov:

Personalised medicine requires a paradigm shift in the concept of evidence

Ugeskr Læger 2019;181:Vo718o48o

Personalised medicine via a central biobank will require the introduction of a paradigm shift navigating by detecting pattern and correlation, instead of evidence of limited use in our complex environment. Accordingly, to make sense, population-wide data comprising "deep phenotype" including genetics, must be submitted to evaluation by "machine intelligence", based on "unsupervised learning" – the algorithm concomitantly improving its power with increasing mass of data. A system not based on deductive logic cannot be checked by logic but must be taken on face value.

KORRESPONDANCE: Bjarne Sigurd. E-mail: sigurd@dadlnet.dk

ANTAGET: 8. januar 2019

PUBLICERET PÅ UGESKRIFTET.DK: 4. marts 2019

INTERESSEKONFLIKTER: ingen. Forfatternes ICMJE-formularer er tilgængelige sammen med artiklen på Ugeskriftet.dk

LITTERATUR

1. Ullum H, Herson C, Wewer U et al. Personlig medicin vil involvere alle specialer, og det handler om andet end genomet. *Ugeskr Læger* 2018;180:V69884.
2. Kennedy-Martin T, Curtis S, Faries D et al. A literature review on the representativeness of randomized controlled trial samples and implications for the external validity of trial results. *Trials* 2015;16:495.
3. Krittanawong C, Zhang H, Wang Z et al. Artificial intelligence in precision cardiovascular medicine. *J Am Coll Cardiol* 2017;69:2657-64.
4. Camerer CF, Dreber A, Holzmeister F et al. Evaluating the replicability of social science experiments in Nature and Science between 2010 and 2015. *Nat Hum Behav* 2018;2:637-44.
5. Stone JV. Bayes' rule. Sebtel Press, 2013.
6. Valiant L. Probably approximately correct. Basic Books, 2013.
7. Mesko B. A guide to artificial intelligence in healthcare. The Medical Futurist, 2017.
8. Silver A. The future is here – AlphaZero learns chess. Chess News, 2017. <https://en.chessbase.com/post/the-future-is-here-alphazero-learns-chess> (25. jan 2019).
9. Silver D, Hubert T, Schrittwieser J et al. Mastering chess and shogi by self-play with a general reinforcement learning algorithm. DeepMind, 2017.
10. Klingenberg T. Ny undersøkelse: «Den sorte boksen» er det største hinderet for å ta i bruk kunstig intelligens. *Teknisk Ukeblad* 2017;10:34.
11. Park DH, Hendricks LA, Akata Z et al. Multimodal explanations: justifying decisions and pointing to the evidence. arXiv:1802.08129v1.
12. Keiding N. Diskussion om Food and Drug Administrations rolle, når tiderne skifter. *Ugeskr Læger* 2018;180:V11170868.
13. Mesko B. The role of artificial intelligence in precision medicine. *Expert Rev Precis Med Drug Dev* 2017;2:239-41.
14. Lovett L. Artificial intelligence is getting the hype, now what are the applications? HIMSS, 2018. <https://www.mobihealthnews.com/content/artificial-intelligence-getting-hype-now-what-are-applications> (25. jan 2019).