AI i sundhedsvæsenet

- omsorg, ansvar og faglighed



DET ETISKE RÅD

Indhold

DET ETISKE RÅD

Dato: 04-11-2026
M: kontakt@etiskraad.dk
W: www.etiskraad.dk

3 Introduktion

5 Hvad er AI?

7 Hvor anvendes AI i sundhed i dag?

10 AI i praksis – etiske dilemmaer og udfordringer

12 Etiske udfordringer for brug af AI i sundhed

13 Bias

15 Forklarlighed

20 Ansvar

21 Ansvar og kunstig intelligens i sundhedsvæsenet

26 Fremtiden for AI i sundhed

28 Noter

Introduktion

Kunstig intelligens (AI) er rykket fra forskningsprojekter og pilotforsøg til konkrete kliniske og administrative løsninger. Evalueringen af de danske AI-signaturprojekter[1] peger på både resultater og udfordringer. Potentialer i effektivisering og kvalitet – samt barrierer i form af dataadgang, jura og uklare rollefordelinger. Forventningerne til, at kunstig intelligens kan bidrage til mere præcis diagnostik, tidligere opsporing, bedre prioritering af ressourcer og mere individualiseret behandling, er dog intakte, og AI spiller en fremtrædende rolle i regeringens strategi for life science frem mod 2030.[2]

Med regeringsgrundlaget fra 2026 har kunstig intelligens desuden fået sit eget kapitel i et regeringsgrundlag – med initiativer af direkte betydning for sundhedsområdet, herunder en digital sundhedsassistent, en AI-accelerationsfond til udrulning i den offentlige sektor og et mål om at frigøre mindst 30.000 offentlige årsværk.[3]

Teknisk ydeevne er kun én af forudsætningerne for, at kunstig intelligens kan skabe varig værdi i sundhedsvæsenet. Stanford AI Index Report 2026 fremhæver desuden behovet for klinisk evidens, governance og etiske rammer.[4]

> **Er et sundhedsvæsen bedre, hvis det er mere effektivt? Hvis færre fejl begås? Hvis flere patienter behandles hurtigt? Eller hvis omsorgen styrkes og ligheden øges?**

Det giver anledning til en række etiske spørgsmål. For selv om hensigten med at udvikle og anvende kunstig intelligens i sundhedsvæsenet typisk er at forbedre det, er det langt fra givet, hvad det vil sige at forbedre et sundhedsvæsen. Er et sundhedsvæsen bedre, hvis det er mere effektivt? Hvis færre fejl begås? Hvis flere patienter behandles hurtigt? Eller hvis omsorgen styrkes og ligheden øges? Og hvordan skal man vurdere situationer, hvor kunstig intelligens styrker nogle af disse værdier, men samtidig svækker andre?

Samtidig rejser teknologien spørgsmål om sundhedsvæsenets kerneværdier og kultur. Hvordan påvirkes omsorgen, relationen mellem patient og sundhedsprofessionelle og de professionelles faglige identitet, hvis beslutninger i stigende grad understøttes af værktøjer, som ikke kan forklares i almindeligt sprog? Og hvad betyder det for tillid og nærvær, hvis klinikere føler sig for afhængige af systemer, de ikke har tid eller kompetencer til at efterprøve?

Spørgsmålene har ingen entydige svar. De berører grundlæggende forestillinger om sundhedsvæsenets formål, de værdier, det bør styres af, og den måde, ansvar, tillid og faglighed skal forstås på i en stadig mere digital praksis.

Dette dokument beskæftiger sig med et hjørne af området - primært med anvendelsen af datadrevet, opgaveafgrænset kunstig intelligens (fx maskinlæring) i sundhedsvæsenets kliniske og administrative praksis. Fokus er ikke på generativ AI i bred forstand, men på de løsninger, der allerede er – eller er på vej – i drift, og som rejser konkrete etiske spørgsmål om omsorg, ansvar og faglighed.

Formålet med introduktionen er ikke at fremlægge Det Etiske Råds samlede stillingtagen til de etiske dilemmaer, som anvendelsen af AI i sundhedsvæsenet rejser. Teksten skal derimod gerne give et overblik over anvendelser og centrale etiske problemstillinger og fungere som afsæt for refleksion og offentlig debat. Målgruppen er patienter, sundhedsprofessionelle, ledere, myndigheder og politikere.

Et gennemgående spørgsmål er, hvordan kunstig intelligens kan understøtte og kvalificere faglighed og menneskelig vurdering uden at erstatte dem. Derfor retter teksten opmærksomheden mod de etiske hensyn, der kan være relevante, når AI indføres i sundhedsvæsenet. Når teksten peger på sådanne hensyn og spørgsmål, er der ikke tale om Det Etiske Råds formelle anbefalinger.

Teksten skriver sig dermed ind i en international ramme, hvor blandt andre UNESCO og Verdenssundhedsorganisationen (WHO) har opstillet etiske principper for udvikling og anvendelse af kunstig intelligens, herunder på sundhedsområdet.[5]

Hvad er AI?

ChatGPT blev lanceret offentligt som en forskningsversion i november 2022. Ved udgangen af juli 2025 havde tjenesten ifølge OpenAI mere end 700 millioner ugentlige aktive brugere på verdensplan.[6] Dertil kommer konkurrerende tjenester som Claude, Gemini og Copilot.

Men disse store sprogmodeller (GenAI Chatbots) er blot det nyeste skud på stammen indenfor det, vi bredt set betegner som kunstig intelligens eller i daglig tale AI.

Der findes ikke en entydig definition på begrebet. Udtrykket blev introduceret i 1955 af den amerikanske forsker John McCarthy, som beskrev kunstig intelligens som videnskaben og teknikken til at skabe maskiner, der udviser adfærd, som ville blive betragtet som intelligent, hvis et menneske gjorde det samme. I dag bruges begrebet typisk om computersystemer, der kan udføre opgaver, som tidligere har krævet menneskelig intelligens, såsom mønstergenkendelse, sproglig forståelse, læring, problemløsning eller beslutningsstøtte.

Fra ekspertsystemer til moderne maskinlæring

Logikbaserede ekspertsystemer (1970–80'erne)
Tidlige AI-systemer byggede på faste "hvis så"-regler, udledt af menneskelige eksperters viden. De var gennemsigtige, men kunne kun håndtere de situationer, der var indkodet på forhånd. De var svære at udbygge og håndterede dårligt usikkerhed, hvilket førte til stagnation og den såkaldte AI-vinter sidst i 1980'erne.

Maskinlæring (2000'erne og frem)
I stedet for faste regler lærer systemerne mønstre ud fra store datamængder. Det gør dem robuste over for variation og i stand til at løse komplekse opgaver som billedgenkendelse, sprogforståelse og anbefalinger. Til gengæld kræver de meget data og regnekraft og kan være uigennemsigtige i deres beslutninger.

Superviseret læring
En central gren af maskinlæring, hvor modellerne trænes på data med et kendt facit. Fx kan en algoritme trænes på tusindvis af mammografier, hvor eksperter har markeret "tegn på kræft/ikke tegn". Modellen justerer sine parametre og bliver gradvist mere præcis i sine klassifikationer.

Dybe neurale netværk
Avancerede modeller inspireret af hjernens opbygning. De består af mange lag af neuroner og kan lære meget komplekse mønstre i billeder, tekst og signaler. Teknologien har drevet gennembrud i bl.a. selvkørende biler, sprogmodeller og medicinsk billedanalyse. De kræver store datamængder og kan være vanskelige at forstå og forklare, hvilket har øget fokus på forklarlighed og biashåndtering.

Generativ AI
I stedet for blot at genkende mønstre lærer disse modeller at skabe nyt indhold – tekst, billeder, lyd eller kode – ved at forudsige det mest sandsynlige næste element ud fra enorme mængder træningsdata. Store sprogmodeller og billedgeneratorer kan på den måde føre samtaler, skrive tekster eller producere billeder ud fra en simpel instruktion (en "prompt"). Teknologien åbner for nye anvendelser inden for bl.a. kommunikation, undervisning og kreativt arbejde, men indholdet kan virke overbevisende uden at være korrekt, og det rejser spørgsmål om ophavsret, misinformation og ansvar for det, der genereres.

Hvor anvendes AI i sundhed i dag

Hvor anvendes AI i sundhed i dag

Kunstig intelligens er på kort tid blevet en integreret del af mange funktioner i sundhedsvæsenet. Nogle af disse anvendelser er synlige for klinikere og patienter, mens andre arbejder i baggrunden og allerede påvirker både arbejdsprocesser, effektivitet og kvalitet. I det følgende gives et overblik over de væsentligste anvendelsesområder

Diagnostik og billedanalyse

AI er længst i anvendelse inden for diagnostik, især billedanalyse. I Danmark bruges AI til at finde kræftforandringer i mammografier og til at sortere normale fund fra, så radiologer kan fokusere på de svære sager. På Herlev og Gentofte Hospital har en AI-algoritme forbedret den diagnostiske kvalitet, reduceret falsk positive og nedbragt antallet af kvinder, der skal indkaldes igen – samtidig med at radiologernes læsearbejde blev reduceret med omkring en tredjedel.[7] (se også ***CASE: AI i praksis – etiske dilemmaer og udfordringer ved AI i brystkræftscreening***)

I Region Nordjylland fungerer AI som første læser af lavrisikobilleder, hvilket letter presset på radiologer.[8]

I akutområdet bruges AI til at frikende røntgenskader uden brud, hvilket ifølge Region Nordjylland reducerer ventetid, minimerer oversete frakturer og gør det muligt for sygeplejersker og yngre læger at afslutte flere forløb selv.[9]

Prognose og tidlig risikovurdering

AI anvendes i stigende grad til at forudsige sygdomsudvikling og identificere patienter i risiko for forværring. Teknologierne kan overvåge vitale data kontinuerligt – både på hospitalet og i patientens eget hjem – og advare sundhedspersonalet, før der opstår kliniske symptomer.

I Region Nordjylland bruges AI-baserede overvågningsløsninger til at opdage forværring tidligt og loftmonterede sensorer til at forudsige og forebygge fald hos ældre patienter.[10]

AI som klinisk beslutningsstøtte

Mange AI-løsninger fungerer som beslutningsstøtte frem for automatisering. De hjælper klinikere med at sortere store informationsmængder, navigere i komplekse forløb, foreslå behandlingsmuligheder og udarbejde mere konsistente behandlingsplaner. Region Nordjylland beskriver i sit AI-katalog en ambition om at anvende AI til blandt andet at understøtte udarbejdelsen af behandlingsplaner og foreslå mulige behandlingsmetoder. Kataloget beskriver desuden konkrete løsninger, der kan hjælpe klinikere med at identificere og prioritere patienter med mulige alvorlige fund.[11] Store sprogmodeller er også begyndt at indgå i klinisk praksis ved at formulere journaltekster, epikriser og patientforklaringer.

Patient- og borgerrettet AI

AI møder borgere direkte gennem wearables, smartphones og sundhedsapps. Disse teknologier opsamler sundhedsdata i realtid, giver personlig rådgivning, screener symptomer og hjælper med navigering i sundhedssystemet. Ure og telefoner fungerer i praksis som uautoriserede screeningsværktøjer, fx ved at registrere hjerterytmeanomalier. Borgere og patienter møder i stigende grad AI, når de søger sundhedsinformation.

I fem undersøgte kategorier af sundhedsrelaterede søgninger udløste mellem 84 og 92 procent et AI-genereret resumé fra Google, ifølge Stanford AI Index Report 2026. Og antallet af videnskabelige publikationer om patienters opfattelse af AI i sundhedsvæsenet er tidoblet fra 2020 til 2025.[12] Nogle bruger direkte AI-chatbots som rådgivere i sundhedsspørgsmål[13] som denne personlige beretning illustrerer,[14] og i januar 2026 introducerede OpenAI ChatGPT Health som et særskilt tilbud til private brugere.[15]

Drift og ressourcestyring

AI bruges til at optimere patientflow og driftsprocesser. Teknologierne hjælper med visitation, kapacitetsstyring, prioritering og ressourcedisponering – alt sammen områder, der presser sundhedsvæsenet. I Sønderborg Kommune bruges AI-løsningen RefCase, der ved hjælp af sprogteknologi giver et hurtigt overblik over sagsoplysninger i borgerens sag og identificerer data fra lignende sager. Ifølge evalueringen kan løsningen bidrage til hurtigere sagsbehandling i sager med mange helbredsoplysninger.[16]

Administration og kommunikation

Kunstig intelligens anvendes også til administrative og kommunikative opgaver i den offentlige sektor. Selv om de følgende eksempler ikke alle stammer fra sundhedsvæsenet, illustrerer de anvendelser, som kan være relevante i sundhedsadministrative sammenhænge. I Norddjurs Kommune bruges SmartMail til sortering og journalisering af e-mails. Roskilde Kommune har udviklet talerobotten Kommune-Kiri, som kan besvare en række spørgsmål, når borgere ringer til kommunens hovednummer. I Sønderborg Kommune anvendes AI desuden til at understøtte sagsbehandleres arbejde med aktindsigt, blandt andet ved at identificere oplysninger, som muligvis skal undtages fra udlevering.[17]

AI i praksis – etiske dilemmaer og udfordringer

10

AI i praksis – etiske dilemmaer og udfordringer ved AI i brystkræftscreening
(En case hentet fra faktaboksen: Brystkræftscreening som ansvarslaboratorium)

I Danmark tilbydes alle kvinder mellem 50 og 69 år screening for brystkræft hvert andet år. Et retrospektivt før-efter-studie fra Region Hovedstaden sammenlignede screeninger af 60.751 kvinder før AI blev indført med screeninger af 58.246 kvinder efter indførelsen.[18] Screeninger, som AI-systemet vurderede som sandsynligt normale, blev læst af én seniorradiolog, mens de øvrige fortsat blev læst af to radiologer – nu med AI-støtte. Hidtil var alle screeninger blevet læst af to radiologer. Ved uenighed blev sagen afgjort enten på en konsensuskonference eller ved en tredje radiolog. Med den nye arbejdsdeling blev der fundet flere kræfttilfælde (0,82 % mod 0,70 % – svarende til 12 flere pr. 10.000 screeninger), samtidig med at både genindkaldelser og falsk positive svar faldt markant, og antallet af radiologiske læsninger af screeningsundersøgelser blev reduceret med omkring en tredjedel. Stigningen i fund rummer dog både muligheden for tidligere opsporing og spørgsmål om overdiagnostik. En større andel af fundene var ikke-invasiv brystkræft (DCIS), som måske aldrig ville have udviklet sig. Omvendt var en større andel af de fundne invasive kræfttilfælde små (44,9 % mod 36,6 %), hvilket taler for tidligere opsporing. Studiets forfattere betegner selv resultaterne som foreløbige, og screeningsintervallet var længere i AI-perioden (median 993 dage mod 845 dage), hvilket i sig selv kan øge antallet af fund.

Dorthe Brogård Kristensen, professor, Institut for Virksomhedsledelse, Syddansk Universitet, påpeger i en kronik,[19] hvorfor AI i brystkræftscreening ikke er en entydigt positiv nyhed, men også rummer følgende etiske dilemmaer og udfordringer:

Fejl og kvalitet

- AI kan både finde flere kræfttilfælde og skabe nye typer fejl. Ifølge Kristensens interviews med radiologer kan AI fx reagere på ar, implantater og forkalkninger som tegn på kræft.
- Kristensen peger i sin kronik på, at der mangler generelle retningslinjer for radiologernes kvalitetskontrol af AI's vurderinger.

Overdiagnostik

- AI opdager flere forstadier og knuder, som måske aldrig ville have udviklet sig – med risiko for overdiagnostik og overbehandling. I studiet fra Region Hovedstaden faldt antallet af falske alarmer og genindkaldelser dog samtidig markant.

Ansvar

- Uklart hvem der har ansvaret, når både AI og radiolog tager fejl.

Arbejdsgange og kompetencer

- Radiologens rolle udvides med vurdering og kvalitetskontrol af AI-systemets resultater.
- Risiko for kompetencetab, mens nyt "usynligt" arbejde opstår.

Økonomi og implementering

- AI kan være dyrere per screening end manuel vurdering.[20]
- Kræver ansvarlig implementering uden politisk eller kommercielt pres.

Etiske udfordringer for brug af AI i sundhed

12

Der er både nationalt og internationalt store forhåbninger til, at kunstig intelligens kan forbedre sundhedsvæsenet. Samtidig har WHO fremhævet en række etiske risici og styringsmæssige udfordringer, som bør håndteres, når AI udvikles og anvendes på sundhedsområdet. Det gælder blandt andet hensynet til autonomi, sikkerhed, gennemsigtighed, ansvar og lighed.[21]

Det rejser spørgsmålet om, hvordan AI kan styrke sundhedsvæsenet, uden at utilsigtede virkninger samtidig svækker eksempelvis omsorg, faglighed, lighed eller patientsikkerhed. Denne problemstilling kommer også til udtryk i casen om AI i brystkræftscreening, hvor forbedringer i nogle resultater ledsages af nye spørgsmål om blandt andet overdiagnostik, arbejdsdeling og ansvar.

I det følgende fokuseres på tre problemstillinger, som står centralt i den etiske diskussion om anvendelsen af kunstig intelligens i sundhedsvæsenet: bias, forklarlighed og ansvar.

Bias

Udtrykket 'bias' stammer oprindelig fra fransk, hvor det betegnede en hældning eller noget skråt. I tidlige versioner af kuglespillet 'boules' blev det fx brugt som en teknisk betegnelse for de kugler, som vejede mere i den ene side, så de trillede skævt.[22] Ordet bruges i dag om en tendens, skævhed eller forudindtagethed i al almindelighed.

En af de mest diskuterede problemstillinger ved brug af kunstig intelligens i sundhedsvæsenet er bias. I 2017 viste forskere fra Stanford, at en algoritme kunne klassificere billeder af hudforandringer på niveau med dermatologer.[23] Studiet undersøgte imidlertid ikke, om algoritmens præstation varierede mellem forskellige hudfarver. Efterfølgende pegede andre forskere på en generel risiko. Hvis mennesker med mørk hud er utilstrækkeligt repræsenteret i træningsdata, kan en model præstere dårligere for denne patientgruppe. Det kan føre til senere eller fejlagtige diagnoser og dermed bidrage til ulighed i sundhed.[24]

For at vurdere, om algoritmisk bias er etisk problematisk, kan man især se på biasens karakter, årsager og virkninger i klinisk praksis.

Biasens karakter

Her spiller det både en rolle, hvor stor biasen er, og hvem den rammer. En mindre forskel i nøjagtighed vurderes ofte anderledes end en stor. Derudover kan bias mod grupper defineret ved fx race eller køn anses som mere problematisk end bias, der skyldes mangel på data om en meget sjælden sygdomsvariant. I Stanford-eksemplet vil nogle tillægge det særlig etisk vægt, at bias rammer mørkhudede.

Biasens årsager

Bias kan have flere oprindelser. Den kan være udtryk for en bevidst diskriminerende præference hos udviklerne. I Stanford-eksemplet vil det hypotetisk svare til at forskerne bevidst var biased mod mennesker med mørk hud og derfor udviklede algoritmen som de gjorde.

Oftere skyldes det dog ubevidste valg – fx brug af data, der er lettest tilgængelige, selv om de ikke er repræsentative. Endelig kan bias opstå af praktiske hensyn, hvor udviklere vælger at skabe en algoritme, selv om de ved, at tilgængelige data ikke er optimalt balancerede, fordi alternativet er ingen algoritme overhovedet.

Biasens virkninger i klinisk praksis

En algoritmes bias, målt på testdata, siger ikke i sig selv, hvordan biasen vil virke i klinisk praksis. Det afhænger af samspillet mellem systemet og det sundhedsper-

sonale, hvis beslutninger det skal understøtte. Undersøgelser tyder på, at klinikere ofte følger forslag fra beslutningsstøttesystemer – også når forslagene er vildledende – mens andre omvendt kan være overdrevent skeptiske, netop fordi forslaget kommer fra en maskine.[25] Spørgsmålet om biasens kliniske gennemslagskraft hænger derfor tæt sammen med automatiseringsbias, som behandles nærmere i afsnittet om ansvar.

Det diskuteres også, om kunstig intelligens overhovedet er mere biased end mennesker.

Menneskelig bias vs. algoritmisk bias

Det diskuteres også, om kunstig intelligens overhovedet er mere biased end mennesker. Et studie, der sammenlignede AI-chatbots svar på kliniske patientvignetter med tidligere publicerede data om klinikeres svar på samme vignetter, fandt, at både klinikere og AI-chatbots gav forskellige anbefalinger afhængigt af patientens køn, race/etnicitet og socioøkonomisk status – om end mønstrene ikke var identiske. Det peger på, at menneskelig klinisk praksis ikke nødvendigvis er mindre biased end en AI-understøttet praksis, men studiets forfattere understreger selv, at det er uklart, om AI-værktøjer forstærker eller modvirker klinikeres eksisterende bias.[26] Læger er jo også blot mennesker med alle de begrænsninger, som følger med, sådan som det fx illustreres af en undersøgelse, som tilsyneladende viser, at kvinder har større risiko end mænd for et dårligt resultat af kirurgiske indgreb, hvis kirurgen er en mand snarere end en kvinde.[27,28] Hvis man på en eller anden måde kan sammenligne den kunstige og naturlige, menneskelige intelligens' bias, rejser sådanne eksempler et spørgsmål om, hvorvidt der egentlig er nogen grund til at være kritisk over for eksistensen af bias hos en kunstig intelligens, hvis den 'blot' er lig med eller måske endda 'mindre' end biasen hos mennesker.

Der er dog også nogen, som har argumenteret for, at der er grund til at tro, at den endelige bias vil være mindre hos mennesker end hos kunstig intelligens. Det gælder fx Søren Holm og Thomas Ploug, som peger på, at der er en række formelle og uformelle bias-korrigerende mekanismer i klinisk praksis.[29] Teamwork, mulighed for kritik fra kollegaer, krav om samtykke, etisk uddannelse samt formelle klage- og autorisationssystemer – kan reducere effekten af menneskelig bias. Sådanne mekanismer findes ikke nødvendigvis i samme grad for AI-systemer.

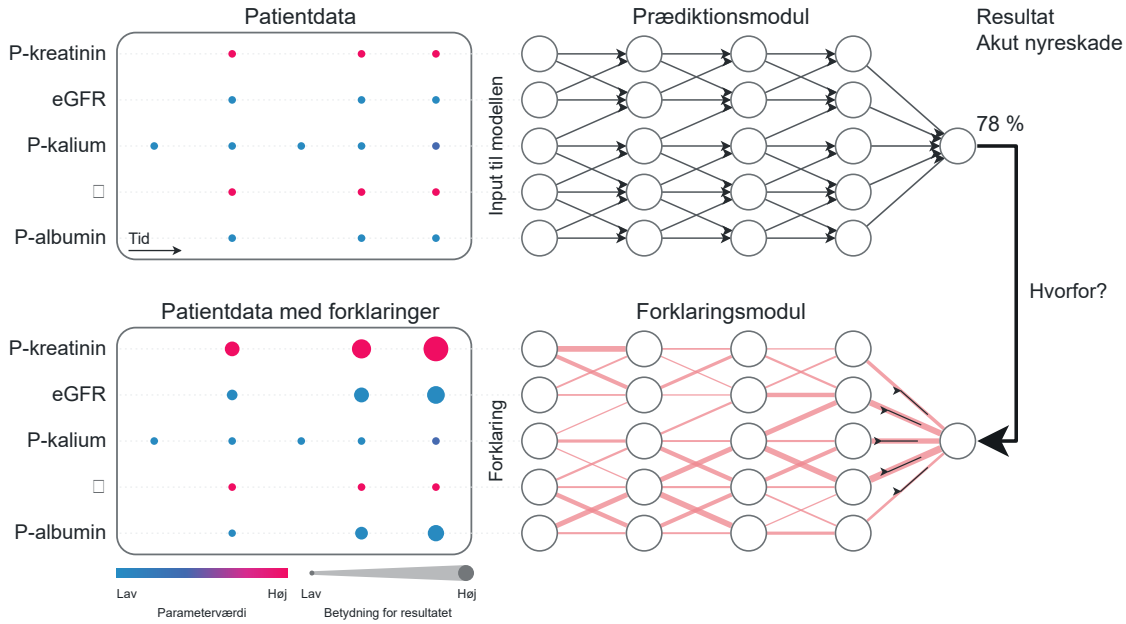
Opsummering

Vurderingen af bias ved brug af AI i sundhed afhænger bl.a. af biasens størrelse og retning, dens årsager samt de faktiske kliniske konsekvenser. Samtidig er bias ikke et problem, der er unikt for AI; det findes overalt, hvor data ikke repræsenterer alle grupper lige godt. Derfor er bias relevant at overveje på tværs af alle typer AI i sundhedsvæsenet – uanset om formålet er at diagnosticere hudkræft, psykiatriske lidelser eller at understøtte praktiserende læger i hverdagen.

Forklarlighed

Forklarlighed er sammen med bias et af de mest omdiskuterede etiske spørgsmål ved brug af kunstig intelligens i sundhedsvæsenet. Forklarlighed handler om, hvorvidt man kan forstå, hvordan et AI-system bevæger sig fra input (fx et scanningsbillede) til output (fx en diagnose). For moderne maskinlæring er dette særligt udfordrende, fordi algoritmerne justerer sig selv på baggrund af store datamængder og komplekse vægtninger, som mennesker typisk ikke kan gennemskue. Resultatet er såkaldte *black box-modeller*.

Denne uigennemsigtighed har både givet anledning til etiske bekymringer og til udviklingen af *forklarlig AI* (xAI). I evalueringen af AI-signaturprojekterne beskrives en løsning på udfordringen således: ”Transparens og forklarlighed bør bygges ind i løsningen, og det skal være muligt at gå tilbage i tid og genskabe grundlaget for, at en AI-model opførte sig som den gjorde.”[30]



Figur 1. Oversigt over xAI-EWS-systemet. *Data fra den enkelte patients elektroniske patientjournal bruges som input i TCN-prædiktionsmodulet. På baggrund af disse data estimerer modellen risikoen for akut nyreskade (AKI), her illustreret med 78 %. DTD-forklaringsmodulet viser derefter, hvilken betydning de enkelte inputvariable har for modellens risikovurdering.*

P: plasma; eGFR: estimeret glomerulær filtrationsrate; DTD: dyb Taylor-dekomposition (deep Taylor decomposition); TCN: temporalt konvolutionsnetværk (temporal convolutional network); xAI-EWS: system til tidlig varsling baseret på forklarlig kunstig intelligens.

Kilde: *Rentegnet og oversat til dansk efter figur 1 i Lauritsen et al. (2020), Explainable artificial intelligence model to predict acute critical illness from electronic health records, Nature Communications, 11, artikel 3852. Originalen er udgivet under CC BY 4.0.*

Nye teknikker gør det muligt.[31] En AI-løsning kan designes så den viser, hvilke faktorer der påvirker algoritmens vurdering mest. Dermed kan det sundhedsfaglige personale forstå, hvad en model bygger på, og bedre vurdere, om resultatet giver klinisk mening. Grundideen kan illustreres med et diagram fra en artikel, som beskriver, hvordan danske forskere har udviklet en form for forklarlig kunstig intelligens, som kan være med til at forudsige akut kritisk sygdom som fx nyresvigt eller blodforgiftning (sepsis) på baggrund af data fra en elektronisk patientjournal.[32]

Som det fremgår af den øverste del af diagrammet på forrige side, udgøres algoritmens input af patientdata i form af målinger af en lang række faktorer ('P-creatinine', 'Kidney eGFR', 'P-potassium', 'P-albumin' osv.) over tid. Efter at være blevet behandlet af algoritmen ('prediction module') konkluderer systemet så, at der er 78 % risiko for akut nyresvigt. Det, som gør systemet forklarligt eller gennemskueligt, er nu, at det ikke nøjes med at præsentere dette output som en rå kendsgerning, man enten må tro på eller lade være med at tro på, men også forklarer eller giver et bud på, hvor stor en betydning, de forskellige inputfaktorer har haft for, at algoritmen er kommet frem til det pågældende resultat. I det illustrerede tilfælde kan man fx se, at P-creatinin har haft stor betydning for resultatet, og tanken er så, at en sådan forklaring vil kunne øge lægens tillid til systemet, bl.a. fordi det gør det lettere for lægen at gøre brug af sin egen viden til at vurdere, hvorvidt forudsigelsen giver klinisk mening.

Er forklarlighed vigtigt? Argumenter for og imod

Udviklingen af forklarlige former for maskinlæring er for en stor del drevet af en bekymring om, hvorvidt det vil være etisk forsvarligt at anvende maskinlæringssystemer i fx sundhedsvæsenet, hvis man ikke kan forklare, hvordan de kommer frem til deres resultater. Men er det overhovedet så vigtigt at kunne forklare det? Og hvis det er, hvad er så egentlig grunden til det? I dette afsnit beskrives to forskellige perspektiver på disse centrale etiske spørgsmål med afsæt i to konkrete tekster.[33]

Den første tekst, som *nedtoner* betydningen af forklarlighed, er "Artificial Intelligence and Black-Box Medical Decisions: Accuracy versus Explainability",[34] der er udgivet af det amerikanske The Hastings Center og skrevet af professor Alex John London, som blandt andet er tilknyttet University of Pittsburgh's Center for Bioethics and Health Law. Den anden tekst, som argumenterer for betydningen af forklarlighed, har titlen "Artificial intelligence and patient-centered decision making",[35] og er skrevet af de to danske filosoffer Jens Christian Bjerring og Jacob Busch fra Aarhus Universitet.

Et argument imod forklarlighed som alment krav

Ifølge Alex John London bør der ikke stilles et universelt krav om forklarlighed. Hans hovedpointer er:

1. Klinisk effekt bør veje tungere end teoretisk forståelse. Hvis en uigennemsigtig model kan opdage kræft tidligere, er det vigtigere end at kunne forklare dens interne logik.
2. Moderne medicin anvender allerede ateoretisk viden. Midler som aspirin og litium blev brugt i årtier uden kendt virkningsmekanisme.
3. For stor tillid til teorier kan være farlig. Historiske eksempler som radikal mastektomi og højdosis kemoterapi blev drevet af plausible teorier, der senere viste sig skadelige.
4. Klinikere arbejder allerede med uigennemsigtige processer. Lægers intuitive skøn bygger også på uigennemsigtige og associative erfaringer.

London afviser også, at forklarlighed er en nødvendig forudsætning for at opdage bias. I hans eksempel med en algoritme, der giver astmatikere lavere dødsrisiko ved lungebetændelse, viser analysen, at resultatet afspejler eksisterende behandlingstraditioner. Modellen havde ikke blot lavet en beregningsfejl, men lært en statistisk sammenhæng, som skyldtes behandlingspraksis. Sammenhængen kunne derfor blive klinisk vildledende, hvis den blev fortolket som udtryk for, at astma i sig selv sænkede risikoen. For London er det derfor vigtigere at sikre korrekt datagrundlag, empirisk validering og kontekstspecifik anvendelse, end at insistere på forklarlighed.

Et argument for forklarlighed

Modsat London argumenterer de to danske filosoffer Jens Christian Bjerring og Jacob Busch for, at forklarlighed er vigtigt, i deres artikel ”Artificial intelligence and patient-centered decision making”. I artiklen antager Bjerring og Busch for argumentets skyld, at det er bevist, at visse former for kunstig intelligens er mere pålidelige og præcise end mennesker, og diskuterer på den baggrund, om det i en sådan situation alligevel vil være problematisk at anvende kunstig intelligens, hvis man ikke kan forklare, hvorfor en bestemt diagnose eller behandling foreslås.

På den ene side argumenterer de for, at sundhedspersonalet vil have en umiddelbar videns- eller erkendelsesmæssig pligt til at lytte til den kunstige intelligens, hvis det virkelig er bevist, at den er mere pålidelig og præcis end de fremmeste eksperter.

Som de ser det, svarer det egentlig blot til, at en hvilken som helst læge har pligt til at lytte til de fremmeste eksperter inden for et givet felt, hvis lægen selv er i tvivl eller usikker.

På den anden side argumenterer de imidlertid også for, at brugen af uigennemsigtige former for kunstig intelligens sandsynligvis vil komme i konflikt med det de kalder patientcentreret medicin. Ved patientcentreret medicin forstår de en sundhedsfaglig praksis, som er styret af to idealer:

1. For det første bør patienter betragtes som personer, hvis værdier, overbevisninger og psykosociale kontekst har stor betydning for deres sundhed.
2. For det andet bør patienter betragtes som ligeværdige parter i den sundhedsmæssige beslutningsproces i den forstand, at deres ønsker, behov og viden bør lyttes til og inddrages.

Selv hvis AI er mere præcis end læger, mener de, at uigennemsigtige modeller truer det andet ideal om kommunikation, forklaring og informeret samtykke.

Selv hvis AI er mere præcis end læger, mener de, at uigennemsigtige modeller truer det andet ideal om kommunikation, forklaring og informeret samtykke. Manglende forklarlighed er derfor ikke kun en teknisk udfordring; den er også en relationel udfordring. Når klinikere ikke kan forklare, hvorfor en anbefaling ser ud, som den gør, risikerer det at svække dialogen, patientens oplevelse af at blive taget alvorligt, og den professionelles mulighed for at udøve omsorg gennem forklaring og involvering, hvilket i sidste ende kan udfordre tilliden mellem det sundhedsfaglige personale og patienten.

Hvis lægen fx ikke kan forklare, hvorfor en algoritme giver Marie 92 % risiko for aggressiv brystkræft, eller hvorfor hendes foretrukne behandling vurderes som mindre succesfuld, bliver det sværere for patienten at træffe et oplyst valg.

De påpeger desuden:

1. At almindelig uigennemsigtighed i medicin (fx ukendt mekanisme bag et lægemiddel) ofte kan kompenseres med viden om studier, data og testpopulationer. Dette er langt sværere med komplekse AI systemer.
2. At forklarlighed styrker tilliden både mellem patient og system og mellem patient og læge. Manglende forklarlighed kan gøre patientens tillid til lægen endnu mere afgørende – og mere sårbar.

Afslutningsvis skal der kort anføres et par ekstra argumenter fra en anden tekst, som supplerer Bjerring og Buschs overvejelser. I artiklen "Explainability for artificial intelligence in health care: a multidisciplinary perspective"[36] gennemgår forfatterne de udfordringer, som manglen på forklarlighed kan give anledning til, fra en række perspektiver. Et rent teknisk, et juridisk, et medicinsk og et patientperspektiv. Og her synes særligt deres overvejelser fra et medicinsk perspektiv at være relevante. Som de ser det, er det især i forbindelse med uenigheder mellem en kunstig intelligens og en kliniker, at forklarlighed er vigtigt. Hvis lægen kender de faktorer, AI har lagt vægt på, kan lægen bedre vurdere, om anbefalingen giver mening i den konkrete kliniske situation. Forklarlighed kan dermed øge både klinisk kvalitet og faglig tillid.

Omsorg og relation

Sundhedsvæsenets opgave er ikke kun at diagnosticere og behandle, men også at drage omsorg. Omsorg udøves i høj grad gennem relationen mellem patient og sundhedsprofessionel – gennem nærvær, forklaring og fælles beslutningstagen. Det rejser spørgsmålet om, hvad der sker med omsorgen, når flere beslutninger understøttes af systemer, som de sundhedsprofessionelle ikke fuldt ud kan forstå, forklare eller påvirke. Oplever klinikere, at de hverken kan gennemskue eller påvirke de værktøjer, de arbejder med, kan det udfordre deres faglige identitet og deres relation til patienterne – i værste fald i retning af større afkobling og oplevet ansvarsløshed.

Omsorg udøves i høj grad gennem relationen mellem patient og sundhedsprofessionel – gennem nærvær, forklaring og fælles beslutningstagen.

Omvendt kan teknologien, hvis den frigør tid fra rutineopgaver, potentielt skabe mere rum til netop nærvær og samtale. Om det ene eller det andet bliver resultatet, afgøres ikke af teknologien selv, men af den måde, den implementeres på. Det Etiske Råd har tidligere behandlet omsorg i sundhedsvæsenet mere indgående i en selvstændig redegørelse (2022),[37] som denne problemstilling knytter an til. Men det rejser også spørgsmål om, hvad der sker med omsorgen når selv små og umiddelbart ubetydelige omsorgs- eller andre opgaver, som tidligere blev varetaget af mennesker bliver overdraget til AI-systemer som fx chatbots baseret på generativ AI.

Ansvar

En tredje problemstilling, som ofte diskuteres i relation til kunstig intelligens, angår ansvar. Ansvar er centralt i relationer mellem mennesker. Den ansvarlige er den, andre kan rose, belønne eller kritisere, hvis noget går bedre eller dårligere end forventet. Ansvar kan udøves både i nære relationer – hvor det typisk kommer til udtryk som ros eller bebrejdelse – og i formelle systemer som retssystemet, hvor ansvar placeres og sanktioneres på institutionaliseret vis.

Et formål med at tilskrive ansvar er at muliggøre samarbejde. Roller i samfundet fungerer, fordi bestemte personer forbindes med bestemte opgaver. Uden denne tillid og klare rolleforventninger ville organisering være langt mere besværlig. Det er dog et åbent spørgsmål, hvad der bestemmer den type ansvar, der følger en rolle. Er det noget naturgivent – som faderansvar – eller noget socialt tildelt, som afhænger af normer og forventninger? Og kan to personer i samme rolle have forskelligt ansvar afhængigt af deres evner?

Evner, fri vilje og ansvar

Evner spiller en vigtig rolle i mange former for ansvarstilskrivelse. Spædbørn holdes fx ikke ansvarlige for handlinger, der skader eller forstyrrer andre. Dette forklares enten med manglende fri vilje eller med, at ansvar kun giver mening, hvis individet kan ændre adfærd som reaktion på kritik eller sanktion. Ansvar kan altså forstås ud fra enten handlingens forårsagelse (fri vilje) eller dens fremtidige konsekvenser (om sanktioner kan påvirke adfærd). Begge perspektiver indgår i almindelig ansvarstænkning.

Oplevelsen af ansvar

Ansvar kan opleves meget forskelligt. For nogle er det en byrde, for andre naturligt. Usikkert ansvar kan være særligt belastende, fordi man ikke ved, hvad man vil blive holdt ansvarlig for – eller føler sig ansvarlig for mere, end ens rolle egentlig kræver. Uklar ansvarsfordeling kan også føre til uhensigtsmæssig adfærd, fordi man bruger energi på at undgå at blive skyldigtgjort for noget, man ikke ved, om man hæfter for.

Kultur og professionel identitet

Ansvar er tæt knyttet til den selvforståelse, som sundhedsprofessionelle har af deres rolle. Hvis man i stigende grad skal stå på mål for anbefalinger, man ikke kan forklare, kan det fremkalde usikkerhed, oplevet ansvarsløshed eller distancering fra kerneopgaven. Det kan i sidste ende udfordre omsorgens kvalitet, fordi nærvær, forklaring og begrundelse traditionelt er en del af omsorgens udtryk i klinikken.

Ansvar og kunstig intelligens i sundhedsvæsenet

21

I sundhedsvæsenet bliver spørgsmålet centralt. Hvem er ansvarlig for en given AI løsning – især hvis den fejler? Ansvaret hviler ikke alene på klinikeren og institutionen, men også på dem, der udvikler, markedsfører og implementerer løsningerne – et forhold, som både WHO og EU's AI-forordning fremhæver, men som let bliver usynligt i den daglige klinik.[38] Det drejer sig om, hvorvidt eksisterende juridiske og professionelle ansvarsmodeller kan håndtere AI, eller om der opstår huller, når nye teknologier indføres.

Et vigtigt aspekt er AI'ens grad af selvstændighed

AI som beslutningsstøtte vs. selvstændig beslutningstager

- Når AI anvendes som beslutningsstøtte, træffer sundhedspersonen fortsat den konkrete kliniske beslutning. Hvordan det juridiske ansvar fordeles, afhænger imidlertid af gældende sundheds-, produkt- og erstatningsret og kan ikke udledes af AI-forordningen alene. Men hvis AI i praksis er mere pålidelig end lægen, kan det udfordre denne ansvarsfordeling.
- Uanset hvad, vil der altid være mennesker (udviklere, producenter, institutioner), som bærer et bagvedliggende ansvar for systemet.

Automatiseringsbias

I praksis kan grænsen mellem støtte og beslutning udviskes. Sundhedspersonale kan komme til at følge AI anbefalinger ukritisk, fordi de tvivler på egen dømmekraft – særligt hvis ansvarsplaceringen er uklar. Dette kan betyde, at AI reelt fungerer som beslutningstager, selvom den formelt blot er rådgiver. Danske feltstudier på radiologiske afdelinger viser, at netop dette – uklarhed om, hvordan man gennemskuer fejl, og hvordan ansvaret placeres – er blandt de spørgsmål, klinikere selv rejser, når AI indføres i deres arbejde.[39] Omvendt kan nogle være overdrevent skeptiske, netop fordi anbefalingen kommer fra en computer.

AI, der lærer efter implementering

I dansk klinisk praksis anvendes i dag overvejende såkaldte *låste AI modeller*, hvor algoritmen ikke ændrer sig i drift, men kun opdateres gennem kontrollerede versioner og fornyet validering.

Danske feltstudier på radiologiske afdelinger viser, at netop dette – uklarhed om, hvordan man gennemskuer fejl, og hvordan ansvaret placeres – er blandt de spørgsmål, klinikere selv rejser, når AI indføres i deres arbejde.

Nogle AI-systemer kan imidlertid i princippet opdatere deres modeller løbende baseret på nye data. Diskussionen af sådanne systemer vedrører derfor primært fremadrettede og principielle spørgsmål om ansvar, governance og patientsikkerhed.

Dette rejser spørgsmål om:

- hvor ofte systemet skal testes og valideres,
- hvem der er ansvarlig for ændringer i modellen,
- og hvornår det er forsvarligt at lade AI lære i drift, især hvis fejl kan få alvorlige konsekvenser.

Klinisk case: Brystkræftscreening som ansvarslaboratorium

Vi kan illustrere ovenstående ansvarsdilemmaer i praksis ved at se på casen om AI i brystkræftscreening:

1. Rolleforskydning og arbejdsdeling. Som Kristensen peger på, kan radiologens rolle forskydes, så vurdering og kontrol af AI-systemets resultater bliver en større del af arbejdet. Ansvar flyttes, men rammer, retningslinjer og kompetenceopbygning følger ikke nødvendigvis med. Det skaber den belastende uklarhed, som ansvarsteorien advarer imod: Hvem hæfter, hvis både AI og kliniker tager fejl – og hvad er det tilstrækkelige faglige grundlag for at fravige AI?
2. Evner og ansvar hænger sammen: Radiologerne, der validerer AI's vurderinger, forventes at kunne gennemskue og korrigere systemets fejl uden konsoliderede standarder for den kontrol. Når evner og forventet ansvar ikke matcher, bliver ansvarstilskrivelsen etisk problematisk.
3. Automatiseringsbias i praksis: Når AI identificerer mønstre eller knuder, som mennesker ikke har set før, kan presset for at følge AI øges. Beslutningsstøtte risikerer at blive beslutningsoverdragelse, så AI de facto bliver medbeslutningstager – uden tilsvarende tydelig ansvarsplacering.
4. Dynamisk ansvar ved implementering: Forskellen mellem foreløbige før-efter-resultater og dokumentation af langsigtede kliniske effekter viser, at ansvar må forstås løbende. Kontinuerlig validering, revision, monitorering af overdiagnostik/overbehandling og klare procedurer for, hvordan man dokumenteret kan gå imod AI anbefalinger.

Implikation: For at gøre ansvar reelt og bæredygtigt i klinikken kræves klare rolledefinitioner, eksplicit ansvarsfordeling mellem producent, hospital og kliniker, målrettet uddannelse (særligt i kvalitetskontrol af AI), samt styringsrammer for iterative modelændringer. Uden dette risikerer man, at ansvar enten koncentreres urimeligt hos den enkelte radiolog eller forsvinder i gråzoner mellem system og menneske.

Øvrige udfordringer

Ud over de tre hovedproblemstillinger er der en række andre udfordringer, som fortjener opmærksomhed, når kunstig intelligens tages i brug i sundhedsvæsenet. Enkelte af dem nævnes her kort:

Deskilling

Et gennemgående tema i teknologietikken er risikoen for afkvalificering. Spørgsmålet er sjældent, om deskilling vil finde sted, når en teknologi overtager en opgave — det gør den stort set altid, i en eller anden form. Det interessante etiske spørgsmål er, hvor det sker, hvordan det sker, og i hvilket tempo — og om der er tid nok til, at nye kompetencer kan nå at erstatte de gamle.

Når teknologi overtager opgaver, hvor færdigheder opbygges og vedligeholdes – fx granskning af billeder eller udarbejdelse af kliniske retningslinjer – risikerer man, at de professionelle mister eller aldrig opbygger netop de kompetencer, der skal til for at kontrollere teknologien og træde til, når den fejler. Udfordringen er størst for yngre klinikere, som ikke når at opbygge erfaringen, før teknologien tages i brug.

Når teknologi overtager opgaver, hvor færdigheder opbygges og vedligeholdes – fx granskning af billeder eller udarbejdelse af kliniske retningslinjer – risikerer man, at de professionelle mister eller aldrig opbygger netop de kompetencer, der skal til for at kontrollere teknologien og træde til, når den fejler.

Lighed i sundhed

Kunstig intelligens kan både mindske og forstærke ulighed i sundhed. Bias-problematikken er den mest direkte forbindelse, men også adgangen til teknologien og evnen til at bruge den kan fordele sig skævt. Det Etiske Råd arbejder parallelt med lighed i sundhed i et selvstændigt projekt.

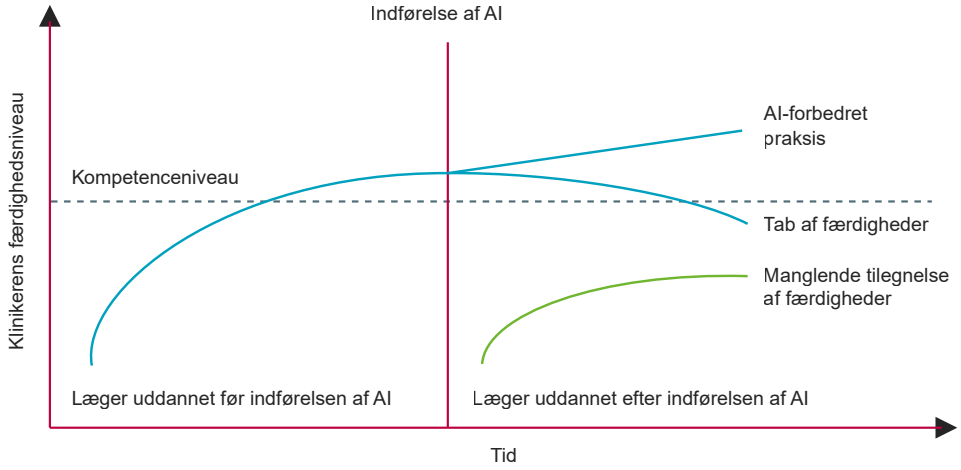
Borgerrettet AI uden for sundhedsvæsenets rammer

Wearables, sundhedsapps og sprogmodeller fungerer i praksis som screenings- og rådgivningsværktøjer uden for det regulerede sundhedsvæsen. Det rejser spørgsmål om kvalitet og ansvar – og om, hvordan sundhedsvæsenet skal forholde sig til de forventninger og bekymringer, borgerne møder op med.

Økonomi og prioritering

AI-løsninger er ikke gratis, og gevinsterne skal dokumenteres. Der er risiko for, at investeringsbeslutninger drives af politisk og kommercielt pres snarere end af evidens – og for at teknologien indføres på områder, hvor behovet ikke er størst.

Når færdigheder forsvinder – deskilling



Rentegnet og oversat til dansk efter en illustration i Tyler M. Berzin og Eric J. Topols artikel »Preserving clinical skills in the age of AI assistance«, The Lancet.

Når kunstig intelligens overtager opgaver, hvor faglige færdigheder opbygges og vedligeholdes – som granskning af røntgenbilleder – risikerer man, at de sundhedsprofessionelle mister eller aldrig opbygger netop de kompetencer, der skal til for at gennemskue teknologiens fejl og træde til, når den svigter.

Udviklingen går dog ikke kun én vej. Figuren herover illustrerer tre mulige forløb. AI kan understøtte klinikernes arbejde og forbedre deres praksis (»AI-forbedret praksis«). Samtidig kan tidligere indlærte færdigheder svækkes, hvis de ikke længere bruges og vedligeholdes (»tab af færdigheder«). For læger, der uddannes med AI fra starten, er en mulig risiko, at visse færdigheder aldrig bliver fuldt udviklet (»manglende tilegnelse af færdigheder«). Forløbene kan forekomme side om side og vil blandt andet afhænge af, hvordan AI bruges i uddannelsen og den kliniske praksis.[40]

Forskningen er endnu sparsom, men ét studie peger på et muligt tab af færdigheder. I et observationsstudie fra fire polske centre sammenlignede forskerne kikkertundersøgelser af tyktarmen udført uden AI før og efter, at centrene var begyndt at bruge AI regelmæssigt. Andelen af undersøgelser, hvor lægerne fandt adenomer – en vævsforandring, som kan være et forstadie til tarmkræft – faldt fra 28,4 til 22,4 procent. Forfatterne ser det som et muligt tegn på tab af færdigheder, men studiet kan ikke afgøre, om AI var årsagen.[41]

Deskilling er dermed ikke kun et spørgsmål om patientsikkerhed, men også om fagidentitet og arbejdsmiljø. Det taler for, at uddannelse i kritisk kvalitetskontrol af AI må følge med teknologien – ikke som et tillæg, men som en forudsætning.

Fremtiden for AI i sundhed

Kunstig intelligens rummer betydelige potentialer for bedre kvalitet, mere præcis diagnostik, mere sammenhængende forløb og bedre udnyttelse af sundhedsvæsenets ressourcer. Samtidig viser både danske erfaringer og internationale studier, at teknologien også kan skabe nye former for bias, gøre beslutningsgrundlaget mindre gennemsigtigt og forskyde ansvar på måder, der udfordrer både faglighed, tillid og patientsikkerhed.

De tre etiske kerneproblemstillinger – bias, forklarlighed og ansvar – er derfor ikke separate hensyn, men tæt forbundne. Bias kan underminere retfærdighed og faglig kvalitet; mangel på forklarlighed kan svække patientens mulighed for at træffe oplyste beslutninger og klinikerens mulighed for at udøve fagligt skøn; og uklar ansvarsfordeling kan skabe utryghed, hæmme læring og i sidste ende gå ud over patienterne.

De tre etiske kerneproblemstillinger – bias, forklarlighed og ansvar – er derfor ikke separate hensyn, men tæt forbundne.

Mammografieksemplet rejser spørgsmålet, om AI kan betragtes som et neutralt redskab, eller om teknologien samtidig ændrer faglige roller, beslutningsprocesser og arbejdsgange. Derfor kræver en ansvarlig anvendelse af AI, at implementeringen sker med omtanke. Der skal være klar ansvarsfordeling, løbende validering, fokus på kompetencer, gennemsigtighed hvor det er muligt, og systematiske vurderinger af både gavnlige og utilsigtede effekter.

AI i sundhedsvæsenet skal ikke erstatte faglighed og menneskelig vurdering, men udvide og kvalificere dem.

AI i sundhedsvæsenet skal ikke erstatte faglighed og menneskelig vurdering, men udvide og kvalificere dem. Når teknologien udvikles og anvendes med respekt for grundlæggende værdier som kvalitet, lighed, faglig dømmekraft og patientens selvbestemmelse, kan AI blive et værdifuldt redskab i et moderne sundhedsvæsen. Men uden klare etiske rammer risikerer vi, at teknologien svækker netop de værdier, den skulle styrke. Det er derfor afgørende, at udviklingen af AI i sundhedsvæsenet følges af en tilsvarende udvikling i etisk refleksion, regulering og professionel stillingtagen.

Opsummering

På baggrund af ovenstående tegner der sig følgende centrale temaer i udvikling, indkøb og anvendelse af kunstig intelligens i sundhedsvæsenet:

1. **Bias og ulighed.** Fungerer AI-løsninger lige godt på tværs af patientgrupper, og hvordan kan skævheder i data, udvikling og anvendelse identificeres? Det handler også om, hvorvidt teknologien kan forstærke eller mindske eksisterende uligheder i sundhed.
2. **Etiske overvejelser ved implementering.** Indførelsen af AI rejser spørgsmål om konsekvenserne for omsorg, faglighed og lighed, og hvordan disse hensyn afvejes – både i klinisk praksis og i sundhedsvæsenets øvrige opgaver. Dansk forskning i vurdering af AI-løsningers værdi fremhæver også betydningen af at vælge løsninger til og fra på et oplyst grundlag.[42]
3. **Validering i klinisk praksis.** AI-løsningernes funktion og konsekvenser, herunder overdiagnostik og andre utilsigtede effekter, kan ikke vurderes alene i laboratoriet. Det rejser spørgsmål om, hvordan løsningerne følges over tid i klinisk praksis, og hvilkeklinisk relevante endemål de evalueres på.
4. **Ansvar og beslutningskompetence.** Anvendelsen af AI rejser spørgsmål om, hvordan ansvarsfordelingen mellem producent, institution og kliniker afklares før ibrugtagning. Hertil kommer klinikeres mulighed for at fravige en AI-anbefaling, og hvordan sådanne beslutninger dokumenteres og anerkendes som legitime.
5. **Faglige kompetencer.** AI kan påvirke sundhedsprofessionelles færdigheder over tid. Det rejser spørgsmål om, hvordan kompetencer vedligeholdes og udvikles, og hvilken rolle uddannelse i kritisk vurdering af AI spiller.
6. **Gennemsigtighed og forklarlighed.** Spørgsmålet er, i hvilken grad AI-resultater kan forklares, og om forklaringerne hjælper klinikere og patienter med at vurdere, om et resultat giver mening i den konkrete kliniske situation.

Forslag til videre læsning: WHO's *Ethics and Governance of Artificial Intelligence for Health* (2021), WHO's vejledning om store multimodale modeller i sundhedsvæsenet (2024) og UNESCO's *Recommendation on the Ethics of Artificial Intelligence* (2021).

Noter

1 https://digst.dk/media/sgjl0dd0/evalueringen-af-signaturprojekterne-2024.pdf - side 2

2 https://www.em.dk/Media/638660486816578445/Strategi%20for%20life%20science_2024.10.31_WEB.pdf

3 Link til regeringsgrundlag – Side 59-63 https://stm.dk/statsministeriet/publikationer/det-politiske-grundlag-for-firkloeverregeringen/

4 https://hai.stanford.edu/assets/files/ai_index_report_2026.pdf - Stanford AI Index Report 2026 – side 255

5 UNESCO (2021): Recommendation on the Ethics of Artificial Intelligence, herunder politikområdet om sundhed og social trivsel https://unesdoc.unesco.org/ark:/48223/pf0000380455; WHO (2021): Ethics and Governance of Artificial Intelligence for Health: WHO Guidance. Genève: World Health Organization.https://www.who.int/publications/i/item/9789240029200

6 https://cdn.openai.com/pdf/a253471f-8260-40c6-a2cc-aa93fe9f142e/economic-research-chatgpt-usage-paper.pdf

7 Kunstig intelligens i sundhed – eksempler på danske AI-løsninger, FADL's Forlag. Side 16 – Kapitel 1: Kunstig intelligens i sundhed, Andreas Pihl, Aryan Dargahi og Thiusius R. Savarimuthu

8 Kunstig intelligens i sundhed – eksempler på danske AI-løsninger, FADL's Forlag. Side 27 – Kapitel 2: Fremtidens sundhedsvæsen kræver digital transformation og kunstig intelligens, Jannik Zeuthen, Klaus Larsen og Mickael Bech

9 Kunstig intelligens i sundhed – eksempler på danske AI-løsninger, FADL's Forlag. Side 28 – Kapitel 2: Fremtidens sundhedsvæsen kræver digital transformation og kunstig intelligens, Jannik Zeuthen, Klaus Larsen og Mickael Bech

10 Kunstig intelligens i sundhed – eksempler på danske AI-løsninger, FADL's Forlag. Side 28 – Kapitel 2: Fremtidens sundhedsvæsen kræver digital transformation og kunstig intelligens, Jannik Zeuthen, Klaus Larsen og Mickael Bech

11 https://rn.dk/-/media/Rn_dk/Nyheder/2025/AI-katalog-januar-2025.pdf - side 4

12 https://hai.stanford.edu/assets/files/ai_index_report_2026.pdf - Stanford AI Index Report 2026 s. 279-281.

13 https://academic.oup.com/eurpub/article/35/6/1092/8323093

14 https://politiken.dk/debat/debatindlaeg/art10481419/Jeg-forstod-f%C3%B8rst-min-kr%C3%A6ftdiagnose-da-Chat-GPT-fik-adgang-til-mit-pr%C3%B8vesvar

15 https://openai.com/da-DK/index/introducing-chatgpt-health/

16 https://digst.dk/media/sgjl0dd0/evalueringen-af-signaturprojekterne-2024.pdf - se case på side 21 fra Sønderborg Kommune - Hurtigere sagsbehandling af kropsbårne hjælpemidler

17 https://digst.dk/media/sgjl0dd0/evalueringen-af-signaturprojekterne-2024.pdf side 62

18 https://pubs.rsna.org/doi/10.1148/radiol.232479

19 https://www.altinget.dk/digital/artikel/professor-brystkraeftscreening-med-ai-giver-hurtigere-resultater-men-det-oeger-risikoen-for-fejl

20 AI-løsningen koster ifølge Kristensen 24 kr. pr. screeningsbillede mod 16 kr. for en vurdering udført af en læge alene – en ekstraudgift, der skal vejes op mod den sparede arbejdstid.

21 World Health Organization (WHO) (2021): Ethics and Governance of Artificial Intelligence for Health: WHO Guidance. Genève: WHO, særligt kapitel 5 om etiske principper for AI i sundhed. https://www.who.int/publications/i/item/9789240029200

22 https://www.etymonline.com/search?q=bias

23 Nature: "Dermatologist-level classification of skin cancer with deep neural networks" https://www.nature.com/articles/nature21056.epdf?author_access_token=8oxIcYWf5UNrNpHsUHd2StRgN0jAjWel9jnR3ZoTv0NXpMHRAJy8Qn10ys2O4tuPakXos4UhQAFZ750CsBNMMsISFHIKinKDMKjShCpHIlYPYUHhNzkn6pSnOCt0Ftf6

24 Adamson et al. (2018): "Machine Learning and Health Care Disparities in Dermatology" https://jamanetwork.com/journals/jamadermatology/article-abstract/2688587

25 Gaube, S. m.fl. (2021): "Do as AI say: susceptibility in deployment of clinical decision-aids", npj Digital Medicine 4, 31

26 Kim, J., Cai, Z.R., Chen, M.L. et al. (2023): "Assessing Biases in Medical Decisions via Clinician and AI Chatbot Responses to Patient Vignettes" - https://jamanetwork.com/journals/jamanetworkopen/fullarticle/2810775

27 Wallis et al. (2021): "Association of Surgeon-Patient Sex Concordance With Postoperative Outcomes", JAMA Surgery.

28 https://www.theguardian.com/society/2022/jan/04/women-more-likely-die-operation-male-surgeon-study

29 Holm & Ploug (2020): "The right to refuse diagnostics and treatment planning by artificial intelligence", Medicine, Health Care and Philosophy.

30 https://digst.dk/media/sgjl0dd0/evalueringen-af-signaturprojekterne-2024.pdf - side 46

31 Hansen, Lars Kai et al (ed.) (2019): Explainable AI: Interpreting, Explaining and Visualising Deep Learning.

32 Lauritsen et al. (2020): 'Explainable artificial intelligence model to predict acute critical illness from electronic health records', Nature Communications 11, 3852.

33 WHO (2021): Ethics and Governance of Artificial Intelligence for Health, Genève. Opstiller seks etiske principper for AI i sundhed: beskyttelse af autonomi; menneskelig trivsel og sikkerhed; gennemsigtighed og forklarlighed; ansvar og ansvarlighed; inklusion og lighed; samt responsive og bæredygtige systemer. Suppleret i 2024 med særskilt vejledning om generativ AI (store multimodale modeller).

34 https://pubmed.ncbi.nlm.nih.gov/30790315/

35 https://link.springer.com/article/10.1007%2Fs13347-019-00391-6

36 https://bmcmedinformdecismak.biomedcentral.com/articles/10.1186/s12911-020-01332-6

37 https://www.etiskraad.dk/udgivelser/indretning-af-sundhedsvaesenet/omsorg-i-sundhedsvaesenet

38 Europa-Parlamentets og Rådets forordning (EU) 2024/1689 (AI-forordningen), art. 16 (forpligtelser for udbydere af højrisiko-AI-systemer), art. 26 (forpligtelser for idriftsættere) og art. 6, stk. 1, jf. bilag I (klassificering af AI i medicinsk udstyr som højrisiko). https://eur-lex.europa.eu/eli/reg/2024/1689/2026-07-27/eng

39 Boesen, M. m.fl. (2025): "Radiologisk AI-testcenter (RAIT)", i A. Pihl, A. Dargahi & T.R. Savarimuthu (red.): Kunstig intelligens i sundhed, FADL's Forlag, kap. 12. Bygger på etnografiske feltstudier på danske radiologiske afdelinger (jf. Galsgaard et al., 2023).

40 https://www.thelancet.com/journals/lancet/article/PIIS0140-6736(25)02075-6/abstract

41 Budzyń, K. et al. (2025): "Endoscopist deskilling risk after exposure to artificial intelligence in colonoscopy: a multicentre, observational study", The Lancet Gastroenterology & Hepatology, 10(10), 896–903. https://doi.org/10.1016/S2468-1253(25)00133-5

42 Fasterholdt, I., Kidholm, K. & Gerdes, A. (2025): "MAS-AI – vurdering af værdien af AI-løsninger", i Kunstig intelligens i sundhed, FADL's Forlag, kap. 17.