
Kritisk veiskille for integrering av kunstig intelligens

KRONIKK

DAMOUN NASSEHI

damoun.nassehi@uib.no

Damoun Nassehi er fastlege og førsteamanuensis ved Institutt for global helse og samfunnsmedisin, Universitetet i Bergen.

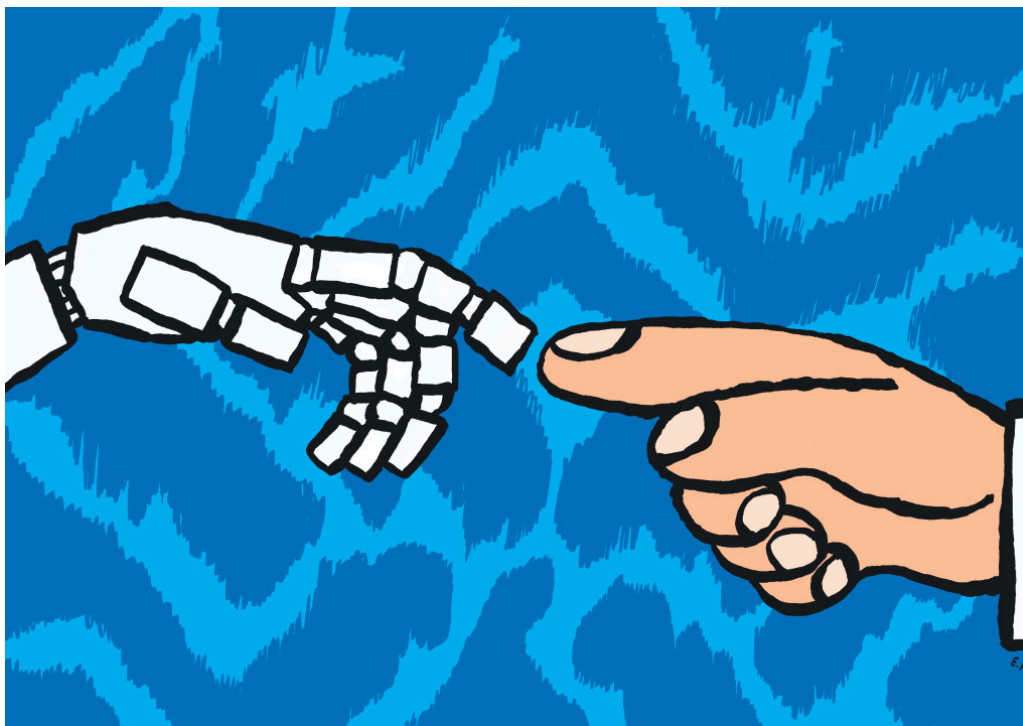
Forfatteren har fylt ut ICMJE-skjemaet og oppgir følgende interessekonflikter: Han er tidligere medlem av Teknologirådet og har investert i globale indeksfond som inkluderer aksjer i teknologiselskaper som driver med kunstig intelligens.

MICHAEL RIEGLER

Michael Riegler er professor ved OsloMet – storbyuniversitetet og forsker ved SimulaMet.

Forfatteren har fylt ut ICMJE-skjemaet og oppgir ingen interessekonflikter.

Vi trenger en helhetlig tilnærming for å sikre at kunstig intelligens kommer både pasienter og helsepersonell til gode.



Illustrasjon: Espen Friberg

Kunstig intelligens (KI) er i ferd med å endre helsevesenet globalt, og Norge er intet unntak. En studie publisert i *International Journal of Environmental Research and Public Health* viste at antallet artikler om KI innen medisin har mer enn fordoblet seg fra 2015 til 2022 (1). Denne veksten reflekterer ikke bare potensialet KI har til å transformere helsetjenester, men også de komplekse utfordringene knyttet til implementeringen av teknologien. Teknologiske fremskritt har gjort det mulig å bruke KI til alt fra bildeanalyse til prediksjon av sykdomsutfall, men uten et solid rammeverk kan det oppstå betydelige etiske og praktiske problemer.

I Norge står vi nå ved et kritisk veiskille. Vi må utvikle en strategi som sikrer at integrering av KI i helsevesenet ikke bare fokuserer på effektivitetsgevinster, men også tar hensyn til våre grunnleggende verdier og etiske prinsipper. Dette er særlig viktig i lys av de mange eksemplene på problematisk forskning og overdrevne påstander innen feltet. For eksempel ble det i en omdiskutert studie hevdet at sykdommer kunne diagnostiseres basert på bilder av tungen. Denne studien viste seg å ha alvorlige metodologiske mangler, både når det gjelder medisinsk metodologi og KI-teknologi (2).

I denne kronikken tar vi utgangspunkt i den pågående internasjonale debatten om kunstig intelligens innen medisin (3), men vårt hovedfokus er å diskutere de unike utfordringene og mulighetene som oppstår når KI skal implementeres i norsk helsevesen.

«En sentral utfordring i utviklingen av KI-systemer for helsevesenet er hvordan vi kan balansere effektivitet med sentrale verdier som likeverd og respekt for pasientautonomi»

Balansering av verdier i norsk kontekst

En sentral utfordring i utviklingen av KI-systemer for helsevesenet er hvordan vi kan balansere effektivitet med sentrale verdier som likeverd og respekt for pasientautonomi. Norge har en sterk tradisjon for solidaritet i helsevesenet, hvor alle, uavhengig av sosioøkonomisk bakgrunn, har rett til lik tilgang på helsetjenester. Dette må reflekteres i KI-systemene som implementeres i norsk helsevesen. En KI-modell som forbedrer diagnostikk ved å redusere ventetider, må samtidig ivareta pasientens rett til å bli sett og hørt som et individ.

Et illustrerende eksempel er en eldre pasient som går til legen med diffuse symptomer. Pasienten har et stort behov for å dele sin historie, uttrykke sine bekymringer og føle seg forstått. Samtidig er legen under tidspress for å stille en diagnose og komme frem til en behandlingsplan. I en slik situasjon kan en KI-modell bidra til å optimalisere pasientflyten og diagnostikkprosessen ved å analysere store datamengder raskere enn legen alene er i stand til. Likevel er det fare for at modellen kan overse nyanser i pasientens fortelling eller mistolke viktige nonverbale signaler som kunne ha vært avgjørende for riktig diagnose.

Dette eksempelet illustrerer behovet for at KI-systemer i helsevesenet ikke bare fokuserer på algoritmisk effektivitet, men også ivaretar den menneskelige dimensjonen ved pasientbehandlingen. Teknologien må være et verktøy som styrker legen, ikke erstatter de mellommenneskelige aspektene. I Norge har vi muligheten til å utvikle en nasjonal konsensus om hvilke verdier som skal prioriteres i utviklingen av KI for helsesektoren. Dette krever en bred og inkluderende prosess, hvor både helsepersonell, pasienter, teknologer og etiske eksperter blir hørt. Målet må være å etablere et rammeverk som reflekterer våre samfunnsverdier, samtidig som det utnytter KI-teknologiens potensial til å forbedre helsetjenester.

Håndtering av bias og norsk tilpasning

En av de største utfordringene ved implementering av KI er risikoen for bias i modellene som utvikles. Bias i KI-systemer er ikke bare et teknisk problem, men også et etisk og klinisk problem som kan ha direkte konsekvenser for pasientbehandlingen. Et eksempel på dette er rasebaserte justeringer i lungefunksjonstester (Pulmonary function tests, PFT), der det lenge har vært praksis å anvende ulike «normale» verdier for lungefunksjon basert på pasientens rase. Dette har ført til bekymring for at slike justeringer kan undervurdere alvorligheten av lungesykdommer hos svarte pasienter, noe som igjen kan føre til forsinket diagnose og behandling [\(4\)](#).

I Norge må vi være spesielt oppmerksomme på hvordan bias kan påvirke behandlingen av sårbare grupper, som den samiske befolkningen, innvandrere og pasienter med ulike sosioøkonomiske bakgrunner. Bias i KI-modeller kan forsterke allerede eksisterende ulikheter i helset utfall dersom datasettet som brukes til å trene modellene, ikke representerer hele befolkningen. For

eksempel, hvis en KI-modell utvikles basert på data fra en befolkning der enkelte grupper er underrepresentert, kan resultatene være skjeve og føre til dårligere behandling for disse gruppene.

For å motvirke bias og sikre rettferdig implementering av KI i norsk helsevesen, foreslår vi flere tiltak.

For det første må vi sikre representativ datainnsamling. Datasett brukt til å trene KI-modeller, må inkludere alle befolkningsgrupper i Norge, inkludert etniske minoriteter, ulike aldersgrupper, kjønn og sosioøkonomiske bakgrunner. Det må iverksettes spesifikke datainnsamlingstiltak for å inkludere underrepresenterte grupper.

For det andre er det nødvendig med regelmessig revisjon. KI-modeller må regelmessig revideres for å avdekke og korrigere skjevheter. Dette innebærer at det bør etableres et nasjonalt system for å overvåke og evaluere ytelsen til KI-modeller i klinisk praksis.

Når det gjelder etiske vurderinger, bør det opprettes tverrfaglige panel bestående av klinikere, etikere, pasientrepresentanter og KI-eksperter for å vurdere potensielle bias og etiske implikasjoner før modellene implementeres.

Dessuten må KI-modeller utviklet utenfor Norge, evalueres og tilpasses norske forhold før de tas i bruk. Dette innebærer ikke bare språklig oversettelse, men også justering etter norske kliniske retningslinjer, helseprioriteringer og kulturelle normer.

Til sist må helsepersonell og pasienter informeres om KI-modellenes begrensninger, inkludert mulige bias, slik at de kan ta informerte beslutninger.

Ved å implementere disse tiltakene kan vi utvikle et KI-system som er rettferdig og robust, og som ikke bare forbedrer kvaliteten på helsetjenestene, men også styrker tilliten til KI blant både pasienter og helsepersonell.

Interdisiplinært samarbeid

For å lykkes med å integrere KI i norsk helsevesen, er det avgjørende å legge til rette for interdisiplinært samarbeid. KI-systemer som brukes i klinisk praksis, må være teknologisk avanserte, men også etisk forsvarlige og klinisk relevante. Dette krever samarbeid mellom teknologer, klinikere, etikere, jurister og pasientrepresentanter.

En modell for interdisiplinært samarbeid kan være å etablere tverrfaglige team som involverer alle relevante aktører i KI-utvikling og implementering, fra idéfase til klinisk bruk. Disse teamene kan jobbe med alt fra teknologiske utfordringer til å sikre at systemene overholder de juridiske og etiske standardene som kreves i helsevesenet. Parallelt med dette bør det opprettes rådgivende etiske komitéer som regelmessig vurderer etiske spørsmål knyttet til KI.

For å sikre at KI-løsninger faktisk er relevante for norsk klinisk praksis, er det også nødvendig å utdanne og ansette fagpersoner med kompetanse innen både medisin og KI. Dette kan bidra til at KI-systemer tilpasses norske kliniske behov, og at helsepersonell er i stand til å bruke teknologien på en forsvarlig måte.

Videre må vi inkludere pasienter og pasientrepresentanter i utviklingsprosessen for KI-systemer. Dette vil bidra til at teknologien møter reelle behov og respekterer pasientautonomi.

Langsiktige implikasjoner og et verdidrevet rammeverk

Integreringen av KI i det norske helsevesenet vil utvilsomt ha dyptgripende og langsiktige konsekvenser for både pasientbehandling og forholdet mellom helsepersonell og pasienter. Etter hvert som KI får en mer sentral rolle i beslutningsprosesser, må vi nøye vurdere hvordan vi kan bevare pasientens tillit til helsevesenet.

En viktig del av denne prosessen er å utvikle et bredt etisk rammeverk som reflekterer de verdiene som er viktige i norsk helsetjeneste. Dette inkluderer respekt for pasientautonomi, hvor pasienter skal forbli aktive deltakere i sin egen behandling, selv når KI blir en mer integrert del av beslutningsprosessen. Ikke skade-prinsippet må også være en grunnleggende rettesnor for utvikling og implementering av KI, slik at teknologien ikke fører til utilsiktede negative konsekvenser.

«KI-systemer skal ikke erstatte relasjonen mellom helsepersonell og pasienter, men heller forsterke den»

For å ivareta de mellommenneskelige dimensjonene i helsetjenesten, bør vi også inkludere omsorgsetikk i vårt etiske rammeverk. KI-systemer skal ikke erstatte relasjonen mellom helsepersonell og pasienter, men heller forsterke den. Dette vil være avgjørende for å sikre at pasienter opplever seg selv som likeverdige og aktive deltakere i egen behandling, selv når KI-teknologi blir en del av behandlingsprosessen.

Transparens og forklarbarhet

Et viktig aspekt ved implementering av KI er å sikre at systemene er transparente og forklarbare. For at KI skal være en pålitelig del av helsevesenet, må både helsepersonell og pasienter forstå hvordan systemene fungerer og hvordan de kommer frem til sine konklusjoner. Uten tilstrekkelig forklarbarhet kan KI skape mistillit og usikkerhet både blant helsepersonell og pasienter.

For å oppnå dette foreslår vi implementering av *model cards* som en standard praksis i norsk helsesektor. Disse dokumentene, som kan sammenlignes med sikkerhetsdatablader i industrien, gir en omfattende oversikt over en KI-

modellens ytelse, opplæringsdata, metodikk og potensielle begrensninger. Ved å gi helsepersonell tilgang til denne informasjonen, muliggjør vi informerte beslutninger om når og hvordan KI-modeller bør brukes i klinisk praksis. Dette bidrar ikke bare til økt transparens, men også til en mer ansvarlig og etisk bruk av KI i helsevesenet.

Videre må vi legge vekt på å evaluere resonneringen bak KI-modellenes beslutninger. Nyere forskning har vist at selv avanserte modeller, som GPT-4 V, kan gi korrekte svar basert på feilaktige resonnementer (5). I en norsk helsekontekst, hvor pasientenes sikkerhet og tilliten til helsesystemet står sentralt, er det kritisk at vi ikke bare vurderer nøyaktigheten av KI-modellenes output, men også kvaliteten på deres resonnement.

En annen viktig faktor er kommunikasjon av usikkerhet (6). KI-modeller i norsk helsevesen bør være i stand til å formidle graden av usikkerhet i sine prediksjoner og anbefalinger. Dette vil gi helsepersonell et mer nyansert beslutningsgrunnlag og bidra til å opprettholde en sunn skepsis til KI-assisterte beslutninger.

«Etisk vurdering bør være en integrert del av KI-planen, og vi foreslår at det etableres mekanismer for kontinuerlig evaluering og justering av KI-systemer i bruk»

Evaluering av Helsedirektoratets KI-plan

Helsedirektoratets KI-plan er et viktig skritt mot systematisk bruk av KI i norsk helsevesen, men vi mener at planen kunne gått enda lenger når det gjelder å definere et tydelig etisk rammeverk for KI-implementering. Etisk vurdering bør være en integrert del av KI-planen, og vi foreslår at det etableres mekanismer for kontinuerlig evaluering og justering av KI-systemer i bruk.

Videre anbefaler vi at planen styrkes med konkrete tiltak for å balansere ulike hensyn, som effektivitet og personvern, samt likhet i tilgang og individualisert behandling. Dette vil sikre at KI-løsningene ikke bare forbedrer effektiviteten, men også ivaretar pasientenes rettigheter og interesser.

Implementeringen av KI i det norske helsevesenet representerer både en stor mulighet og en betydelig utfordring. En vellykket integrering krever mer enn bare teknologisk ekspertise – det krever et omfattende etisk og tverrfaglig rammeverk som sikrer at teknologien brukes på en rettferdig og ansvarlig måte. Ved å utvikle KI-løsninger som respekterer norske verdier og ivaretar pasientenes autonomi, kan vi sikre at helsetjenesten både blir mer effektiv og mer pasientfokuset.

REFERENCES

1. Karpov OE, Pitsik EN, Kurkin SA et al. Analysis of Publication Activity and Research Trends in the Field of AI Medical Applications: Network Approach.

Int J Environ Res Public Health 2023; 20: 5335. [PubMed][CrossRef]

2. Nassehi D. Kunstig intelligens og pseudovitenskap er en farlig kombinasjon. Dagens Medisin 20.8.2024.

<https://www.dagensmedisin.no/forskning-kunstig-intelligens-medisin/kunstig-intelligens-og-pseudovitenskap-er-en-farlig-kombinasjon/647826> Lest 23.8.2024.

3. Yu K-H, Healey E, Leong T-Y et al. Medical Artificial Intelligence and Human Values. N Engl J Med 2024; 390: 1895–904. [PubMed][CrossRef]

4. Braun L, Fausto-Sterling A, Fullwiley D et al. Racial categories in medical practice: how useful are they? PLoS Med 2007; 4: e271. [PubMed][CrossRef]

5. Jin Q, Chen F, Zhou Y et al. Hidden flaws behind expert-level accuracy of multimodal GPT-4 vision in medicine. NPJ Digit Med 2024; 7: 190. [PubMed][CrossRef]

6. Huang Y, Song J, Wang Z et al. Look Before You Leap: An Exploratory Study of Uncertainty Measurement for Large Language Models. arXiv 2023 doi: 10.48550/arXiv.2307.10236.. [CrossRef]

Publisert: 21. november 2024. Tidsskr Nor Legeforen. DOI: 10.4045/tidsskr.24.0445

Mottatt 23.8.2024, første revisjon innsendt 3.9.2024, godkjent 16.9.2024.

Opphavsrett: © Tidsskriftet 2026 Lastet ned fra tidsskriftet.no 23. mars 2026.