

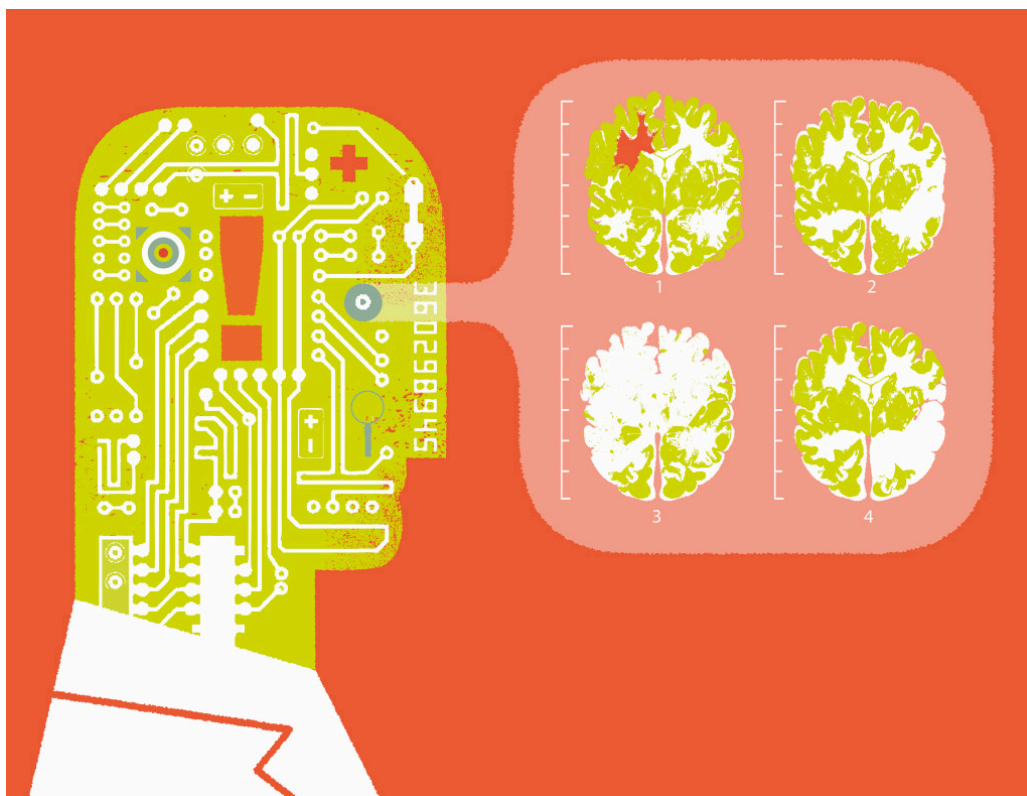
Mye hype, men lite kunnskap om kunstig intelligens

FRA ANDRE TIDSSKRIFTER

KETIL SLAGSTAD

Tidsskriftet

Kunnskapsgrunnlaget for implementering av kunstig intelligens i medisinsk diagnostikk er spinkelt.



Illustrasjon: Gillian Blease / NTB Scanpix

Det er knyttet enorme økonomiske interesser til utvikling og bruk av kunstig intelligens i fremtidige helsetilbud, og i nyhetsartikler blir teknologien ofte framstilt som revolusjonerende. Men hva er kunnskapsgrunnlaget for innføring av kunstig intelligens i helsetjenesten?

Dyplæringsalgoritmer innebærer at man gjennom dype nevralt nettverk identifiserer hvilke kjennetegn som er karakteristiske i et gitt klinisk scenario, for eksempel et funn på et røntgenbilde, uten at nettverket på forhånd blir instruert om hvilke tegn det skal lete etter. I en ny systematisk oversiktsartikkel publisert i tidsskriftet BMJ ble den diagnostiske treffsikkerheten til dyplæringsalgoritmer sammenliknet med den til erfarne klinikere (1). Analysen omfattet 81 ikke-randomiserte studier som var publisert og ti randomiserte studier som var registrert, hvorav to var publisert.

Bare ni av de 81 studiene var prospektive, og kun seks av disse ble utført under virkelighetstro forhold. Flesteparten av studiene tilhørte feltet radiologi, fulgt av oftalmologi, dermatologi, gastroenterologi og histopatologi. I 61 av de 81 sammendragene ble kunstig intelligens beskrevet som sammenliknbart med eller bedre enn klinisk bedømming. Kun ni sammendrag nevnte at det var nødvendig med fremtidige prospektive studier. I nesten ingen av studiene var datasettene tilgjengelige. Risikoen for systematiske skjevheter ble bedømt som høy i 58 av de 81 studiene.

– Dette er en viktig studie som gir et edruelig bilde av hvor man er i utviklingen av kunstig intelligens i medisinsk diagnostikk, sier Ishita Barua, som er lege og stipendiat ved Universitetet i Oslo og Oslo universitetssykehus, der hun forsker på kunstig intelligens i endoskopi.

– Kun et fåtall studier er prospektive, randomiserte eller utført i en ekte klinisk setting, og flertallet av studiene har høy risiko for systematiske skjevheter. Kunnskapsgrunnlaget for å implementere kunstig intelligens i klinisk praksis er dermed tynt og begrenset, sier Barua.

– Datagrunnlaget og kodene bak algoritmene ble kun gjort tilgjengelige for andre forskere i svært få studier. Dette gjør det vanskelig å reprodusere funnene, sier hun.

– En annen mangel er at oversiktsartikkelen kun omfatter bruk av dyplæring i medisinsk bildediagnostikk, og ikke tar for seg andre typer kunstig intelligens. Mange beslutningsstøtteverktøy i klinisk praksis er for eksempel basert på maskinlæring, som blant annet blir brukt til å identifisere atrieflimmer på EKG, sier Barua, som etterlyser flere prospektive studier for å undersøke effekt av kunstig intelligens som beslutningsstøtte i sykehus.

LITTERATUR

1. Nagendran M, Chen Y, Lovejoy CA et al. Artificial intelligence versus clinicians: systematic review of design, reporting standards, and claims of deep learning studies. BMJ 2020; 368: m689. [PubMed][CrossRef]

Publisert: 4. juni 2020. Tidsskr Nor Legeforen. DOI: 10.4045/tidsskr.20.0354
Opphavsrett: © Tidsskriftet 2026 Lastet ned fra tidsskriftet.no 17. februar 2026.