



UiO • Universitetet i Oslo

Hvordan kan kunstig intelligens kvalitetsforbedre og effektivisere radiologenes arbeidsoppgaver

Erik Haavardsholm

Masteroppgave
Erfaringsbasert master i helseadministrasjon
Høst 2020

Avdeling for helseledelse og helseøkonomi
Institutt for helse og samfunn
Det medisinske fakultet, Universitetet i Oslo

2020

Hvordan kan kunstig intelligens kvalitetsforbedre og effektivisere radiologenes arbeidsoppgaver

Erik Haavardsholm

<http://duo.uio.no>

Trykk: Reprosentralen UiO

Sammendrag

Radiologi er en medisinsk spesialitet som over tid har utviklet seg gjennom store teknologiske fremskritt. Mulighetene innen bildediagnostikk har skapt nye behov og en etterspørsel som har økt, og det rapporteres om underkapasitet både globalt og i Norge. Med datateknologiske fremskritt har kunstig intelligens (artificial intelligence, heretter kalt AI) hatt en betydelig utvikling de senere årene og blir forespeilet som den neste store revolusjonene innen bildediagnostikk. Programmene AI er bygget opp som nevralt nettverk med muligheter for å lære og utføre komplekse oppgaver som tidligere kun har vært utført av mennesker. Det er en generell optimisme for hva kunstig intelligens kan bidra med av diagnostiske verktøy, men samtidig har det utløst en bekymring og skepsis innen fagmiljøene for at datamaskiner overtar oppgaver som i dag hører under fagmiljøene.

Denne oppgaven har som mål å belyse muligheter kunstig intelligens gir innen bildediagnostiske verktøy, samt vurdere begrensningene og utfordringene som følger med. Hvilke effektiviserende og kvalitetsforbedrende effekter som denne teknologien kan gi for faget radiologi og hvilke utfordringer må belyses før implementering. Oppgaven diskuterer hvordan fagmiljøet kan forholde seg til dette. Problemstillinger om hvordan forhold innen faglig forsvarlighet, jus, personvern økonomi og tekniske utfordringer må vurderes, drøftes.

Metoden er basert på litteraturgjennomgang. Som utgangspunkt er det gjort søk gjennom systematiske oversiktsartikler supplert med ytterligere enkeltartikler for å belyse problemstillingene i oppgaven.

Litteraturen viser mange lovende resultater ved testing av ulike AI-algoritmer som diagnostiske verktøy innen ulike fagområder av radiologien. En gjennomgående tendens er at det i litteraturen etterlyses bedre og mer standardiserte studier som kan validere programmene på en vitenskapelig måte. AI har utvilsomt et stort potensial innen medisinsk bildediagnostikk og vi har trolig bare sett starten. AI vil trolig kunne gi effektive og gode verktøy som avlaster radiologer for tidkrevende screening-arbeid, og som kan bidra med analyse av bildefunn, kvantifisering og forløpsvurdering. Med mulighetene AI gir for å hente ut karakteristika fra bildene i store volumer, samt integrasjon mot annen diagnostikk fra vevsprøver, genetisk profil osv., åpner det seg muligheter innen integrert diagnostikk og individtilpasset diagnostikk og behandling. Fagmedisinske miljøer bør ha en proaktiv og konstruktiv tilnærming til AI for å være med på å forme og påvirke hvordan denne teknologien implementeres og hvilken rolle radiologene skal ha. Integrering av AI som et fagområde innen radiologi med egne læringsmål i LIS-utdanningen kan være en god start.

Innhold

BEGREPER OG DEFINISJONER	1
1. INNLEDNING.....	3
1.1 Kunstig intelligens – en forespeilet revolusjon av radiologien	3
1.2 Problemstilling	4
2. BAKGRUNN	5
2.1 Radiologi	5
2.2 Radiologenes arbeidsoppgaver	7
3. METODE.....	9
3.1 Struktur på litteratursøket	9
3.2 Funn ved litteratursøket.....	9
4. RESULTATER.....	11
4.1 Hva er kunstig intelligens	11
4.1.1 Maskinlæring	11
4.1.2 Deep learning	11
4.2 Kan AI løse fremtidens kapasitetsutfordringer	13
4.2.1 Kapasitetsutfordring og ressurser	13
4.2.2 Standardisering.....	14
4.3 AI i en radiologisk avdeling.....	15
4.3.1 Bildeopptak	15
4.3.2 Preprosessering	15
4.3.3 Bildeframstilling.....	15
4.3.4 Bildetolkning.....	16
4.3.5 Rapportering.....	16
4.3.6 Integrert diagnostikk	16
4.4 AI-verktøy i radiologisk diagnostikk	18
4.4.1 Computer aided detection	18
4.4.2 Segmentering	18
4.4.3 Radiomics	19
4.4.4 Nevroradiologi.....	21

4.4.5 Øre-nese-hals-radiologi	23
4.4.6 Thoraxradiologi og brystdiagnostikk	24
4.4.7 Abdominal- og urologisk radiologi.....	26
4.4.8 Muskel-skjelett-radiologi.....	27
4.4.9 Andre områder	27
5. DISKUSJON	28
5.1 Effektivisering og kvalitetsforbedring.....	28
5.2 Ledelse og implementering	29
5.2.1 Juridiske utfordringer	29
5.2.2 Personvern	29
5.2.3 Effektivisering i lys av økonomiske hensyn	30
5.2.4 Endringsmotstand	31
5.2.5 Strategier for bruk av AI	32
5.3 Tekniske utfordringer	34
5.3.1 Datatilgjengelighet	34
5.3.2 Transparens	34
5.3.3 Overførbarhet.....	34
5.4 Faget radiologi	36
5.4.1 Radiologenes fremtid	36
5.4.2 Fagmiljøets tilnærming.....	37
6. KONKLUSJON	38
7. REFERANSER.....	39

Begreper og definisjoner

AI: Artificial intelligence, som oversatt til norsk omtales som «kunstig intelligens». Innen radiologi er dette dataprogramvarer/algoritmer som er i stand til å utføre oppgaver som radiologene tradisjonelt har måtte gjøre.

Applikasjon: En programvare som har som formål å forbedre en arbeidsprosess, hjelp i bildetolkning, fremstilling av undersøkelser på hensiktsmessig måte, reduksjon av stråledose, bedret og effektivisert bildeopptak, med mer.

Biopsi: Å hente ut vevsprøve fra kroppen som så sendes til patolog for diagnostikk.

CAD: Computer aided detection. En programvare som hjelper til med å oppdage patologi i radiologiske bilder.

CT: Computer tomografi. Bildedannende teknikk ved hjelp av ioniserende røntgenstråler i et roterende rør, som gi snittbilder gjennom kroppen.

«Deep learning»: Er et engelsk begrep som kan direkte oversettes til dyp læring, i oppgaven forkortet til «DL». Dette er en gren av maskinlæring, som igjen er et område innen kunstig intelligens.

Histologi: Vevsdiagnostikk ved mikroskopisk undersøkelse av kroppsvev, gjerne tatt i forbindelse med diagnostikk av uavklart tumor eller annen sykdomstilstand i aktuelle organ.

Interobservatørvariabilitet: Variasjon i hvordan to ulike radiologer oppdager, tolker og formidler patologi i radiologiske undersøkelser.

Intraobservatørvariabilitet: Variasjon i hvordan en enkelt radiolog oppdager, tolker og formidler patologi i radiologiske undersøkelser.

Intervensjon: Prosedyrer som innebærer tilgang til kroppens hulrom eller anatomiske strukturer via tilgang gjennom hud eller andre kroppsåpninger.

LIS: Lege i spesialisering.

Maskinlæring: Et område innen kunstig intelligens.

Mammografi: Røntgenundersøkelse av brystvev.

MSK-radiologi: Muskel-skjelett radiologi.

MR: Magnetisk resonanstomografi. Bildedannende modalitet ved hjelp av magnetfelt og radiobølger.

PACS (picture archiving and communication system): Radiologenes bildearkivsystem der radiologene henter opp og bearbeider bildene de skal tolke.

Piksel: en av mange små punkter/bildeelementer som til sammen utgjør et bilde på en datamaskin.

Radiomics: Fagområde som handler om å ekstrahere data og karakteristika fra medisinske bildematerialer, som for eksempel signalintensitet, form, tekstur, bølgelengde osv.

Radiolog: Lege med spesialistgodkjenning innen fagområdet/spesialiteten radiologi.

Radiograf: Profesjon med teknisk utdanning for drift av medisinsk bildedannende utstyr. Det er radiografene som utfører bildeopptaket ved de fleste undersøkelsene.

Triage: Rangering av alvorlighetsgrad innen medisinske tilstander.

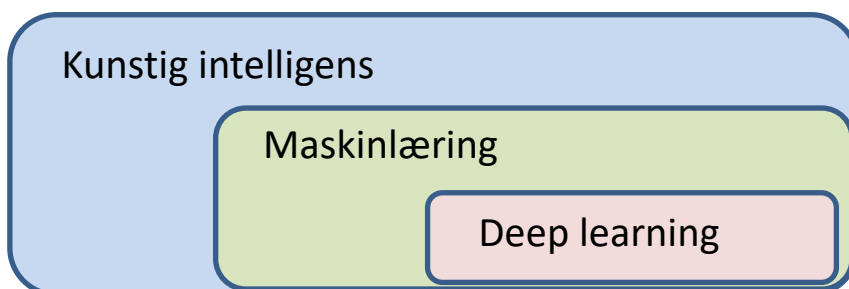
UL: Ultralyd er en bildedannende teknikk ved hjelp av lydbølger og ekko.

ØNH-radiologi: Øre-nese-hals radiologi.

1. Innledning

1.1 Kunstig intelligens – en forespeilet revolusjon av radiologien

Kunstig intelligens (heretter kalt AI, artificial intelligence) er et fagområde innen datateknologi der avanserte algoritmer kan utføre avanserte oppgaver som tidligere har vært forbeholdt mennesker. AI er et bredt begrep som dekker et stort område, der maskinlæring er et område innen AI og deep learning igjen er et område innen maskinlæring (figur 1.1). Det er «deep learning» med nevralt nettverk som har vist seg å være den banebrytende teknologien for utvikling av programmer til medisinsk bildediagnostikk. I senere tid har AI gjort store fremskritt og blir av mange sett på som radiologiens neste kvantesprang. Avanserte AI-programmer vil kunne detektere patologi, hjelpe til med kvantifisering og klassifisering samt forslag til diagnose. De mest optimistiske på vegne av AI går så langt som å hevde at radiologene vil være overflødige i fremtiden(1, 2). Andre mer nøkterne mener realistiske mål er å bedre effektiviteten noe som resulterer i at radiologene kan beskrive flere undersøkelser, i tillegg til en kvalitetsøkning i form av mindre oversett patologi, mer presis diagnostisering og redusert inter- og intraobservatørvariabilitet. Det har vært en stor entusiasme for dette fagfeltet de siste 10 årene der antall publikasjoner har økt dramatisk(3). Samtidig som det er stor optimisme for hvilke muligheter dette gir er det også utfordringer som må belyses. Det er viktig at de radiologiske miljøene orienterer seg om hvilken teknologi som til enhver tid er tilgjengelig og hvilke muligheter og begrensninger det gir. Det overordnede målet må være å yte best mulig diagnostisk kvalitet for pasientene samt utnytte tilgjengelige ressurser best mulig. På veien mot dette målet må man sikre trygg og god implementering av teknologien, tenke gjennom og definert arbeidsoppgavene til radiologene og utrede konsekvensene av endrede arbeidsoppgaver. Juridiske og etiske problemstillinger må også stå sentralt. Det er nødvendig at de radiologiske miljøene har en visjon for videre utvikling av faget og hvordan de i best mulig grad kan bruke sine kunnskaper dersom kunstig intelligens overtar enkelte av deres arbeidsoppgaver.



Figur 1.1 Skjematisk oversikt over begrepene innen kunstig intelligens.

1.2 Problemstilling

Oppgaven har som mål å belyse hva AI er og hvilken rolle den har i radiologi samt belyse flere områder der AI kan effektivisere og kvalitetsforbedre radiologenes arbeidsoppgaver. Oppgaven gir en oversikt over hvor langt utviklingen er kommet, samt aktuelle studier som viser hvordan AI kan brukes innen ulike fagfelt av radiologien per i dag. Samtidig beskriver oppgaven potensielle muligheter for fremtiden, behov og ønsker for radiologiske verktøy innen AI og belyser aktuelle problemstillinger og utfordringer. Hvordan kan radiologer være med på å påvirke utviklingen og deres eget fag og arbeidssituasjon? Hvordan skal det radiologiske fagmiljøet ta imot AI og integrere dette i sin profesjon? Hvilke fordeler og ulemper vil kunne følge ved implementering av avanserte AI-baserte bildetolkningsverktøy? Hvilke utfordringer og fallgruver må man være observante på når man skal implementere og ta i bruk denne teknologien? Har ledere forutsetninger og kunnskaper og ikke minst faglig støtte i sine institusjoner til å implementere fremtidige AI-programvarer på en trygg og god måte? Kan ledere overbevise skeptikerne om at AI er et effektivt og kvalitetsforbedrende hjelpemiddel til radiologene? Kan man integrere AI som en del av LIS-utdanningen for å forberede kunnskapsgrunnlaget og gjøre nye radiologer rustet til å ta i bruk og kritisk kunne vurdere AI i sin arbeidshverdag?

Oppgaven har som ambisjon å gi en oversikt over ovenfor nevnte temaer. Selv om jeg innenfor oppgavens rammer må behandle noen områder til dels overfladisk, forsøker jeg å gjøre en bred nok tilnærming til at temaet kan gi en hjelp til radiologer og ledere i radiologiske avdelinger og institutter, samt andre som er involvert i drift av helsetjenesten, til å få en oversikt over aktuelle problemstillinger knyttet til AI i radiologi. Temaet er stort og oppgaven vil således gi et innblikk i enkelte områder der AI kan ha en plass i bildediagnostikken. Ved å øke kunnskapsnivået og belyse utfordringer vil man forhåpentligvis ha bedre forutsetninger for å gjøre gode valg og bruke AI fornuftig i fremtidens radiologi.

2. Bakgrunn

2.1 Radiologi

Radiologi er en medisinsk spesialitet og et essensielt fagfelt innen medisinsk diagnostikk som bidrar med viktige diagnostiske undersøkelser innen alle medisinske fagområder. Fra radiologiens opprinnelse ved Wilhelm Conrad Röntgens oppdagelse av røntgenstråler i 1895 til dagens radiologiske avdelinger har faget gjennomgått en stor utvikling. Allerede i 1917 ble fagfeltet etablert som en egen spesialitet. Fra 1940 begynte man også gradvis å ta i bruk UL (ultral lyd) i medisinsk diagnostikk men først på -60 og -70 tallet denne teknologien tatt i mer utstrakt bruk i Norge. Intervensjonsradiologi gjorde sin inntreden allerede på 1960-tallet med angiografi (kontrastundersøkelse av blodårene) og etter hvert er mer avansert intervensjon som utblokking av tette arterier, stansing av blødninger (embolisering) og flere andre inngrep muliggjort. MR (magnetisk resonanstomografi) og CT (computertomografi), to av de viktigste modalitetene innen radiologi i dag, startet på 1970/80-tallet. Disse to modalitetene revolusjonerte de diagnostiske mulighetene innen radiologien da man for første gang fikk fremstilt tverrsnitt gjennom kroppen samt med tredimensjonale fremstillinger. Mange av de tidligere konvensjonelle røntgenundersøkelsene ble etter hvert erstattet med CT og MR som ga en betydelig økt diagnostisk presisjon. Fra analoge bildeopptak av konvensjonelle røntgenbilder fikk vi etter hvert digitale bildeopptak og bildebehandling (tidlig 2000-tallet). Bildekvaliteten og mulighetene innen bildediagnostikk har forbedret seg drastisk med den teknologiske utviklingen. Avanserte undersøkelser tar stadig kortere tid å gjennomføre og de gir mer og bedre diagnostisk informasjon. Vi har fått muligheten til å diagnostisere flere sykdomstilstander enn vi tidligere kunne. Kliniske leger har mulighet til å bekrefte eller avkrefte medisinske tilstander med større sikkerhet som følge av bedret bildediagnostikk. Raskere avklaring av kliniske problemstillinger gir muligheter for raskere iverksetting av riktig behandling og potensielt kortere liggetid på sykehus. Sparte liggedøgn er ønskelig både som kostnadsbesparelse og for at pasientene skal unngå unødvendig forlenget sykehusopphold. Som følge av dette har omfanget av undersøkelser gradvis økt og i tillegg er mange av undersøkelsene mer avanserte og inneholder større mengder bildematerialer enn tidligere(4). Logistikken gjennom helsevesenet er sårbart for flaskehals og er avhengig av effektiv radiologisk bildediagnostikk for å unngå forsinkelser i pasientens tilferdig behandling. Stadig økende bruk av radiologisk diagnostikk fører økt arbeidsmengde for radiologene som skal tolke undersøkelsene, en utfordring som settes på spissen når behovet og etterspørsel øker mer enn det som tilbudet av radiologikapasitet kan økes.

Radiologene er leger med spesialistutdanning. For å bli radiolog må man etter endt medisinstudium og turnustjeneste/LIS1 gjennom en strukturert spesialistutdannelse ved en eller flere radiologiske avdelinger, tilsammen i ca. 5 år. I løpet av disse årene skal leger i spesialisering (LIS) tilegne seg en rekke ferdigheter og kunnskap innen de ulike fagfeltene av radiologi. Dette gjøres i praksis ved å arbeide på en radiologisk avdeling under supervisjon, deltagelse i vakt, veiledning fra spesialister samt deltagelse på nasjonale fagkurs.

Ved større sykehus og private klinikker er ofte radiologene subspecialiserte innen snevre radiologiske fagfelt, hovedsakelig organseksjonert, som for eksempel muskel/skjelettradiologi, abdominalradiologi, nevradiologi, intervensjon med mer. Dette har vist seg å være hensiktsmessig for å kunne fordype seg og holde seg oppdatert med den høyeste faglige kompetansen som de kliniske avdelingene trenger når de henviser pasientene. Kunnskapsfeltet og mulighetene innen radiologi er blitt så stort at subspecialisering er nødvendig og hensiktsmessig der man har store nok avdelinger/institutter til å kunne praktisk organisere og tilby dette. Det er imidlertid ingen formelle subspecialiteter i Norge per i dag, men det er etablert underforeninger av norsk radiologisk forening innen syv subspecialiteter(5).

På mindre sykehus med små radiologiske avdelinger vil radiologene gjerne arbeide mer som generalister, da det ikke er praktisk gjennomførbart å seksjonere i radiologiske subspecialiteter.

Å bli radiolog, spesielt med faglige subspecialiserte kvalifikasjoner, er en omfattende og langvarig utdanning. Dette gjør det ressurskrevende og tidkrevende å øke radiologkapasiteten.

Radiologiske avdelinger har utfordringer med underkapasitet av radiologer i forhold til arbeidsmengden, en trend som har forverret seg over tid(4). Når mengden av undersøkelser øker, samtidig som kravene til kvalitet øker vil dette gi økt press på radiologene for å øke effektiviteten. Dersom vi over tid når et skjæringspunkt der arbeidsbelastningen per radiolog går utover kvaliteten og forsvarligheten, har vi en betydelig utfordring. Håpet er at AI kan bidra til å demme opp for en slik utvikling.

2.2 Radiologenes arbeidsoppgaver

For å vite hvor man kan effektivisere og forbedre må man først ha oversikt over hvilke oppgaver radiologene har. Dette vil variere noe fra sted til sted og avhengig av størrelsen på sykehuset, om det er et universitetssykehus, lokalsykehus eller privat institutt.

Første trinnet i arbeidsflyten til radiologene er vurdering av henvisningen. Det er noe ulike praksis ved avdelinger i hvilken grad radiologene involveres i dette, men ved mer komplekse og ressurskrevende undersøkelser som CT, MR, UL og intervensjonsprosedyrer er det stort sett radiologer som vurderer hvordan undersøkelsen skal utføres, om det skal endres til en annen modalitet, hvilken protokoll som skal brukes og hvor raskt undersøkelsen må bli gjennomført. Denne vurderingen baserer seg på flere informasjonskilder, som henvisningstekst, samtale med henviser, laboratoriesvar, journaldokumenter med mer. Denne integrerte informasjonsmengden gir et helhetsbilde av den kliniske kontekst for angjeldende pasient og er således grunnlag for å velge riktig undersøkelsesmetode.

Neste trinn utgjør radiologens viktigste oppgave, å granske, tolke og beskrive bilder fra ulike modaliteter som røntgen, CT, MR og ultralyd, samt utføre intervensjonsprosedyrer. Bildene blir vanligvis tatt av radiografer som sørger for at bildeopptaket er gjort i henhold til forhåndsgitte protokoll. Bildene overføres deretter til PACS (picture archiving and communication system) for radiologisk vurdering. Ved ultralyd er det stort sett radiologene selv som gjør bildeopptaket, ettersom dette er en dynamisk undersøkelse der det er vanskeligere å standardisere bildeopptaket. Når bilder er tolket og beskrevet signeres beskrivelsen av radiologen eller blir kontrastert av ytterligere en radiolog, som kvalitetssikring. Dette krever ressurser ettersom en ekstra radiolog må granske bildene på nytt. I en viss grad mottar de fleste radiologiske avdelinger undersøkelser som er utført andre steder men ønskes vurdert på nytt med hensyn til en spesifikk problemstilling, såkalt re-granskning.

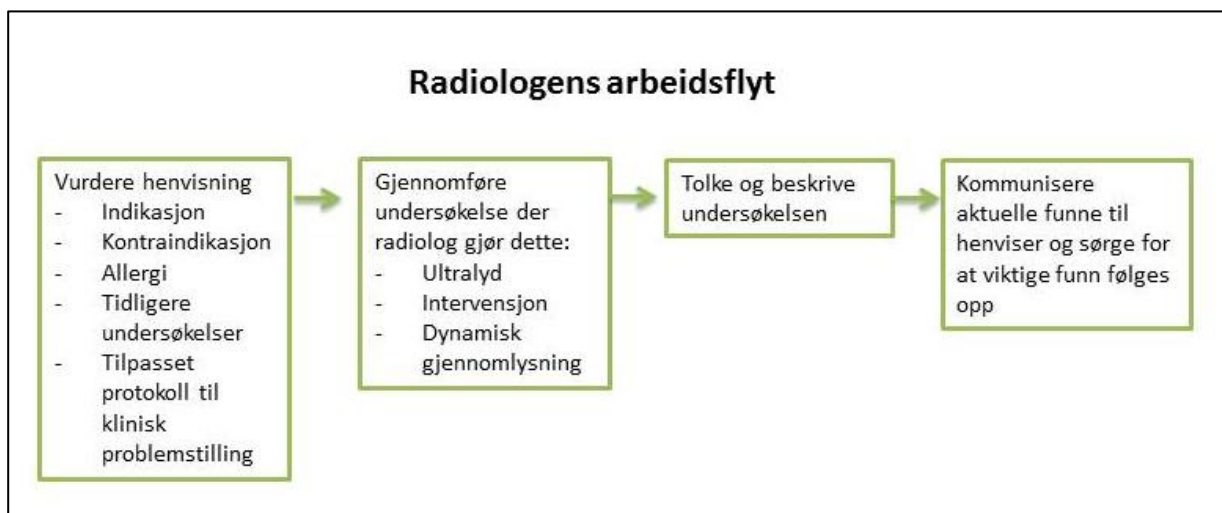
Bildefunnene skal også kommuniseres til kliniker via en rapport/journalnotat og noen ganger i tillegg gjøres dette muntlig til når det er viktige funn som krever behandlingsmessige tiltak raskt. I andre tilfeller er det nyttig at henviser ser gjennom undersøkelsene sammen med radiologen som en forberedelse til kirurgi eller planlegging av annen behandling. Kliniker kan også ha behov for forklaring eller presisering av funn som er beskrevet. Kommunikasjonen mellom henviser og radiolog er derfor viktig i den diagnostiske prosessen.

Et viktig ledd i prosessen er at radiologene demonstrerer bilder for kliniske avdelinger, på undersøkelser på aktuelle pasienter, samt deltar i multidisiplinære møter der flere spesialiteter

møtes for å diskutere pasientens utredning og diagnostikk, videre behandling, kontroll og prognose. Disse multidisiplinære møtene kan bestå av f. eks av onkolog, radiolog, kirurg og patolog. I tillegg til arbeidet med tolkning av bilder har radiologene en viktig funksjon i å kommunisere funnet til henvisende instans samt vurdere supplerende undersøkelser eller videre kontroller. God kommunikasjon mellom kliniker og radiolog bidrar til gjensidig utveksling av viktig pasientinformasjon mellom radiolog og kliniker. Dette øker sannsynligheten for at pasienten får riktig undersøkelse som gir best mulig diagnostikk og dermed riktig og god behandling.

I tillegg til ovennevnte utfører radiologene en rekke intervensjonsprosedyrer, noe som innebærer bildeveiledede invasive prosedyrer for diagnostikk og behandling. Eksempler på dette er innleggelse av dren for å tappe ulike væskeansamlinger, injeksjon av kontrast eller terapeutiske substanser i ulike hulrom av kroppen, utblokkning av tette blodårer samt annen diagnostikk og behandling av karlidelser via arteriell tilgang. Det gjøres ofte en rekke ulike biopsier, bildeveiledet med ultralyd eller CT for å sikre vevsprøver til histologisk undersøkelse hos patolog, noe som ofte er gullstandarden i diagnostikken ved for eksempel kreftdiagnostikk.

Sist men ikke minst, å undervise, veilede og supervisere leger i spesialisering samt andre kolleger er en sentral men ressurskrevende oppgave for radiologene. For enkelte radiologer vil også kvalitetsforbedring, pasientsikkerhet, revisjon av protokoller med mer, være oppgaver som også krever tid og ressurser.



Figur 2.1 Trinnene i radiologenes arbeidsflyt fra mottatt henvisning til aktuelle funn er formidlet tilbake til henviser.

3. Metode

3.1 Struktur på litteratursøket

Denne oppgaven er basert på systematiske litteratursøk. Søket er gjennomført via Embase, Medline og Pubmed med kombinasjon av søkeordene: «artificial intelligence», «machine learning», «deep learning», «neural network», «automation», «radiology», «radiolog», «radiologist».

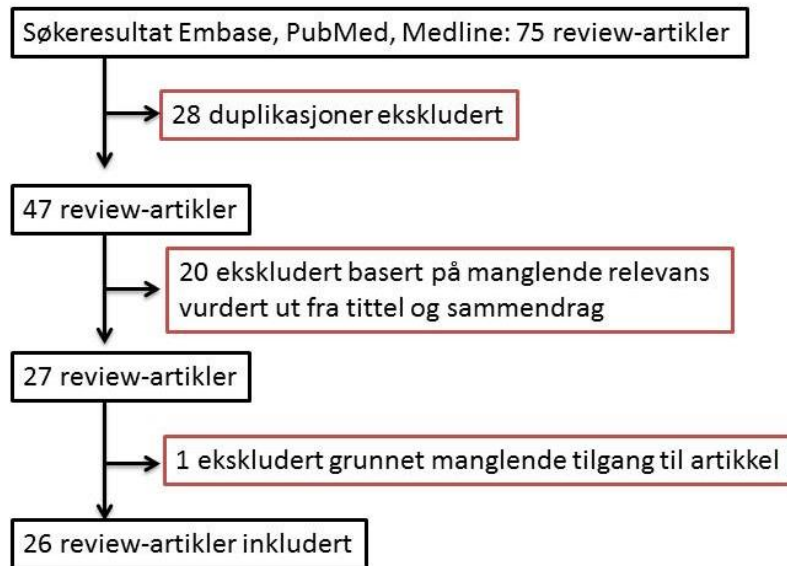
Søket er avgrenset til engelsk, norsk, svensk og dansk. Tidsperioden for utgivelser er avgrenset til 2017 frem til juni 2020. Valget om å snevre inn tidsperioden til kun de nyeste publikasjonene er begrunnet med at denne oppgaven ønsker å gjøre opp status vedrørende hvor man står i dag.

Det foregår en rask utvikling innen AI og mye har skjedd på ganske få år. Når det i tillegg er et bredt spekter og stort antall av publiserte artikler over de siste årene vurderte jeg at det var mest relevant å fokusere på de siste 3-4 årene. Nyere review-artikler inkluderer ofte aktuell kunnskap fra tidligere review-artikler. Søket er filtrert på systematiske oversiktsartikler. Overlapping mellom treff i de søkte databasene ble filtrert og deretter ble aktuelle artikler inkludert eller ekskludert ut fra en vurdering hvorvidt artiklene kunne belyse problemstillingen for oppgaven. Utover søket i overordnede kunnskapsdatabaser er det gjort spesifikke søk på andre temaer oppgaven belyser. Det er også brukt relevante primærartikler som review-artiklene refererer til (såkalt snøball-metoden). Artiklene er sortert etter organspesifikke fagområder av det radiologiske tema de beskriver. Denne oppgaven har ikke som målsetning å gi en komplett oversikt over alle områder der AI er testet i radiologisk bildediagnostikk, ettersom litteraturmengden er for omfattende for en slik oppgave. Det er i stedet for fokusert på enkelte studier innen ulike fagområder av radiologien som var aktuelle for å eksemplifisere hvilke muligheter AI gir og ikke minst belyse aktuelle utfordringer som utgangspunkt for en konstruktiv diskusjon.

3.2 Funn ved litteratursøket

Søkestrategien over systematiske oversiktsartikler ga 34 treff i Embase, 18 treff i PubMed og 23 treff i Medline. Når duplikasjoner var filtrert ut var summen av artikler 47. Etter en gjennomgang av artiklene ble 26 inkludert og 20 ekskludert, vurdert ut fra resyme/abstract om artikkelen var relevant for denne oppgaven. Inkluderte artikler er gjennomgått og sortert etter hvilke områder av

radiologi de omtaler, artikler som omtaler AI-teknologi med potensial for å kunne påvirke radiologers arbeidsoppgaver.



Figur 3.1 Litteratursøk i systematiske oversiktsartikler.

4. Resultater

Innledningsvis presenterer jeg en oversikt over teknologien AI og spesifikt hvilke områder av AI som kan bidra med effektive verktøy innen radiologi. De neste kapitlene beskriver kapasitetsutfordringer og peker skjematisk på flere områder i arbeidsflyten på en radiologisk avdeling der AI kan ha en funksjon. Hovedfokuset er rettet mot siste kapittel der det gis en oversikt over diagnostiske radiologiske AI-verktøy sortert på organsystem.

4.1 Hva er kunstig intelligens

Vitenskapen rundt AI handler om å utvikle systemer som kan utføre oppgaver som opprinnelig krever menneskelig intelligens for å utføre, som for eksempel å lære og løse problemer(6). AI skal kunne utføre oppgavene like bra eller bedre og raskere enn det mennesker kan(7). AI er et bredt begrep som omfatter teknologi som kan brukes innen mange områder av samfunnet, for eksempel selvkjørende biler, datamaskiner som spiller spill, nummerskiltgjenkjenning, talegjenkjenning, chat-bot, ansiktsgjenkjenning samt mye mer. AI er en teknologi som John McCarthy og kolleger startet med allerede i 1956 med sitt prosjekt Dartmouth Summer Research Project(6).

4.1.1 Maskinlæring

Et av områdene innen AI er maskinlæring. Dette er algoritmer trenes til å lære mønstre fra datainput, til forskjell fra eksplisitt programmering som enklere, tradisjonelle programmer er avhengig av (8). Maskinlæring gjør AI-systemet i stand til å trekke ut informasjon fra rådata og lære av erfaringene underveis(9). I tolkningen av et bilde blir algoritmen matet med karakteristika den skal vurdere ved bildene og vil da ved hjelp av statistisk analyse vurdere og karakterisere bildet. Programmene har evne til å lære og justere seg for å utføre oppgavene stadig bedre etter hvert som de får mer data å jobbe med. Dette er litt sammenlignbart med en radiolog som blir bedre av desto større volum av undersøkelser vedkommende undersøker. En datamaskin med AI har i motsetning til en radiolog evne til å håndtere et tilnærmet ubegrenset volum av data.

4.1.2 Deep learning

Et spesielt område innen maskinlæring er «deep learning». Her er programmene bygget opp av nevralt nettverk i flere lag. Dette er teknologiske løsninger som er inspirert av biologiske nevralt nettverk vi har i hjernen. Måten disse nevralt nettverkene er bygget opp med flere dypere lag kalles konvolusjonelle nettverk. De ulike lagene i de nevralt nettverkene prosesserer innkommende informasjon og kommer så ut med et resultat/output. Konvolusjonelle nettverk brukes mye i bildeanalyse innen AI og har vist seg å være den mest lovende metoden for bildeanalyse(10, 11). Det har vist seg at øk dybde i de konvolusjonelle nettverkene, det vil si at de

består av mange lag, og store treningsdatamaterialer gir bedrer prestasjon av en algoritme (12, 13). Et bilde bestående av mange små piksler vil analyseres av de øverste lagene av nettverket som henter ut informasjon og prosesserer videre i de dypere lagene (14). Det gjøres repetitive analyser av bildet med ulike filtre/konvolusjoner som i sum vil gi informasjon om geometrisk form, struktur osv. Dette bygger opp et hierarki av karakteristika for det aktuelle bildet og det avgis et resultat/output i form av f.eks. et diagnoseforslag. For å lære seg disse oppgavene må nettverket trenes med et stort antall bilder der fasiten er gitt, såkalt supervisert læring med merkede data. Supervisert læring er foreløpig den mest brukte metoden for utvikling av AI-algoritmer i radiologi(15). Treningen av nettverket starter med tilfeldig valgte filtre og enkelte utgangsverdier og vil ved gjentatte forsøk prøve og feile frem til verdiene er justert og optimalisert opp mot treningsbildene. Nettverket repeterer disse analysene og forsøkene tusener til millioner av ganger for å optimalisere og justere parameterne ved algoritmen fra innkommende data til resultatet. Nettverket lærer seg selv hvordan det skal komme frem til den gitte diagnosen ut fra bildene. Etter nettverket er trent opp må det testes mot nye bilder for å verifisere at det er godt nok. Det har vist seg utfordrende å skaffe nok treningsbilder med fasit til slike nevrale nettverk i radiologi. Det er nødvendig med kvalifiserte fagpersoner til å klassifisere og kvalitetssikre bildene som skal brukes i treningen, altså supervisert læring av AI-algoritmen. Når en radiolog vurderer bilder og finner patologi, f.eks. en tumor vil vedkommende klassifisere denne ut fra karakteristika som størrelse, signal/tetthet, kontrastoppladning, form osv. Slike forhåndsbestemte karakteristika benytter ikke de nevrale nettverkene seg av. De lærer seg sin egen måte å komme frem til svaret på. Altså vil de selv finne frem til et hierarki av karakteristika som definerer lesjonen. Hvordan de nevrale nettverkene kommer frem til dette har man per nå ikke fullstendig oversikt over og det er dette som er kalt «black box». Nettverket mottar bildeinformasjon og avgir en diagnose uten at vi vet hvordan den kommer frem til svaret. Nevrale nettverk har vært brukt i flere tiår, men det er først de seneste årene man har hatt muligheten til å trene store nevrale nettverk som følge av at man har fått tilgang til store kvanta med merket data, rimeligere og kraftigere datamaskiner samt nye og forbedrede design av «deep learning»-algoritmer(16). «Radiomics» er et engelsk begrep som i stor grad oppsummerer hva AI-baserte algoritmer i radiologien gjør. Begrepet handler om hvordan store kvantitative bildekarakteristika hentes ut fra medisinske bilder (radiologiske undersøkelser i dette tilfellet) ved hjelp av AI-algoritmer, som igjen kan kombineres med kliniske eller biologiske mål(17). Resultatet og målet er verktøy som kan brukes til diagnostisk beslutningsstøtte. Når det senere i oppgaven refereres til AI-algoritmer er det her snakk om algoritmer som bruker «deep learning».

4.2 Kan AI løse fremtidens kapasitetsutfordringer

4.2.1 Kapasitetsutfordring og ressurser

Antall radiologiske undersøkelser per innbyggerantall har økt betydelig i de fleste land de siste 25 årene, basert på data fra OECD(18, 19). I USA har antall MR-undersøkelser per 1000 innbyggere økt fra 34 i 1995 til 128 i 2019(18), i for eksempel Danmark har økt fra 21 i 2003 til 89 i 2019. Samme trenden ser man for CT undersøkelser der økningen i USA har økt fra 79 i 1995 til 279 i 2019, i Danmark fra 43 i 2003 til 145 i 2019(19). Når man i tillegg vet at undersøkelsene er blitt mer avanserte, inneholder flere sekvenser og dermed større bildematerialer, vil den reelle økningen av antall bilder være betydelig større. Antall radiologer per innbyggere og per antall CT/MR varierer mye mellom ulike land(20). De nordiske landene har bedre tilgang på radiologer i forhold til land som for eksempel Storbritannia og Japan.

Enhver radiologisk avdeling vil ha mål om kort ventetid på undersøkelse og beskrivelse for å ivareta pasientsikkerhet og et diagnostisk godt tilbud. Ettersom utviklingen tenderer mot økt behov for radiologiske undersøkelser, samtidig som radiologene sliter med å holde tritt med denne økningen i arbeidsmengde vil man enten måtte redusere antall undersøkelser som utføres, øke antall radiologer, forskyve oppgaver til andre profesjoner eller lage gode verktøy som effektiviserer radiologenes arbeidshverdag. Ved å lage gode hjelpemidler for radiologene i form av programmer som kan hjelpe til med deteksjon, karakterisering, segmentering og klassifisering, vil radiologene kunne jobbe mer effektivt og bearbeide større mengder bildematerialer uten at det går utover forsvarligheten. Samtidig vil man ivareta faglig høy kvalitet på tolkningen av bildene. Det er lite sannsynlig at mengden undersøkelser vil gå ned ettersom stadig flere medisinske tilstander og kontrollregimer, f.eks. ulike kreftsykdommer, krever stadig mer radiologisk oppfølging. Radiologene bruker særlig mye tid på å gjennomgå store bildematerialer ved utredning og kontrollundersøkelser ved kreftforløp. Av alle bildene i en undersøkelse, som kan inneholde flere hunder bilder, vil det i snitt være få av bildene der det er aktuell patologi. Antall bilder per undersøkelse som må screenes for å finne patologi har økt gradvis over tid, og dette kan føre til at det blir stadig mindre tid per undersøkelse med derav økende risiko for å overse patologi. En effektiv AI-applikasjon kan imidlertid gjennomgå store datamengder på kort tid uten at kvaliteten forringes. Forutsetningen er at applikasjonen er treffsikker og har høy kvalitet. Ved å bruke AI til screening-arbeidet kan radiologene bruke sin kompetanse mer effektivt.

Radiologer er høyt utdannet og kostbar arbeidskraft i et begrenset antall. Samfunnet har brukt store ressurser på utdanning av denne arbeidskraften. At man ønsker å bruke denne

kompetansen så effektivt som mulig samtidig som man ivaretar kvaliteten på arbeidet som utføres er naturlig både med hensyn på pasientsikkerhet og ressursfordeling men også i et samfunnsøkonomisk perspektiv. Man vil bruke radiologene til oppgavene som krever deres faglige ekspertise og lange medisinske utdannelse. Et overordnet mål for anvendelse av AI-verktøy i radiologi må være avlastning av tidkrevende oppgaver for å styrke diagnostikk basert på faglige kriterier.

4.2.2 Standardisering

Det er ønskelig å tilstrebe større grad av standardisering i radiologien med mindre inter- og intraobservatørvariabilitet. Det er et klart mål å ha minst mulig variasjon i tolkningen mellom to radiologer og i variasjonen av hvordan samme radiolog tolker bilder over tid. AI vil trolig kunne gi verktøy som reduserer slik intra- og interobservatørvariabilitet. Ved bedret standardisering vil man også kunne oppnå høyere og mer forutsigbar kvalitet på de radiologiske tjenestene.

Helsetjenesten ønsker å tilby likeverdige tjenester uavhengig av hvor man er bosatt. Der som man har effektive AI-programmer som løfter kvaliteten på bildediagnostikken vil dette komme pasientene til gode. Dette forutsetter at tilgjengeligheten til disse AI-programmene også er lik.

4.3 AI i en radiologisk avdeling

AI gir et bredt spekter av muligheter innen effektivisering og kvalitetsforbedring innen arbeidsflyten på en radiologisk avdeling. Kapittelet gir en skjematisk oversikt over enkelte områder der AI kan bidra med effektivisering av oppgaver på en radiologisk avdeling(6), summert opp i figur 4.1.

4.3.1 Bildeopptak

Ved både røntgen, CT, MR og ultralyd er det flere faktorer som avgjør hvor godt bildeopptaket blir. Pasienten må posisjoneres riktig. Bildeopptaket må i enkelte tilfeller synkroniseres med pustefrekvens eller hjertefrekvens og det finnes programmer som korrigerer for bevegelsesartefakter. Metall i proteser, andre implantater eller andre fremmedlegemer kan korrigeres for gjennom programmer og måter man gjennomfører undersøkelser. Informasjon og instruksjoner må formidles til pasienten på språk som pasienten forstår. Det skal vurderes kontrastmengde og kontrastfase når bildeopptaket gjøres. Alle disse trinnene er områder der AI kan være nyttige hjelpemidler for optimalisering. Et godt bildeopptak legger grunnlaget for de neste trinnene i arbeidsflyten.

4.3.2 Preprosessering

Når bildeopptaket er gjort må bildene prosesseres og det lages rekonstruksjoner i ønskede plan i 3 dimensjoner der det er aktuelt. Det legges på ønskede filtre i henhold til gjeldende protokoller. I preprosesseringen bearbeides rådataene fra bildeopptaket for å bedre kvaliteten og redusere støy.

4.3.3 Bildeframstilling

Bildene som sendes til bildefremvisningsprogrammet PACS må så henges opp på en hensiktsmessig måte, for å gi best mulig oversikt over relevante bilder, sammenlignet med eldre aktuelle undersøkelser. I tillegg er det ønskelig at snittplanene mellom nye og gamle undersøkelser er synkronisert, slik at radiologen til enhver tid har fremstilt samme område av kroppen når vedkommende ser gjennom bildene. Dette gjør sammenligningen enklere, mer sømløs og tidsbesparende. Her kan AI-algoritmer bidra til å automatisere hengningsprotokoller, hente opp relevante eldre undersøkelser fra arkivet og presentere bildene på en hensiktsmessig måte. Ved omfattende og komplekse undersøkelser med mange serier/sekvenser er dette være en tidkrevende oppgave som krever radiologressurser, som kan brukes bedre til tolkning og klinisk vurdering av bildene.

4.3.4 Bildetolkning

Dette er i kjernen av radiologenes oppgaver. Først og fremst må foreliggende patologi oppdages, noe som innebærer at radiologen må se gjennom alle bildene til aktuelle undersøkelser. For en CT eller MR undersøkelse innebærer dette ofte flere hundre og i enkelte tilfeller flere tusen bilder i en enkelt undersøkelse. En slik kvantitativ gjennomgang av store bildematerialer kan åpenbart effektiviseres med AI-programmer. Når patologien er detektert må den segmenteres gjennom kvantifisering av størrelse, form, avgrensning mot omkringliggende strukturer, tetthet/signal, kontrastoppladning. Dette danner grunnlaget for sannsynlig diagnose og differensialdiagnoser inkludert stadie-inndeling (staging) der det er aktuelt. Staging vil si at sykdomsutbredelsen klassifiseres i henhold til fastsatte kategorier, hyppig brukt f.eks. i kreftutredning. I tillegg vurderes behov for supplerende diagnostikk eller ytterligere radiologiske undersøkelser for avklaring av eventuelle funn, samt foreslå oppfølging eller kontroll av funn.

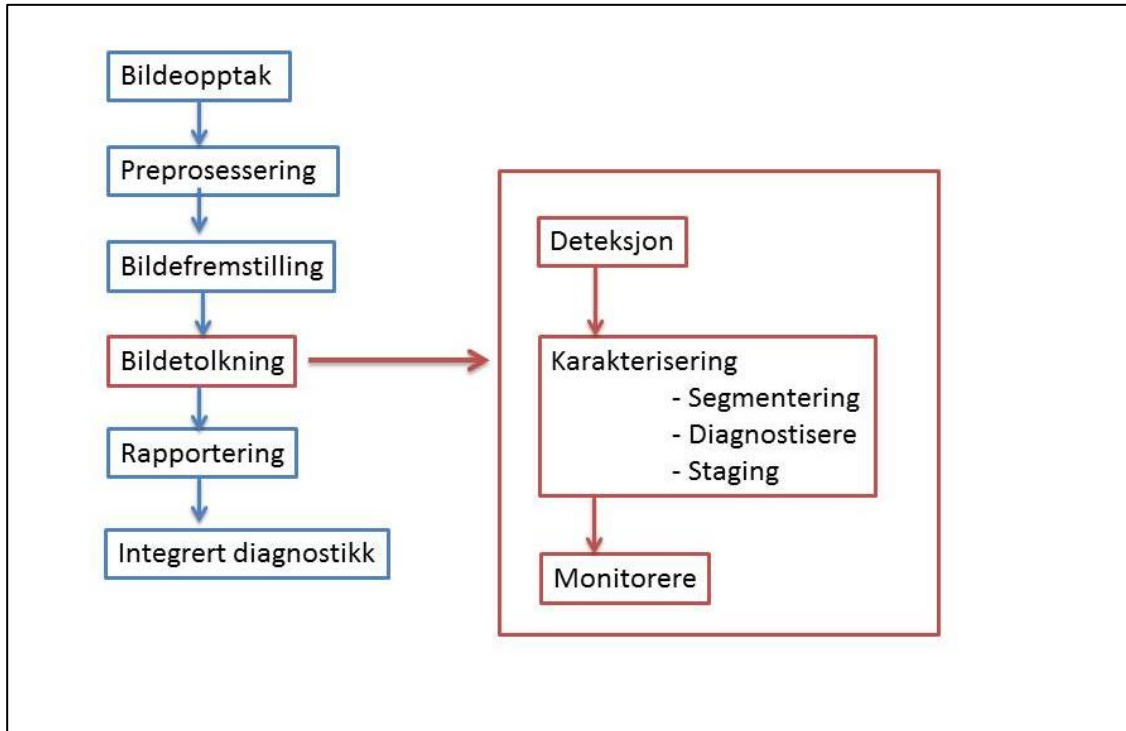
4.3.5 Rapportering

På bakgrunn av tolkningene rapporteres funnene i en radiologisk rapport, et journaldokument som blir tilgjengelig i pasientens journal. Rapporten skal ha en logisk struktur og må inneholde relevante bildefunn som er formidlet på en måte som gjør at kliniker oppfatter disse korrekt, for dermed å gi pasienten riktig og god behandling. Standardisert rapportering i beskrivelser er et tema som har vært aktuelt de siste årene og stadig flere tar dette i bruk. Mens radiologenes beskrivelser delvis kan være subjektive med unntak av noen allmenngyldige krav til oppsett og form, vil en AI- applikasjon standardisere beskrivelsene og gi mer forutsigbar struktur på informasjonen som formidles. Talegjenkjenning er et verktøy som brukes i de fleste radiologiske avdelinger og institutter. Talegjenkjenning innebærer at radiologene leser inn beskrivelsen i en diktafon, hvorpå en talegjenkjenningsprogramvare overfører dette til et tekstdokument fortløpende. Rapporten vil således være tilgjengelig som en tekstfil i journalen straks radiologen er ferdig med diktatet. Dette er en tydelig forbedring fra tidligere der vanlig praksis var at det ble opprettet et diktat som i ettertid ble skrevet inn i journalsystemet av en sekretær. Dette kunne ta timer til dager og faren for at viktige diagnostiske opplysninger ikke ble videreført til henviser var større.

4.3.6 Integrert diagnostikk

Integrert diagnostikk refererer til en multidisiplinær synkronisering av diagnostisk informasjon for å øke presisjonen i diagnostikk. Ved å kombinere f.eks. histologisvar fra patolog med radiologisk bildeinformasjon (radiomics) og eventuelt genetisk profil åpnes muligheter for presisjonsmedisin på et nytt nivå. AI-algoritmer gir muligheter for å kombinere disse ulike informasjonskildene i store

datamengder og dermed selv finne karakteristika og mønstre som ikke tidligere har vært mulig. Her åpner det seg muligheter for å evaluere alvorlighetsgrad og prognose i større grad enn man har kunnet frem til nå. Man vil potensielt kunne vurdere grad av risiko og tilpasse behandlingstilbud mer individuelt til hver pasient.



Figur 4.1 Trinnene fra bildeopptak til diagnostikk der AI kan ha en rolle i forbedring og effektivisering.

4.4 AI-verktøy i radiologisk diagnostikk

4.4.1 Computer aided detection

Tradisjonell CAD (computer aided detection/diagnosis) har vært utviklet fra 1980-tallet og har vært brukt som et hjelpemiddel i deteksjon og karakteristikk av lungenoduli (knuter i lungevevet) samt knuter/oppfyllinger i brystvevet ved mammografi(21). CAD har vært brukt som et diagnostisk hjelpemiddel av radiologer for å unngå å overse lesjoner, samt hjelp til kvantitativ analyse av aktuelle lesjoner. Det har med andre ord ikke erstattet radiologens oppgave men fungert mer som en «second opinion» og et ekstra sikkerhetsnett for å oppdage potensiell patologi. CAD-programvarene gjør en kvantitativ analyse av bildedataene og peker på områder som den tolker som avvikende fra normal anatomi. Radiologen må gjennomgå funnene og vurdere om det er aktuell patologi eller falsk positiv. Radiologene tar fortsatt den endelige avgjørelsen når det skal konkluderes i svarrapporten. Synergieffekten kommer når CAD har sin styrke i å gjennomgå store bildematerialer mens radiologen har den faglige og intellektuelle kapasiteten til å vurdere og fortolke funnene. Som følge av dette har CAD blitt tatt i bruk i klinisk diagnostikk og mange leverandører tilbyr slike programvarer. I de to tidligere nevnte kliniske eksemplene, brystdiagnostikk og lungenodusdeteksjon er CAD egnede verktøy(22, 23).

Via teknologiske fremskritt med AI ser man potensialet for en videreutvikling av verktøy som CAD. CADe (computer aided detection) er det tradisjonelle CAD-systemene som assisterer i deteksjon av patologi/abnormiteter. CADx-systemer referer til AI systemer som assisterer i diagnosesetting, altså karakterisering av patologi systemet detekterer(24). AI gir muligheter for at programmene kan sammenligne med tidligere undersøkelser, beregne volum og størrelsesendring/vekstrate, samt gjøre en risikovurdering for malignitet av aktuelle lesjon. DL-baserte CAD-systemer har vist seg mer effektive enn tradisjonelle CAD-systemer(22). Det er forventet at man etter hvert vil få tilgjengelige programmer for både deteksjon og vurdering av malignitetsrisiko av lungenoduli(25). Flere kliniske AI-systemer har blitt godkjent til klinisk bruk. R2 Image Checker system ble FDA-godkjent i USA i allerede 1998 og er et støtteverktøy for deteksjon i mammografiscreening(26). QuantX Advanced system er et annet eksempel på et godkjent program i klinisk bruk som assisterer i brystkreftdiagnostikk, et CADx-system(6).

4.4.2 Segmentering

AI for automatisert segmentering av ulike organer er gjennomgått i en systematisk oversiktsartikkel fra 2019 (27). Artikkelen viser at det er flest studier som tar for seg segmentering innen nevreradiologi. Segmentering er brukt for vurdering av multippel sclerose, gliomer,

hvitsubstanslesjoner, slag, med mer. Innen thoraxradiologi er segmentering brukt ved kardiovaskulære sykdommer, hjerteinfarkt, og lungesykdommer. Ved muskel-skjelettradiologi er columna (rygggraden) området der segmentering er mest brukt, for eksempel ved detektering av metastaser eller annen kreft. Andre områder av skjelettet er også evaluert med hensyn på tilstander som osteoporose, artritt, artrose, fraktur, prolaps, degenerasjon med mer. Innen abdominalradiologi er det gjort studier på segmentering av organer som lever, pankreas, milt, nyrer, prostata, colon og karstrukturer. Ved brystradiologi var segmentering brukt til deteksjon og evaluering av lesjoner i bryst. Fettvevssegmentering var brukt til evaluering av fettinnhold intraabdominalt, subcutant og omkring organer i evaluering av tilstander som fedme, diabetes, metabolsk syndrom og artrose. Artikkelen gjør ikke opp status vedrørende hvordan disse AI algoritmene presterer, men gir en oversikt over områder der AI er under utvikling og områder der AI kan ha en plass i fremtidens radiologi.

4.4.3 Radiomics

Radiomics handler om hvordan man gjennom AI kan hente ut data og karakteristika fra bildene, som man visuelt ikke kan gjøre med det menneskelige øyet(28). I radiologien er det behov for å kunne skille farlige funn fra ufarlige funn og kunne karakterisere funnene ut fra en sannsynlighet/risiko. I mange tilfeller er det nødvendig med vevsprøve fra en oppfylling/tumor for å fastslå endelig diagnose, samt vurdere prognose og behandlingsmuligheter. Radiomics kan gi økt presisjon i bildediagnostikk, differensiere benigne fra maligne lesjoner, predikere behandlingsrespons og prognose(3), som et bidrag til persontilpasset medisin, et tema som er stadig mer aktuelt.

I tabellen under (tabell 4.1) gis en skjematisk oversikt over områder der AI kan bidra i ulike organseksjoner av radiologi.

Organseksjon	Muligheter med AI-verktøy
Nevroradiologi	<p>Triage av CT-caput.</p> <p>Segmentering, volumberegning, deteksjon og diagnoseassistanse.</p> <p>Vurdering av demens.</p> <p>Diagnostisere og kvantifisere karokklusjoner.</p> <p>Diagnostisering av hjerneslag.</p> <p>Kvantifisere og segmentere MS utbredelse.</p> <p>Hypofysediagnostikk ved MR.</p> <p>Differensiering og karakterisering av hjernetumores.</p> <p>Evaluerer av spinal stenose.</p>
ØNH-radiologi	<p>Planlegging av stråleterapi ved kreft i hode/halsregion. Prediksjon av prognose og bivirkninger.</p> <p>Vurdere osteoporose og annen patologi ved maxillofacial radiologi.</p> <p>Vurdering av oppfyllinger av thyroidea ved UL, karakterisering og diagnostisering.</p>
Thoraxradiologi	<p>Deteksjon av noduli ved CT. Karakterisering, segmentering og risikovurdering av lesjoner.</p> <p>Segmentering og vurdering av kardiovaskulære og pulmonale sykdomstilstander.</p> <p>Segmentering og diagnostisering ved CT/MR/UL hjerte.</p> <p>Prognosesetting ved lungetumores.</p> <p>Deteksjon av tuberkulose ved røntgen thorax.</p>
Brystradiologi	<p>Screening mammogrammer, sile ut negative og detektere mulig patologi.</p> <p>CAD for deteksjon. Detektere mikrokalk.</p> <p>Segmentering og prediksjon av overlevelse.</p> <p>Radiomics og persontilpasset medisin.</p>
Abdominalradiologi	<p>Segmentering og klassifisering av leverlesjoner og HCC.</p> <p>Vurdering av fettlever.</p> <p>CT-colografi: Deteksjon av polypper/veggforandringer.</p>
Urologisk-radiologi	<p>Deteksjon og karakteristikk av prostatalesjoner ved multiparametrisk MR.</p>
MSK-radiologi	<p>Vurdering av skjelettalder.</p> <p>Deteksjon av metastaser i skjelett.</p> <p>Deteksjon av fraktur.</p> <p>Vurdere osteoporose, artritt og artrose.</p>
Administrativt	<p>Språkprosessering. Fremheve funn i beskrivelser, kategorisere og merke undersøkelser. Generere strukturerte beskrivelser.</p>

Tabell 4.1 Sammenfatning av områder der AI kan gi diagnostiske verktøy innen radiologi

4.4.4 Nevroradiologi

Case triage er et eksempel på hvordan AI-program kjører i bakgrunnen og legger de mest kritiske undersøkelsene øverst på listen, slik at radiologene får en prioritert rekkefølge på hva som må beskrives. Dette kan forebygge at viktige undersøkelser med kritisk funn må vente unødvendig. I 2018 ble det publisert en studie i Lancet (29) der man så på et stort antall CT-caput (hode) retrospektivt der det ble laget en DL-algoritme som skulle identifisere alvorlig patologi som intrakraniell blødning, masseeffekt, skallefrakturer og midtlinjeforskyvning. Studien viste at algoritmen presist identifiserte alvorlig patologi som trengte raskt diagnostikk - lovende resultater for bruk av en slik algoritme i triagering av CT-caput.

AI-applikasjoners nytteverdi og innflytelse på arbeidsoppgavene til radiologene innen nevreradiologi er undersøkt i en review fra 2020 (30). De har avgrenset søket til perioden 2017-2019. Konklusjonen er at AI innen nevreradiologi ikke bare er i utviklingsfasen, men at flere applikasjoner er i klinisk bruk. De fleste fungerer som støtteverktøy for radiologene i deteksjon og tolkning, samt utvider radiologens muligheter innen diagnostikk. Eksempelvis finnes det tilgjengelige applikasjoner som gjør beregning av hjernevolum sammenlignet opp mot en normativ database, noe som er nyttig i diagnostikk av demens og Alzheimer. Det er applikasjoner for deteksjon av hjerneslag og blødning på CT, deteksjon av abnormiteter i hjernevevet, deteksjon av karokklusjoner samt anatomisk segmentering i bildene. Det finnes applikasjoner som tolker bildefunn og foreslår diagnoser, samt oppretter strukturerte rapporter/beskrivelser med aktuelle bildefunn. Foreløpig er det relativt snevre AI-applikasjoner som kan utføre eller støtte radiologene i enkelte av oppgavene, men de kan foreløpig ikke erstatte radiologene.

Ved hjerneslag er det avgjørende for behandlingen at pasienten kommer raskt til sykehus og får utført rask diagnostikk for iverksetting av nødvendig behandling som trombolyse (blodpropløsende medikamenter) eller endovaskulær trombektomi (fjerning av blodpropp ved intervensjon via arterier)(31-33). I denne diagnostikken brukes CT eller MR, oftest CT. Den radiologiske vurderingen av bildene er subjektiv og avhengig av ekspertise, der det kan være ulike grader av forsinkelser i tolkningsprosessen og fare for å overse patologi. En systematisk oversikt fra 2020(31) så på hvordan AI-applikasjoner kan assistere radiologene i diagnostisering av hjerneslag som følge av okklusjon av cerebrale arterier. De gikk gjennom studier fra 2014-2019. I studien er det vurdert flere AI-applikasjoner for CT-diagnostikk som er tilgjengelige på markedet, og det varierer hva disse kan tilby av automatisert perfusjon, bildeanalyse av angiografi og deteksjon av akutt slag. Konklusjonen er at AI kan forbedre deteksjon og gi raskere triagering ved karokklusjon i forbindelse med hjerneslag, men at applikasjonene foreløpig egner seg som

supplement og hjelpemiddel for radiologene i diagnostikken, ikke som en erstatning. De kommenterer at det er relativt store sprik i resultatene av ulike algoritmer og sensitiviteten er til tider lav, median 68 %. AI-applikasjonene har på nåværende tidspunkt sine svakheter og forbedringspotensialer. Årsaker til feiltolkning av algoritmen beskrives å være bevegelsesartefakter, anatomiske variasjoner som slyngede kar, timingen på kontrastfasen og kontrastdose samt eldre skader på hjernevevet. Samtidig etterlyser de studier med større grad av standardisering i vurderingen av hvordan prestasjonene til AI-applikasjonene måles.

Sammenligning av prestasjoner mellom en AI-algoritme mot radiolog i tolkning av MR bilder ved mistanke om akutt hjerneslag er undersøkt i en studie fra 2020 (34). Studien viste signifikant bedre sensitivitet for AI-algoritmen, men lik spesifisitet sammenlignet med radiologisk tolkning.

AI har også vist seg nyttig i kvantitativ volumetrisk måling av cerebrale lesjoner ved MR undersøkelser og dermed sykdomsutbredelse hos MS-pasienter (35).

Differensiering mellom primært CNS-lymfom og glioblastom ved hjelp av en AI-applikasjon er vurdert i en studie fra 2018 (36). De to sykdomstilstandene kan ha relativt likt utseende på radiologiske undersøkelser, men ganske ulike behandlingsregimer. Man har derfor sett på mulighetene for å differensiere mellom disse diagnosene uten å måtte ta biopsi/vevsprøve for histologisk verifisert diagnose. Konklusjonen er at det foreløpig er relativt få studier som har vurdert AI til bruken av dette, men at de få studiene som er gjort viser lovende resultater og har potensial til å assistere radiologer i vanskelige kasus. I studiene som sammenlignet radiolog mot AI-algoritme presterte AI-algoritmen bedre eller like bra som radiologene.

AI kan brukes til å differensiere hypofysitt fra hypofysemetastaserings ved behandling med immunterapi ved kreft(37). Studien konkluderer med at algoritmen kan være nyttig i klinisk bruk som et verktøy i beslutningsstøtte ved aktuelle problemstilling.

4.4.5 Øre-nese-hals-radiologi

Diagnostiseringen av maligne thyroideaknuter på ultralyd ved hjelp av CAD er vurdert i en systematisk oversikt fra 2019(38). De konkluderer med at tradisjonelle CAD og DL-basert CAD presterer godt i forhold til radiologer hvorav DL-basert CAD har like god sensitivitet og spesifisitet som radiolog. I en systematisk oversikt fra 2019 har man også sett på CAD-systemer ved diagnostisering av thyroideaknuter på ultralyd(39). Konklusjonen deres er at sensitiviteten av CAD-systemene er like gode som diagnostisering av radiologer, men spesifisiteten var lavere enn hos erfarne radiologer. I denne studien ble det også påpekt at CAD-systemene brukte en del tid på å prosessere og tolke bildene, lengre tid enn radiologene. Altså var det i dette tilfellet ingen tidsbesparelse å hente. De mener systemene har en god vei å gå før de kan erstatte radiologene. En systematisk oversikt fra 2019 så på bruken av CAD i diagnostisering og karakterisering av thyroideaknuter ved ultralyd(40). Studien så på hvilken rolle CAD kan ha i å assistere radiologene og om bruken av CAD kan redusere behovet for vevsprøve/ FNAC (fin-nål aspirasjons cytologi). Studien konkluderte med at CAD presterer mer nøyaktig i deteksjon av benigne knuter, men har samme nøyaktighet som radiologer i å detektere maligne knuter, altså kan CAD redusere behovet for FNAC, men foreløpig ikke erstatte det ettersom FNAC fortsatt er gullstandard for å påvise malignitet. De påpeker at CAD kan være en nyttig metode som «second opinion» for radiologene i klinisk praksis.

AI-algoritmer er også testet ut i planlegging av stråleterapi ved hode/halskreft, samt evnen til å predikere utfall, som metastasering og residiv ved bruk av radiomics (41). Studien adresserer utfordringer AI-algoritmen får når ulike utstyrsparemetere ikke er standardisert, som rørstrøm, rørspenning samt kontrastmengde og -fase. I konklusjonen nevner de stråleplanlegging og dosekalkulering som tidkrevende oppgaver AI-algoritmer kan automatisere. Trolig kan AI gi bedre verktøy for modellering og prediksjon av bivirkninger samt prognose, som et skritt i retning av persontilpasset og prediktiv medisin. Samtidig som de påpeker at AI gir lovende utsikter til applikasjoner, er det foreløpig behov for mer kunnskap og evidens før disse kan brukes i klinisk praksis.

Tolkning av MR-bilder ved hode/hals-kreft ved hjelp av AI er testet ut en studie (42). De konkluderer med at det er for lite litteratur på dette feltet på nåværende tidspunkt og at metodiske variasjoner mellom studier gjør det vanskelig å sammenfatte og sammenligne resultater.

En systematisk oversikt fra 2020 sett på studier som vurderer AI-algoritmer i dental- og maxillofacial radiologi(43). Algoritmene som ble evaluert hadde egenskaper som å markere

anatomiske landemerker, diagnostisering av osteoporose, klassifikasjon og segmentering av cyster og tumorer samt identifikasjon av periodontitt. Konklusjonen er at det er variasjon i diagnostisk prestasjon mellom de ulike AI-algortimene og det trengs ytterligere forskning og validering før applikasjonene kan implementeres i klinisk praksis.

4.4.6 Thoraxradiologi og brystdiagnostikk

I Norge har vi et screeningprogram for brystkreft, noe som innebærer at det genereres store mengder mammogrammer. Å ha bildeverktøy som screener disse bildene og peker på suspekte områder vil være nyttig avlastning for radiologene.

Mammografiscreening med DL-AI-algoritme som screening-verktøy for å sile ut normale mammogrammer er undersøkt i en svenske studie fra 2020(44). Undersøkelsen konkluderer med at ved bruk av AI-algoritmen kan hvert femte screening-mammogram ekskluderes fra radiologisk granskning uten å overse noen cancere, samtidig som antall falskt positive også reduseres. Med hensyn på at normal praksis er at mange undersøkelser dobbeltgranskes vil besparelsen i kostnader og radiologisk kapasitet være stor. Studien påpeker videre at man vil forvente stadig bedre prestasjoner fra slike AI-algoritmer, noe som vil kunne automatisere en stor del av screening-arbeidet med mammogrammer i fremtiden. Kreftregisteret i Norge er i gang med et lignende prosjekt. Målet deres er å utvikle en AI-basert metode for å sile ut 70% av negative undersøkelser, der de 30% resterende skal tydes på ordinær måte med to radiologer (45).

Studier har vist gode resultater ved bruk av «deep learning»-algoritmer for deteksjon av mikrokalk i brystlesjoner – en prediktiv indikator for tidlig brystkreft(46). Bruken av AI for å hente ut ulike biomarkører for diagnostikk og prognosevurdering ved brystkreft, deriblant radiomics er vurdert i en studie fra 2018(47). De konkluderer med at potensialet er stort for radiomics i retning av økt presisjon og persontilpasset medisin. Det er behov for store datasett som spenner over ulike sentre og land for å lage gode nok databaser til utvikling av slike algoritmer. Å skaffe slike datamengder er en stor utfordring.

Bildeanalyse av MR-bryst ved hjelp av AI-algoritmer i kreftdiagnostikk er undersøkt(48). Studiens konklusjon er at applikasjonene foreløpig er det stykke unna klinisk implementering, og at ytterligere utvikling er nødvendig for å øke prestasjonsnivået. AI er brukt til brystkreftevaluering ved automatisert segmentering av «mest kontrastoppladende tumorvev» før og etter første syklus med kjemoterapi, en studie fra 2018(49). Studien viste at AI kan predikere residivfri overlevelse ved brystkreft.

Ved CT-thorax i utredning av ulike sykdomstilstander vil det ofte dukke opp fortetninger eller knuter i lungene som i medisinsk terminologi kalles noduli. Avhengig av karakteristika er det en del av disse som skal følges opp med regelmessige kontroller med CT-thorax. Aktuelle forandringer følges ofte med flere påfølgende undersøkelser der det vurderes om de har endret seg i størrelse og om det har tilkommet nye forandringer. I gjennomgangen av primærundersøkelsen og oppfølgingsundersøkelsene må radiologene gjennomgå store bildematerialer og det vil være noe varierende hvor flinke radiologene er til å fange opp aktuelle lungenoduli. Dette er en oppgave som kan effektiviseres og kvalitetssikres ved at man bruker hjelpemidler som CAD og mer avanserte DL-algoritmer, noe som i flere studier har vist seg nyttig i utredning og oppfølging av lungenoduli(22, 50). En oversiktsartikkel fra 2020 så på DL-algoritmer til bruk i deteksjon av lungenoduli ved CT(12). Studien konkluderer med at det finnes flere metoder for å designe applikasjoner, men av undersøkte metoder fant de ikke en definert metode (gullstandard) som presterte godt i alle trinnene i prosessen ved automatisert nodulusdeteksjon.

Deteksjon av lungetuberkulose ved bruk av AI-algoritme er undersøkt i en studie fra 2019(51). I studien påpekes det at de fleste studiene er utviklingsstudier og at det er få kliniske studier. I tillegg er det store sprik i rapportert presisjon på evaluerte AI-algoritmer. Det bemerkes også flere svakheter i studiene med potensiell overestimering av prediktiv styrke av algoritmene. Konklusjonen er at AI-baserte algoritmer er lovende, men mer forskning og kliniske studier er nødvendig i fremtiden.

AI-algoritmer viser også lovende resultater i bruk ved prognosesetting, noe som er testet ut ved CT tatt av pasienter med ikke-småcellet lungecancer der pasienter ble behandlet med enten stråleterapi eller kirurgi (52). Konklusjonen deres er at DL kan brukes i mortalitetsrisiko stratifisering og at ytterligere forskning og utvikling bør gjøres på dette feltet.

Innen radiologisk hjertediagnostikk gir AI flere muligheter. I en artikkel fra 2019 gis en skjematisk oversikt områder der AI kan brukes(53). Segmenterings-algoritmer ved MR kan kvantifisere ende-systolisk og ende-diastolisk volum, ejeksjonsfraksjon og hjertekammerstørrelse med tilhørende hjertemuskelvolum. AI-algoritmer har vist gode resultater i diagnostisering av trange partier i kransearteriene (som forsyner hjertemuskulaturen med blod) ved nukleærmedisinske myocard-perfusjons-teknikker(54). Kalsiumscore (kvantifisering av mengden kalk i kransearteriene) kan segmenteres fra CT-hjerte ved hjelp av AI. Flere studier har sett på ultralydundersøkelse av hjertet der AI-algoritmer utfører oppgaver som å segmentere, identifisere og kvantifisere struktur og funksjon samt oppdage patologi(55, 56).

4.4.7 Abdominal- og urologisk radiologi

CAD brukes også i deteksjon av polypper ved CT-colografi, en undersøkelse der pasienten etter en dedikert CT-protokoll får undersøkt tykktarmen og radiologen kan virtuelt gå gjennom tykktarmens lumen for å oppdage polypper eller tumor. CAD er et hjelpemiddel for økt deteksjon, men gir samtidig mange falske positive funn(57).

AI kan brukes for å segmentere lever med hensyn på fettinfiltrasjon og potensielt leverskade som følge av non-alkoholisk steatohepatitt (58). Studien viste at algoritmen ga gode resultater i forhold til manuell vurdering.

Segmentering og klassifisering av leverlesjoner og hepatocellulært carcinom (HCC) ved bruk av AI er vurdert i en systematisk oversikt fra 2019 (59). Studien så på algoritmenes evne til å skille HCC fra andre leverlesjoner, skille HCC fra levercirrhose (leverfibrose) og utvikling av nye tumorer. Studien konkluderer med at det er tilgjengelige AI-algoritmer som har evne til å utføre de ovenfor nevnte oppgavene men at studiene har svakheter og det mangler gode nok data og studier for å gjøre en god nok vurdering av nøyaktighet og sensitivitet. Anbefalingen er at det bør gjøres multisenterstudier som inkluderer større datasett.

Multiparametrisk MR for deteksjon av prostatacancer har blitt et veletablert diagnostisk verktøy i klinisk praksis. Tolkning av MR-sekvensene krever høy kompetanse hos radiologene, en subspecialisert oppgave innen radiologi. Avhengig av kompetansenivå er det en risiko for interobservatørvariabilitet i hvordan bildene tolkes. AI-algoritmer er forventet å øke sensitiviteten i deteksjon og redusere interobservatørvariabilitet, der flere studier har vist lovende resultater (60-62). Ved å kombinere og korrelere bildekarakteristika med karakteristika fra patologi/histologi, som ledd i integrert diagnostikk, ser man potensialet i å øke presisjonen til AI-algortimene med hensyn på deteksjon, klassifisering og prognostisering(60). For videre utvikling innen dette feltet er det behov for større og bedre datasett med kvalitetssikret merking.

4.4.8 Muskel-skjelett-radiologi

Flere studier viser at AI-algoritmer gir muligheter innen diagnostikk av sykdomstilstander i ryggen som segmentering, CAD, og sykdomsklassifisering (63-66). Man har for eksempel sett på AI som hjelpemiddel for å diagnostisere og vurdering av forløp ved spinal stenose, samt preoperativ planlegging og prediksjon av postoperativt resultat. AI har vist seg som et effektivt hjelpemiddel i beslutningsstøtte for klinikere.

Vurdering av skjelettalder kan utføres av AI-algoritmer (67). Ved enkelte indikasjoner tar man røntgen av ulike deler av skjelettet for å estimere skjelettalder som ved modenhet av skjelettet, sammenlignet mot kronologisk alder. Konklusjonen er at AI gir muligheter for standardisering og redusert interobservatørvariabilitet, noe som er gunstig i disse vurderingene. Det er et effektiviseringspotensial ved tidsbesparelse av disse arbeidsoppgavene ved bruk av AI. Vesentlig variasjon i hvordan de ulike studiene rapporterer prestasjonene til algoritmene gjør det vanskelig å sammenligne dem.

4.4.9 Andre områder

Språkprosessering, i engelsk litteratur omtalt som «natural language processing», er et område som har vært forsket mye på innen AI. Det handler om at programmet gjenkjenner teksten i f.eks. en radiologisk beskrivelse. Dette kan brukes til å fremheve viktige radiologiske funn i beskrivelsene, men også kategorisere/merke undersøkelsene, noe som er viktig for videre utvikling av AI-programvarer for bildediagnostikk(68). Som tidligere nevnt trenger disse programmene store treningsdatasett med merkede/kategoriserte undersøkelser der fasiten er gitt. Det kan generelt kategorisere og generere store kohorter for videre forskning.

En systematisk oversikt over tilgjengelig litteratur med hensyn på potensialet for implementering av AI-basert bildeanalyseverktøy i klinisk praksis har klassifisert litteraturen i utviklingsfaser(69). Det konkluderes med at resultatene er lovende, men fremdeles umodne og i tidlige faser, med hensyn på klinisk implementering.

5. Diskusjon

5.1 Effektivisering og kvalitetsforbedring

Denne oppgaven søker å belyse ulike fagområder av radiologien der det er lovende resultater med hensyn til utvikling av AI-algoritmer som kan bistå radiologene i deteksjon, klassifikasjon, kvantifisering og segmentering, samt forslag til diagnostisering. Dette kan gi økt standardisering, mindre variasjon i tolkninger og dermed bedre kvalitet på den radiologiske diagnostikken samtidig som effektiviteten bedres. AI vil kunne fungere som et sikkerhetsnett og en «assistent» som fanger opp det radiologen kan ha oversett, som henter ut kvantitative data og som gir forslag til diagnose.

Kvalitetsforbedring er noe alle profesjoner vil ha som mål, men kvalitetsforbedring i medisinsk bildediagnostikk er vanskelig å måle(70). Det er diskutert et bredt spekter av kvalitetsforbedringsindikatorer i radiologi i en artikkel fra 2010(70). Noen av disse er kvantitative som antall undersøkelser beskrevet, ventetid på undersøkelse og beskrivelse av undersøkelse, undersøkelsestid og antall undersøkelser per maskin. Andre mer kvalitative indikatorer er falsk positive og falsk negative funn, korrelasjon av radiologiske funn mot patologiske funn og komplikasjoner. Disse kvalitetsindikatorerne kan man tenke seg at kan forbedres ved bruk av AI-algoritmer som er velfungerende og presise, men i hvilken grad kan være vanskelig å vurdere på forhånd av anskaffelse og implementering.

Det foregår utstrakt utviklingsarbeid innen AI men det er utfordrende å lage gode verktøy ettersom kompleksiteten i medisinske bilder er stor, samtidig som kravene til nøyaktighet og presisjon må være høye for at man skal kunne ta de i klinisk bruk. Til tross for lovende resultater er det viktig å ta hensyn til at teknologien er i relativt tidlig fase. Foreløpig er det varierende resultater i prestasjonene til AI-algoritmer som har vært testet ut. Noen presterer like godt og bedre enn radiologer, mens det på andre ikke er like gode resultater. Det er behov for bedre og mer standardisert forskning for å eliminere svakheter og dokumentere effekten av nye AI-programmer (71). Det er også behov for flere kliniske studier før flere av algoritmene kan implementeres i klinisk drift med den nødvendige sikkerhet og kvalitet. Det finnes allerede AI-algoritmer og CAD som har vært i klinisk bruk i flere år, primært for deteksjonsoppgaver som f.eks. brystdiagnostikk, lungenoduli og colonpolypper(6, 22, 23, 26). AI er forespeilet en sentral rolle i mammografi-screening der studier har vist gode resultater (44). Felles for oppgavene der AI presterer godt er at det er relativt snevre oppgaver, noe som foreløpig er egnet til AI-algoritmene. Hvor lang tid det vil ta å lage robuste diagnostiske algoritmer for mer komplekse oppgaver er usikkert.

I medisinen har vi frem til nå vært vant til å stole på menneskelige vurderinger og avgjørelser som tas på bakgrunn av tilgjengelig og personlig ervervet evidensbasert kunnskap. Når man skal implementere AI-applikasjoner som skal overta eller assistere legene i kognitive oppgaver knyttet til diagnostikk, prognosevurdering og behandling, reiser det noen utfordringer som bør belyses. Dette diskuteres i det følgende.

5.2 Ledelse og implementering

5.2.1 Juridiske utfordringer

Når ny teknologi utvider mulighetene for hva datamaskiner kan utføre oppstår det samtidig nye juridiske problemstillinger. Hvem står ansvarlig om AI gjør feil? Om AI-algoritmen skal kunne brukes som et selvstendig diagnostisk verktøy uten vurdering av radiolog, er det da leverandøren som er ansvarlig? Dersom AI-programvaren kun skal være et hjelpemiddel til radiologen, vil radiologen fortsatt sitte med den avgjørende vurderingen og ansvaret(72). Dette kan gjøre at leverandørene av AI-applikasjoner vegrer seg for å levere applikasjoner som gjør selvstendig diagnostikk med mål om å erstatte radiologene, ettersom de vil risikere å bli holdt erstatningsansvarlige. Det vil også være problemstillinger der AI og radiolog tolker forskjellig. Radiologen kan være usikker på sin tolkning av et bildefunn, men velger å stole på AI fremfor egen vurdering. Kan radiologen holdes ansvarlig for en eventuell feiltolkning av undersøkelsen? Det er ingen enkle svar på disse spørsmålene, og både et godt lovverk og etter hvert rettspraksis som utvikles i takt med implementering, blir sentrale forhold. Denne problemstillingen ble diskutert i en juridisk artikkel fra 2019 fra USA(73). Noe av problemet er manglende rettspraksis på dette feltet foreløpig. Som utgangspunkt mener forfatterne av artikkelen at det tryggeste for en lege vil være å bruke AI som et beslutningsstøtteverktøy, og at utgangspunktet bør være at man stoler på egne vurderinger i henhold til gjeldende god praksis og retningslinjer. Om man i fremtiden får så gode AI-verktøy at vurderingen til AI presterer bedre enn den gjennomsnittlige lege/radiolog, kan da radiologen bli holdt ansvarlig for å ikke ha fulgt anbefalingen fra AI? Vil god praksis være å stole på AI-vurderingen, fremfor den subjektive vurderingen til radiologen? De juridiske problemstillingene er et omfattende og komplekst felt og vil ikke behandles ytterligere innenfor rammene av denne oppgaven, men ved implementering av AI må man være klar over og ha vurdert problemstillingene knyttet til dette.

5.2.2 Personvern

Datasikkerhet/personvern og god håndtering av pasientdata er en etisk utfordring når man skal hente ut store mengder medisinske data fra radiologiske bildearkiver og journaler(24). Det jobbes

med å utvikle løsninger på dette. For eksempel å bruke krypterte data eller trene algoritmer lokalt fremfor å hente ut bildedata til eksterne enheter. Det er regelverk for hvilke data som kan hentes ut og hvordan de skal håndteres. Det må vurderes om samtykke fra pasienter må innhentes og dataene må anonymiseres(7).

5.2.3 Effektivisering i lys av økonomiske hensyn

Et viktig spørsmål man bør stille seg før implementering av en AI-applikasjon er i hvor stor grad programmet faktisk gir økt effektivitet og/eller kvalitetsforbedring. Kostnadseffektivitet kan skapes i flere dimensjoner; tjenesten kan bli billigere, den kan få økt kvalitet til samme pris, eventuelt kan økt kvalitet også aksepteres med økt kostnad dersom man er villig til å betale mer for høyere kvalitet.

Kan man stole på at AI har fanget opp alle aktuelle lesjoner ved AI-scanning etter for eksempel skjelettmetastaser, eller må radiologen uansett se gjennom bildene like grundig som de vanligvis ville gjort? I så fall er ikke effektiviseringspotensialet åpenbart i forhold til kostnaden ved anskaffelse av AI-programvaren. Kvalitetsforbedringen vil imidlertid kunne oppnås ved at programvaren oppdager lesjoner radiologen eventuelt kunne oversett, og således fungere som en sekundærgransker.

Vil radiologene oftere komme i konflikt med programvarens tolkninger og dermed tvile på sin egen vurdering? Vil dette gjøre det vanskeligere å konkludere, eller blir det enklere? Dette vil komme an på situasjonen. Svarene og vurderingene i disse spørsmålene vil nok være avhengig av hva slags AI-applikasjoner det gjelder og hvilke oppgaver de er ment å utføre.

Økonomien vil være en viktig faktor ved både anskaffelse og drift av AI-programmer i bildediagnostikk. Det ligger sterke kommersielle interesser bak utviklingen med en økonomisk drivkraft i tillegg til en motivasjon om å skape gode verktøy og hjelpemidler innen helsetjenesten. Man må vurdere om investeringen kan forsvares i forhold til forventet nytte. Vil nytten være størst ved store sentre med store volumer av undersøkelser, eller kan mindre avdelinger/institutter også forsvare innkjøp av denne teknologien? Dette vet man først når man har tilgjengelige priser på aktuelle produkter, et estimat på hva produktene kan gi av økt effektivitet, og der prisene er definert i et åpent, konkurrerende marked. Forventet kvalitetsforbedring og verdsetting av dette er vanskelig i praksis.

Det finnes litteratur som har sett på økonomisk nytteverdi av CAD i brystkreftscreening i en klinisk kontekst (74). Studien viste fordeler og ulemper med CAD sammenlignet med dobbelgranskning av to radiologer. Fordelen var deteksjonsrate av kreft samt tidsbesparelse ved at kun en radiolog trenger å vurdere bildene i tillegg til CAD, noe som også ga en potensiell økonomisk gevinst. Tidsbesparelsen var begrenset av at radiologene brukte lengre tid på å tolke bildene som var vurdert av CAD. Samtidig som at man fikk en tidsbesparelse ved tolkning av bildene, ble flere pasienter re-innkalt til supplerende undersøkelser ved bruk av CAD. Balansen mellom kostnader ved anskaffelse og drift av en CAD-applikasjon, lavere kostnader ved redusert dobbelgranskning med mulig noe økt re-innkalling av pasienter til supplerende undersøkelser, er krevende og må belyses. Kost-nytte effektiviteten har enkelte studier vist å være ugunstig med hensyn på implementering av CAD(75), med det nevnte perspektivet, mens andre har konkludert med en gunstig kost-nytte effekt (76).

Ved vurdering av økonomiske hensyn må det vurderes konkret ved hver anskaffelse hvor stort omfang av undersøkelser AI kan avlaste radiologene med, effektiviteten og treffsikkerheten til AI-applikasjonen, tidsbesparelsen for radiologene, pris på anskaffelse og vedlikehold samt kvalitetsforbedringsaspektet. I et økonomisk perspektiv bør man ha klart for seg hva målet skal være for implementering av AI-applikasjonen, samtidig bør det være en realistisk formening om hvor man skal hente ut gevinst og i hvilken grad.

5.2.4 Endringsmotstand

Motstand mot endring er et kjent og viktig tema. Dette gjelder ikke bare innen radiologi. Ved endringsprosesser vil det gjerne være varierende holdninger fra de som blir berørt av endringen. Noen kan være positive og andre mer skeptiske eller negative. Ledere bør være forberedt på motstand ved implementering av AI-applikasjoner som overtar eller assisterer i enkelte av radiologenes oppgaver.

Flere forhold kan gjøre seg gjeldende. Det kan dreie seg om rasjonelle spørsmål og innvendinger basert på uavklarte faglige premisser. Slike motforestillinger må tas seriøst, og kan gi stor verdi i forebygging av uforutsette problemer med nye løsninger. Men motstand kan også være mindre rasjonelt begrunnet og skyldes individuelle eller miljøbasert uvilje til endring. Slike perspektiver vil ofte kreve forskjellige tilnærminger, avhengig av hvilke mennesketyper som er involvert, individuelle interesser, miljø og virksomhet.

Motstand fra medisinske spesialister kan også ha utgangspunkt i en frykt for å miste sitt fag, sin eksklusive rolle eller arbeidsoppgaver og således oppleves som en trussel. AI kommer inn som en

faktor som kan utfordre stabiliteten og forutsigbarheten i et miljø. De fleste vil oppleve å miste noe i en endringsprosess, men ved å legitimere dette tapet vil de som påvirkes lettere ta inn over seg løsningsperspektivene(77). Utfordringen er å håndtere dette på en god måte der radiologene involveres, samt sørge for god informasjon og nødvendig opplæring. Man bør sørge for at radiologene opplever mestring i bruk av avanserte AI-verktøy, noe som også kan bidra til å redusere motstanden.

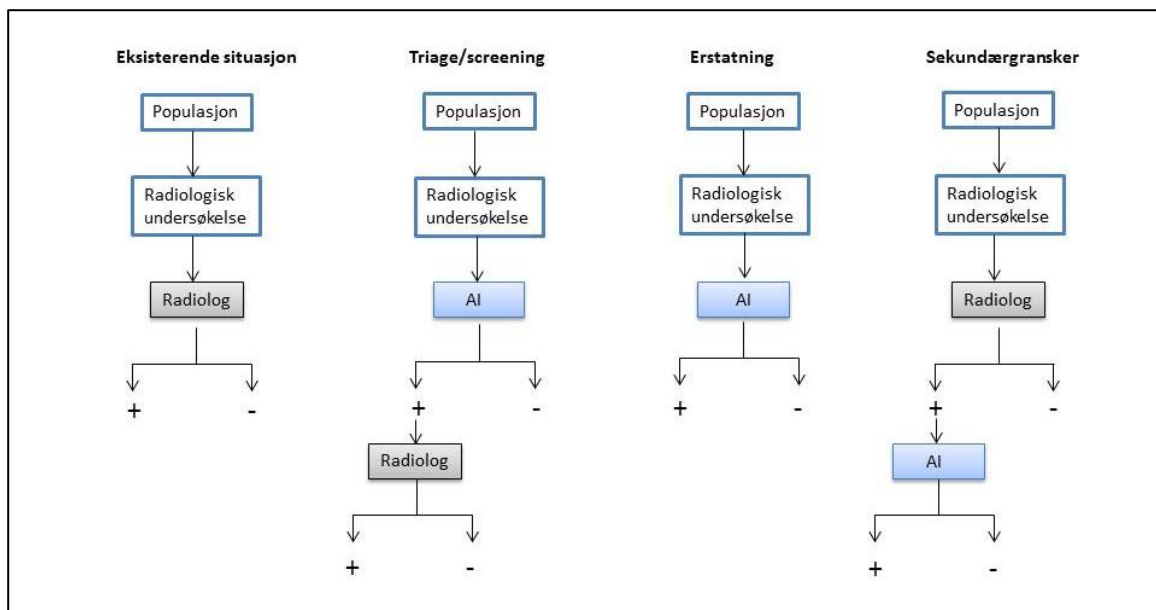
Å møte motstanden med god og evidensbasert informasjon er avgjørende for å få med seg radiologene i en konstruktiv endringsprosess der effektivitet og kvalitetsforbedring er målsetningen. Radiologer er vant til endringer og teknologiske fremskritt, noe historien har vist, men de må overbevises med evidensbasert dokumentasjon. Radiologenes kunnskap og erfaring er verdifull ved implementering og de har førstehåndskunnskap om hvor skoen trykker og hvilke verktøy man har behov for. Ved å gi radiologene bedre innsikt i hva AI kan bidra med og fremheve fordelene for radiologene kan man forhåpentligvis øke entusiasmen og redusere motstanden.

5.2.5 Strategier for bruk av AI

Det er flere måter man kan bruke AI i klinisk praksis(15). Mens man i dagens praksis tolker undersøkelsene med radiolog som kommer med et resultat/svar, vil AI kunne brukes som en triage/screening ved at undersøkelsene filtreres gjennom primærgranskning av AI, hvorpå radiolog gransker undersøkelser med positive funn. Om AI skal erstatte radiologene vil undersøkelsene tolkes av AI som gir et selvstendig svar/resultat. Alternativt kan AI brukes som et tillegg, en «second opinion» til radiologen. Undersøkelsen tolkes av radiolog og positive funn kjøres gjennom AI-algoritmen som kommer med en vurdering. Alternativene illustrert i figur 5.1, inspirert fra artikkelen til Tang A. med medarbeidere fra 2018(15).

Implementeringen bør gjøres på en brukervennlig måte, helst integrert i PACS som radiologene allerede bruker. Da vil terskelen senkes for å ta i bruk de nye verktøyene. Om radiologene må bruke flere ulike programmer samtidig vil det kunne oppleves tungvint og tidkrevende, noe man ønsker å unngå. En praktisk og brukervennlig integrering vil trolig lette overgangen til å ta AI-programmene i bruk.

Det vil være hensiktsmessig å starte forsiktig ut med velprøvde og godt dokumenterte AI-verktøy, slik at man kommer godt ut fra starten. Ved en vellykket implementering, der radiologer og annet helsepersonell ser gevinsten, øker sjansen for å bedre mottakelse ved neste implementering. Man kan også dokumentere effekten og gevinsten ved å gjennomføre et forbedringsprosjekt der man ser på gitte parametere før og etter implementering.



Figur 5.1 Ulike måter å inkludere AI-applikasjoner i klinisk drift.

5.3 Tekniske utfordringer

5.3.1 Datatilgjengelighet

En utfordring knyttet til utvikling av effektive AI-algoritmer er tilgangen til store merkede datamengder(27). Det er behov for å bygge store databaser med representative datasett som kan brukes i trening og utvikling av algoritmene. Dette er foreløpig en av de store begrensningene i utviklingen av AI-algoritmer til klinisk bruk. Å lage store databaser med merkede undersøkelser, det vil si en fasit, krever at kvalifiserte radiologer må gjennomgå slike store bildematerialer, noe som er tidkrevende og kostbart. Spesielt utfordrende er det med sjeldne sykdommer/diagnoser der man ikke har tilgjengelige store datamengder og der kompetansen blant radiologer også er begrenset(24). En annen problemstilling er om disse merkede bildene har riktig fasit. Man må spørre seg om radiologene har satt riktig diagnose/fasit på undersøkelsene. I mange tilfeller vil radiologene være i tvil og man baserer resultatet på sannsynlighet av en eller flere diagnoser, altså vil det i mange tilfeller ikke være helt opplagt et korrekt fasitsvar. Vi vet også at medisinske sannheter er ferskvare, noe som betyr at man hele veien må korrigere og justere fasiten i henhold til oppdatert evidensbasert medisinsk forskning. Vil disse algoritmene ha evnen til å tilpasse seg slike endringer?

5.3.2 Transparens

En viktig problemstilling ved implementering er utfordringen med «black box», altså at vi ikke vet eksakt hvordan algoritmen kommer frem til en gitt konklusjon. Når en radiolog har gjort en feilvurdering av en undersøkelse, kan man gå tilbake og undersøke hendelsesforløpet. Man kan da vurdere om radiologen hadde adekvat informasjon fra henvisningen, om det var manglende kunnskap, misforståelse, forstyrrende elementer som artefakter, høyt arbeidspress eller annet som kunne forklare årsaken til feilen. Radiologen har også mulighet til å forklare og analysere forløp og årsakssammenheng, noe som gjør at man kan sette inn tiltak for å forhindre lignende feil i fremtiden. Dersom man ikke kan evaluere årsaken til at en AI-algoritme gjør feil, blir det problematisk å sette inn korrigerende tiltak. Dessuten blir det vanskelig å forutse hvilke feil AI kan gjøre neste gang. Å gjøre AI-algoritmene mer transparente for å overkomme «black box»-problemet er et felt under utvikling(6, 78).

5.3.3 Overførbarhet

En praktisk utfordring ved utvikling av AI-applikasjoner er at de må kunne fungere godt selv om det er variasjoner i hvordan undersøkelsene er utført. Det er multiple variabler rundt bildeopptaksprotokoll og tekniske parametere for hvordan bildeopptaket gjøres. Ulike produsenter av radiologisk utstyr har forskjellige metoder og teknologi, noe som vil gi ulik

oppløsning, signal-contrast-ratio, støy. Det kan være variasjoner i snittykkelse, filtre som brukes og sekvenser som utføres, samt justeringer av protokoller fortløpende lokalt på hver enkelt avdeling. En undersøkelsesprotokoll for en gitt problemstilling vil således være dynamisk og stadig endre seg etter hvert som bedre metoder blir tilgjengelige og økt kunnskap erverves. En AI-algoritme må kunne tilpasse seg disse variasjonene og ha et godt nok treningsmateriale som gjør algoritmen robust under ulike forhold og forutsetninger. En AI-algoritme kan prestere godt ut fra treningsdata den har blitt presentert for, men kan vise seg å ikke prestere godt når den tas i bruk ved et annet sykehus/institutt med en annen demografi(15). Det kan være behov for lokale tilpasninger og trening før den presterer godt nok. Slike tekniske aspekter kan ikke alene vurderes av lokale avdelingsledere, men det bør være fagmiljøer på overordnet nivå i helseforetak/helsemyndigheter med AI- og radiologisk kompetanse, som kan bistå og være rådgivere for god implementering. Det vil være en fordel med kompetanse også lokalt på radiologiske avdelinger for å kunne stille de kritiske spørsmålene og være med på lokale tilpasninger samt vurdere fordeler og ulemper med de ulike AI-produktene.

5.4 Faget radiologi

5.4.1 Radiologenes fremtid

Den lovende og raske utviklingen av AI innen medisinsk bildediagnostikk de siste årene har ført med seg en rekke spekulasjoner om radiologenes fremtid. Noen har gått så langt som å hevde at radiologene i stor grad vil bli erstattet av datamaskiner(2, 79). Pendelen har svingt fra en betydelig optimisme, til mer nøkterne forventninger. Man har innsett at AI har sine åpenbare fordeler og potensiale, men samtidig begrensninger. At radiologene fullstendig skal erstattes av avanserte datamaskiner virker usannsynlig, i hvert fall innen overskuelig fremtid, tatt i betraktning spekteret av arbeidsoppgaver radiologene har. Når det etter hvert kommer gode AI-applikasjoner som kan overta enkelte av oppgavene i bildetolkningsarbeidet, eventuelt effektivisere disse oppgavene, vil radiologene fremdeles ha en viktig rolle i intervensjonsprosedyrer, multidisiplinære møter, kommunikasjon med pasienter og andre kolleger, og på den måten ha en sentral plass i helhetsvurderingen av pasientens bildediagnostikk, videre forløp og utredning vurdert ut fra pasienten kliniske kontekst(3, 72). I en slik vurdering samles trådene fra mange informasjonskilder, som pasientens alder og komorbiditet (andre samtidige sykdommer og helsetilstand), pasientens motivasjon og ønske for videre utredning eller behandling, utbredelse og utvikling av sykdomstilstanden over tid, samt andre faktorer. Dette er komplekse vurderinger som krever en integrasjon av medisinsk kunnskap, erfaring, empati, realisme, moral, etikk, samfunnsøkonomisk vurdering og tilgjengelighet av behandlings-/utredningstilbud. Foreløpig er ikke AI i nærheten av å kunne utføre slike komplekse kognitive vurderinger, som til syvende og sist er kjernen i arbeidet legene må gjøre når de skal optimalisere og individtilpasse tilbudet for pasientene. Her ligger radiologenes styrke i forhold til AI, noe som bør fremheves i en fremtidig strategi ved integrasjon av AI i radiologien(80). Man skal heller ikke undervurdere verdien av menneskelig kontakt og støtte som radiologer kan gi andre kolleger i helsevesenet og ikke minst til pasientene.

Radiologer går en spennende tid i møte der fagfeltet står foran betydelige endringer og teknologiske fremskritt. AI gir muligheter for gode verktøy innen bildediagnostikk som vil bli stadig bedre og bedre. Det er positivt at disse verktøyene vil kunne avlaste radiologene for ressurskrevende kvantitativ gjennomgang av bildemateriale, noe som vil frigjøre tid til kognitivt utfordrende oppgaver som karakterisering, vurdering, klinisk diskusjon og interaksjon med andre medisinske profesjoner samt undervisning av leger i spesialisering.

Arbeidsoppgavene, arbeidsflyten og arbeidshverdagen for radiologene har helt fra fagets opprinnelse vært i konstant endring, men i økende grad med teknologisk utvikling og nyvinninger.

Å reflektere over radiologenes fremtid som et statisk begrep blir således kunstig, for situasjonen om 10 år er forskjellig fra i dag og om 20 år ytterligere endret. Interessant er det imidlertid å reflektere over hvordan man ser for seg trendene i utviklingen, i hvert fall i nær fremtid. Å ha en overordnet oversikt over hva som finnes av tilgjengelig teknologi og hva det jobbes med å utvikle, gir en pekepinn på hvordan radiologer og radiologiske avdelinger bør forberede seg på og tilrettelegge virksomheten sin for å best mulig kunne ta i bruk verktøy som vil kunne gi en gevinst i form av effektivitet, kvalitetsforbedring og økonomisk besparelse.

5.4.2 Fagmiljøets tilnærming

Det radiologiske fagmiljøet trenger en overordnet strategi og visjon for å ta i bruk AI. Å etablere samarbeid mellom virksomheter og radiologiske miljøer for deling av erfaringer vedrørende bruk, optimalisering og implementering av tilgjengelige AI-applikasjoner kunne være både nødvendig og trolig verdifullt for å kunne gjøre gode valg av applikasjoner samt fasilitere en gitt implementering. Man kan i slike samarbeid dra nytte av hverandres erfaringer og trekke i samme retning som et samlet radiologisk miljø. Å ta i bruk AI-applikasjoner som er effektive gi konkurransefortrinn i forhold til de som ikke tar dette i bruk. For pasientene som helhet vil det være gunstig at det tilbys et likeverdig radiologisk tilbud med hensyn på tilgjengelighet, kvalitet og pris. Om det også blir tilgjengelige gode applikasjoner for karakterisering med diagnoseforslag, kan det potensielt gi et faglig løft til mindre og mellomstore radiologiske institutter og avdelinger der man ikke har de samme ressursene som seksjonerte/subspesialiserte radiologer gir, slik man har på større sykehus. Dette er igjen i trå med visjonen om å levere et helhetlig og likeverdig tilbud til befolkningen.

I lys av utviklingen innen AI og hvilken signifikant rolle dette vil spille innen radiologien for fremtiden, bør AI inkluderes som en del av utdannelsen av nye radiologer som et ledd i spesialiseringen. Radiologene trenger utdanning i hva AI er og hvordan det kan brukes i radiologi. Det er nødvendig for å kunne bruke verktøyene hensiktsmessig, forstå verdien og samtidig feilkilder. Ved revisjon og implementering av en helt ny LIS-utdanning i Norge, som ble iverksatt 2019, handler ikke et eneste av læringsmålene om AI innen radiologi. Dette er tankevekkende og mulig en indikasjon på at erkjennelsen av hvilken betydning og hvilke muligheter AI har, foreløpig ikke er etablert innen fagmiljøet. Det vil trolig være mer hensiktsmessig med en aktiv og fremoverlent tilnærming for på den måten å være med på å forme og påvirke AI innen radiologien. I LIS utdanningen bør man inkludere AI som en sentral teoretisk del, for eksempel kurset om modalitetslære. Her vil man kunne undervise i AI på lik linje som strålefysikk, MR, CT og røntgen og dermed sørge for at alle nyutdannede radiologer får en basiskunnskap i AI. Man kan også se for seg nye profesjoner med hybridkompetanse mellom radiologi og datateknologi.

6. Konklusjon

Oppgaven søker å belyse hvordan AI gir muligheter innen ulike fagområder av radiologien, men ikke alle modaliteter og fagområder er behandlet utfyllende innenfor denne oppgavens rammer. Litteraturen viser utvilsomt lovende og spennende resultater og det investeres store midler fra multinasjonale selskaper innen utvikling av AI-programmer. Vi har bare sett begynnelsen og AI vil prestere stadig bedre og trolig innen mer avanserte oppgaver i fremtiden, så det er liten tvil om at AI vil få en vesentlig plass i fremtidens radiologi. Det er behov for fremtidige gode studier som evaluerer de ulike applikasjonene som kommer på markedet og studiene bør være utformet slik at man lettere kan sammenligne de ulike produktene. Man trenger studier som evaluerer kvalitet, effektivitet og svakheter som følge av bruk av AI- applikasjoner i et helhetlig perspektiv.

Økonomiske forhold må også belyses på en stringent måte. Det radiologiske miljøet må orientere seg og holde seg oppdatert på hvilke muligheter som åpner, og ikke minst ha en visjon for hvilken rolle radiologene skal ha som en sentral brikke i den diagnostiske broen mellom radiolog og kliniker. Man må selvsagt ha fokus på utfordringer ved implementering. Forst og fremst må man ivareta forsvarlighet og pasientsikkerhet, men også problemstillinger knyttet til jus, personvern, økonomi, og tekniske utfordringer må tas hensyn til. Ledere må være forberedt på skepsis og endringsmotstand, noe som må møtes med god evidensbasert kunnskap og en tydelig fremheving av gevinstene teknologien gir for både pasienter, klinikere og radiologer. Slik kan man få med seg radiologene som delaktige ressurser i en implementeringsprosess. For radiologene gir AI lovende fremtidsutsikter med hensyn på effektivisering og mer gunstig ressursbruk ved å redusere tidkrevende screening-arbeid og dermed overkomme den stadig økende mengden undersøkelser. Faglig vil AI trolig gjøre radiologene bedre, ikke erstatte dem ved at verktøy som assisterer dem i diagnostikk og deteksjon innebærer et kvalitetsforbedrende og effektiviserende tiltak. AI åpner muligheter for økt presisjon og persontilpasset diagnostikk, noe som vil kunne gjøre arbeidsoppgavene til radiologene mer spennende og videreutviklet faget. Mange studier peker på at man fortsatt er i en tidlig fase, men gir også klare indikasjoner på hva vi kan forvente oss fremover. AI bør inkluderes som et fagområde innen utdanningen av nye LIS i radiologi for både å forberede kommende radiologer og øke kompetansen, noe som vil gjøre dem bedre rustet til å ta i bruk og være med på utvikling av fremtidens radiologiske verktøy. Ved å omfavne og samtidig involvere seg i utviklingen kan radiologene være med å forme sin egen fremtid og fagområde. En eventuell frykt for AI bør erstattes med nysgjerrighet og kunnskapssøken for å forstå mulighetene og begrensningene.

7. Referanser

1. Krittanawong C. The rise of artificial intelligence and the uncertain future for physicians. *Eur J Intern Med.* 2018;48:e13-e4.
2. Chockley K, Emanuel E. The End of Radiology? Three Threats to the Future Practice of Radiology. *J Am Coll Radiol.* 2016;13(12 Pt A):1415-20.
3. Pesapane F, Codari M, Sardanelli F. Artificial intelligence in medical imaging: threat or opportunity? Radiologists again at the forefront of innovation in medicine. *Eur Radiol Exp.* 2018;2(1):35.
4. McDonald RJ, Schwartz KM, Eckel LJ, Diehn FE, Hunt CH, Bartholmai BJ, et al. The effects of changes in utilization and technological advancements of cross-sectional imaging on radiologist workload. *Acad Radiol.* 2015;22(9):1191-8.
5. <https://www.legeforeningen.no/foreningsledd/fagmed/norsk-radiologisk-forening/underforeninger/>.
6. El Naqa I, Haider MA, Giger ML, Ten Haken RK. Artificial Intelligence: reshaping the practice of radiological sciences in the 21st century. *Br J Radiol.* 2020;93(1106):20190855.
7. Artificial intelligence and medical imaging 2018: French Radiology Community white paper. *Diagn Interv Imaging.* 2018;99(11):727-42.
8. Chartrand G, Cheng PM, Vorontsov E, Drozdal M, Turcotte S, Pal CJ, et al. Deep Learning: A Primer for Radiologists. *RadioGraphics.* 2017;37(7):2113-31.
9. McBee MP, Awan OA, Colucci AT, Ghobadi CW, Kadom N, Kansagra AP, et al. Deep Learning in Radiology. *Acad Radiol.* 2018;25(11):1472-80.
10. Hinton GE, Salakhutdinov RR. Reducing the Dimensionality of Data with Neural Networks. *Science.* 2006;313(5786):504-7.
11. LeCun Y, Bengio Y, Hinton G. Deep learning. *Nature.* 2015;521(7553):436-44.
12. Halder A, Dey D, Sadhu AK. Lung Nodule Detection from Feature Engineering to Deep Learning in Thoracic CT Images: a Comprehensive Review. *J Digit Imaging.* 2020;33(3):655-77.
13. Tajbakhsh N SK. Comparing two classes of end-to-end machine-learning models in lung nodule detection and classification: MTANNs vs. CNNs. *Pattern Recognition.* 2017;63:11.
14. Abildgaard A HE, Sakinis T, Roterud H, Bjørnerud A, Beyer M et al. Vil radiologer bli erstattet av kunstig intelligens? *Tidsskr Nor Legeforen.* 2018;138(17):3.
15. Tang A, Tam R, Cadrin-Chênevert A, Guest W, Chong J, Barfett J, et al. Canadian Association of Radiologists White Paper on Artificial Intelligence in Radiology. *Canadian Association of Radiologists Journal.* 2018;69(2):120-35.
16. Fourcade A, Khonsari RH. Deep learning in medical image analysis: A third eye for doctors. *J Stomatol Oral Maxillofac Surg.* 2019;120(4):279-88.
17. Avanzo M, Stancanello J, El Naqa I. Beyond imaging: The promise of radiomics. *Phys Med.* 2017;38:122-39.
18. OECD. <https://data.oecd.org/healthcare/magnetic-resonance-imaging-mri-exams.htm>.
19. OECD. <https://data.oecd.org/healthcare/computed-tomography-ct-exams.htm#indicator-chart>.
20. Nakajima Y, Yamada K, Imamura K, Kobayashi K. Radiologist supply and workload: international comparison--Working Group of Japanese College of Radiology. *Radiation medicine.* 2008;26:455-65.
21. Doi K. Computer-aided diagnosis in medical imaging: Historical review, current status and future potential. *Computerized Medical Imaging and Graphics.* 2007;31(4):14.
22. Cheng JZ, Ni D, Chou YH, Qin J, Tiu CM, Chang YC, et al. Computer-Aided Diagnosis with Deep Learning Architecture: Applications to Breast Lesions in US Images and Pulmonary Nodules in CT Scans. *Sci Rep.* 2016;6:24454.

23. Kooi T, Litjens G, van Ginneken B, Gubern-Mérida A, Sánchez CI, Mann R, et al. Large scale deep learning for computer aided detection of mammographic lesions. *Med Image Anal.* 2017;35:303-12.
24. Hosny A, Parmar C, Quackenbush J, Schwartz LH, Aerts H. Artificial intelligence in radiology. *Nat Rev Cancer.* 2018;18(8):500-10.
25. Causey JL, Zhang J, Ma S, Jiang B, Qualls JA, Politte DG, et al. Highly accurate model for prediction of lung nodule malignancy with CT scans. *Sci Rep.* 2018;8(1):9286.
26. Roehrig J, Doi T, Hasegawa A, Hunt B, Marshall J, Romsdahl H, et al. Clinical Results with R2 Imagechecker System. In: Karssemeijer N, Thijssen M, Hendriks J, van Erning L, editors. *Digital Mammography: Nijmegen, 1998.* Dordrecht: Springer Netherlands; 1998. p. 395-400.
27. Lenchik L, Heacock L, Weaver AA, Boutin RD, Cook TS, Itri J, et al. Automated Segmentation of Tissues Using CT and MRI: A Systematic Review. *Acad Radiol.* 2019;26(12):1695-706.
28. Lambin P, Leijenaar RTH, Deist TM, Peerlings J, de Jong EEC, van Timmeren J, et al. Radiomics: the bridge between medical imaging and personalized medicine. *Nat Rev Clin Oncol.* 2017;14(12):749-62.
29. Chilamkurthy S, Ghosh R, Tanamala S, Biviji M, Campeau NG, Venugopal VK, et al. Deep learning algorithms for detection of critical findings in head CT scans: a retrospective study. *Lancet.* 2018;392(10162):2388-96.
30. Olthof AW, van Ooijen PMA, Rezazade Mehrizi MH. Promises of artificial intelligence in neuroradiology: a systematic technographic review. *Neuroradiology.* 2020.
31. Murray NM, Unberath M, Hager GD, Hui FK. Artificial intelligence to diagnose ischemic stroke and identify large vessel occlusions: a systematic review. *J Neurointerv Surg.* 2020;12(2):156-64.
32. Albers GW, Marks MP, Kemp S, Christensen S, Tsai JP, Ortega-Gutierrez S, et al. Thrombectomy for Stroke at 6 to 16 Hours with Selection by Perfusion Imaging. *New England Journal of Medicine.* 2018;378(8):708-18.
33. Nogueira RG, Jadhav AP, Haussen DC, Bonafe A, Budzik RF, Bhuva P, et al. Thrombectomy 6 to 24 Hours after Stroke with a Mismatch between Deficit and Infarct. *New England Journal of Medicine.* 2017;378(1):11-21.
34. Lee H, Lee E-J, Ham S, Lee H-B, Lee JS, Kwon SU, et al. Machine Learning Approach to Identify Stroke Within 4.5 Hours. *Stroke.* 2020;51(3):860-6.
35. Brugnara G, Isensee F, Neuberger U, Bonekamp D, Petersen J, Diem R, et al. Automated volumetric assessment with artificial neural networks might enable a more accurate assessment of disease burden in patients with multiple sclerosis. *European Radiology.* 2020;30(4):2356-64.
36. Nguyen AV, Blears EE, Ross E, Lall RR, Ortega-Barnett J. Machine learning applications for the differentiation of primary central nervous system lymphoma from glioblastoma on imaging: a systematic review and meta-analysis. *Neurosurg Focus.* 2018;45(5):E5.
37. Mekki A, Derclé L, Lichtenstein P, Nasser G, Marabelle A, Champiat S, et al. Machine learning defined diagnostic criteria for differentiating pituitary metastasis from autoimmune hypophysitis in patients undergoing immune checkpoint blockade therapy. *Eur J Cancer.* 2019;119:44-56.
38. Xu L, Gao J, Wang Q, Yin J, Yu P, Bai B, et al. Computer-Aided Diagnosis Systems in Diagnosing Malignant Thyroid Nodules on Ultrasonography: A Systematic Review and Meta-Analysis. *European Thyroid Journal.* 2019.
39. Zhao WJ, Fu LR, Huang ZM, Zhu JQ, Ma BY. Effectiveness evaluation of computer-aided diagnosis system for the diagnosis of thyroid nodules on ultrasound: A systematic review and meta-analysis. *Medicine (Baltimore).* 2019;98(32):e16379.
40. Hesson N. The Role of Computer-Aided Diagnosis (CAD) in the Diagnosis and Characterization of Thyroid Nodules when Combined with Conventional Ultrasound: A Literature Review. *Journal of Medical Imaging and Radiation Sciences.* 2019;50(3).

41. Giraud P, Giraud P, Gasnier A, El Ayachy R, Kreps S, Foy JP, et al. Radiomics and Machine Learning for Radiotherapy in Head and Neck Cancers. *Front Oncol.* 2019;9:174.
42. Jethanandani A, Lin TA, Volpe S, Elhalawani H, Mohamed ASR, Yang P, et al. Exploring Applications of Radiomics in Magnetic Resonance Imaging of Head and Neck Cancer: A Systematic Review. *Front Oncol.* 2018;8:131.
43. Hung K, Montalvao C, Tanaka R, Kawai T, Bornstein MM. The use and performance of artificial intelligence applications in dental and maxillofacial radiology: A systematic review. *Dentomaxillofac Radiol.* 2020;49(1):20190107.
44. Lång K, Dustler M, Dahlblom V, Åkesson A, Andersson I, Zackrisson S. Identifying normal mammograms in a large screening population using artificial intelligence. *European Radiology.* 2020.
45. <https://www.kreftregisteret.no/Forskning/Prosjekter/maskinlaring-i-mammografiprogrammet/>.
46. Wang J, Yang X, Cai H, Tan W, Jin C, Li L. Discrimination of Breast Cancer with Microcalcifications on Mammography by Deep Learning. *Sci Rep.* 2016;6:27327.
47. Weaver O, Leung JWT. Biomarkers and Imaging of Breast Cancer. *AJR Am J Roentgenol.* 2018;210(2):271-8.
48. Codari M, Schiaffino S, Sardanelli F, Trimboli RM. Artificial Intelligence for Breast MRI in 2008-2018: A Systematic Mapping Review. *AJR Am J Roentgenol.* 2019;212(2):280-92.
49. Drukker K, Li H, Antropova N, Edwards A, Papaioannou J, Giger ML. Most-enhancing tumor volume by MRI radiomics predicts recurrence-free survival “early on” in neoadjuvant treatment of breast cancer. *Cancer Imaging.* 2018;18(1):12.
50. Hua KL, Hsu CH, Hidayati SC, Cheng WH, Chen YJ. Computer-aided classification of lung nodules on computed tomography images via deep learning technique. *Onco Targets Ther.* 2015;8:2015-22.
51. Harris M, Qi A, Jeagal L, Torabi N, Menzies D, Korobitsyn A, et al. A systematic review of the diagnostic accuracy of artificial intelligence-based computer programs to analyze chest x-rays for pulmonary tuberculosis. *PLoS One.* 2019;14(9):e0221339.
52. Hosny A, Parmar C, Coroller TP, Grossmann P, Zeleznik R, Kumar A, et al. Deep learning for lung cancer prognostication: A retrospective multi-cohort radiomics study. *PLOS Medicine.* 2018;15(11):e1002711.
53. Petersen SE, Abdulkareem M, Leiner T. Artificial Intelligence Will Transform Cardiac Imaging-Opportunities and Challenges. *Front Cardiovasc Med.* 2019;6:133-.
54. Betancur J, Commandeur F, Motlagh M, Sharir T, Einstein AJ, Bokhari S, et al. Deep Learning for Prediction of Obstructive Disease From Fast Myocardial Perfusion SPECT: A Multicenter Study. *JACC Cardiovasc Imaging.* 2018;11(11):1654-63.
55. Zhang J, Gajjala S, Agrawal P, Tison GH, Hallock LA, Beussink-Nelson L, et al. Fully Automated Echocardiogram Interpretation in Clinical Practice. *Circulation.* 2018;138(16):1623-35.
56. Karuzas A, Sablauskas K, Skrodenis L, Verikas D, Rumbinaite E, Zaliaduonyte-Peksiene D, et al. P1465 Artificial intelligence in echocardiography - Steps to automatic cardiac measurements in routine practice. *European Heart Journal.* 2019;40(Supplement_1).
57. Trilisky I, Dachman AH, Wroblewski K, Vannier MW, Horne JM. CT Colonography with Computer-aided Detection: Recognizing the Causes of False-Positive Reader Results. *RadioGraphics.* 2014;34(7):1885-905.
58. Graffy PM, Sandfort V, Summers RM, Pickhardt PJ. Automated Liver Fat Quantification at Nonenhanced Abdominal CT for Population-based Steatosis Assessment. *Radiology.* 2019;293(2):334-42.
59. Azer SA. Deep learning with convolutional neural networks for identification of liver masses and hepatocellular carcinoma: A systematic review. *World J Gastrointest Oncol.* 2019;11(12):1218-30.

60. Harmon SA, Tuncer S, Sanford T, Choyke PL, Türkbey B. Artificial intelligence at the intersection of pathology and radiology in prostate cancer. *Diagn Interv Radiol.* 2019;25(3):183-8.
61. Greer MD, Lay N, Shih JH, Barrett T, Bittencourt LK, Borofsky S, et al. Computer-aided diagnosis prior to conventional interpretation of prostate mpMRI: an international multi-reader study. *Eur Radiol.* 2018;28(10):4407-17.
62. Liu L, Tian Z, Zhang Z, Fei B. Computer-aided Detection of Prostate Cancer with MRI: Technology and Applications. *Acad Radiol.* 2016;23(8):1024-46.
63. Azimi P, Yazdaniyan T, Benzel EC, Aghaei HN, Azhari S, Sadeghi S, et al. A Review on the Use of Artificial Intelligence in Spinal Diseases. *Asian Spine J.* 2020.
64. Chmelik J, Jakubicek R, Walek P, Jan J, Ourednicek P, Lambert L, et al. Deep convolutional neural network-based segmentation and classification of difficult to define metastatic spinal lesions in 3D CT data. *Med Image Anal.* 2018;49:76-88.
65. Lessmann N, van Ginneken B, de Jong PA, Išgum I. Iterative fully convolutional neural networks for automatic vertebra segmentation and identification. *Med Image Anal.* 2019;53:142-55.
66. Lee S, Choe EK, Kang HY, Yoon JW, Kim HS. The exploration of feature extraction and machine learning for predicting bone density from simple spine X-ray images in a Korean population. *Skeletal Radiol.* 2020;49(4):613-8.
67. Dallora AL, Anderberg P, Kvist O, Mendes E, Diaz Ruiz S, Sanmartin Berglund J. Bone age assessment with various machine learning techniques: A systematic literature review and meta-analysis. *PLoS One.* 2019;14(7):e0220242.
68. Sorin V, Barash Y, Konen E, Klang E. Deep Learning for Natural Language Processing in Radiology-Fundamentals and a Systematic Review. *J Am Coll Radiol.* 2020;17(5):639-48.
69. Sollini M, Antunovic L, Chiti A, Kirienko M. Towards clinical application of image mining: a systematic review on artificial intelligence and radiomics. *Eur J Nucl Med Mol Imaging.* 2019;46(13):2656-72.
70. Abujudeh HH, Kaewlai R, Asfaw BA, Thrall JH. Quality Initiatives: Key Performance Indicators for Measuring and Improving Radiology Department Performance. *RadioGraphics.* 2010;30(3):571-80.
71. Nagendran M, Chen Y, Lovejoy CA, Gordon AC, Komorowski M, Harvey H, et al. Artificial intelligence versus clinicians: systematic review of design, reporting standards, and claims of deep learning studies. *Bmj.* 2020;368:m689.
72. Recht M, Bryan RN. Artificial Intelligence: Threat or Boon to Radiologists? *J Am Coll Radiol.* 2017;14(11):1476-80.
73. Price WN, II, Gerke S, Cohen IG. Potential Liability for Physicians Using Artificial Intelligence. *JAMA.* 2019;322(18):1765-6.
74. Masud R, Al-Rei M, Lokker C. Computer-Aided Detection for Breast Cancer Screening in Clinical Settings: Scoping Review. *JMIR Med Inform.* 2019;7(3):e12660.
75. Guerriero C, Gillan MG, Cairns J, Wallis MG, Gilbert FJ. Is computer aided detection (CAD) cost effective in screening mammography? A model based on the CADET II study. *BMC Health Serv Res.* 2011;11:11.
76. Sato M, Kawai M, Nishino Y, Shibuya D, Ohuchi N, Ishibashi T. Cost-effectiveness analysis for breast cancer screening: double reading versus single + CAD reading. *Breast Cancer.* 2014;21(5):532-41.
77. Jacobsen DI. *Organisasjonsendringer og endringsledelse.* Fagbokforlaget, Bergen. 2004.
78. Luo Y, Tseng H-H, Cui S, Wei L, Haken RKT, Naqa IE. Balancing accuracy and interpretability of machine learning approaches for radiation treatment outcomes modeling. *BJR|Open.* 2019;1(1):20190021.
79. Obermeyer Z, Emanuel EJ. Predicting the Future - Big Data, Machine Learning, and Clinical Medicine. *N Engl J Med.* 2016;375(13):1216-9.

80. Liew C. The future of radiology augmented with Artificial Intelligence: A strategy for success. *Eur J Radiol.* 2018;102:152-6.