Bacheloroppgave

NTNU
Norges teknisk-naturvitenskapelige universitet
Fakultet for medisin og helsevitenskap
Institutt for sirkulasjon og bildediagnostikk

Victoria Andersson, Karoline Blomvik og Tonje Christine Øverby

Kunstig intelligens i diagnostiseringen på mammografi

Bacheloroppgave i Radiografi

Veileder: Øystein Olsen

Mai 2021

NTNU
Kunnskap for en bedre verden

Victoria Andersson, Karoline Blomvik og Tonje Christine Øverby

Kunstig intelligens i diagnostiseringen på mammografi

Bacheloroppgave i Radiografi
Veileder: Øystein Olsen
Mai 2021

Norges teknisk-naturvitenskapelige universitet
Fakultet for medisin og helsevitenskap
Institutt for sirkulasjon og bildediagnostikk

NTNU
Kunnskap for en bedre verden

Forord

Bacheloroppgaven er den avsluttende delen av vår 3-årige radiografutdannelse ved NTNU Trondheim våren 2021.

Temaet for oppgaven har vært spennende å jobbe med, og kunstig intelligens er noe vi ser på som svært relevant for fremtidens radiografer. Det har også vært en lærerik prosess å skrive denne oppgaven, og vi tar med oss mye kunnskap om både kunstig intelligens, mammografi og samarbeid videre.

Vi ønsker å takke Øystein Olsen for god veiledning underveis. Han har gjort mye mer enn vi noen gang hadde forventet, og vi er takknemlige for at han har dratt oss gjennom denne oppgaven.

Vi vil også takke dem som har tatt seg tid til å lese gjennom oppgaven og kommet med hjelpsomme tilbakemeldinger.

Trondheim, 27.05.21

Victoria Andersson, Karoline Blomvik og Tonje Christine Øverby

Sammendrag

Bakgrunn Kreft er den ledende dødsårsaken i Norge, og brystkreft er en av kreftdiagnosene med høyest forekomst. Mammografiscreening brukes for å kunne oppdage eventuell kreft tidlig, slik at prognosen blir bedre. Imidlertid er en utfordring med screening over-diagnostisering og falske positive resultater, noe som medfører «unødvendig» tilbakekalling. En slik tilbakekalling vil for pasienten kunne bety økt stress og bekymring knyttet til en eventuell kreftdiagnose. Kunstig intelligens (KI) er blitt forsøkt innført for å kunne sikre at resultatene pasienten får faktisk stemmer, og nyere KI-algoritmer blir stadig testet ut. Målet med oppgaven er å undersøke hvordan KI påvirker diagnostiseringen av brystkreft.

Metode Oppgaven ble gjennomført med litteraturstudie som metode. Både kvalitative og kvantitative studier ble inkludert, og åtte vitenskapelige artikler har blitt valgt ut med tanke på relevans for temaet og problemstillingen.

Resultat Nyere KI-systemer har blitt utprøvd og sammenliknet med gamle. Blant annet har ny, KI-basert Computer-Aided Detection (CAD) blitt sammenliknet med eldre, konvensjonell CAD. En av studiene i oppgaven viste at KI-basert CAD ga færre falske positive resultater enn konvensjonell CAD. Nyere KI-systemer har også blitt sammenliknet med radiologenes evne til å stille riktig diagnose. Flere studier indikerte at de nyere KI-systemene var like gode som eller bedre enn radiologene til å markere riktig på bildene, noe som også kunne føre til reduksjon av falske positive resultater. De nyeste systemene har vist seg å være til god hjelp i tolkning, og med mer trening av systemene vil sensitiviteten og spesifisiteten øke enda mer.

Konklusjon Bruken av nyere KI-systemer vil kunne redusere arbeidsmengden og bedre arbeidsflyten til radiografene. I tillegg vil kommunikasjonen, pasientomsorgen og kvaliteten på helsetjenesten i sin helhet bedres.

Abstract

Background Cancer is the number one cause of death in Norway, and breast cancer is the most frequent cancer diagnosis. Mammography screening is one of the measures used to detect cancer at an early stage, in order to obtain a better prognosis. However, mammography screening is associated with overdiagnosis and false-positive results, both of which indicate a higher recall rate. An “unnecessary” recall can cause the patient stress and worry due to a potential cancer diagnosis. Artificial intelligence (AI) has been introduced to ensure that the results the patient receives are true, and newer AI algorithms are constantly tested. Consequently, the aim of the thesis is to investigate how AI will affect the diagnosing of breast cancer.

Method This thesis was done using a literature study. Both qualitative and quantitative studies were included, and eight scientific articles were selected in regard to relevance of the topic and the aim of the thesis.

Results Newer AI systems have been tested and compared to old ones, including new AI based Computer-Aided Detection (CAD) and older conventional CAD. One of the studies in the thesis showed that AI-CAD reduced false-positive results compared to conventional CAD. The newer AI systems have also been compared to radiologists’ ability to make the correct diagnosis. Several articles showed that the new AI systems were as good as, or better than, the radiologists in marking the images correctly, which too could reduce false-positive results. The newest systems have proven to be helpful in interpretation, and with more training of the systems the sensitivity and specificity will improve even more.

Conclusion The use of newer AI systems can reduce workload and ease workflow for the radiographers. Furthermore, communication, patient care and quality of the healthcare system can improve.

Innhold

Forord ..... **i**

Sammendrag ..... **ii**

Abstract ..... **iii**

Liste over tabeller og figurer ..... **v**

Ordliste ..... **vi**

1 Innledning **1**

1.1 Mammografiscreening 1

1.2 Utfordringer ved screening 1

1.3 KI i medisinsk sammenheng 2

2 Teori **4**

2.1 Bildekvalitet på mammografi 4

2.2 Prediktiv verdi 4

2.3 Kunstig intelligens 5

3 Metode **7**

3.1 Søkeprosessen 7

3.2 Inklusjons- og eksklusjonskriterier 7

3.3 Litteratursøk 8

3.4 Datautvalg 10

3.5 Etiske implikasjoner 10

4 Resultat **11**

4.1 Studie 1 13

4.2 Studie 2 13

4.3 Studie 3 14

4.4 Studie 4 14

4.5 Studie 5 15

4.6 Studie 6 15

4.7 Studie 7 16

4.8 Studie 8 16

5 Diskusjon **17**

5.1 KI alene 17

5.2 KI som assistanse 17

5.3 Reduksjon av falske positive resultater 18

5.4 Kvaliteten på helsetjenesten 19

5.5 Bildekvalitet 19

5.6 Etiske perspektiver ved bruk av KI 19

5.7 Utvikling av CAD 20

5.8 Kritikk av metoden 21

6 Konklusjon **22**

Litteratur **23**

Liste over tabeller og figurer

Tabell 1: Kriterier for søk…….……7

Tabell 2: Litteratursøk……...……..….8

Tabell 3: Inkluderte studier……...……..……11

Tabell 4: Kort oppsummering av studiene…….……..……12

Figur 1…….…….……....5

Figur 2…….……...…..6

Figur 3……...…….10

Ordliste

KI: Kunstig intelligens. Datasystem som kan utføre oppgaver som i utgangspunktet krever menneskelig kunnskap.

Maskinlæring: En datamaskin lærer ved hjelp av mønstre i inndataen.

Dyp læring: Undergruppe av maskinlæring. En datamaskin lærer ved at inndataen filtreres gjennom flere lag.

Kunstige nevrale nettverk: En type samling av data som er basert på hvordan nevronene i den menneskelige hjernen er strukturert. Et nøkkelkonsept innenfor KI.

CNN: Convolutional Neural Network. Dyp læringsalgoritme som kan ta inn et bilde, for deretter å markere viktige områder og skille ulike objekter fra hverandre.

CAD: Computer-Aided Detection. Datasystem som brukes som beslutningsstøtte ved bildetolkning.

Prevalens: Hvor mange som har eller har hatt en bestemt sykdom på et gitt tidspunkt.

Insidens: Antall nye tilfeller av en bestemt sykdom i løpet av en gitt tidsperiode.

ROC: Receiver Operating Characteristic. En ROC-kurve brukes for å vise sammenhengen mellom klinisk sensitivitet og spesifisitet grafisk.

AUC/AUROC: Arealet under (ROC-)kurven. AUC sier noe om hvor god en klassifiseringsmodell er, på en skala fra 0-1. Jo høyere AUC, desto bedre er modellen til å skille mellom klasser, for eksempel sykdom/ikke sykdom.

ROI: Region of Interest. Markeringer av interesseområder i medisinske bilder.

κ: Cohen's kappa. Mål på samsvar mellom flere observatører der det er tatt høyde for tilfeldigheter. $\kappa = 1$ = perfekt samsvar. $\kappa = 0$ = helt tilfeldig.

1 Innledning

Kreft er den ledende dødsårsaken i Norge; et faktum siden 2017, da det for første gang var flere som døde av kreft enn av hjerte- og karsykdommer. Men, selv om flere diagnostiseres med kreft i dag enn for 50 år siden, er dødeligheten på vei ned. Dette gjelder spesielt for brystkreft, den hyppigste kreftformen blant kvinner. Og de som ikke blir friske *fra* kreften, lever lenger *med* kreften (1).

1.1 Mammografiscreening

En årsak til den økte overlevelsen hos bryskreftpasienter er Mammografiprogrammet, som ble innført i Norge i 1996. Gjennom dette programmet inviteres friske kvinner fra 50-69 år annethvert år til screening av begge brystene (2). Andre land har liknende ordninger, og i USA anbefales kvinnene å få utført mammografiundersøkelser hvert år etter fylte 40 (3).

Hensikten med Mammografiprogrammet er å diagnostisere tilstanden i et tidlig stadium. Gjennom programmet vil kvinnene enten få påvist brystkreft eller ikke. Det er likevel ikke så enkelt, da en pasient kan få påvist brystkreft uten nødvendigvis å ha det. 30 av 1000 kvinner som deltar i Mammografiprogrammet blir innkalt til tilleggsundersøkelser. 6 av disse 30 får påvist brystkreft, hvorav 1 antas å være overdiagnostisert (2).

1.2 Utfordringer ved screening

Overdiagnostisering i screening innebærer at noen får påvist en kreftsykdom som de ellers kunne ha levd uberørt med. Likevel kan man ikke vite at nettopp disse pasientene er overdiagnostiserte, og man må ta funnene på alvor. Derfor vil de bli tilbakekalt til tilleggs-undersøkelser på lik linje med de riktig diagnostiserte. Man kan heller ikke vite om tilbakekallingen og eventuelt behandlingen blir "unødvendig", noe som er et stort problem med overdiagnostikk. I tilfeller hvor pasienten blir innkalt til tilleggsundersøkelser etter screeningen, vil hun få beskjed om at det er noen forandringer man tror kan være kreft. Det kalles falske positive resultater når pasienten innkalles til tilleggsundersøkelser, men det ikke detekteres kreft (4).

Det er flere negative sider ved overdiagnostikk og falske positive resultater. En tilbakekalling vil kunne påføre pasienten bekymring, uro og redsel knyttet til å vente på svar og eventuell behandling av sykdommen (2). Videre vil dette kunne medføre flere sykehusbesøk og diagnostiske tester, mulig redusert deltagelse i fremtidige screeninger og økte kostnader for screeningprogrammet (4).

Diagnostiseringen slik den er i dag medfører også noen utfordringer for mammografi-radiografene. Disse er direkte knyttet til overdiagnostisering og feildiagnostisering gjennom falske positive resultater. Man regner med at 24 av 1000 screeningundersøkelser gir falske positive resultater (2), noe som gir økt arbeidsbelastning for radiografene. I tillegg stiller tilbakekalling krav til kommunikasjon med engstelige pasienter, og det kan bli utfordrende å vite hva man kan si og ikke.

Det har blant annet blitt forsøkt innføring av kunstig intelligens (KI) i diagnostiseringen. Hovedpoenget med denne innføringen er å gjøre fordelene med screening enda større enn ulempene, men det vil også medføre noen indirekte konsekvenser, blant annet når det gjelder arbeidssituasjonen til radiografene. Når et nytt system blir innført, må dette gjennomføres gradvis for å kunne aksepteres, da det stilles krav til anskaffelse eller oppdatering av kunnskap hos brukerne av systemet eller andre som berøres av det.

1.3 KI i medisinsk sammenheng

KI brukes i det aller meste av medisinsk teknologi i dag, og både begrepet og teknikken har blitt benyttet i flere tiår. KI i medisin inkluderer for eksempel online bestillinger av legetimer, selvinnsjekk på sykehus eller andre helseinstitusjoner, medisinutvikling, helsemonitorering og behandling av medisinske data (5). I bildediagnostikk benyttes KI blant annet til rekonstruering av CT-bilder gjennom autorekonstruksjon (maskinen rekonstruerer selv) (6, s. 63-70) og bolus tracking for å tilpasse kontrastdelay til pasient (7, s. 101-2). På konvensjonell røntgen har KI vært i bruk helt siden 1980, da Automatic Exposure Control (AEC) ble innført for å tilpasse kV-verdiene til pasientene (8, s. 2). Siden den gang har definisjonen av KI endret seg. Man vil i dag kanskje ikke definere AEC som KI, og KI knyttes nå mer direkte til underkategoriene maskinlæring og dyp læring.

Innen mammografi er KI også blitt tatt i bruk, og Computer-Aided Detection (CAD), et slags samarbeid mellom radiologen og maskinen, har vært vanlig i mange år. Dog har dette ikke vist seg å gi en forbedring av diagnostisk nøyaktighet. CAD kan på samme måte som AEC kalles en eldre teknologi, og bør videreutvikles for at man skal kunne utnytte systemet til det fulle. Helt siden CAD ble tatt i bruk har hastigheten på utviklingen gradvis økt, blant annet på grunn av fremskrittene i dyp læring (9). Derfor er det sannsynlig at nyere KI-systemer vil innføres i diagnostiseringen på mammografi, og vi ønsker å undersøke hvilke konsekvenser det vil føre med seg. Følgende problemstilling er derfor formulert: Hvordan påvirker kunstig intelligens diagnostiseringen av brystkreft?

2 Teori

2.1 Bildekvalitet på mammografi

I motsetning til konvensjonell røntgen er mammografi skreddersydd for å avbilde bløtvev. Komprimering av brystene er viktig for å fremstille strukturer og detaljer i vevet, samt å bedre kontrastoppløsningen (10, s. 380). I tillegg reduseres spredt stråling og dermed støy i bildet (11, s. 752). En studie kom frem til at radiologene var svært lite tolerante for endring i kontrast i bildene. De var mer tolerante for endringer som kvantestøy og tap av romlig oppløsning, men heller ikke disse endringene ble godtatt i stor grad. Generelt sett stiller radiologene høye krav til bildekvalitet i mammografibilder (12).

2.2 Prediktiv verdi

Sensitivitet og spesifisitet sier noe om hvor stor sannsynligheten er for at man får riktig resultat fra en test. Der sensitivitet er sannsynligheten for å påvise en bestemt sykdom hos pasienter som faktisk bærer sykdommen, er spesifisitet sannsynligheten for at man får negativt svar på en test når man er frisk. Jo høyere sensitivitet og spesifisitet, desto mer kan man stole på at testen gir riktig svar. Dette er ikke det samme som prediktiv verdi, da sistnevnte sier noe om hvor stor sannsynligheten er for at det gitte testresultatet stemmer overens med pasientens faktiske tilstand (frisk eller syk). Dersom pasienten er syk og testen har gitt positivt resultat, har den en positiv prediktiv verdi. Denne kan regnes ut ved ligning 1. Dersom pasienten er frisk og testen har gitt negativt resultat, gir det en negativ prediktiv verdi som kan regnes ut gjennom ligning 2 (13).

$$Positiv\ prediktiv\ verdi = \frac{sensitivitet \cdot prevalens}{sensitivitet \cdot prevalens + (1 - spesifisitet)(1 - prevalens)} \qquad [1]$$

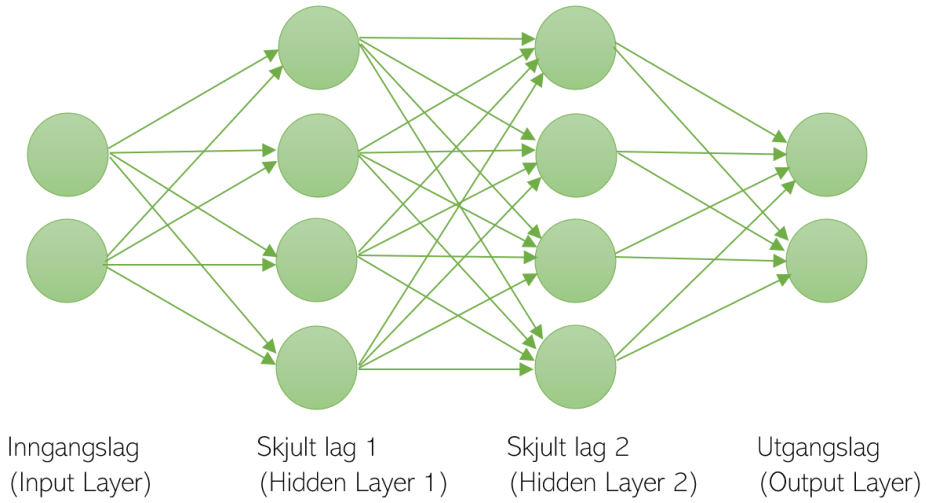
$$Negativ\ prediktiv\ verdi = \frac{spesifisitet(1 - prevalens)}{(1 - sensitivitet)prevalens + spesifisitet(1 - prevalens)} \qquad [2]$$

Prevalensen i en befolkning vil ha betydning for den prediktive verdien, da lav prevalens gir større sannsynlighet for falske positive resultater enn høy prevalens (13).

2.3 Kunstig intelligens

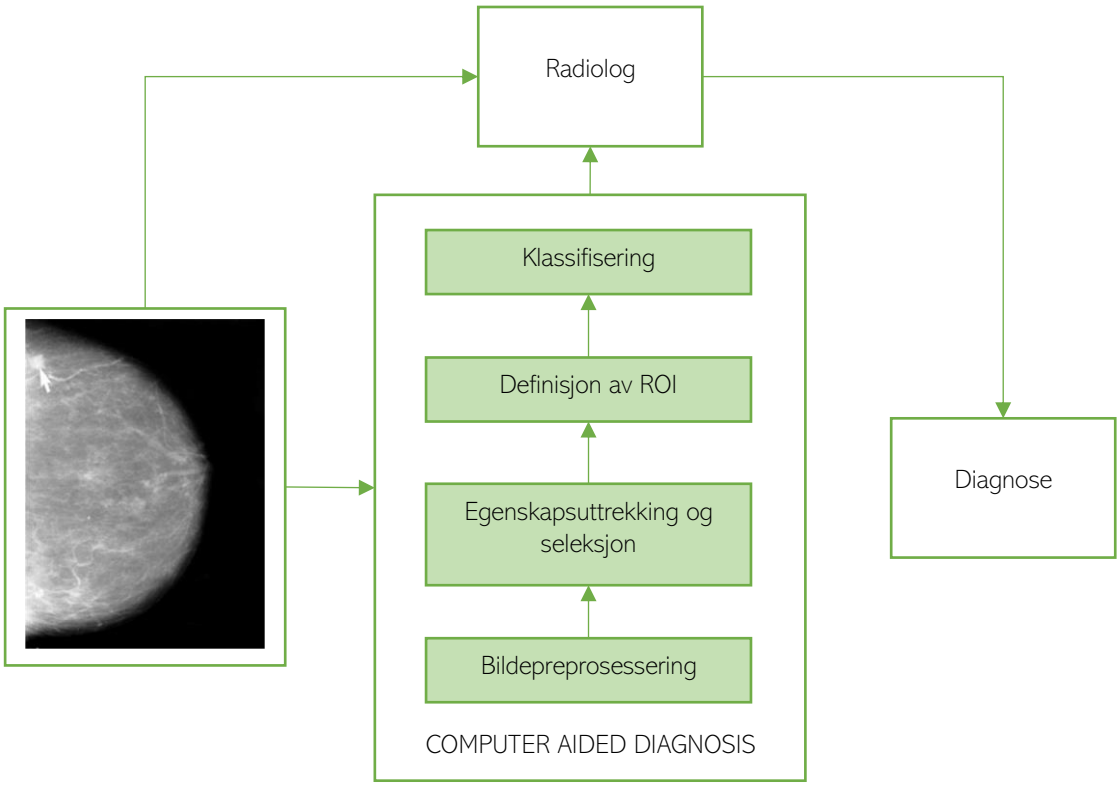
KI er et bredt og avansert begrep, og kan defineres som systemer som kan utføre oppgaver som i utgangspunktet krever menneskelig kunnskap (14). Man kan dele KI inn i flere undergrupper, men grovt sett snakker man om maskinlæring og dyp læring. Begge disse undergruppene er organisert i såkalte kunstige nevrale nettverk (figur 1), en type samling av data som er basert på hvordan nevronene i den menneskelige hjernen er strukturert. "Nevronene" i de kunstige nevrale nettverkene kalles noder (15).

Maskinlæring baserer seg på algoritmer som finner unike mønstre i inndataen som blir prosessert, uten behov for komplisert programmering. Ved hjelp av disse mønstrene kan maskinen lære å ta bedre beslutninger senere ved samme type data (16). Dyp læring er en underkategori av maskinlæring, hvor maskinen lærer fra erfaring. Dette skjer ved at data sendes gjennom mange skjulte lag av prosesseringsenheter, og for hvert lag den gjennomgår forbedres beslutningen som er basert på forrige lag. Til slutt skal maskinen ha lært seg hvordan dataen kan kategoriseres (14). Selv om man har mye kunnskap om hvordan disse systemene er bygd opp, vil de i noen tilfeller bli så kompliserte at ikke engang fagpersoner forstår hva maskinen beslutter. Dette gjør at KI ofte sammenliknes med en svart boks (17).



Figur 1. Enkel fremstilling av et kunstig nevralt nettverk basert på dyp læring med to skjulte lag. Dataen sendes inn i inngangslaget, og i forbindelsene (pilene) vektes dataen. Dataen filtreres gjennom de skjulte lagene og resultatet får vi i utgangslaget.

CAD er et av de første KI-systemene som ble tatt i bruk som assistanse i tolkning av mammografibilder. Her benyttes algoritmer for å analysere bildene for mulige lesjoner og unormale forkalkninger. Systemet vil markere områder hvor det trolig ligger en lesjon, og dermed blir det tydeligere for radiologene hvor de skal vie sin oppmerksomhet. CAD vil ikke kunne detektere alle typer kreft, men har fungert som en beslutningsstøtte for radiologene (11, s. 754). Se figur 2 for illustrasjon av hvordan CAD brukes i diagnostisering (18). Nye KI-systemer utvikles stadig, både som assistanse, men også for bruk alene. I tilfeller hvor KI brukes alene vil ikke radiologene ta del i beslutningen (19).



Figur 2. Prinsippet med Computer-Aided Diagnosis (hele diagnostiseringsprosessen). Computer-Aided Detection går under punktene "egenskapsuttrekking og seleksjon" og "definisjon av ROI".

3 Metode

Vi har valgt litteraturstudie som metode for å besvare problemstillingen, og en kombinasjon av kvalitative og kvantitative studier er inkludert. Kvalitative studier formidler hvordan KI-systemene faktisk fungerer i praksis, og kvantitative studier sikrer at resultatene ikke er tilfeldige.

3.1 Søkeprosessen

For å finne relevant litteratur har vi brukt medisinske og vitenskapelige databaser – PubMed, Medline og ScienceDirect. Disse søkemotorene er anerkjente i forskning, og derfor anser vi funnene som pålitelige. De er også anbefalt av flere universitetsbiblioteker (20-21). Vi har brukt flere søkemotorer for å kunne dekke et bredere spekter av informasjon. Noen søkemotorer ga flere relevante funn enn andre, noe som kommer frem i tabell 2.

I artikkelsøket ble ulike søkeord kombinert med *and* for å finne artikler som inneholdt alle søkeordene, og dermed snevre inn og spesifisere søket. MeSH, et emneordsystem med synonymer til medisinske ord på norsk og engelsk, ble brukt for å finne forslag til liknende søkeord. Disse to faktorene bidro til at antall treff ble redusert.

3.2 Inklusjons- og eksklusjonskriterier

Tabell 1 presenterer hvilke inklusjons- og eksklusjonskriterier som ble satt i søket.

Tabell 1: Kriterier for søk

Inklusjon	*Eksklusjon*
Artikler utgitt de siste 10 årene	Artikler eldre enn 10 år
Researchartikler	Reviewartikler
Artikler som omhandler mammografi	Artikler som kun omhandler det tekniske bak KI
Artikler som sier noe om hvordan KI kan påvirke radiologer og/eller pasienter på mammografi	Artikler som omhandler KI innenfor andre områder i tolkning (ikke mammografi)

Gjennom søket ønsket vi å finne artikler som var utgitt i løpet av de siste fem årene, men opptil 10 år gamle artikler kunne godtas dersom relevansen var spesielt høy. Med høy relevans menes artikler som, uavhengig av publiseringsdato, sier noe om inntrykket man har av KI, og avspeiler hvordan synet kan ha endret seg. Derfor ble 10 år gamle artikler inkludert i alle søkene, forutsatt at antall treff ikke overskred 250. Likevel ble alle artikler eldre enn 10 år ekskludert. Dette er fordi KI som system er i rivende utvikling, og eldre artikler enn dette simpelthen blir utdaterte. Ved mer enn 250 treff i søket, ble kun artikler utgitt de siste fem årene inkludert. Tallet 250 er valgt for at det skulle være mulig å gå gjennom alle treff i søkene.

3.3 Litteratursøk

Tabell 2 gir en oversikt over hvilke databaser og søkeord som ble brukt, og antall treff dette ga. Det kommer også frem antall artikler som ble vurdert og inkludert ved hvert treff. Kolonnen "Inkludert" ser ut til å inneholde svært mange artikler, men det stemmer ikke overens med det faktiske utvalgte antallet. Dette skyldes duplikater, da flere søk har gitt samme resultat.

Tabell 2: Litteratursøk

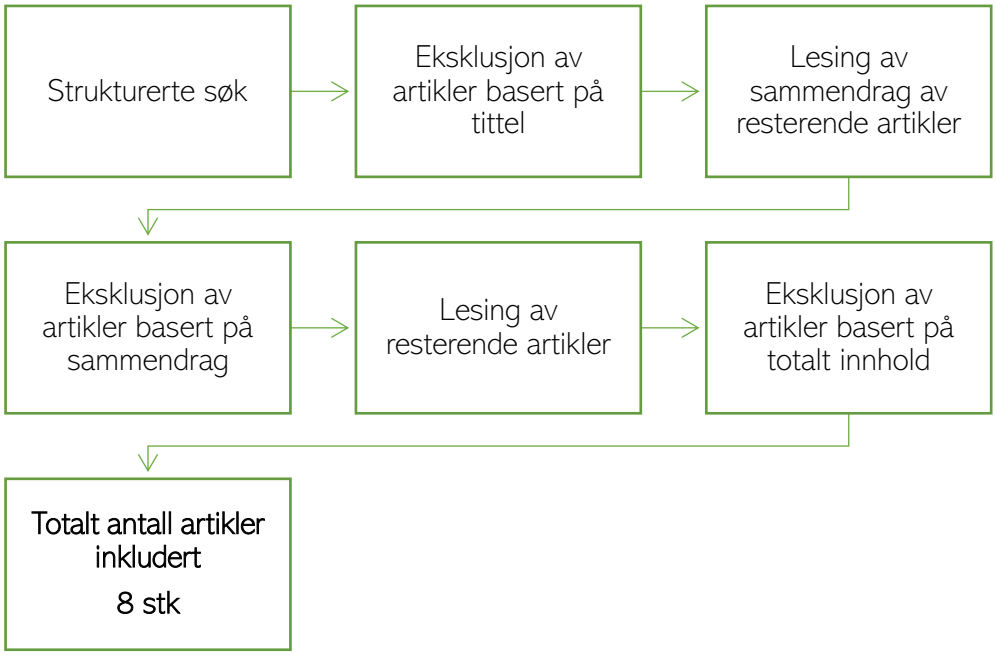
Søkeord	*Treff*		*Vurdert*	*Inkludert*
Mammography	PubMed	14 047	–	–
	ScienceDirect	14 901	–	–
	Medline	13 972	–	–
Mammographer	PubMed	3 504	–	–
	ScienceDirect	170	1	–
	Medline	13	1	1
Radiologist	PubMed	32 024	–	–
	ScienceDirect	71 415	–	–
	Medline	10 374	–	–
Breast cancer	PubMed	208 388	–	–
	ScienceDirect	254 274	–	–
	Medline	157 980	–	–
Artificial intelligence	PubMed	82 471	–	–
	ScienceDirect	86 514	–	–
	Medline	18551	–	–
Deep learning	PubMed	22 055	–	–
	ScienceDirect	172 624	–	–
	Medline	17 536	–	–
CAD	PubMed	30 542	–	–
	ScienceDirect	99 165	–	–
	Medline	25 534	–	–

Tabell 2: Litteratursøk fortsettelse

Søkeord	*Treff*		*Vurdert*	*Inkludert*
Digital imaging	PubMed	37 425	–	–
	ScienceDirect	374 583	–	–
	Medline	1 989	–	–
Mammographic imaging	PubMed	2 567	–	–
	ScienceDirect	4 512	–	–
	Medline	61	–	–
Patient experience	PubMed	220 092	–	–
	ScienceDirect	669 636	–	–
	Medline	6 184	–	–
Density	PubMed	374 364	–	–
	ScienceDirect	1 000 000 +	–	–
	Medline	356 628	–	–
False-positive	PubMed	23 870	–	–
	ScienceDirect	275 063	–	–
	Medline	23 884	–	–
Deep learning *and* mammographer	PubMed	58	5	1
	ScienceDirect	12	1	–
	Medline	2	1	1
Mammography *and* CAD *and* digital imaging	PubMed	133	7	2
	ScienceDirect	428	–	–
	Medline	1	–	–
Artificial intelligence *and* mammographic imaging *and* patient experience	PubMed	1	1	–
	ScienceDirect	120	7	1
	Medline	–	–	–
Deep learning *and* mammography *and* density	PubMed	38	2	1
	ScienceDirect	386	–	–
	Medline	39	7	2
Mammography *and* artificial intelligence *and* breast cancer *and* radiologist	PubMed	142	16	6
	ScienceDirect	272	–	–
	Medline	47	11	3
Mammography *and* artificial intelligence *and* breast cancer *and* CAD	PubMed	76	9	1
	ScienceDirect	202	2	–
	Medline	12	2	1
False-positive *and* artificial intelligence *and* mammography	PubMed	61	6	2
	ScienceDirect	266 (2016-2021)	–	–
	Medline	54	4	3

3.4 Datautvalg

For å få nok perspektiv på temaet uten å overskride omfanget for oppgaven, ønsket vi i utgangspunktet å inkludere mellom seks og ti artikler. Sammendraget i alle artikler som ble vurdert basert på tittel, ble gjennomlest. Deretter ble det evaluert hvilke av disse som faktisk var relevante for problemstillingen. De artiklene som ble ansett som aktuelle ble gjennomlest i sin helhet, og resten ble forkastet. Siden ble det gjort en ny vurdering av artiklene, og totalt ble åtte artikler valgt som grunnlag for oppgaven som vist i figur 3.



Figur 3. Flytskjema over hvordan vi har gått frem for å avgjøre hvilke artikler som skulle inkluderes og ekskluderes.

3.5 Etiske implikasjoner

I en litteraturstudie er ikke det etiske aspektet sentralt, da man ikke benytter seg av deltakeres personlige synspunkter og erfaringer. Det var ingen deltakere å beskytte, og derfor trengte vi ikke å søke om forhåndsgodkjenning fra Regionale komiteer for medisinsk og helsefaglig forskningsetikk eller Norsk senter for forskningsdata. Likevel må det være fokus på etikken rundt innsamling av data og artiklene som velges ut. Det er også viktig å kreditere forfatterne av artiklene for arbeidet de har gjort, og ikke ta æren for noe vi ikke har produsert selv.

4 Resultat

For å få en oversikt over resultatene, er det først utformet to tabeller. Tabell 3 presenterer de artiklene som ble inkludert, og tildeler hver av dem et nummer slik at det blir enklere å holde kontroll på de ulike artiklene senere i oppgaven. Artiklene vil videre omtales som studier, og med bruk av disse numrene. Tabell 4 gir en oversikt over studienes mål, metode, resultat og konklusjon.

Tabell 3: Inkluderte studier

Studie	*Referanse*
1	Rodriguez-Ruiz A, Lång K, Gubern-Merida A, Broeders M, Gennaro G, Clauser P et al. Stand-Alone Artificial Intelligence for Breast Cancer Detection in Mammography: Comparison With 101 Radiologists. J Natl Cancer Inst. 2019;111(9):916-22.
2	Rodriguez-Ruiz A, Krupinski E, Mordang J-J, Schilling K, Haywang-Köbrunner SH, Sechopoulos I et al. Detection of Breast Cancer with Mammography: Effect of an Artificial Intelligence Support System. Radiology. 2019;290(2):305-14.
3	Mohamed AA, Berg WA, Peng H, Luo Y, Jankowitz RC, Wu S. A deep learning method for classifying mammographic breast density categories. Med Phys. 2018;45(1):314-21.
4	Lehman CD, Wellman RD, Buist DSM, Kerlikowske K, Tosteson ANA, Miglioretti DL. Diagnostic Accuracy of Digital Screening Mammography With and Without Computer-Aided Detection. JAMA Intern Med. 2015;175(11):1828-37.
5	Lehman CD, Yala A, Schuster T, Dontchos B, Bahl M, Swanson K et al. Mammographic Breast Density Assessment Using Deep Learning: Clinical Implementation. Radiology. 2019;290(1):52-58.
6	Kim H-E, Kim HH, Han B-K, Kim KH, Han K, Nam H et al. Changes in cancer detection and false-positive recall in mammography using artificial intelligence: a retrospective, multireader study. Lancet Digit Health. 2020;2(3):e138-48.
7	Mayo RC, Kent D, Sen LC, Kapoor M, Leung JWT, Watanabe AT. Reduction of False-Positive Markings on Mammograms: a Retrospective Comparison Study Using an Artificial Intelligence-Based CAD. J Digit Imaging. 2019;32(4):618-24.
8	Watanabe AT, Lim V, Vu HX, Chim R, Weise E, Liu J et al. Improved Cancer Detection Using Artificial Intelligence: a Retrospective Evaluation of Missed Cancers on Mammography. J Digit Imaging. 2019;32(4):625-37.

Tabell 4: Kort oppsummering av studiene

Studie	*Mål med studien*	*Metode*	*Resultat*	*Konklusjon*
1	Undersøke KI-systemets evne til å detektere brystkreft på egenhånd, sammenliknet med radiologenes evne til det samme.	Det ble sammenliknet data fra mammografiundersøkelser hvor 101 radiologer hadde tolket bildene, og data fra et KI-system som tolket de samme bildene uavhengig av hverandre.	KI-systemet hadde en sensitivitet/spesifisitet på 0,840, og gjennomsnittet hos radiologene lå på 0,814 (AUROC). Dette gir en differanse på 0,026.	Ytelsen til KI-systemet var tilnærmet lik radiologenes ytelse.
2	Undersøke radiologenes evne til å tolke mammografibilder med og uten assistanse fra et KI-system.	Radiologene tolket halvparten av mammografibildene med KI-systemet og resten uten. Resultatene ble sammenliknet.	Radiologenes AUC økte fra 0,87 til 0,89 med assistanse fra KI-systemet. Sensitiviteten og spesifisiteten økte.	Radiologene forbedret evnen til å gjenkjenne brystkreft ved bruk av KI-systemet, uten at det krevde ekstra tid i tolkningen.
3	Undersøke om et system basert på dyp læring kan skille mellom ulike typer brysttetthet.	Det ble konstruert en CNN-basert modell koblet til et stort antall mammografibilder, og deretter ble det evaluert hvordan systemet klarte å klassifisere brysttetthet i bildene.	Nøyaktigheten i klassifiseringen økte med antall treningsdata.	Systemet hadde høy evne til å klassifisere ulike brysttettheter.
4	Måle mammografi-screeningens ytelse med og uten CAD.	Det ble sammenliknet hvor nøyaktig tolkning av bilder var med og uten CAD. 271 radiologer tolket bildene.	Screeningytelsen ble ikke forbedret med CAD på noen av de målte parameterne. Sensitiviteten ble redusert for bilder tolket med kontra uten CAD.	CAD forbedret ikke diagnostisk nøyaktighet på mammografi.
5	Utvikle en algoritme basert på dyp læring for å vurdere brysttetthet på mammografi.	Et CNN-system ble trent til å evaluere brysttetthet basert på erfarne radiologers opprinnelige tolkning av 41 479 bilder. Den resulterende algoritmen ble testet på nye bilder og radiologene (testsett). 500 av bildene i testsettet ble tolket av nye radiologer (reader study-sett) før algoritmen ble implementert i klinisk praksis, hvor en siste gruppe radiologer kategoriserte bildene med bruk av algoritmen.	<u>Opprinnelige radiologer:</u> - 9 % hovedsakelig fettvev - 50 % fibroglandulært vev - 36 % heterogent tett - 5 % ekstremt tett <u>KI i klinisk implementering:</u> - 6 % hovedsakelig fettvev - 52 % fibroglandulært vev - 40 % heterogent tett - 2 % ekstremt tett	KI-systemets vurdering av brysttettheten stemte godt overens med radiologenes vurdering av det samme bildet.
6	Utvikle en KI-algoritme for diagnostisering av brystkreft og undersøke om dette kan hjelpe radiologene ved å forbedre diagnostisk nøyaktighet.	En CNN-basert KI-algoritme ble utviklet. Radiologene vurderte bildene i to omganger, først uten KI, så med.	Radiologene alene hadde lavere AUROC enn KI-systemet alene. Med KI-assistanse hadde radiologene en høyere AUROC enn radiologene alene, dog lavere enn KI alene.	KI-assistanse medførte økt diagnostisk nøyaktighet hos radiologene.
7	Finne ut om KI-basert CAD kan redusere falske positive per bilde, sammenliknet med konvensjonell CAD.	Antallet markerte ROI for de to systemene ble sammenliknet for sensitivitet og spesifisitet.	KI-basert CAD ga en reduksjon i falske positive sammenliknet med konvensjonell CAD, uten at sensitiviteten ble redusert. Spesifisiteten var høyere ved KI-basert CAD enn ved konvensjonell CAD.	Reduksjonen av falske positive kan resultere i mindre tid brukt til tolkning av bilder.
8	Undersøke om KI-basert CAD kan hjelpe radiologene med å forbedre sensitiviteten i mammografi-screening og kreftdeteksjon.	Alle bildene som ble tolket hadde i utgangspunktet blitt vurdert som negative tidligere. Bildene ble vurdert på nytt med og uten bruk av KI-basert CAD.	Radiologene hadde en gjennomsnittlig økt deteksjonsrate ved hjelp fra KI-basert CAD. Det var en økning i AUROC med bruk av systemet.	Bruken av systemet kan forbedre arbeidsflyten, medføre tidligere oppdagelse av kreft og redusere antall falske negative mammografibilder.

4.1 Studie 1

KI-systemet alene hadde en høyere AUC enn 62 av de 101 radiologene, samt høyere sensitivitet enn 55 av 95 radiologer. KI-systemet var alltid dårligere enn den beste radiologen. Radiologene hadde mulighet til å sammenlikne med tidligere mammografiundersøkelser, noe KI-systemet ikke hadde. Dette var en ulempe med KI-systemet.

Det ble konkludert med at KI-systemet alene var like godt i tolkning som en gjennomsnittlig radiolog. Et slikt KI-system kan påvirke arbeidsflyten, men man må vurdere hvilken betydning en slik implementering har for omsorgen av pasientene.

4.2 Studie 2

Radiologenes AUC var gjennomsnittlig høyere med KI-assistanse enn uten (henholdsvis 0,89 kontra 0,87). Sensitiviteten økte med KI-assistanse (86 % kontra 83 %), og spesifisiteten så ut til å bli forbedret (79 % kontra 77 %). Tid brukt til tolkning for hvert kasus var omtrent lik med og uten assistanse (149 sekunder kontra 146 sekunder). AUC ved KI-systemet alene var liknende radiologenes gjennomsnittlige AUC (0,89 kontra 0,87).

Konklusjonen var at radiologene forbedret sin evne til å gjenkjenne kreft på mammografi når de brukte et KI-system som assistanse, uten å behøve ekstra tid til tolkning. KI-systemet kan hjelpe radiologene med hvor de skal vie sin oppmerksomhet i tolkningen.

4.3 Studie 3

Det ble testet to modeller, hvor den ene ble trent fra bunnen av, og den andre var trent på forhånd. Begge modellene ble trent opp med økende antall bilder (500 til 7000). I modellen trent fra bunnen av økte AUC-verdiene (og dermed presisjonen) fra henholdsvis 0,9081 til 0,9421. I den forhåndstrente modellen holdt AUC-verdiene seg relativt stabile (0,9265 kontra 0,9243). Etter å ha fjernet støyete og potensielt feilkategoriserte bilder, økte AUC til 0,9882 for den utrente modellen og 0,9857 for den forhåndstrente modellen – begge statistisk signifikant høyere enn når hele bildedatasettet ble brukt.

Studien viste at modellene ga høy presisjon i klassifiseringen av to brysttetthetskategorier (spredt tetthet og heterogen tetthet), noe som var vanskelig for en radiolog å skille mellom.

4.4 Studie 4

Screeningytelsen ble ikke forbedret med CAD på noen av de vurderte parameterne. Sensitiviteten var 85,3 % med CAD og 87,3 % uten, og spesifisiteten var 91,6 % med og 91,4 % uten. Det var ingen forskjell i kreftdeteksjonsrate. Radiologene som tolket både med og uten CAD, hadde en betydelig redusert sensitivitet for bildene tolket med CAD kontra uten.

Forfatterne konkluderte med at CAD ikke forbedret diagnostisk nøyaktighet i mammografi, og at kostnadene for CAD (400 millioner amerikanske dollar årlig) var for høye i forhold til de fordelene det ga kvinnene. De pekte derimot til slutt på at CAD kan forbedre mammografi på sikt gjennom mer og riktig trening av systemet.

4.5 Studie 5

Se tabell 3 for sammenlikning av vurdering av brysttetthet mellom dyp læring (KI) og radiologene. KI-modellen viste godt samsvar med radiologenes vurdering av bildene i testsettet ($\kappa = 0,67$), og med radiologene i reader study-settet ($\kappa = 0,78$). Det var også godt samsvar med radiologene i klinisk implementering-settet ($\kappa = 0,85$). Når det gjaldt den binære kategoriseringen av tette og ikke-tette bryst, ble 94 % av vurderingene til KI-systemet akseptert av radiologene som tolket dette.

Forfatterne konkluderte med at denne KI-modellen basert på dyp læring kan brukes til å kategorisere og vurdere brysttetthet i mammografibilder på samme nivå som erfarne mammografiradiologer.

4.6 Studie 6

KI-systemet alene hadde en AUROC på 0,940, mens radiologene hadde en AUROC på 0,810. Når radiologene benyttet KI-systemet som assistanse i tolkningen økte AUROC til 0,881. Forbedringen i AUROC var mer synlig ved tette bryst enn ved fettrike bryst. KI-systemet var bedre til å detektere tumormasser, deformasjon og asymmetri enn radiologene.

KI-algoritmen, trent med et stort antall (mammografi-)data, viste bedre evne til å gjenkjenne brystkreft enn radiologene. Den statistisk signifikante forbedringen i radiologenes ytelse med KI-assistanse viste at systemet er et effektivt verktøy i tolkningen.

4.7 Studie 7

Totalt sett var det en 69 % reduksjon i falske positive per bilde ved bruk av KI-basert CAD kontra konvensjonell CAD, og ingen reduksjon i sensitivitet. Det var en 83 % reduksjon i falske positive markeringer ved forkalkninger og en 56 % reduksjon ved masser. Nesten halvparten (48 %) av alle tilfellene viste *ingen* KI-basert CAD-markeringer, og bare 17 % viste *ingen* konvensjonelle CAD-markeringer.

Det ble konkludert med at en 69 % reduksjon av falske positive per bilde kan resultere i en 17 % reduksjon av tid brukt til tolkning per kasus. I tillegg hadde en reduksjon av tilbakekallinger grunnet falske positive resultater flere direkte sosiale og økonomiske fordeler, eksempelvis angst hos pasientene som blir tilbakekalt og kostnader knyttet til screeningen.

4.8 Studie 8

Ved bruk av KI-basert CAD var det en 11 % forbedring i kreftdeteksjonsrate for radiologene som deltok i studien, uavhengig av deres erfaring. Likevel var det radiologene med minst erfaring som hadde størst utbytte av assistanse fra KI-basert CAD ved deteksjon. Når det gjaldt tilbakekalling, var det forskjell hos radiologene med ulik erfaring. Radiologene i gruppe 1 (3 års erfaring) hadde en økt tilbakekallingsrate på ca. 1 %, de i gruppe 2 (3 års erfaring) hadde en reduksjon på 6,25 % og de i gruppe 3-7 (3-42 års erfaring) hadde ingen forandring. Ved bruk av KI-basert CAD økte AUROC for gjennomsnittet av alle radiologene fra 0,760 til 0,815, noe som utgjør en 7,2 % økning i AUROC.

Studien viste at det KI-baserte CAD-systemet kan være til fordel for radiologer i tolkning av mammografibilder. Bruk av KI i praksis kan potensielt bedre arbeidsflyten, medføre tidligere kreftdeteksjon og redusere antall falske negative bilder.

5 Diskusjon

Formålet med denne oppgaven var å undersøke hvordan KI påvirker diagnostiseringen av brystkreft, og vi har sett at KI brukes både alene og som assistanse for radiologene. Flere systemer har vist seg å prestere like godt som eller bedre enn radiologene, men et samspill mellom KI og radiolog ser ut til å gi best resultat. Et av hovedfunnene i oppgaven er reduksjon av falske positive resultater ved bruk av KI. I tillegg har det vært interessant å se på hvilken påvirkning bruken av systemene vil ha på kvaliteten på helsetjenesten, og hvilke krav KI-systemene stiller til bildekvalitet. En økt bruk av KI er knyttet til noen etiske perspektiver, som bør vurderes før man kan gi en datamaskin ansvar for deler av diagnostiseringen. Til slutt har utviklingen av CAD vært verdt å merke seg.

5.1 KI alene

Noen av studiene anga at KI alene er tilnærmet like god som en gjennomsnittlig radiolog i tolkning, og både studie 1 og 5 pekte på dette. Med tilnærmet like god mener vi at sensitiviteten og spesifisiteten hos KI og radiolog er omtrent lik. Studie 5 diskuterte hvor godt KI-systemet var til å klassifisere brysttettheter, og forfatterne kom frem til at systemet kunne sammenliknes med erfarne radiologer. Dette temaet har også studie 3 undersøkt, men disse forfatterne kom frem til at KI var spesielt god til å klassifisere brysttettheter som var vanskelig for radiologene å skille. Både studie 3 og 5 benyttet seg av dyp læring, men studie 3 fokuserte på å trene opp modellene, og det kan si noe om at disse KI-algoritmene bare blir bedre og bedre med mer trening.

5.2 KI som assistanse

Studiene 2, 6, 7 og 8 undersøkte bruken av KI som assistanse for radiologene. De kom frem til at screeningytelsen ble bedre enn det den ville vært uten bruk av KI. Studie 4 vurderte også bruk av KI som assistanse, men den er et unntak her. Med bedre screeningytelse mener vi hovedsakelig en reduksjon av falske positive resultater. Studie 6 vurderte både bruk av KI alene og som assistanse, samt radiologenes tolkning alene. De kom frem til noe interessant: av de tre parameterne var det KI alene som hadde høyest AUROC, og KI som assistanse hadde høyere AUROC enn radiolog alene, dog lavere enn KI alene. Det kan tyde på at KI alene gir mer falske

positive resultater enn radiologene alene, og at det sannsynligvis er KI som assistanse som er det beste alternativet – noe studie 2, 7 og 8 også indikerte.

Selv de studiene fant ut at systemene er tilnærmet like gode som en gjennomsnittlig radiolog, fortalte om at KI kan fungere veivisende for radiologene med minst erfaring. Disse radiologene vil ha et mer positivt utbytte enn dem med middels til mye erfaring. For eksempel kan KI-systemer styre oppmerksomheten deres mot bilder med stor sannsynlighet for kreft (studie 2), og på den måten kan radiologene slippe å bruke mye tid på å tolke bildene med lav sannsynlighet for kreft. Dette er en måte å redusere antall falske positive resultater på.

5.3 Reduksjon av falske positive resultater

For mammografiradiografene vil bruk av nyere KI-systemer kunne bety en løsning på flere av de utfordringene screeningen normalt sett byr på. Først og fremst vil et redusert antall falske positive resultater ifølge studie 7 medføre færre tilbakekalte pasienter. Dette vil gi redusert arbeidsmengde (studie 2) og bedret arbeidsflyt (studie 1 og 7). Det vil være et mindre antall kvinner som skal innom mammografilabben, da 24 av 1000 blir borte (2), og radiografene vil få mer tid til dem som kommer tilbake. Mer tid til pasientene vil kunne resultere i bedret kommunikasjon og pasientomsorg.

En reduksjon i antall falske positive resultater vil også direkte påvirke pasientene. Først og fremst vil de kvinnene som ikke blir tilbakekalt unngå å oppleve bekymring, uro og stress (2) grunnet usikkerhet om en eventuell kreftdiagnose (studie 2 og 7). Dette gjelder spesielt de kvinnene som tidligere kanskje hadde blitt tilbakekalt. Studie 2 anga at KI ikke gir økt tolkningstid. Det kan bety at tiden det tar før pasientene blir diagnostisert ikke påvirkes, og behandlingen kan komme i gang like raskt som uten bruk av KI. Ved at resultatene på screeningen blir sikrere, vil pasientene i større grad kunne stole på at svarene screeningen gir er riktige.

5.4 Kvaliteten på helsetjenesten

Når pasienten kommer på mammografiscreening, er det radiografer hun møter. Dersom radiologene feiltolker et bilde, og en pasient som ikke faktisk har kreft blir tilbakekalt, kan radiografen bli oppfattet som den ansvarlige. Ved å bruke nyere KI-systemer, og dermed redusere antall falske positive resultater, vil man sannsynligvis kunne redusere nettopp dette (studie 7). Når pasienten kan stole mer på at de resultatene hun får faktisk er sanne, vil dette også kunne styrke kvaliteten på helsetjenesten i sin helhet.

5.5 Bildekvalitet

Kvaliteten på mammografibildene vil kunne påvirke hvordan bildene tolkes. Radiologer har høye krav til bildekvalitet, og aksepterer i liten grad endringer i kontrast, romlig oppløsning og støy (12). Studie 3 viste til at KI-systemet fikk en økt AUC etter at støyete og potensielt feilkategoriserte bilder ble fjernet, og gjennom flertallet av studiene i litteratursøket er KI-systemer blitt sett på som like gode som eller bedre enn radiologene. Grunnet dette er det sannsynlig at KI-systemet kan ha liknende krav til bildekvalitet som radiologene har. I tillegg er KI-systemene utviklet av mennesker, noe som kan antyde at de i utgangspunktet har liknende «tankegang». Dette kan si noe om at bildekvaliteten har betydning for hvor korrekt KI-systemene vil fungere, og bildekvaliteten kan dermed også ha en innvirkning på pasientens utfall. Derfor vil trolig radiografenes jobb for å tilstrebe reproduserbare bilder av høy kvalitet fortsatt være en viktig del av hverdagen.

5.6 Etiske perspektiver ved bruk av KI

Flere av studiene påpekte at KI kan bidra til å bedre arbeidsflyten. Man burde likevel huske på pasienten og hvordan hun kan påvirkes av dette, noe studie 1 poengterte. I de tilfellene der KI blir benyttet som assistanse, kan det være enkelt for uerfarne radiologer å ukritisk godta avgjørelsen til KI-systemet uten å stole på sin egen kunnskap (studie 4).

Dagens KI-systemer kan i noen tilfeller betraktes som en svart boks fordi de blir for kompliserte for menneskelig forståelse (17). Da dukker det opp noen etiske spørsmål som bør tas stilling til. Hvem har ansvaret når en datamaskin utfører en oppgave som i utgangspunktet hadde krevd

kunnskapen til et menneske med en bestemt autorisasjon – for eksempel en radiolog? Hvordan skal pasienten forholde seg til dette når mennesket ikke lenger har kontroll? Selv om profesjonen alltid må sørge for forsvarlig bruk, kan det være vanskelig å vite hvem som har det spesifikke ansvaret, nettopp fordi det er en datamaskin som har utført oppgaven. Det er utfordrende å ta stilling til disse aspektene, og derfor vil kanskje en økt bruk av KI heller være en kost/nytte-vurdering. For eksempel kan en vesentlig reduksjon av falske positive resultater være en såpass stor og viktig fordel at nytten av systemet overveier de nevnte etiske problemene.

5.7 Utvikling av CAD

For å forstå hvordan utviklingen av CAD påvirker hvor godt systemet fungerer, kan en eldre studie gi et innblikk. Et eksempel er studie 4, som inneholder gammel data, med tall fra 2003-2012. I studien ble ytelsen til mammografiscreening med og uten CAD målt ved at systemet assisterte radiologene. CAD-systemet ga ingen forbedret effekt på diagnostiseringen; den ble tvert imot redusert i noen tilfeller. Dette viser at et gammelt CAD-system ikke fungerer bedre enn radiologene på egenhånd.

Imidlertid indikerer studien at CAD har blitt tatt mer i bruk de siste årene. I 2003 ble CAD brukt i 5 % av tilfellene, mens i 2012 hadde bruken økt til 83 %. Selv om bruken økte, fungerte ikke systemene optimalt i 2012. Derfor burde systemet forbedres kontinuerlig for å bedre den diagnostiske nøyaktigheten, slik at økt bruk vil være hensiktsmessig. Et annet perspektiv er at man må kunne lære seg å utnytte systemet til det fulle, slik at det ikke blir en belastning. Bruken av et slikt gammelt CAD-system kan føre til større sannsynlighet for falske positive resultater, og radiologene må bruke tid på å dobbeltsjekke hvorvidt markeringen stemmer med det som faktisk er sant.

5.8 Kritikk av metoden

Selv om en litteraturstudie bidrar til vid kunnskap om KI, finnes det usikkerhetsmomenter knyttet til denne metoden. En av disse er hvorvidt alle relevante studier har blitt inkludert i oppgaven. Blant annet kan relevante studier ha blitt oversett på grunn av feilaktige eller manglende søkeord, eller fordi databaser som er brukt ikke var dekkende. Både søk og inklusjon av studier kan ha vært preget av subjektive meninger. Studiene kan ha blitt feiltolket og blitt tillagt resultater som ikke stemmer med de faktiske resultatene.

Studiene som er brukt i oppgaven er fra flere verdensdeler. Dette gir et mangfoldig innblikk i bruken av KI-systemer verden over. Det mangler dog norske studier, og derfor er det usikkert hvorvidt resultatene som er funnet er overførbare til Norge. Imidlertid ser det ut til at helsevesenet i de ulike landene studiene har blitt gjennomført i, er strukturert på liknende måte, og vi kan derfor si at resultatene våre er gyldige også i Norge. En svakhet i oppgaven er utvalget av relevante studier, og om resultatene ville vært annerledes dersom vi hadde funnet og inkludert andre studier.

6 Konklusjon

Bruk av KI-systemer på mammografi har økt, og nye systemer vil sannsynligvis innføres. Slik systemene er i dag, fungerer de best som en assistanse for radiologene i markering av bilder og diagnostisering av brystkreft. En av de største fordelene med bruk av nyere KI-systemer i diagnostiseringen er en reduksjon av falske positive resultater. Dette vil kunne redusere arbeidsmengden og bedre arbeidsflyten til radiografene. I tillegg kan kommunikasjonen, pasientomsorgen og kvaliteten på helsetjenesten i sin helhet bedres. Bruken av systemene kan stille like høye krav til bildekvalitet som tidligere, og radiografenes jobb med billedtaking blir ikke mindre viktig nå.

Litteratur

1. Kreftregisteret. Brystkreft [Internett]. Oslo: Kreftregisteret; 21.01.21 [oppdatert 21.01.21; hentet 22.02.21]. Tilgjengelig fra: https://www.kreftregisteret.no/Temasider/kreftformer/Brystkreft/

2. Kreftregisteret. Fordeler og ulemper ved mammografiscreening [Internett]. Oslo: Kreftregisteret; 13.01.21 [oppdatert 13.01.21; hentet 23.02.21]. Tilgjengelig fra: https://www.kreftregisteret.no/screening/mammografiprogrammet/Fordeler-og-ulemper/

3. American Breast Cancer Foundation. Mammograms [Internett]. Columbia: American Breast Cancer Foundation; 2021 [hentet 23.02.21]. Tilgjengelig fra: https://bit.ly/3r9niVn

4. Roman M, Skaane P, Hofvind S. The cumulative risk of false-positive screening results across screening centres in the Norwegian Breast Cancer Screening Program. Eur J Radiol. 2014;83(9):1639-44.

5. Amisha, Malik P, Pathania M, Rathaur VK. Overview of artificial intelligence in medicine. J Family Med Prim Care. 2019;8(7):2328-31.

6. Kusk MW. Multislice CT – billedkvalitet, dosis & teknik. 1. utg. Frederiksberg: Radiografiens forlag; 2011.

7. Prokop M, van der Molen AJ. Patient Preparation and Contrast Media Application. I: Galanski M, Prokop M, red. Spiral and Multislice Computed Tomography of the Body [Internett]. Stuttgart: Thieme; 2003. s. 83-108 [hentet 19.03.21]. Tilgjengelig fra: http://search.ebscohost.com/login.aspx?direct=true&db=nlebk&AN=533989&site=ehost-live&ebv=EB&ppid=pp_Cover

8. Hardy M, Harvey H. Artificial intelligence in diagnostic imaging: impact on the radiography profession. Br J Radiol. 2020;93(1108):20190840.

9. Geras KJ, Mann RM, Moy L. Artificial Intelligence for Mammography and Digital Breast Tomosynthesis: Current Concepts and Future Perspectives. Radiology. 2019;293(2):246-59.

10. Bushong SC. Radiologic Science for Technologists: Physics, Biology, and Protection. 10. utg. St. Louis: Elsevier; 2013.

11. Lampignano JP, Kendrick LE. Bontrager's Textbook of Radiographic Positioning and Related Anatomy. 9. utg. St. Louis: Elsevier; 2018.

12. Boita J, van Engen RE, Mackenzie A, Tingberg A, Bosmans H, Bolejko A et al. How does image quality affect radiologists' perceived ability for image interpretation and lesion detection in digital mammography? Eur Radiol. 2021;:9 s.

13. Lydersen S. Hva er sannsynligheten for riktig resultat av en diagnostisk test? Tidsskr Nor Legeforen. 2017;137(18):3 s.

14. Shaikh K, Krishnan S, Thanki R. Artificial Intelligence in Breast Cancer Early Detection and Diagnosis [Internett]. Cham: Springer Nature Switzerland; 2021 [hentet 02.03.21]. Tilgjengelig fra: https://doi.org/10.1007/978-3-030-59208-0

15. Dvergsdal H. nevralt nettverk [Internett]. Oslo: Store norske leksikon; 14.02.09 [oppdatert 11.11.19; hentet 10.03.21]. Tilgjengelig fra: https://snl.no/nevralt_nettverk

16. Tennøe T. Kunstig intelligens: smart eller skremmende? [Internett]. Oslo: Teknologirådet; 26.01.17 [oppdatert 26.01.17; hentet 02.03.21]. Tilgjengelig fra: https://teknologiradet.no/kunstig-intelligens-smart-eller-skremmende/

17. Carter SM, Rogers W, Win KT, Frazer H, Richards B, Houssami N. The ethical, legal and social implications of using artificial intelligence systems in breast cancer care. Breast. 2020;49:25-32.

18. Stoitsis J, Valavanis I, Mougiakakou SG, Golemati S, Nikita A, Nikita KS. Computer aided diagnosis based on medical image processing and artificial intelligence methods. Nucl Instrum Methods Phys Res A. 2006;569(2):591-95.

19. Sechopoulos I, Mann RM. Stand-alone artificial intelligence – The future of breast cancer screening? Breast. 2020;49:254-60.

20. Universitetet i Oslo. Flere databaser [Internett]. Oslo: UiO Universitetsbiblioteket; 08.12.17 [oppdatert 06.06.19; hentet 11.03.21]. Tilgjengelig fra: https://www.ub.uio.no/fag/laring-utdanning/fagdidaktikk/ressurser/

21. Norges teknisk-naturvitenskapelige universitet. Databaser/artikler [Internett]. Trondheim: NTNU Universitetsbiblioteket; 2021 [hentet 11.03.21]. Tilgjengelig fra: https://www.ntnu.no/blogger/ub-mh/finn-litteratur/databaser-artikler/

NTNU

Kunnskap for en bedre verden