

Health Technology Assessment (HTA)

Artificiell intelligens vid bröstcancerscreening med mammografi

**[Artificial intelligence
in breast cancer screening
with mammography]**

TITEL

ARTIFICIELL INTELLIGENS VID BRÖSTCANCERSCREENING MED MAMMOGRAFI
[Artificial intelligence in breast cancer screening with mammography]

HTA syd

Avdelningen för forskning och utbildning
Skånes universitetssjukhus
Region Skåne
Södra sjukvårdsregionen

Sakkunniggrupp

Maria Erngrund, överläkare, bröstkirurgi, Kirurgkliniken, Blekingesjukhuset

Kristina Lång, överläkare, docent, bröststradiologi, Unilabs, Skånes universitetssjukhus

Maria Madestam, överläkare, bröststradiologi, Unilabs, Helsingborgs lasarett

Martin Malmberg, överläkare, PhD, VO hema onk-strål, VE onkologi och strålningsfysik,
Skånes universitetssjukhus

Projektorganisation, se Appendix A

Citera denna rapport enligt följande:

HTA syd. Artificiell intelligens inom bröstcancerscreening med mammografi [Artificial intelligence in breast cancer screening with mammography]. Lund: Region Skåne. 2024:1. 93 s. [hämtad dag-mån-år]

Tillgänglig via: <https://vardgivare.skane.se/kompetens-utveckling/sakkunniggrupper/hta-skane/#110365>

HTA syd: Rapport 2024:1
ISBN: 978-91-989049-0-1

Publiceringsdatum: 2024-06-05

Health Technology Assessment, HTA

En HTA-rapport utvärderar vilka positiva och negativa effekter en behandling eller annan insats kan ha för patienter. Förkortningen HTA står för Health Technology Assessment. Utifrån en aktuell frågeställning inom hälso- och sjukvården görs en systematisk litteraturoversikt. Tillförlitligheten i det vetenskapliga underlaget bedöms, enligt en noggrann och transparent metod. Även etiska, hälsoekonomiska och organisatoriska aspekter som rör frågan tas upp. I arbetet deltar alltid ämnessakkunniga. Rapporten granskas av minst två oberoende experter.

Den färdiga HTA-rapporten fungerar som ett underlag för beslut om införande eller avveckling och kan också peka på kunskapsluckor. Rapporten innehåller inte några rekommendationer.

Innehållsförteckning

Sammanfattning	6
English short summary	8
Rapportens innehåll	10
Förkortningar	11
1 Bakgrund	12
2 Metoder och material	13
2.1 Frågeställning	13
2.1.1 PICO	13
2.2 Litteratursökning och urvalsprocess.....	14
2.3 Praxisundersökning	15
2.4 Organisation, ekonomi och etik.....	15
3 Samlad bedömning av klinisk evidens	16
3.1 Litteratursökning och urvalsprocess.....	16
3.1.1 PRISMA.....	18
3.2 Beskrivning av inkluderade artiklar.....	18
3.2.1 Originalartiklar.....	18
3.2.2 Systematiska översikter.....	27
3.3 Resultat från inkluderade artiklar	27
3.3.1 Utfallsmått O1: Cancerdetektion.....	27
3.3.2 Utfallsmått O2: Återkallandefrekvens	32
3.3.3 Utfallsmått O3: Intervallcancer	37
3.3.4 Utfallsmått O4: TNM-stadium	37
3.3.5 Utfallsmått O5: Bröstcancermortalitet.....	40
3.3.6 Utfallsmått O6: Radiologbelastning	40
3.3.7 Utfallsmått O7: Livskvalitet	43
3.4 Analys av effektmått.....	43
3.5 Evidensgradering.....	44
3.6 Sammanställning av kunskapsläget.....	50
4 Riktlinjer och rekommendationer	51
5 Praxisundersökning	52
5.1 Bröstcancerscreening med mammografi i Södra sjukvårdsregionen	52
6 Organisatoriska aspekter	54
7 Ekonomiska aspekter	55
8 Etiska aspekter	59
9 Identifierade kunskapsluckor	61
10 Diskussion	61
11 Referenser	65
12 Appendix	69
Appendix A: Projektorganisation.....	69
Appendix B: Sökstrategier och databaser.....	72

Appendix C: Inkluderade artiklar	79
Appendix D: Exkluderade artiklar	81
Appendix E: Pågående studier	87
Appendix F: Ekonomiska aspekter	90
Underlag för beräkning av kostnader	90

Sammanfattning

Bröstcancer är den vanligaste cancerdiagnosen bland kvinnor. Bröstcancerscreening med mammografi erbjuds alla kvinnor i Sverige i åldrarna 40–74 år. Syftet är att bidra till minskad dödlighet i bröstcancer genom att ge möjlighet till tidig upptäckt och behandling. Mammografibilderna från screeningen granskas idag i Södra sjukvårdsregionen av två radiologer (röntgenläkare). Om bilderna ger misstanke om cancer återkallas kvinnan för ytterligare utredning och provtagning. Även bland de kvinnor som deltar i screeningprogrammet upptäcks och diagnosticeras enstaka cancerfall mellan screeningtillfällena och detta kallas intervallcancer.

Den tekniska utvecklingen med datorgenomförd bildanalys och artificiell intelligens (AI) för tolkning av radiologiskt bildmaterial har de senaste åren gjort stora framsteg, och flera olika modeller för användning av AI vid mammografi har utvecklats och erbjuds idag som olika produkter. AI inom bildanalys vid bröstcancerscreening med mammografi kan användas på flera sätt, a) som granskare, för att ersätta en eller flera radiologer. När AI helt ersätter alla radiologer i hela processen kallas det för stand-alone AI b) för triage, för att selektera vilka bilder som ska granskas av en eller flera radiologer och c) som granskningsstöd, det vill säga att radiologen får stöd i sitt arbete genom att AI sätter en riskpoäng och/eller en markering på bilder med misstänkt förändring.

Socialstyrelsens senaste rekommendation från 2023 om bröstcancerscreening beskrev införande av AI-stöd vid bildtolkning som en möjlig framtida utveckling av screeningprogrammet.

Forskningsområdet präglas av snabba framsteg och hög forskningsaktivitet. Intresset är stort från sjukvården att ta del av de potentiella kliniska fördelarna av AI i screeningprogrammet. Frågan om effekterna av användning av AI vid bröstcancerscreening med mammografi belyses därför i denna HTA-rapport.

Efter genomgång av den vetenskapliga litteraturen och bedömning av de ingående studierna kan följande slutsatser dras gällande användning av AI vid bröstcancerscreening med mammografi:

- Användning av AI som stand-alone vid bröstcancerscreening med mammografi ger lägre eller likvärdig cancerdetektion jämfört med dubbelgranskning utförd av radiologer (låg tillförlitlighet ⊕⊕○○).
- Användning av AI i kombination med radiolog vid bröstcancerscreening med mammografi ger likvärdig eller högre cancerdetektion jämfört med dubbelgranskning utförd av radiologer (måttlig tillförlitlighet ⊕⊕⊕○).

- Det går inte att dra någon slutsats om användning av AI som stand-alone vid bröstcancerscreening med mammografi avseende återkallandefrekvens jämfört med dubbelgranskning utförd av radiologer (mycket låg tillförlitlighet ⊕○○○).
- Användning av AI i kombination med radiolog vid bröstcancerscreening med mammografi ger likvärdig återkallandefrekvensen jämfört med dubbelgranskning utförd av radiologer (låg tillförlitlighet ⊕⊕○○).
- Det går inte att dra någon slutsats om hur användning av AI vid bröstcancerscreening med mammografi påverkar intervallcancerfrekvensen jämfört med dubbelgranskning utförd av radiologer (mycket låg tillförlitlighet ⊕○○○).
- Det går inte att dra någon slutsats om hur användning av AI vid bröstcancerscreening med mammografi påverkar detektion av cancer i olika stadier jämfört med dubbelgranskning utförd av radiologer (mycket låg tillförlitlighet ⊕○○○).
- Användning av AI vid bröstcancerscreening med mammografi minskar radiologbelastningen jämfört med dubbelgranskning utförd av radiologer (låg tillförlitlighet ⊕⊕○○).

De övriga utfallsmåtten i rapporten – livskvalitet och bröstcancer mortalitet – är inte studerade i någon av de inkluderade artiklarna.

Att införa AI inom bröstcancerscreening med mammografi medför flera etiska aspekter att beakta. Dels vad gäller generaliserbarhet och överförbarhet från den data algoritmen har tränats på, dels legalt och tekniskt kring hantering av de stora datamängderna. Detta kan eventuellt påverka den personliga integriteten på olika sätt och även den kliniska professionens roll och tilltro till sin kunskap. En annan etisk utmaning är osäkerheten kring risken för att ökad cancerdetektion inte leder till minskad bröstcancer mortalitet. Betydelsen av detta är oklar och individer kan ha olika preferenser.

Användning av AI i bröstcancerscreening med mammografi medför inte någon kostnadsbesparing jämfört med nuvarande rutin med dubbelgranskning utförd av radiologer. Detta utifrån de vetenskapliga studiernas resultat om skillnader i andel konsensusbedömningar och återkallande i fyra olika scenarier samt alternativa antaganden om pris per granskad kvinna för användning av AI. Eftersom de olika scenarierna pekar på att användning av AI i bröstcancerscreening med mammografi medför ökade kostnader på kort sikt, blir det viktigt att säkerställa att AI leder till hälsovinster eller andra värden för vården som kan motivera dessa kostnadsökningar.

English short summary

Breast cancer is the most common cancer diagnosis among women. Breast cancer screening with mammography is offered to women in Sweden aged 40–74. The aim of the screening program is to reduce breast cancer mortality by enabling early detection and treatment. The mammography images from the screening are read by two radiologists. If the images raise suspicion of cancer, the woman is recalled for further investigation and testing. Even among women participating in the screening program, some cancer cases are diagnosed between screening occasions, and this is defined as interval cancers.

Technological advancements in computer-performed image analysis, and artificial intelligence (AI) for interpreting radiological images have made significant progress in recent years, and several different models for using AI in mammography have been developed and are now being offered as different products. In breast cancer screening with mammography AI as a tool in image analysis can be used in several ways, a) as a reader, to replace one or more radiologists. When AI completely replaces all radiologists in the entire process, it refers to stand-alone AI b) in triage, to select which images are to be read by one or more radiologists and c) as reading support, i.e. that the radiologist is supported in the work by AI indicating a risk score and mark the area on the images with suspicious changes.

The latest recommendation regarding breast cancer screening from the Swedish National Board of Health and Welfare (Socialstyrelsen) from 2023 described the introduction of AI support in image interpretation as a potential future development of the screening program.

The research field is characterized by rapid advancements and high research activity. There is significant interest from the healthcare sector in utilizing the potential clinical benefits of AI within the screening program. Therefore, this HTA report examines the effects of using AI in mammography screening.

After reviewing the scientific literature, and assessing the included studies, the following conclusions can be drawn regarding the use of AI in mammography screening:

- Use of stand-alone AI in breast cancer screening with mammography render a lower or equivalent cancer detection rate compared to double reading by radiologists (low certainty of evidence ⊕⊕○○).
- Use of AI in combination with radiologist reading in breast cancer screening with mammography provides equivalent or higher cancer detection rates compared to double reading by radiologists (moderate certainty of evidence ⊕⊕⊕○).

- It is not possible to draw any conclusion regarding how stand-alone AI in breast cancer screening with mammography would affect number of recalls compared to double reading by radiologists (very low certainty of evidence ⊕○○○).
- Use of AI in combination with radiologist in breast cancer screening with mammography render similar recall rate compared to double reading by radiologists (low certainty of evidence ⊕⊕○○).
- It is not possible to draw any conclusion regarding how the use of AI in breast cancer screening with mammography affects the interval cancer rate compared to double reading by radiologists (very low certainty of evidence ⊕○○○).
- It is not possible to draw any conclusion about how the use of AI in breast cancer screening with mammography affects the detection of cancer at different stages compared to double reading by radiologists (very low certainty of evidence ⊕○○○).
- Use of AI in breast cancer screening with mammography reduces the workload of radiologists compared to double reading by radiologists (low certainty of evidence ⊕⊕○○).

The other outcome measures in the report – quality of life and breast cancer mortality – have not been studied in any of the included articles.

Introducing AI in breast cancer screening with mammography brings several ethical aspects to consider. Both in terms of generalizability and transferability from that data that the algorithm has been trained on, and also legally and technically around handling the large data sets. Altogether, this may possibly affect personal integrity in different ways, and challenge the role of the clinician including their trust in their knowledge. Another ethical challenge is the uncertainty about the risk of increased cancer detection rate doesn't necessarily cause lower breast cancer mortality. The implication of this is unclear and individuals might have diverse opinions about this.

Use of AI in breast cancer screening with mammography does not result in any cost savings compared to the current routine of double reading by radiologists. This is based on the results of the scientific studies on differences in the percentage of consensus assessments and recall in four different scenarios and different assumptions about the cost per woman for the use of AI. Since the scenarios points towards increased costs in the short term, it becomes important to ensure that AI then leads to health gains or other values for care which may motivate additional costs.

Rapportens innehåll

- Metodbeskrivning
- PICO
- Uttömmande litteratursökning
- Flödesschema
- Relevansbedömning
- Kvalitetsgranskning
- Tabelldata
- Sammanvägning av resultat
- Metaanalys
- Narrativ analys
- Evidensgradering
- Sammanfattning
- Ekonomi
- Praxisundersökning
- Organisation
- Etik
- Pågående studier
- Exkluderade studier
- Sakkunniggrupp
- Extern granskning
- Kunskapsluckor
- Jävsdeklarationer

Förkortningar

Förkortning	Förklaring
AI	Artificiell intelligens
CBTST	Córdoba Breast Tomosynthesis Screening Trial
DCIS	Duktal Cancer in Situ
DR	Double Reading (dubbelgranskning)
ECIBC	European Commission Initiative on Breast Cancer
GRADE	Grading of Recommendations Assessment, Development and Evaluation
HTA	Health Technology Assessment
MASAI	Mammography Screening with Artificial Intelligence
NRSI	Non-Randomized Studies of Interventions
PACS	Picture Archiving and Communication System
PICO	Patient or population, Intervention, Comparison or control, Outcome
PRISMA	Preferred Reporting Items for Systematic reviews and Meta-Analyses
QUADAS	Quality Assessment of Diagnostic Accuracy Studies
RCC	Regionala cancercentrum
RCT	Randomized Controlled Trial
RoB	Risk-of-Bias tool for randomized trials
ROBINS-I	Risk Of Bias In Non-randomized Studies of Interventions
ROBIS	Risk Of Bias in Systematic reviews
SBU	Statens beredning för medicinsk och social utvärdering
SR	Systematic Review (systematisk översikt)
TNM	Tumör, Nod (lymfkörtel), Metastas (dottertumör)
VE	Enhetsnamn inom verksamhetsområde (VO)
VO	Verksamhetsområde

1 Bakgrund

Bröstcancer är den vanligaste cancerdiagnosen bland kvinnor. Under 2021 drabbades 8 619 kvinnor i Sverige (Cancerfonden 2023). I dag finns goda möjligheter att behandla sjukdomen med kirurgi, antihormonella tableter, cellgifter, immunterapi och radioterapi. Den relativa femårsöverlevnaden är hög: 92,8 %. Eftersom bröstcancer är relativt vanligt avled ändå så många som 1 326 kvinnor på grund av bröstcancer under 2021 (Cancerfonden 2023).

En god prognos har ett starkt samband med tidig diagnos, det vill säga att tumören upptäcks medan den fortfarande är liten och inte har spritt sig vidare i kroppen. Små tumörer ger sällan symtom och upptäcks inte genom klinisk undersökning med bröstpalpation. Däremot kan de ofta ses vid en radiologisk undersökning, och bröstcancerscreening med mammografi erbjuds i nuläget alla kvinnor mellan 40 och 74 år i Sverige. Av den bröstcancer som diagnosticeras hos kvinnor i den åldersgruppen i Sverige idag upptäcks cirka 60–65 % genom bröstcancerscreening med mammografi (NKBC 2024).

Den tekniska utvecklingen med datorgenomförd bildanalys och artificiell intelligens (AI) för tolkning av radiologiskt bildmaterial har de senaste åren gjort stora framsteg, och flera olika produkter för användning av AI vid mammografi har utvecklats. Till HTA syd har det inkommit en fråga om vilka effekter användning av AI har vid granskning av bilder vid bröstcancerscreening med mammografi jämfört med dagens standardundersökning, som innebär att två radiologer gör en individuell bedömning av de mammografibilder som tagits vid ett screeningtillfälle, detta benämns vanligen dubbelgranskning utförd av radiolog.

Inom bildanalys är maskininlärning det vanligaste sättet att använda AI. Detta görs genom att i flera olika dataset med bilder träna upp algoritmer och med neurala nätverk finna mönster. Genom dessa mönster lär sig algoritmen hur bilden ska tolkas. Ju fler bilder från olika populationer och ju mer AI-algoritmen får träna på bilderna, desto mer förbättras resultaten. Det finns idag flera olika algoritmer tränade för att upptäcka cancer utifrån bilder tagna med olika tekniker vid bröstcancerscreening med mammografi. Hur dessa algoritmer används och kalibreras skiljer sig åt och kan styras på olika sätt. AI inom bildanalys är således inte en specifik produkt eller ett enskilt definierat sätt att arbeta.

AI inom bildanalys vid bröstcancerscreening med mammografi används på flera sätt, a) som granskare, för att ersätta en eller flera radiologer. När AI helt ersätter alla radiologer i hela processen kallas det för stand-alone AI b) för triage, för att selektera vilka bilder som ska granskas av en eller flera radiologer och c) som granskningsstöd, det vill säga att radiologen får stöd i sitt arbete genom att AI sätter en riskpoäng och/eller en markering på bilder med misstänkt förändring.

Denna rapport utvärderar vilken evidens som finns för AI inom bröstcancerscreening med mammografi vad gäller förmåga att upptäcka bröstcancer utan att öka antalet onödiga utredningar.

Den beaktar även radiologbelastning samt hur bröstcancermortalitet och behandling av bröstcancer skulle kunna komma att påverkas genom att titta på så kallade prognostiska och behandlingsprediktiva faktorer. Rapporten innehåller också en analys av ekonomiska aspekter kring ett införande av AI inom bröstcancerscreening med mammografi i Region Skåne och övriga delar av Södra sjukvårdsregionen, samt tar upp etiska aspekter av användning av AI vid bröstcancerscreening med mammografi.

2 Metoder och material

2.1 Frågeställning

Vad har AI-användning vid granskning av bilder inom bröstcancerscreening med mammografi för effekt jämfört med dagens standard med dubbelgranskning utförd av radiologer utan tillgång till AI?

2.1.1 PICO

PICO utformades i samråd inom projektgruppen. Även patientföreträdare har varit delaktig.

Tabell 1. Beskrivning av studiens PICO

PICO	Beskrivning
P	Personer som genomgår bröstcancerscreening med mammografi
I	Artificiell intelligens
C	Nuvarande screeningmetod (dubbelgranskning utförd av radiologer)
O	O1 Cancerdetektion O2 Återkallandefrekvens O3 Intervallcancer O4 TNM-stadium O5 Bröstcancermortalitet O6 Radiologbelastning O7 Livskvalitet
P=Patients, I=Intervention, C=Comparison, O=Outcome	

Tabell 2. Avgränsningar i PICO

Komponent	Avgränsning
Språk	Engelska, svenska, danska, norska

2.2 Litteratursökning och urvalsprocess

Sökstrategierna utformades av informationsspecialister på HTA syd i samråd med projektets sakkunniga och projektledare. De systematiska litteratursökningarna utfördes under mars 2024 i databaserna MEDLINE (via Ovid), Embase (via Ovid), Cochrane Library samt CINAHL (via EBSCO). Ingen begränsning i tid gjordes. Fullständiga sökstrategier finns i Appendix B.

Kompletterande sökningar gjordes i Google Scholar och i referens- och citeringslistor till relevanta artiklar. Ingen uppdaterad sökning gjordes, då HTA-rapporten slutfördes inom tre månader efter sökningarna. Däremot skapades automatiska uppdateringar från MEDLINE, Embase, Cochrane och CINAHL, som granskades på veckobasis fram till 3 maj 2024.

Sökningar efter HTA-rapporter gjordes på HTA-relaterade organisationers webbsidor i mars 2024. Sökningar efter kliniska riktlinjer gjordes i svenska och internationella källor i februari och mars 2024. Sökningar efter pågående kliniska studier gjordes i relevanta databaser i april 2024. För detaljer, se Appendix B.

Dubbletrensning gjordes i EndNote med Bramers metod (Bramer 2016). Två informationsspecialister gjorde sedan i screeningverktyget Rayyan ett första urval av artiklar som uppfyllde PICO baserat på information från titel och abstrakt. Meningsskiljaktigheter löstes genom konsensusförfarande eller hänfördes till projektledarna. Projektledarna gjorde ytterligare ett urval baserat på titel- och abstraktgranskning i Rayyan och relevansgranskade sedan de återstående artiklarna i fulltext. Varje bedömning gjordes av minst två projektledare oberoende av varandra. Meningsskiljaktigheter löstes genom konsensusförfarande eller hänfördes till projektgruppen.

Projektgruppen bedömde risken för bias för de relevanta artiklarna med olika granskningsmallar enligt HTA-metodik, så som den beskrivs i metodboken framtagen av Statens beredning för medicinsk och social utvärdering (SBU 2023): För randomiserade kontrollerade studier (RCT) användes RoB2 och för systematiska översikter användes Snabbstar följt av ROBIS. För övriga inkluderade artiklar gjordes ett övervägande mellan att använda ROBINS-I och QUADAS-2. Studierna som ligger till grund för dessa artiklar har en diagnostisk ansats och ofta är det primära utfallsmåttet cancerdetektion. Båda mallarna prövades på en subgrupp av utfallsmått och artiklar. Bedömningen blev att samma artiklar föll ut med väsentligen samma risk för bias för respektive utfallsmått oavsett mall. Valet att använda QUADAS-2 avgjordes av att studierna som främsta syfte hade att kartlägga hur bra

interventionen är på att diagnosticera snarare än vad interventionen leder till för långsiktiga patientnära effekter.

2.3 Praxisundersökning

Medarbetare inom för frågan relevant verksamhet i Södra sjukvårdsregionen tillfrågades om hur bröstcancerscreening med mammografi fungerar i de olika regionerna. Sakkunniga och projektledare sammanfattade processen.

2.4 Organisation, ekonomi och etik

Avseende de organisatoriska och ekonomiska aspekterna tillfrågades medarbetare inom för frågan relevant verksamhet i Södra sjukvårdsregionen vad verksamheter och ledning kan komma att vilja beakta vid ett eventuellt framtida införande. Sakkunniga och projektledare för HTA-analysen sammanfattade den inhämtade informationen och detta redovisas i avsnitt 6.

Rapporten redovisar också en kostnadskonsekvensanalys som omfattar resursanvändning, som enligt två svenska studier (Dembrower 2023, Lång 2023) möjligen påverkas vid användning av AI i bröstcancerscreening med mammografi. Avsnittet kommenterar även en brittisk modellanalys av kostnadseffektivitet för användning av AI inom bröstcancerscreening med mammografi jämfört med sedvanlig dubbelgranskning utförd av två radiologer (Vargas-Palacios 2023). Den ekonomiska analysen redovisas i avsnitt 7.

Genomgången av etiska aspekter tar sin utgångspunkt i SBU:s vägledning (SBU 2021) samt diskussioner inom projektgruppen och med externt metodstöd inom medicinsk etik för att identifiera centrala aspekter av stöd av AI i samband med bröstcancerscreening med mammografi.

Utgångspunkterna för att identifiera etiska aspekter rör frågor kring tillståndets svårighetsgrad och varaktighet, i vilken utsträckning den förändrade undersökningsmetoden påverkar hälsa och livskvalitet samt vilka kunskapsluckor som finns. I dessa diskussioner har även resonemang kring jämlikhet och rättvisa, autonomi, integritet, eventuell påverkan på tredje part, värderingar och särintressen samt forskningsetiska aspekter förts. Genomgången återfinns i avsnitt 8,

3 Samlad bedömning av klinisk evidens

3.1 Litteratursökning och urvalsprocess

De systematiska databassökningarna resulterade i totalt 2 138 träffar (MEDLINE 985, Embase 753, Cochrane CENTRAL 90 samt CINAHL 310). Ytterligare 46 artiklar påträffades genom sökning i andra källor. Efter dubblettrensning återstod 1 489 unika träffar. Efter första urvalet återstod 93 artiklar vars abstrakt granskades av projektgruppen. 60 artiklar relevansgranskades i fulltext. Av dessa befanns 19 artiklar vara relevanta. Efter bedömning av risk för bias återstod tio originalartiklar med som mest måttlig risk för bias: en RCT och nio icke-randomiserade studier (NRSI), figur 1 och 2. Totalt sju systematiska översikter som stämde med PICO identifierades. Efter att systematiska översikter med hög risk för bias enligt Snabbstar sorterats bort, kvarstod två systematiska översikter. Dessa granskades med avseende på risk för bias enligt ROBIS. Båda bedömdes ha måttlig risk för bias.

Study	Risk of bias domains				Overall
	D1	D2	D3	D4	
Dahlblom 2021	+	-	+	+	-
Dembrower 2023	-	+	+	+	-
Elhakim 2023	+	+	+	+	+
Heywang-Köbrunner 2023	+	-	+	+	-
Hickman 2023	+	+	+	X	X
Kühl 2023	+	+	+	-	-
Larsen 2022	+	+	+	+	+
Lauritzen 2022	+	+	+	+	+
Marinovich 2023	X	+	+	+	X
Raya-Povedano 2021	+	+	+	+	+
Romero-Martín 2022	+	-	+	+	-

Domains:
D1: Patient selection.
D2: Index test.
D3: Reference standard.
D4: Flow & timing.

Judgement
X High
- Some concerns
+ Low

Figur 1. Sammanfattning av risk för bias i inkluderade artiklar, icke-randomiserade studier granskade med QUADAS-2 (SBU 2023). Risken för bias bedömdes lika för alla utfallsmått i de icke-randomiserade studierna.

		Risk of bias domains					
		D1	D2	D3	D4	D5	Overall
Study	Lång 2023 O1	+	+	+	+	+	+
	Lång 2023 O2	+	+	+	+	+	+
	Lång 2023 O4	+	+	+	+	+	+
	Lång 2023 O6	+	-	+	-	-	-

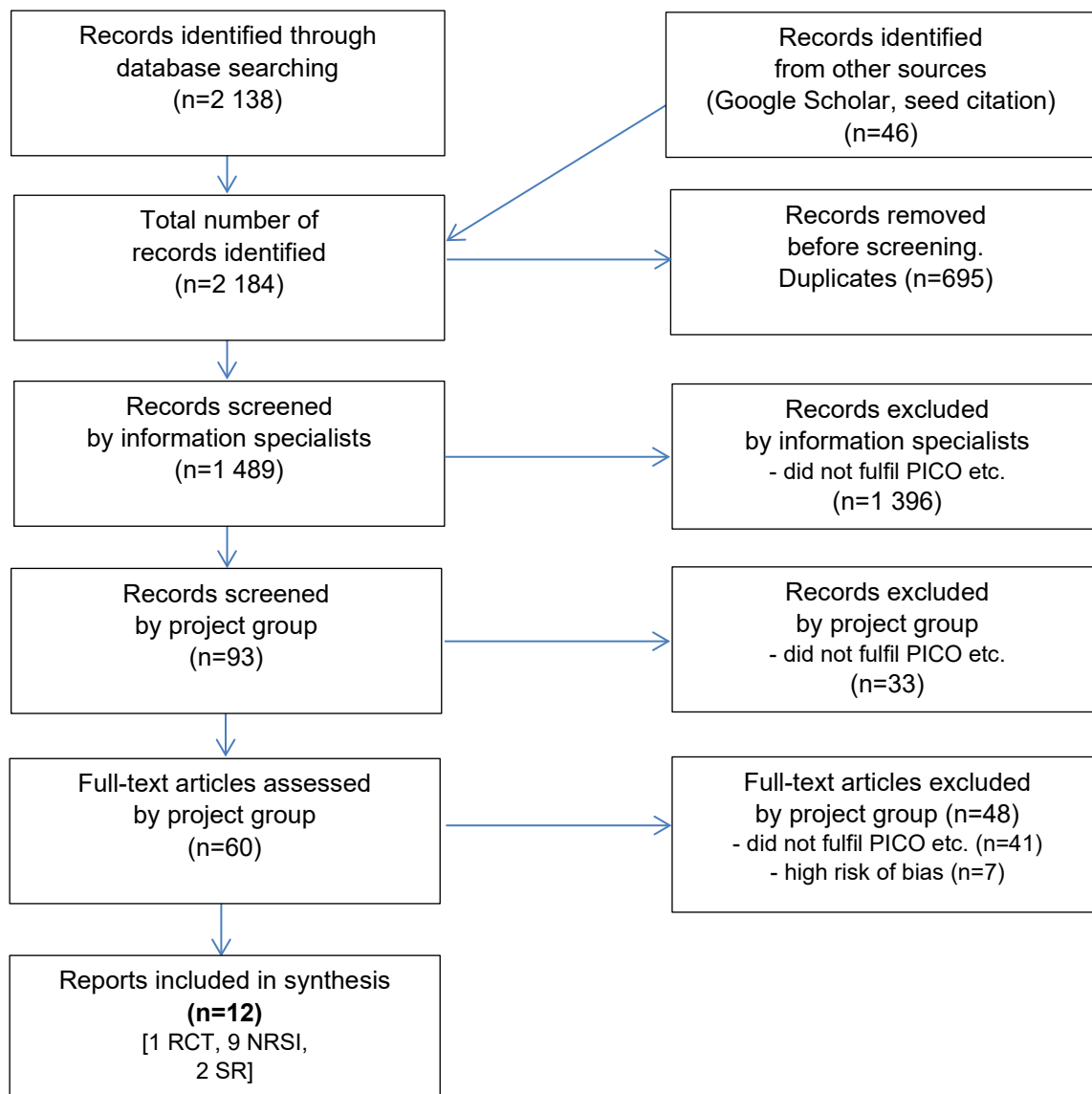
Domains:
D1: Bias arising from the randomization process.
D2: Bias due to deviations from intended intervention.
D3: Bias due to missing outcome data.
D4: Bias in measurement of the outcome.
D5: Bias in selection of the reported result.

Judgement
- Some concerns
+ Low

Figur 2. Sammanfattning av risk för bias i inkluderad randomiserad klinisk studie granskad med RoB2 (SBU 2023).

Inga nya, relevanta artiklar hittades genom automatiska uppdateringar av sökningarna. För detaljer om sökresultatet, se flödesschema nedan, figur 3, enligt PRISMA-riktlinjerna (baserat på Moher 2009). Referenser till inkluderade respektive exkluderade artiklar återfinns i Appendix C och D. Elva relevanta, pågående studier i ämnet hittades, bland annat de svenska studierna MASAI från Unilabs/Skånes universitetssjukhus (NCT04838756) och ScreenTrustCAD från Karolinska Institutet/Capio S:t Görans sjukhus i Stockholm (NCT04778670) samt den norska AIMS Norway (NCT06032390). De relevanta pågående studierna finns i Appendix E.

3.1.1 PRISMA



Figur 3. Flödesschema enligt PRISMA 2009

3.2 Beskrivning av inkluderade artiklar

3.2.1 Originalartiklar

De tio originalartiklar som inkluderas i analysen presenteras i tabell 3. Alla inkluderade artiklar är europeiska: tre från Sverige (Dahlblom 2021; Dembrower 2023; Lång 2023), tre från Danmark (Elhakim 2023; Kühl 2023; Lauritzen 2022), två från Spanien (Raya-Povedano 2021; Romero-Martín 2022) och en vardera från Norge (Larsen 2022) och Tyskland (Heywang-Köbrunner 2023). Det bör noteras att det är samma population som studeras i samtliga danska studier, men AI-produkter och

scenarier som testas skiljer sig åt. Även de två spanska studierna studerar samma population. Lång (2023) är den enda randomiserade, kontrollerade studien. Dembrower (2023) är en prospektiv, non-inferiority studie, medan övriga anges vara retrospektiva studier. De retrospektiva studierna presenterar dock data som är prospektivt insamlade från kohorter inom bröstcancerscreening med mammografi.

Som redan nämnts kan AI användas på flera olika sätt inom bröstcancerscreening med mammografi. Tabell 4 beskriver hur AI använts i de olika inkluderade studierna. I rapporten definieras AI som granskare när man använt AI och låtit den bedöma mammografibilden. AI:n kan då ersätta en eller två radiologers bedömning. AI-triage innebär att AI använts för att sortera bilder. Bilder med låg risk granskas antingen av endast en, eller ingen radiolog, medan bilder med hög risk innebär antingen att två radiologer granskar bilderna, eller att kvinnan återkallas för ytterligare undersökningar. AI som granskningsstöd innebär att granskande radiolog under något steg i screeningprocessen har fått stöd av AI-produkten för att identifiera en avvikelse i bildmaterialet, genom exempelvis ett markerat område i bilden. I några studier har man också använt AI i olika kombinationer, till exempel både för triage och som granskningsstöd.

Tabell 3. Inkluderade studier

Författare År Land Studienamn	Studietyp	Antal screeningtillfällen (individer)	Studerad population	Åldersspann (medelålder)	Intervention (AI)	Kontroll (mammografi)	Utfallsmått
Dahlblom 2021 Sverige	Retrospektiv kohortstudie	14 768 (14 768)	Screeningpopulation från Malmö Breast Tomosynthesis Screening Trial (2010–2015)	N/A (57 år)	Transpara™ (1.7.0, ScreenPoint Medical) AI som enskild granskare	Dubbelgranskning utförd av radiologer (oblindat)	Cancerdetektion Intervallcancer
Dembrower 2023 Sverige ScreenTrustCAD	Prospektiv non- inferiority studie	55 581 (55 581)	Screeningpopulation (2021–2022)	40–74 (median 55) år	Lunit Insight MMG™ (1.1.6, Lunit) AI som enskild granskare, AI som granskningsstöd	Dubbelgranskning utförd av radiologer (blindad AI/ radiologer, öppen vid konsensus- diskussioner)	Cancerdetektion Återkallandefrekvens TNM-stadium Radiologbelastning
Elhakim 2023 Danmark	Retrospektiv kohortstudie	257 671 (153 372)	Screeningpopulation (2014–2018) ^a	50–69 (59,3) år ^a	Transpara™ (1.7.0, ScreenPoint Medical) AI som enskild granskare	Dubbelgranskning utförd av radiologer (blindade, men oblindade vid konsensus)	Cancerdetektion Återkallandefrekvens Intervallcancer TNM-stadium
Heywang-Köbrunner 2023 Tyskland	Retrospektiv kohortstudie	17 884 (17 884)	Screeningpopulation (2018)	50–69 år (N/A)	ProFound AI™ (iCAD) AI som enskild granskare	Dubbelgranskning utförd av radiologer (blindade). Inga konsensus- diskussioner	Cancerdetektion

Författare År Land Studienamn	Studietyp	Antal screeningtillfällen (individer)	Studerad population	Åldersspann (medelålder)	Intervention (AI)	Kontroll (mammografi)	Utfallsmått
Kühl 2023 Danmark	Retrospektiv kohortstudie	249 402 (149 495)	Screeningpopulation (2014–2018) ^a	50–69 (59,3) år ^a	Lunit Insight MMG™ (1.1.6, Lunit) AI som enskild granskare	Dubbelgranskning utförd av radiologer (blindade, men oblindade vid konsensus	Cancerdetektion Återkallandefrekvens Intervallcancer TNM-stadium
Larsen 2022 Norge	Retrospektiv kohortstudie	122 969 (47 877)	Screeningpopulation (2009–2018)	50–69 år (N/A)	Transpara™ (1.7.0, ScreenPoint Medical). Riskpoäng 1–0 samt kontinuerligt rå riskpoäng. AI triage, AI som enskild granskare	Dubbelgranskning utförd av radiologer (blindade), konsensusdiskussion	Cancerdetektion Återkallandefrekvens Radiologbelastning
Lauritzen 2022 Danmark	Retrospektiv kohortstudie	114 421 (114 421)	Screeningpopulation (2014–2015) ^a	50–69 (59) år	Transpara™ (1.7.0, ScreenPoint Medical) AI triage, AI som enskild granskare	Dubbelgranskning utförd av radiologer (blindade), konsensusdiskussion	Cancerdetektion Intervallcancer Radiologbelastning
Lång 2023 Sverige MASAI	RCT	I: 39 996 C: 40 024	Screeningpopulation (2021–2022)	40–80 (median 54) år ^b	Transpara™ (1.7.0, ScreenPoint Medical) AI triage, AI som granskningsstöd	Dubbelgranskning utförd av radiologer (ej blindade), konsensusdiskussion	Cancerdetektion Återkallandefrekvens TNM-stadium Radiologbelastning

Författare År Land Studienamn	Studietyp	Antal screeningtillfällen (individer)	Studerad population	Åldersspann (medelålder)	Intervention (AI)	Kontroll (mammografi)	Utfallsmått
Raya-Povedano 2021 Spanien CBTST	Retrospektiv kohortstudie	15 987 (15 986)	Screeningpopulation från Córdoba Tomosynthesis Screening Trial 2015–2016 ^c	50–69 (58) år	Transpara™ (1.6.0) AI triage, AI som granskningsstöd, AI som enskild granskare	Dubbelgranskning utförd av radiologer (blindade, ingen konsensusdiskussion)	Cancerdetektion Återkallandefrekvens Radiologbelastning
Romero-Martín 2022 Spanien CBTST	Retrospektiv kohortstudie	15 999 (15 998)	Screeningpopulation från Córdoba Tomosynthesis Screening Trial 2015–2016 ^c	50–69 (58) år	Transpara™ (1.7.0) AI enskild granskare	Dubbelgranskning utförd av radiologer (blindade, ingen konsensusdiskussion)	Cancerdetektion Återkallandefrekvens

^a Samma, eller delvis samma, population i Elhakim 2023, Kühl 2023 och Lauritzen 2022. Inkluderar även kvinnor med bröstcancer i sjukhistorien eller genetisk predisposition över 70 års ålder.

^b Inkluderar även kvinnor 74–80 år med bröstcancer i sjukhistorien eller genetisk predisposition.

^c Samma population i Raya-Povedano 2021 och Romero-Martín 2022.

Tabell 4. Beskrivning av metoder och tillvägagångssätt i respektive inkluderad studie

Författare År Land Studienamn	Studietyp	Intervention scenarier	Definition friska	Definition sjuka och källa till uppföljningsdata
Dahlblom 2021 Sverige	Retrospektiv kohortstudie	1: AI som granskare. Stand-alone. AI riskpoäng ≥ 62 på 100-gradig skala + förändring korrekt lokaliserad.	Cancerfri uppföljning fram till nästa screeningtillfälle efter 18–24 månader	Histopatologisk diagnos (invasiv eller DCIS) som följd av utredning efter screeningtillfället. Intervallcancer inom 18–24 månader. Källa: Ej beskrivet i aktuell studie, hänvisar till tidigare.
		2: AI som granskare. Stand-alone. AI riskpoäng < 62 på 100-gradig skala + förändring korrekt lokaliserad.		
		3: AI som granskare. Stand-alone AI förändring korrekt lokaliserad, oavsett riskpoäng.		
Dembrower 2023 Sverige	Prospektiv non-inferiority studie	1: AI som granskare. Stand-alone. AI riskpoäng $\geq 53,4$ på 100-gradig skala.		Histopatologisk diagnos (invasiv eller DCIS) inom tre månader från screeningtillfället. Källa: Svenska cancerregistret och journalgenomgång.
		2: AI som granskare, kombinerad med R1 ^a . AI riskpoäng $\geq 53,4$ på 100-gradig skala.		
		3: AI som granskare, kombinerad med R1 och R2 ^a . AI riskpoäng $\geq 53,4$ på 100-gradig skala.		
Elhakim 2023 Danmark	Retrospektiv kohortstudie	1: AI som granskare. Stand-alone. AI matchad mot R1 sensitivitet.	Cancerfri uppföljning fram till nästa screeningtillfälle efter 24 månader	Histopatologisk diagnos (invasiv eller DCIS) som följd av utredning efter screeningtillfället. Intervallcancer inom 24 månader. Källa: Dansk Kvalitetsdatabas for Mammografiscreening (DKMS) och Danish Breast Cancer Cooperative Group (DBCG).
		2: AI som granskare. Stand-alone AI matchad mot R1 specificitet.		
		3: AI som granskare, kombinerad med R2. AI matchad mot R1s sensitivitet.		
		4: AI som granskare, kombinerad med R2. AI matchad mot R1s specificitet.		

Författare År Land Studienamn	Studietyp	Intervention scenarier	Definition friska	Definition sjuka och källa till uppföljningsdata
Heywang-Köbrunner 2023 Tyskland	Retrospektiv kohortstudie	1: AI som granskare. Stand-alone. AI-CAD >30 riskpoäng.	Cancerfri uppföljning mot register 2022 (>36 månader)	Histopatologisk diagnos (invasiv eller DCIS) som en följd av utredning efter screeningtillfället. Intervallcancer efter kontroll mot cancerregister, ej angiven tidsram. Källa: Cancerregister (ospecificerat).
		2: AI som granskare, kombinerad med R1. AI-CAD >30 riskpoäng.		
		3: AI som granskare, kombinerad med R2. AI-CAD >30 riskpoäng på 100-gradig skala.		
Kühl 2023 Danmark	Retrospektiv kohortstudie	1: AI som granskare. Stand-alone. AI matchad mot R1s sensitivitet.	Cancerfri uppföljning fram till nästa screeningtillfälle efter 24 månader	Dokumenterad diagnos (invasiv eller DCIS) som en följd av utredning efter screeningtillfället. Intervallcancer inom 24 månader. Källa: Dansk Kvalitetsdatabas for Mammografiscreening (DKMS) och Danish Breast Cancer Cooperative Group) DBCG.
		2: AI som granskare. Stand-alone. AI matchad mot R1s specificitet.		
Larsen 2022 Norge	Retrospektiv kohortstudie	1: AI som granskare stand-alone, kombinerad med R1. AI matchad mot medel-återkallandefrekvensen (5,8 %).	Cancerfri uppföljning fram till nästa screeningtillfälle efter 24 månader	Dokumenterad diagnos (invasiv eller DCIS) som en följd av utredning efter screeningtillfället. Intervallcancer inom 24 månader. Källa: Framgår inte.
		2: AI som granskare, kombinerad med R1. AI riskpoäng 10.		
		3: AI som triage. AI-riskpoäng 1–5 ingen vidare åtgärd. AI-riskpoäng 6–10 granskas av R1 och R2.		
		4: AI som triage. AI-riskpoäng 1–7 ingen vidare åtgärd. AI-riskpoäng 8–10 granskas av R1 och R2.		

Författare År Land Studienamn	Studietyp	Intervention scenarier	Definition friska	Definition sjuka och källa till uppföljningsdata
		5: AI som triage. AI-riskpoäng 1–9 ingen vidare åtgärd. AI-riskpoäng 10 granskas av R1 och R2.		
		6: AI som triage. AI-riskpoäng 1–5 ingen vidare åtgärd. AI-riskpoäng 6–10 granskas av R1.		
		7: AI som triage. AI-riskpoäng 1–5 granskas av R1. AI-riskpoäng 6–10 granskas av R1 och R2.		
		8: AI som triage. AI-riskpoäng 1–7 granskas av R1. AI-riskpoäng 8–10 granskas av R1 och R2.		
		9: AI som triage. AI-riskpoäng 1–5 ingen vidare åtgärd. AI-riskpoäng 6–7.5 granskas av R1. AI-riskpoäng på 7.6–10 granskas av R1 och R2.		
		10: AI som triage. AI-riskpoäng 1–5 ingen vidare åtgärd. AI-riskpoäng 6–7.5 granskas av R1 och R2. AI-riskpoäng 7.6–10 granskas av R1.		
		11: AI som granskare stand-alone. AI matchad mot återkallandefrekvensen i screeningpopulationen efter dubbelgranskning (3,2 %).		

Författare År Land Studienamn	Studietyp	Intervention scenarier	Definition friska	Definition sjuka och källa till uppföljningsdata
Lauritzen 2022 Danmark	Retrospektiv kohortstudie	1: AI som granskare stand-alone. AI matchad mot screening sensitivitet. 2: AI som triage. AI riskpoäng 1<5 ingen vidare åtgärd; AI riskpoäng >5 granskas av R1 och R2; AI matchad mot återkallandefrekvens återkallas direkt.	Cancerfri uppföljning fram till nästa screeningtillfälle efter 24 månader	Dokumenterad diagnos (invasiv eller DCIS) som en följd av utredning efter screeningtillfället. Intervallcancer inom 24 månader. Källa: Danish Pathology Register.
Lång 2023 Sverige MASAI	RCT	AI som triage och granskningsstöd. AI riskpoäng 1–9 granskas av R1, AI riskpoäng 10 + granskas av R1 och R2 Alla bilder som granskades av radiolog innehöll AI riskpoäng och för riskpoäng 8–10 inkluderades också markering.	Ingen cancerdiagnos registrerad 2–17 månader efter screeningtillfället	Histopatologisk diagnos (invasiv eller DCIS) inom 6 månader efter screening. Källa: Svenska cancerregistret, journalgenomgång. (Screening pågick april 2021 till juli 2022, uppföljning i register/journaler skedde senast sep 2022.)
Raya-Povedano 2021 Spanien CBTST	Retrospektiv kohortstudie	AI som triage. AI riskpoäng 1–7 ingen vidare åtgärd; AI riskpoäng 8–10 granskas av R1 och R2; AI 2 % med högst riskpoäng återkallas direkt.	Normal screening och ingen cancer inom 2 år	Histopatologisk diagnos (invasiv eller DCIS) som en följd av utredning efter screeningtillfället (dubbelgranskning mammografi eller tomosyntes). Intervallcancer inom 24 månader. Källa: Journalgenomgång.
Romero-Martín 2022 Spanien CBTST	Retrospektiv kohortstudie	AI som granskare stand-alone. AI riskpoäng >74 på 100-gradig skala + korrekt lokaliserad.	Normal screening och ingen cancer inom 2 år	Histopatologisk diagnos (invasiv eller DCIS) som en följd av utredning efter screeningtillfället (dubbelgranskning mammografi eller tomosyntes). Intervallcancer inom 24 månader. Källa: Journalgenomgång.

^a R1= första radiolog, R2= andra radiolog

3.2.2 Systematiska översikter

Två relevanta systematiska översikter bedömdes ha måttlig risk för bias.

Freeman (2021) är en brittisk systematisk översikt om cancerdetektion, utan rapporterade finansiella intressen, som inkluderar 13 artiklar baserade på 12 studier. Totalt omfattas 131 822 kvinnor som genomgått bröstcancerscreening med mammografi och 36 olika typer av AI-produkter. Dessa artiklar publicerades mellan 2019 och 2021. Artiklarna behandlade AI som a) granskare (n = 9) och b) för triagering (n = 4). Samtliga artiklar ingick i analysen och alla bedömdes ha hög risk för bias samt även problem med tillämplighet. Att inkludera artiklar med hög risk för bias gör resultatet mer osäkert. Majoriteten (94 %) av de 36 studerade AI-produkterna var sämre än två radiologer som bedömer cancerdetektion. Den systematiska översikten finner att AI som enskild granskare var bättre än bedömning av en radiolog för att korrekt detektera cancer.

Hickman (2022) är en brittisk systematisk översikt om cancerdetektion, med rapporterade finansiella intressen, som inkluderar totalt 14 artiklar och 15 studier (varav tre var publicerade som preprints). Dessa studier behandlar tolv unika AI-algoritmer. Artiklarna publicerades mellan 2017 och 2020 och studerar AI som a) granskare (n = 8) och b) vid triagering (n = 7). Alla artiklar, oavsett risk för bias (8 av 14 bedömdes ha en hög risk för bias), ingick i den narrativa analysen. I den primära metaanalysen ingick fem artiklar (n = 185 252 kvinnor). Att inkludera artiklar med hög risk för bias gör resultatet mer osäkert. Den systematiska översikten finner att algoritmer, som del av triagering, kan identifiera mellan 17 och 90 % av normala screeningbilder beroende på förvald nivå. Översikten visar även att AI som enskild granskare och radiologbedömning är likvärdiga på att detektera cancer.

Dessa systematiska översikter bedöms inte helt kunna besvara vår frågeställning eftersom det har tillkommit flera originalpublikationer inom området sedan de publicerades.

3.3 Resultat från inkluderade artiklar

3.3.1 Utfallsmått O1: Cancerdetektion

Bröstcancerscreening med mammografi syftar till att identifiera bröstcancer. Utfallsmåttet cancerdetektion beskriver andelen screeningundersökningar som resulterar i en bröstcancerdiagnos, antingen duktal cancer in situ (DCIS) – vilken innebär en bröstcancer som inte anses ha utvecklat invasiva egenskaper – eller invasiv bröstcancer. I denna rapport presenterar vi även sensitivitet (förmåga att fånga in alla med bröstcancer) och specificitet (förmåga att fria alla utan bröstcancer) vid screeningtillfället. Se tabell 5 för resultat.

De flesta av de ingående studierna som redovisade AI som stand-alone (granskare för att helt ersätta dubbelgranskning utförd av radiologer) visade scenarier med en lägre cancerdetektionsfrekvens för AI (Dahlblom 2021, Dembrower 2023, Elhakim 2023, Heywang-Köbrunner 2023, Larsen 2022). Endast två studier uppvisade scenarier där AI som ersättning för radiologisk dubbelgranskning visade på samma eller högre cancerdetektionsfrekvens (Dahlblom 2021, Romero-Martin 2022).

Samtliga studier som undersökte scenarier med AI som granskare i kombination med granskande radiolog visade minst lika bra cancerdetektion som dubbelgranskning utförd av två radiologer (Dembrower 2023, Elhakim 2023, Heywang-Köbrunner 2023, Larsen 2022). Även de studier där AI användes för triage, med och utan scenarier med kombination AI som granskningsstöd, visade på minst lika hög cancerdetektionsfrekvens jämfört med dubbelgranskning utförd av radiologer, dock beroende av vilket scenario man valt (Lauritzen 2022, Lång 2023, Raya-Povedano 2021).

Eftersom AI-produkterna kan ställas in på olika grader av sensitivitet, kan hög sensitivitet uppnås på bekostnad av låg specificitet om AI används som enskild granskare. Detta illustreras i Larsen 2022 som studerar olika nivåer av sensitivitet. Andra inkluderade artiklar har anpassat sensitiviteten för AI-produkten efter den som verksamheten har med dubbelgranskning utförd av två radiologer, då bibehålls cancerdetektionen (Raya-Povedano 2021, Romero-Martín 2022).

Cancerdetektionen är beroende av hur AI-produkten används (som granskare, triage, granskningsstöd) och vilka gränser för riskpoäng som väljs.

Sammanfattningsvis visar resultaten i de inkluderade artiklarna:

Användning av AI som stand-alone vid bröstcancerscreening med mammografi ger lägre eller likvärdig cancerdetektion jämfört med dubbelgranskning utförd av radiologer.

Användning av AI i kombination med radiolog vid bröstcancerscreening med mammografi ger likvärdig eller högre cancerdetektion jämfört med dubbelgranskning utförd av radiologer.

Tabell 5. Utfallsmått O1. Cancerdetektion

Författare År Land Studienamn	Antal screening- tillfällen (kvinnor)	Cancerdetektion (n)		Sensitivitet (konfidensintervall 95 %)		Specificitet (konfidensintervall 95 %)		Kommentar
		Intervention (AI)	Kontroll (DR)	Intervention (AI)	Kontroll (DR)	Intervention (AI)	Kontroll (DR)	
Dahlblom 2021 Sverige	14 768 (14 768)	1: 0,62 % (91) * 2: – * 3: 0,90 % (133) *	0,64 % (95)	1: 57,6 % * 2: – * 3: 84,1 %*	60,1 %	1: 91,3 %* 2: –* 3: 90,8 %*	–	Scenario 2 ej enskilt relevant, då dessa är inkluderade i scenario 3
Dembrower 2023 Sverige ScreenTrustCAD	55 581 (55 581)	1: 0,44 % (246) * 2: 0,47 % (261) 3: 0,48 % (269)	0,45 % (250)	–	–	–	–	Sensitivitet och specificitet kan ej beräknas, uppföljande data saknas
Elhakim 2023 Danmark	257 671 (153 372) ^a	1: 0,50 % (1 301)* 2: 0,46 % (1 197)* 3: 0,60 % (1 555) 4: 0,59 % (1 522)	0,58 % (1 509)	1: 63,7 % (61,6–65,8)* 2: 58,6 % (56,6–60,0)* 3: 76,2 % (74,3–78,0) 4: 74,6 % (72,6–76,4)	73,9 % (72,0–75,8)	1: 96,5 % (96,4–96,5)* 2: 97,8 % (97,7–97,8)* 3: 97,3 % (97,2–97,3) 4: 97,9 % (97,8–97,9)	97,9 % (97,9–98,0)	Cancerdetektion beräknad utifrån data i artikeln
Heywang- Köbrunner 2023 Tyskland	17 884 (17 884)	1: 0,60 % (107)* 2: 0,64 % (115) 3: 0,65 % (116)	0,64 % (115)	1: 81,7 % (74,2–87,0)* 2: 87,8 % (82,0–92,2) 3: 88,5 % (82,9–92,8)	87,8 % (82,0–92,2)	1: 80,2 % (79,7–80,7)* 2: 75,0 % (74,5–75,5) 3: 75,1 % (74,5–75,6)	87,2 % (68,8–87,6)	

Författare År Land Studienamn	Antal screening- tillfällen (kvinnor)	Cancerdetektion (n)		Sensitivitet (konfidensintervall 95 %)		Specificitet (konfidensintervall 95 %)		Kommentar
		Intervention (AI)	Kontroll (DR)	Intervention (AI)	Kontroll (DR)	Intervention (AI)	Kontroll (DR)	
Kühl 2023 Danmark	249 402 (149 495) ^a	1: 0,52 % (1 292)* 2: 0,51 % (1 273)*	0,59 % (1 475)	1: 63,6 % (64,4–65,6)* 2: 62,6 % (60,5–64,7)*	74,0 % (72,1–75,9)	1: 97,5 % (97,5–97,6)* 2: 97,5 % (97,5–97,6)*	97,8 (97,8–97,9)	Cancerdetektion beräknad utifrån data i artikeln
Larsen 2022 Norge	122 969 (47 877)	1: 0,59 %* 2: 0,59 % 3: 0,60 % 4: 0,58 % 5: 0,53 % 6: 0,53 % 7: 0,61 % 8: 0,61 % 9: 0,59 % 10: 0,53 % 11: 0,45 %*	0,61 % (752)	Scenario 1–4 > 70 % Scenario 5–6 < 70 % Scenario 7–9 > 70 %	78,6 % (75,9–81,1)	–	–	
Lauritzen 2022 Danmark	114 421 (114 421) ^a	1: -* 2: 0,68 % (779)	0,69 % (791)	1: -* 2: 69,7 % (66,9–72,4)	70,8 % (68,0–73,5)	1: 94,9 % (94,7–95,0)* 2: 98,6 % (98,5–98,7)	98,1 % (98,1–98,2)	
Lång 2023 Sverige MASAI	I: 39 996 C: 40 024	0,61 % (244)	0,51 % (203)	–	–	–	–	Sensitivitet och specificitet kan ej beräknas, uppföljande data saknas

Författare År Land Studienamn	Antal screening- tillfällen (kvinnor)	Cancerdetektion (n)		Sensitivitet (konfidensintervall 95 %)		Specificitet (konfidensintervall 95 %)		Kommentar
		Intervention (AI)	Kontroll (DR)	Intervention (AI)	Kontroll (DR)	Intervention (AI)	Kontroll (DR)	
Raya-Povedano 2021 Spanien CBTST	15 987 (15 986) ^b	0,49 % (78)	0,48 % (76)	69,0 % (60,0–76,8)	67,3 % (58,2–75,2)	96,3 % (96,0–96,6)	95,4 % (95,1–96,7)	Specificiteten är beräknad utifrån data i artikeln
Romero-Martín 2022 Spanien CBTST	15 999 (15 998) ^b	0,50 % (80)*	0,48 % (76)	70,8 % (61,8–78,4)*	67,3 % (58,2–75,2)	97,4 % (97,1–97,6)*	95,4 % (95,1–95,7)	Specificiteten är beräknad utifrån data i artikeln

* AI som stand-alone

^a Samma, eller delvis samma population i Elhakim 2023, Kühl 2023 och Lauritzen 2022

^b Samma population i Raya-Povedano 2021 och Romero-Martín 2022

3.3.2 Utfallsmått O2: Återkallandefrekvens

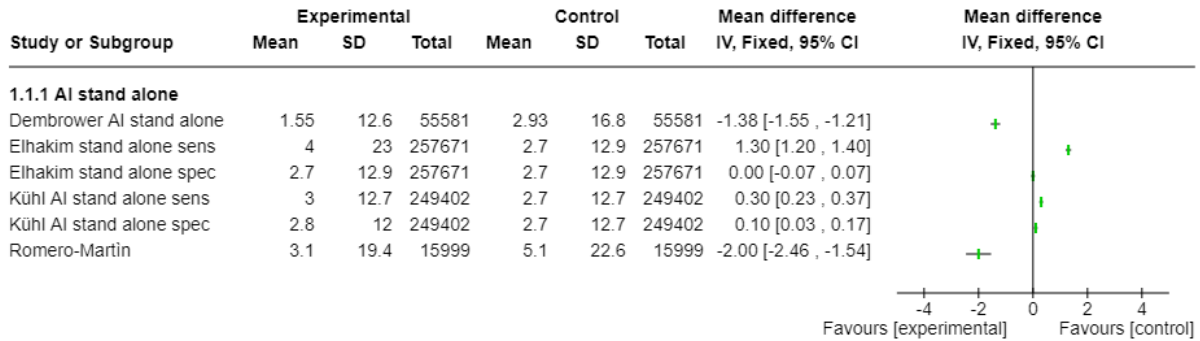
Vid avvikelser noterad vid bröstcancerscreening med mammografi återkallas kvinnan för ytterligare bilddiagnostik och ibland även provtagning. Beslutet att återkalla kvinnan sker ofta efter diskussion mellan två radiologer, så kallat konsensusmöte, vilket är rutin i Södra sjukvårdsregionen. Andelen som återkallas av alla som genomgår bröstcancerscreening med mammografi kallas för återkallandefrekvens. Tabell 6 visar hur användning av AI påverkar återkallandefrekvens samt andel kvinnor som återkallas i onödan, det vill säga frias från cancer efter genomgången utredning. Andelen falskt återkallade speglas också i specificiteten som redovisas under utfallsmått O1: Cancerdetektion (tabell 5). Eftersom de flesta inkluderade studier är retrospektiva, blir utfallsmåttet återkallandefrekvens i dessa studier teoretiskt, det vill säga att studierna rapporterar frekvensen patienter som skulle ha blivit återkallade om AI använts.

Studierna visar stora variationer av återkallandefrekvens avseende både intervention och kontroll i form av dubbelgranskning utförd av radiologer. Som exempel varierar det mellan 2,0 % (Lång 2023) och 13,4 % (Heywang-Köbrunner 2023) för dubbelgranskning utförd av radiolog, där båda studierna använde konsensusmöten. För interventionen är variationen ännu större, (1,6 % Dembrower 2023 och 25,4 % Heywang-Köbrunner 2023). I artiklarna av Raya-Povedano (2021) och Romero-Martín (2022), som utgår från samma studiepopulation, förekommer inte konsensusmöten. Det innebär att om en granskare vill återkalla patienten sker detta oavsett vad den andra granskaren har bedömt.

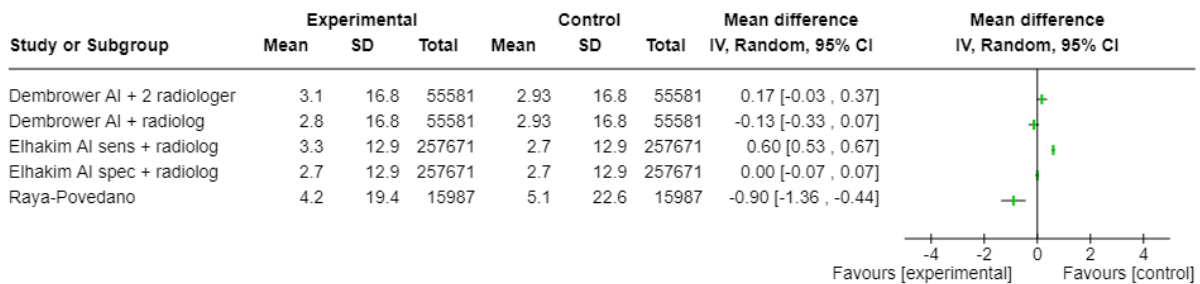
De svenska prospektiva studierna Dembrower (2023) och Lång (2023) uppvisar delvis olika resultat vad gäller återkallandefrekvensen och frekvensen av falskt återkallade. I Dembrower (2023) redovisas att när AI används som stand-alone, minskar andelen återkallade med statistisk signifikans samtidigt som andelen falskt återkallade minskar, dock redovisas inga test för skillnad i falskt återkallade. När AI används som granskare tillsammans med en radiolog (AI+ 1 radiolog), minskar återkallandefrekvensen, dock ej med statistisk signifikans, inklusive de falskt återkallade. När man däremot använder AI som granskare i tillägg till dubbelgranskning (AI + 2 radiologer) ses att återkallandefrekvensen ökar något, dock inte med statistisk signifikans och frekvensen av falskt negativa ökar också något. I Lång (2023) där AI används för triage och som granskningsstöd ökar återkallande något, dock ej statistiskt signifikant, medan frekvensen av falskt återkallade är oförändrad.

För att åskådliggöra variationen avseende återkallandefrekvens mellan artiklarna redovisas resultat och konfidensintervall nedan i forest plots (figur 4–6), uppdelat på olika sätt att använda AI (samt RCT separat). Observera att Heywang-Köbrunner (2023) och Larsen (2022) inte åskådliggörs här då nödvändiga data saknas i artiklarna. Heywang-Köbrunner (2023) visar att återkallandefrekvensen ökar

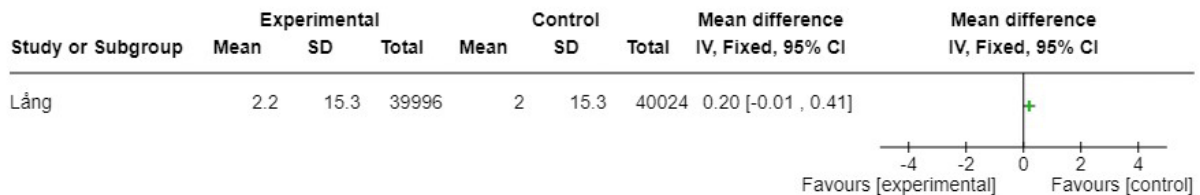
med AI, oavsett hur AI används, medan Larsen (2022) visar minskning av återkallandefrekvens i tio olika scenarier, varav två med AI stand-alone och resterande med AI som triage.



Figur 4. Resultat för återkallandefrekvens för stand-alone AI (%)



Figur 5. Resultat för återkallandefrekvens för AI + radiolog (%)



Figur 6. Resultat för återkallandefrekvens för RCT (%)

Slutsats gällande återkallandefrekvens är således beroende av hur AI används, som granskare, triage eller som granskningsstöd samt vald inställning av AI-produkten, exempelvis gällande sensitivitet och specificitet.

Sammanfattningsvis visar resultaten i de inkluderade artiklarna:

Användning av AI som stand-alone vid bröstcancerscreening med mammografi ger varierande återkallandefrekvens jämfört med dubbelgranskning utförd av radiologer.

Användning av AI i kombination med radiolog vid bröstcancerscreening med mammografi ger likvärdig återkallandefrekvens jämfört med dubbelgranskning utförd av radiologer.

Tabell 6. Utfallsmått O2. Återkallandefrekvens

Författare År Land Studienamn	Antal screening- tillfällen (kvinnor)	Återkallandefrekvens (95 % konfidensintervall)		Falskt återkallad ^a (n)		Kommentarer
		Intervention (AI)	Kontroll (DR)	Intervention (AI)	Kontroll (DR)	
Dembrower 2023 Sverige ScreenTrustCAD	55 581 (55 581)	1: 1,6 % (1,5–1,7)* 2: 2,8 % (2,7–2,9) 3: 3,1 % (3,0–3,2)	2,9 % (2,8– 3,1)	1: 1,1 %* 2: 2,3 % 3: 2,6 %	2,5 %	
Elhakim 2023 Danmark	257 671 (153 372)	1: 4,0 % (3,9–4,1)* 2: 2,7 % (2,6–2,7)* 3: 3,3 % (3,3–3,4) 4: 2,7 % (2,6–2,7)	2,7 % (2,6– 2,7)	1: 4,0 % (10 341)* 2: 2,2 % (5 678)* 3: 2,7 % (7 014) 4: 2,1 % (5 396)	2,1 % (5 352)	
Heywang-Köbrunner 2023 Tyskland	17 884 (17 884)	1: 20,3 %* 2: 25,4 % 3: 25,4 %	13,4 %	1: 19,7 % (3 519)* 2: 24,8 % (4 436) 3: 24,8 % (4 429)	12,7 % (2 275)	Återkallandefrekvens beräknad utifrån antal angivna positiva fynd, inget konsensusmöte inräknat
Kühl 2023 Danmark	249 402 (149 495)	1: 3,0 % (2,9–3,0)* 2: 2,8 % (2,7–2,8)*	2,7 % (2,7– 2,8)	1: 2,4 (6 075)* 2: 2,7 (6 905)*	2,1 % (5 342)	
Larsen 2022 Norge	122 969 (47 877)	1: 2,7 %* 2: 2,8 % 3: 2,6 % 4: 2,1 % 5: 1,2 % 6: 2,2 % 7: 3,1 % 8: 2,9 % 9: 2,5 % 10: 2,3 % 11: 3,2 %*	3,2 %	1: 2,1 % (2 636)* 2: 2,2 % (2 683) 3: 2,0 % (2 466) 4: 1,5 % (1 845) 5: 0,6 % (784) 6: 1,6 % (2 018) 7: 2,4 % (3 002) 8: 2,3 % (2 686) 9: 1,9 % (2 297) 10: 1,8 % (2 187) 11: 2,7 % (3 341)*	2,6 % (3 144)	

Författare År Land Studienamn	Antal screening- tillfällen (kvinnor)	Återkallandefrekvens (95 % konfidensintervall)		Falskt återkallad ^a (n)		Kommentarer
		Intervention (AI)	Kontroll (DR)	Intervention (AI)	Kontroll (DR)	
Lauritzen 2022 Danmark	114 421	1: -* 2: -	2,5 %	1: 5,1 % (5 825)* 2: 1,3 % (1 578)	1,7 % (2 107)	
Lång 2023 Sverige MASAI	I: 39 996 C: 40 024	2,2 % (2,0–2,3)	2,0 % (1,9– 2,2)	1,5 % (617)	1,5 % (614)	
Raya-Povedano 2021 Spanien CBTST	15 987 (15 986)	4,2 % (3,9–4,5)	5,1 % (4,7– 5,4)	3,7 % (593)	4,6 % (731)	
Romero-Martín 2022 Spanien CBTST	15 999 (15 998)	3,1 % (2,8–3,4)*	5,1 % (4,7– 5,4)	2,6 % (413)*	4,6 % (732)	

* AI som stand-alone
^a Antalet med positiva fynd men utredning utan bröstcancerdiagnos efter återkallelse/antalet screenade

3.3.3 Utfallsmått O3: Intervallcancer

Med intervallcancer avser man bröstcancer som diagnosticeras i intervallet mellan två screeningtillfällen (Perry 2006). Det innebär en cancer som inte diagnosticerades i samband med screening (negativt screeningtillfälle, kan ha inkluderat återkallande), utan som upptäcktes till exempel genom att kvinnan själv har känt en förändring i bröstet. Intervallcancer är oundvikliga, men målet är att de ska vara så få som möjligt. Om intervallcancerarna blir för många riskerar värdet av screening att sjunka.

Användning av AI vid bröstcancerscreening med mammografi kan potentiellt leda till såväl färre som fler intervallcancer. Eftersom antalet intervallcancer kan ha stor betydelse för effektiviteten av screeningprogrammet, är detta utfallsmått centralt för att värdera effekterna av AI vid bröstcancerscreening med mammografi.

De två prospektiva studierna Lång (2023) och Dembrower (2023) studerar inte utfallet intervallcancer. Däremot rapporterar Dahlblom (2021), Elhakim (2023), Köhl (2023), Lauritzen (2022), Raya-Povedano (2021) och Romero-Martín (2022) att en viss andel (9,1–24,7 %) av de tumörer som diagnosticerades som intervallcancer retrospektivt kunde identifieras av AI vid den föregående screeningundersökningen. Dock går det inte att säga om det i förlängningen hade lett till att en cancer faktiskt diagnosticerats eftersom studierna är retrospektiva.

Sammanfattningsvis visar resultatet i de inkluderade studierna:

Användning av AI inom bröstcancerscreening med mammografi retrospektivt identifierar en varierande andel av de tumörer som diagnosticerades som intervallcancer vid den föregående screeningundersökningen.

3.3.4 Utfallsmått O4: TNM-stadium

När en bröstcancer diagnosticeras karakteriseras den mycket noggrant utifrån cellbild, storlek och förekomst av lymfkörtel- och distansmetastaser. Detta för att dessa karakteristika är av stark prognostisk betydelse som därmed påverkar val av behandling och dessutom fungerar som ett surrogatmått för överlevnad. Endast ett fåtal av de ingående studierna redovisar detta utfallsmått.

I de retrospektiva studierna går det enbart att studera hur de cancer som AI identifierar förhåller sig inom respektive TNM-stadium för de cancer som radiolog redan identifierat. Utifrån resultaten som redovisas kan man därmed endast dra slutsatser från de två prospektiva studierna. Lång (2023) visar att andelen DCIS (förknippat med mycket god prognos jämfört med invasiv cancer) är högre då AI

används som granskningsstöd än vid jämförande dubbelgranskning utförd av radiolog. Det anges dock ej om resultatet är av statistisk signifikans. Dembrower (2023) visar ingen skillnad mellan AI-granskning plus en radiolog och dubbelgranskning utförd av radiolog gällande andelen invasiva cancrar eller utifrån tumörstorlek eller lymfkörtelmetastaser (se tabell 7).

Sammanfattningsvis visar resultaten från de inkluderade studierna:

Användning av AI vid bröstcancerscreening med mammografi ger varierande resultat avseende vilket TNM-stadium de detekterade cancrarna är i jämfört med dubbelgranskning utförd av radiologer.

Tabell 7. Utfallsmått O4. TNM-stadium

Författare År Land Studienamn	Antal screening- tillfällen (kvinnor)	Andel DCIS av antal diagnosticerade (n)		Andel invasiv cancer av antal diagnosticerade (n)		Andel i storlek > 2 cm % av antal invasiva cancrar (n)		Andel med lymfkörtelmetastas av antal invasiva cancrar (n)		Kommentar
		Intervention (AI)	Kontroll (DR)	Intervention (AI)	Kontroll (DR)	Intervention (AI)	Kontroll (DR)	Intervention (AI)	Kontroll (DR)	
Dembrower 2023 Sverige	55 581 (55 581)	-	-	1: 76,0 % (187)* 2: 76,6 % (200) 3: 76,6 % (206)	78,0 % (195)	1: 41,7 % (78)* 2: 40,5 % (81) 3: 39,8 % (82)	39,4 % (77)			Storlek >2 cm eller positivt lymfkörtel- status i samma grupp delat på antal invasiv cancer
Elhakim 2023 Danmark	257 671 (153 372)	1: 13,8 % (179) 2: 14,3 % (171) 3: 13,3 % (207) 4: 13,3 % (202)	14,0 % (211)	1: 86,2% (1 122)* 2: 85,7% (1 026)* 3: 86,7% (1 348) 4: 86,7% (1 320)	86,0 % (1298)	1: 19,6 % (220)* 2: 19,5 % (200)* 3: 18,5 % (250) 4: 18,0 % (238)	15,9 % (207)	1: 26,3 % (296)* 2: 26,3 % (267)* 3: 25,4 % (343) 4: 24,8 % (328)	24,2 % (314)	
Kühl 2023	249 402 (149 495)	1: 12,9 % (167) 2: 13,0 % (165)	14,0 % (209)	1: 87,0 % (1125)* 2: 87,0 % (1108)*	86,0 % (1268)	1: 20,8 % (234)* 2: 21,4 % (238)	15,2% (193)	1: 27,8 % (313)* 2: 27,8 % (308)	24,1 % (306)	
Lång 2023 Sverige MASAI	I: 39 996 C: 40 024	24,6 % (60)	18,7 % (38)	75,4 % (184)	81,2 % (165)	17,4 % (32)	21,8 % (36)	-	-	

* AI som stand-alone

3.3.5 Utfallsmått O5: Bröstcancer mortalitet

Inga studier identifierades som undersökte om AI vid bröstcancerscreening med mammografi påverkar bröstcancer mortalitet.

3.3.6 Utfallsmått O6: Radiologbelastning

I nuvarande klinisk rutin i Sverige används dubbelgranskning utförd av radiolog (se även avsnitt 5.1) för granskning av alla mammografier som utförs i screeningprogrammet. Vid oenighet anordnas konsensusmöte för att avgöra om kvinnan ska återkallas för utökad utredning. Med radiologbelastning avses hur mycket radiologerna belastas för granskning av screeningbilder.

Belastningen mäts i majoriteten av studierna genom antal bilder som behöver granskas i ett system med AI-användning jämfört med i nuvarande kliniskt rutinsystem med dubbelgranskning (Dembrower 2023, Larsen 2022, Lauritzen 2022, Lång 2023, Raya-Povedano 2021). Generellt sjunker radiologbelastningen när AI används vid bröstcancerscreening med mammografi. Det finns dock stora skillnader i vinster, vilket kan bero på den varierande utformningen av AI-användningen i de olika studierna. De prospektiva studierna (Dembrower 2023, Lång 2023) har även studerat antal konsensusmöten. Lång (2023) påvisar ingen skillnad mellan ett system med eller utan AI, medan Dembrower (2023) visar att antalet konsensusmöten ökar något. Heywang-Köbrunner (2023) anger att antalet konsensusmöten hypotetiskt dubbleras. Det bör noteras att ingen studie har mätt den faktiska tidsåtgången (se tabell 8).

Sammanfattningsvis visar resultaten i de inkluderade artiklarna:

Användning av AI vid bröstcancerscreening med mammografi minskar radiologbelastningen jämfört med dubbelgranskning utförd av radiologer.

Tabell 8. Utfallsmått O6. Radiologbelastning

Författare År Land Studienamn	Antal screening-tillfällen (kvinnor)	Antal (% förändring) granskningar utförda av radiolog		Antal (andel förändring) timmar granskning utföres av radiolog		Andel konsensusmöten (antal)		Kommentar
		Intervention (AI)	Kontroll (DR)	Intervention (AI)	Kontroll (DR)	Intervention (AI)	Kontroll (DR)	
Dembrower 2023	55 581 (55 581)	27 791 (-50 %)	55 581	–	–	1: 5,69 % (3 162)* 2: 8,95 % (4 972) 3: 10,8 % (6 002)	7,38 % (4 104)	
Heywang-Köbrunner 2023 Tyskland	17 884 (17 884)	1: –* 2: – 3: –	–	–	–	1: – 2: 25 % 3: –	12,8 %	Uppskattar att antalet fall som kräver konsensusdiskussioner fördubblas
Larsen 2022	122 969 (47 877)	1: 61 485 (-50 %)* 2: 261 485 (-50 %) 3: 61 485 (-50 %) 4: 86 078 (-70 %) 5: 110 672 (-90 %) 6: 92 227 (-75 %) 7: 30 742 (-25 %) 8: 43 039 (-35 %) 9: 77 470 (-63 %) 10: 77 470 (-63 %) 11: 0 (-100 %)*	122 969	–	–	1: 10,4 % (12 724)* 2: 14,2 % (17 394) 3: 6,5 % (7 955) 4: 5,0 % (5 894) 5: 2,3 % (2 805) 6: 4,3 % (5 301) 7: 7,9 % (9 683) 8: 7,2 % (8 880) 9: 5,6 % (6 934) 10: 5,1 % (6 322) 11: – *	8,8 % (10 790)	Uppskattat antal konsensusmöten
Lauritzen 2022	114 421	1: 0 (-100 %)* 2: 42 836 (-62,6 %)	114 421	–	–	–	–	

Författare År Land Studienamn	Antal screening- tillfällen (kvinnor)	Antal (% förändring) granskningar utförda av radiolog		Antal (andel förändring) timmar granskning utföres av radiolog		Andel konsensusmöten (antal)		Kommentar
		Intervention (AI)	Kontroll (DR)	Intervention (AI)	Kontroll (DR)	Intervention (AI)	Kontroll (DR)	
Lång 2023 Sverige MASAI	I: 39 996 C: 40 024	46 345 (-44,3 %)	83 231	–	–	4 % (1 584)	3.9 % (1 576)	
Raya- Povedano 2021 Spanien CBTST	15 987 (15 986)	9 100 (-71,5 %)*	31 974	63 (-71,5 %)	222	0	0	Ingen tid för granskning mättes i studien. Man skattade ungefärlig tidsåtgång per granskning till 25 sekunder för en mammografiundersökning. Om någon granskare återkallade, blev det återkallat – inga konsensusmöten
* AI som stand-alone								

3.3.7 Utfallsmått O7: Livskvalitet

Inga studier identifierades som undersökte hur AI vid bröstcancerscreening med mammografi påverkar livskvalitet.

3.4 Analys av effektmått

Studierna presenterar resultaten av AI-användning baserat på olika scenarier med varierande gränsvärden för exempelvis sensitivitet och specificitet, vilket förhindrade en sammanlagd analys. Heterogenitet förekom även inom flera aspekter i PICO. Inklusionskriterierna varierade mellan studierna, vilket resulterade i olika populationer (P). AI tillämpades på olika sätt i de olika studiernas scenarier (I), och standardradiologisk granskning (C) skilde sig mellan länderna där studierna utfördes. Dessutom varierade bedömningen av vissa utfallsmått (O) mellan studierna. Därför grundas evidensgraderingen på en narrativ analys enligt etablerade riktlinjer för hur detta ska göras och rapporteras (Campbell 2020). Detta innebär att det inte är möjligt att erhålla någon tydlig numerisk effekt med det befintliga underlaget. Se faktarutan nedan om evidensgradering.

3.5 Evidensgradering

Faktaruta evidensgradering och GRADE

I en systematisk översikt väger man samman resultat från olika studier. Detta kan till exempel göras i en metaanalys när detta är möjligt eller i en narrativ analys. Man gör en sammanvägning för varje utfall som utvärderas. När sammanvägningen är gjord så är det sista steget i den systematiska översikten att bedöma hur tillförlitliga de sammanvägda resultaten är. HTA syd använder likt SBU det internationellt utarbetade GRADE-systemet som ett stöd i bedömningarna (SBU 2023).

Alla bedömningar ska motiveras för transparens så att läsaren kan göra en egen värdering av tillförlitligheten. Det sammanvägda resultatet bedöms med GRADE utifrån fem olika aspekter: 1) Kvalitet 2) Samstämmighet och överensstämmelse 3) Överförbarhet och relevans 4) Precision i data och 5) Publikationssnedvridning

Utifrån detta klassificeras sedan tillförlitligheten i fyra nivåer:

Det sammanvägda resultatet har **hög tillförlitlighet** ⊕⊕⊕⊕. Bedömningen är att resultatet stämmer.

Det sammanvägda resultatet har **måttlig tillförlitlighet** ⊕⊕⊕○. Bedömningen är att det är troligt att resultatet stämmer.

Det sammanvägda resultatet har **låg tillförlitlighet** ⊕⊕○○. Bedömningen är att det är möjligt att resultatet stämmer.

Det sammanvägda resultatet har **mycket låg tillförlitlighet** ⊕○○○. Det går inte att bedöma om resultatet stämmer.

En måttlig tillförlitlighet kan tolkas som att det är troligt att resultatet stämmer medan en låg tillförlitlighet kan tolkas som att det är möjligt att resultatet stämmer. I båda fallen är resultaten osäkra och tolkningen är att det i de flesta fall behövs mera forskning.

När det saknas studier som uppfyller inklusionskriterierna anges "studier saknas" utan klassificering.

Evidensstyrkan för varje utfallsmått redovisas i en separat GRADE-tabell enligt nedan. Observera att flera av artiklarna är baserade på material som används i mer än en artikel, vilket framgår av tabeller i avsnitt 3.3. Eftersom det saknas studier av effekten på bröstcancermortalitet och livskvalitet, kan någon evidensgradering inte göras avseende dessa utfallsmått.

Evidens- gradering enligt GRADE	Utfallsmått: Cancer- detektion	8 NRSI					Studiedesign: kohortstudier ⊕⊕⊕⊕		
O1 AI stand- alone	Sänkande faktorer					Höjande faktorer			Total tillförlitlighet
	Kvalitet	Samstämmighet och överensstämmelse	Överförbarhet och relevans	Precision i data	Publikations- snedvridning	Effektstorlek	Dos-respons samband	Effektunderskattning	
Dahlblom 2021, Dembrower 2023, Elhakim 2023, Heywang- Köbrunner 2023, Kühl 2023, Larsen 2022, Lauritzen 2022, Romero-Martín 2022	0	-1	-1	0	0	0	0	0	⊕⊕○○
Slutsats	Användning av AI som stand-alone vid bröstcancerscreening med mammografi ger lägre eller likvärdig cancerdetektion jämfört med dubbelgranskning utförd av radiologer (låg tillförlitlighet ⊕⊕○○).								
Kommentar	Olika inklusionskriterier (P), flera varianter av intervention (I), olika utformning av dubbelgranskning (C), varierande definitioner av sann cancer (O) bidrar till avdrag för överförbarhet och relevans). Ej samstämmiga resultat. Även flera typer av koppling till de olika företagen i majoriteten av studierna, inga publikationer med negativa resultat. Således bidrar både brister i överförbarhet, relevans, samstämmighet samt viss publikationssnedvridning till avdragen.								

Evidensgradering enligt GRADE	Utfallsmått: Cancer-detektion	1 RCT, 6 NRSI					Studiedesign: RCT och kohortstudier ⊕⊕⊕⊕		
O1 AI i kombination med radiolog	Sänkande faktorer					Höjande faktorer			Total tillförlitlighet
	Kvalitet	Samstämmighet och överensstämmelse	Överförbarhet och relevans	Precision i data	Publicerings-snedvridning	Effektstorlek	Dos-respons samband	Effektunderskattning	
Dembrower 2023, Elhakim 2023, Heywang-Köbrunner 2023, Larsen 2022, Lauritzen 2022, Lång 2023 (RCT), Raya-Povedano 2021	0	0	-1	0	0	0	0	0	⊕⊕⊕○
Slutsats	Användning av AI i kombination med radiolog vid bröstcancerscreening med mammografi ger likvärdig eller högre cancerdetektion jämfört med dubbelgranskning utförd av radiologer (måttlig tillförlitlighet ⊕⊕⊕○).								
Kommentar	Olika inklusionskriterier (P), flera varianter av intervention (I), olika utformning av dubbelgranskning (C), varierande definitioner av sann cancer (O) bidrar till avdrag för överförbarhet och relevans. Även flera typer av koppling till de olika företagen i majoriteten av studierna, inga publikationer med negativa resultat. Således bidrar både brister i överförbarhet och relevans samt viss publikationssnedvridning till avdraget.								

Evidens- gradering enligt GRADE	Utfallsmått: Återkallande- frekvens	4 NRSI				Studiedesign: kohortstudier ⊕⊕⊕⊕			
O2 AI stand- alone	Sänkande faktorer					Höjande faktorer			Total tillförlitlighet
	Kvalitet	Samstämmighet och överensstämmelse	Överförbarhet och relevans	Precision i data	Publications- snedvridning	Effektstorlek	Dos-respons samband	Effektunderskattning	
Dembrower, 2023, Elhakim, 2023, Kühl, 2023, Romero-Martín 2022	-1	-1	-1	0	0	0	0	0	⊕○○○
Slutsats	Det går inte att dra någon slutsats om användning av AI som stand-alone vid bröstcancerscreening med mammografi avseende återkallandefrekvens jämfört med dubbelgranskning utförd av radiologer (mycket låg tillförlitlighet ⊕○○○).								
Kommentar	Studiedesignen i majoriteten av studierna ger ingen möjlighet att dra någon slutsats om den faktiska återkallandefrekvensen, eftersom uppföljningen är teoretisk, vilket ger studierna avdrag för kvalitet. Olika nivåer på sensitiviteten i interventionen ger olika resultat på återkallandefrekvensen (I) vilket ger brister i samstämmigheten. Olika nivåer på återkallandefrekvens i klinisk rutin i olika studier (C) bidrar till avdrag för överförbarhet och relevans. Även flera typer av koppling till de olika företagen i majoriteten av studierna, inga publikationer med negativa resultat. Således bidrar även viss publikationssnedvridning till avdragen.								

Evidens- gradering enligt GRADE	Utfallsmått: Återkallande- frekvens	1 RCT, 8 NRSI				Studiedesign: RCT och kohortstudier: ⊕⊕⊕⊕			
O2 AI i kombi- nation med radiolog	Sänkande faktorer					Höjande faktorer			Total tillförlitlighet
	Kvalitet	Samstämmighet och överensstämmelse	Överförbarhet och relevans	Precision i data	Publikations-snedvridning	Effektstorlek	Dos-respons samband	Effektunderskattning	
Dembrower, 2023, Elhakim, 2023, Heywang-Köbrunner 2023, Köhl, 2023, Larsen 2022, Lauritzen 2022, Lång 2023 (RCT), Raya-Povedano 2021	0	-1	-1	0	0	0	0	0	⊕⊕○○
Slutsats	Användning av AI i kombination med radiolog vid bröstcancerscreening med mammografi ger likvärdig återkallandefrekvensen jämfört med dubbelgranskning utförd av radiologer (låg tillförlitlighet ⊕⊕○○).								
Kommentar	Olika nivåer på sensitiviteten i interventionen ger olika resultat på återkallandefrekvensen (I) vilket ger brister i samstämmigheten. Olika nivåer på återkallandefrekvens i klinisk rutin i olika studier (C) bidrar till avdrag för överförbarhet och relevans. Även flera typer av koppling till de olika företagen i majoriteten av studierna, inga publikationer med negativa resultat. Således bidrar även viss publikations-snedvridning till avdragen.								

Evidensgradering enligt GRADE	Utfallsmått: Intervall-cancerfrekvens	6 NRSI					Studiedesign: kohortstudier: ⊕⊕⊕⊕		
O3	Sänkande faktorer					Höjande faktorer			Total tillförlitlighet
	Kvalitet	Samstämmighet och överensstämmelse	Överförbarhet och relevans	Precision i data	Publications-snedvridning	Effektstorlek	Dos-respons samband	Effektunderskattning	
Dahlblom 2021, Elhakim 2023, Köhl 2023, Lauritzen 2022, Raya-Povedano 2021, Romero-Martín 2022	-1	-1	-1	0	0	0	0	0	⊕○○○
Slutsats	Det går inte att dra någon slutsats om hur användning av AI vid bröstcancerscreening med mammografi påverkar intervallcancerfrekvensen jämfört med dubbelgranskning utförd av radiologer (mycket låg tillförlitlighet ⊕○○○).								
Kommentar	Studierna ej designade för detta utfallsmått, olika definitioner av intervallcancer, AI kan inte ge den finala diagnosen utan endast återkalla. Detta ger brister i kvalitet, överensstämmelse och överförbarhet i studierna. Även flera typer av koppling till de olika företagen i majoriteten av studierna, inga publikationer med negativa resultat. Således bidrar även viss publikationssnedvridning till avdragen.								

Evidensgradering enligt GRADE	Utfallsmått: TNM-stadium	1 RCT, 3 NRSI					Studiedesign: RCT och kohortstudier: ⊕⊕⊕⊕		
O4	Sänkande faktorer					Höjande faktorer			Total tillförlitlighet
	Kvalitet	Samstämmighet och överensstämmelse	Överförbarhet och relevans	Precision i data	Publications-snedvridning	Effektstorlek	Dos-respons samband	Effektunderskattning	
Dembrower 2023, Elhakim 2023, Köhl 2023, Lång 2023 (RCT)	-1	-1	-1	-1	0	0	0	0	⊕○○○
Slutsats	Det går inte att dra någon slutsats om hur användning av AI vid bröstcancerscreening med mammografi påverkar detektion av cancer i olika stadier jämfört med dubbelgranskning utförd av radiologer (mycket låg tillförlitlighet ⊕○○○).								
Kommentar	Studierna ej designade för detta utfallsmått, AI kan inte ge den finala diagnosen, utan endast återkalla. Precisionen är låg, små populationer med cancer. Detta ger brister i kvalitet, överensstämmelse och överförbarhet i studierna. Även flera typer av koppling till de olika företagen i majoriteten av studierna, inga publikationer med negativa resultat. Således bidrar även viss publikationssnedvridning till avdragen. Otillräcklig tillförlitlighet anges för tre avdrag eller mer.								

Evidens- gradering enligt GRADE	Utfallsmått: Radiolog- belastning	1 RCT, 5 NRSI					Studiedesign: RCT och kohortstudier: ⊕⊕⊕⊕		
O6	Sänkande faktorer					Höjande faktorer			Total tillförlitlighet
	Kvalitet	Samstämmighet och överensstämmelse	Överförbarhet och relevans	Precision i data	Publikations- snedvridning	Effektstorlek	Dos-respons samband	Effektunderskattning	
Dembrower 2023, Heywang- Köbrunner 2023, Larsen 2022, Lauritzen 2022, Lång 2023 (RCT), Raya-Povedano 2021	-1	0	0	-1	0	0	0	0	⊕⊕○○
Slutsats	Användning av AI vid bröstcancerscreening med mammografi minskar radiologbelastningen jämfört med dubbelgranskning utförd av radiologer (låg tillförlitlighet ⊕⊕○○).								
Kommentar	Studiedesignen är inte optimal för detta utfallsmått i majoriteten av studierna. Resultaten redovisas i antal bilder som granskas och antal möten, men inte hur mycket tid detta tar. Således avdrag för kvalitet och precision. Även flera typer av koppling till de olika företagen i majoriteten av studierna, inga publikationer med negativa resultat, därför bidrar även viss publikationsnedvridning till avdragen.								

3.6 Sammanställning av kunskapsläget

- Användning av AI som stand-alone vid bröstcancerscreening med mammografi ger lägre eller likvärdig cancerdetektion jämfört med dubbelgranskning utförd av radiologer (låg tillförlitlighet ⊕⊕○○).
- Användning av AI i kombination med radiolog vid bröstcancerscreening med mammografi ger likvärdig eller högre cancerdetektion jämfört med dubbelgranskning utförd av radiologer (måttlig tillförlitlighet ⊕⊕⊕○).
- Det går inte att dra någon slutsats om användning av AI som stand-alone vid bröstcancerscreening med mammografi avseende återkallandefrekvens jämfört med dubbelgranskning utförd av radiologer (mycket låg tillförlitlighet ⊕○○○).
- Användning av AI i kombination med radiolog vid bröstcancerscreening med mammografi ger likvärdig återkallandefrekvensen jämfört med dubbelgranskning utförd av radiologer (låg tillförlitlighet ⊕⊕○○).

- Det går inte att dra någon slutsats om hur användning av AI vid bröstcancerscreening med mammografi påverkar intervallcancerfrekvensen jämfört med dubbelgranskning utförd av radiologer (mycket låg tillförlitlighet ⊕○○○)
- Det går inte att dra någon slutsats om hur användning av AI vid bröstcancerscreening med mammografi påverkar detektion av cancer i olika TNM-stadier jämfört med dubbelgranskning utförd av radiologer (mycket låg tillförlitlighet ⊕○○○).
- Användning av AI vid bröstcancerscreening med mammografi minskar radiologbelastningen jämfört med dubbelgranskning utförd av radiologer (låg tillförlitlighet ⊕⊕○○).

Gällande de övriga utfallsmåtten i rapporten - livskvalitet och bröstcancer mortalitet - är dessa inte studerade i någon av de inkluderade artiklarna.

4 Riktlinjer och rekommendationer

Socialstyrelsens senaste rekommendation om bröstcancerscreening (Socialstyrelsen 2023) beskrev AI-stöd vid bildtolkning som en möjlig framtida utveckling av screeningprogrammet. Socialstyrelsen lyfte dock fram att det pågår studier i Sverige som utvärderar AI. ECIBC är ett centralt organ för europeiska screeningrekommendationer och brukar fortlöpande uppdatera sina rekommendationer. Den senaste uppdateringen är från maj 2024, då ECIBC gav en villkorad rekommendation (med lågt vetenskapligt stöd) att använda AI i dubbelgranskning i stället för dubbelgranskning utan AI samt att inte ersätta dubbelgranskning med enkelgranskning med AI (ECIBC 2024).

Ett första nationellt vårdprogram för bröstcancerscreening har varit ute på remiss och kommer snart att publiceras (Regionala cancercentrum i samverkan [RCC] 2024). Vårdprogrammet sammanfattar att det ännu inte finns någon uppdaterad rekommendation från ECIBC eller Socialstyrelsen om användandet av AI, efter det att resultat från prospektiva studier har blivit publicerade. Man konstaterar även att för att kunna ge en generell rekommendation behövs ytterligare evidens med resultat beträffande intervallcancerfrekvens, kostnadseffektivitet, optimalt screeninggranskningsprotokoll och reproducerbara resultat.

I en gemensam skrivelse (Brady 2024) av radiologiska samfund i Europa, USA, Kanada, Australien och Nya Zeeland framförs potentiella problem och etiska aspekter av användandet av AI i radiologisk verksamhet, och vikten av en säker hantering vid utvärdering och införande betonas. Även Socialstyrelsen och Nationellt vårdprogram för bröstcancerscreening understryker vikten av framtagande av rutiner för hur AI-produkterna ska användas (Socialstyrelsen 2023; RCC 2024).

I vårdprogrammet betonas att om en screeningenhet väljer att införa AI, måste det ske inom ramen för

ett ordnat införande med krav på noggrann monitorering av algoritmens prestanda och kvalitetsindikatorerna för screening.

Det finns inga andra relevanta rekommendationer från europeiska riktlinjer.

Sammanfattningsvis finns för närvarande inga nationella eller internationella riktlinjer som rekommenderar användning av AI som ersättning för en mänsklig granskare vid dubbelgranskning av screeningbilder i bröstcancerscreening med mammografi.

5 Praxisundersökning

5.1 Bröstcancerscreening med mammografi i Södra sjukvårdsregionen

Socialstyrelsen rekommenderar att personer i åldern 40–74 år och som vid födseln tillskrivits det juridiska könet kvinna erbjuds bröstcancerscreening med mammografi med 18–24 månaders intervall. Regionerna bör ha rutiner för att kalla och ta emot transpersoner som ingår i målpopulationen (Socialstyrelsen 2023). För personer med ärftligt ökad risk kan andra intervaller eller metoder vara aktuella. I Region Blekinge, Region Kronoberg och Region Skåne, men inte i Region Halland, ingår även uppföljningskontroller efter genomgången bröstcancerbehandling i screeningens arbetsflöde, även om de inte ingår i screeningpopulationen. Dessa undersökningar har dock en separat undersökningskod.

Screeningundersökningen omfattar två bilder per bröst (kraniokaudal och mediolateral oblik). Bilderna kompletterar varandra. Bildtagningen sker på ett standardiserat sätt enligt särskilda positioneringskriterier. På så vis blir bildmaterialet reproducerbart och jämförelse kan göras mellan screeningomgångarna.

Bilderna sparas i en lista i PACS (Picture Archiving and Communication System) där de presenteras enligt ett standardiserat visningsprotokoll, som inkluderar jämförelse med äldre undersökningar. Genom att jämföra med bilder äldre än föregående undersökningstillfälle kan man lättare upptäcka långsamväxande cancrar. Har kvinnan tidigare röntgats vid en annan mammografienhet kan man som granskare begära åtkomst till dessa bilder. Granskningen sker då först när beställda bilder har lästs in i bildsystemet.

Varje screeningundersökning granskas oberoende av två radiologer som är specialiserade på bedömning av mammografibilder. Studier har visat högre sensitivitet med dubbelgranskning än med enkelgranskning (Taylor 2008). Dubbelgranskningen kan vara blindad eller oblandad. Vid blindad

granskning har andregranskaren inte tillgång till förstegranskarens bedömning. I Region Skåne och Region Blekinge är dubbelgranskningen oblandad medan den i Region Kronoberg är blindad.

Utöver bedömning av bildmaterialet tar den granskande radiologen också del av information som har framkommit vid undersökningstillfället. Detta sker i form av en bröstschablon och frågeformulär med fritext.

	Blodig/serös sekretion	
	Palpabel knöl	
	Indragning	
	Hudförändring	
	Ärr	
	Bortopererat bröst	
	Implantat	
	Reduktionsplastik	

Ny knöl?	<input type="checkbox"/> Dx	<input type="checkbox"/> Sin
Genomgått bröstoperation?	<input type="checkbox"/> Dx	<input type="checkbox"/> Sin
- Malignt?	<input type="checkbox"/> Dx	<input type="checkbox"/> Sin
- Vilket år?	<input type="text"/>	<input type="text"/>
Hormon?	<input type="checkbox"/> Ja	
Mammografi på annan ort?	<input type="checkbox"/> Ja	
Var och när?	<input type="text"/>	
Fri text	<input type="text"/>	

Figur 7. Bild från Nationella vårdprogrammet för bröstcancerscreening (RCC 2024)

I samband med screeningundersökningen gör bildtagande personal en visuell bedömning av bröstet. Fynd som kan påverka granskningen eller kan tala för bröstcancer noteras på en bröstschablon (såsom tidigare kirurgi, hudförändring, indragning, blodig sekretion, eksem, se Figur 7). Kvinnan tillfrågas om symtom som kan tala för bröstcancer (knuta, indragningar, förstorat hårt bröst, tecken på inflammation, blodig eller klar sekretion). Eventuell hormonell behandling noteras. Det finns även ett fält för fritext där annan relevant information kan noteras (till exempel symtomduration, om det finns tidigare bilder på en annan enhet, att kvinnan har fått information om att söka vårdcentral). Vid uppgiven knuta markeras denna ut på bröstet med röntgentät markör, så att de anamnestiska fynden kan kopplas till bilden.

Alla bilder från screeningen granskas sedan av två radiologer, så kallad dubbelgranskning. Varje radiolog tar individuellt ställning till om kvinnan ska återkallas för utökad undersökning. Exempel på orsaker till återkallande kan vara symtom i form av knutor och misstänkt tumör på bilderna, men även tekniskt bristfälliga bilder. Fall där åtminstone en radiolog har velat återkalla patienten diskuteras och konsensus uppnås.

För närvarande används inte AI inom screeningprogrammet i Södra sjukvårdsregionen. Region Halland planerar dock att införa AI vid bröstcancerscreening med mammografi i juni 2024.

Om radiologerna efter dubbelgranskning väljer att återkalla så kommer kvinnan på ett nytt besök. Då tas alltid kompletterande bilder där man ibland använder annan teknik (ultraljud, tomosyntes) för att få mer information. Vid behov tas cell- eller vävnadsprov av radiologen vid samma besök, som också är

den som informerar och diskuterar med kvinnan om fynden och fortsatt planering. Över hälften av de som återkallas får ingen cancerdiagnos. Dessa kvinnor återgår till ordinarie screeningprogram.

6 Organisatoriska aspekter

Om AI skulle införas i bröstcancerscreeningen med mammografi skulle flera organisatoriska aspekter behöva beaktas och åtgärdas. Det finns olika AI-produkter och olika sätt att implementera AI i granskning av bilder inom bröstcancerscreeningen med mammografi med varierande effekt på cancerdetektion, återkallning och granskningsvolym. Olika granskningsprotokoll kan vara aktuella beroende på behov, förutsättningar och önskad effekt. Det finns flera typer av mammografiapparater och bildbehandling som kan påverka AI-produktens bedömningar och som därför behöver testas innan införande. Detta kan göras lokalt eller i en nationell valideringsplattform (VAI-B). Fördelningen av AI-produktens riskstratifiering på den aktuella populationen kan då fastslås och eventuellt justeras. Då AI-produkter kontinuerligt förbättras måste ny utvärdering göras vid algoritmuppdatering, vilket kan göras lokalt. AI-produktens prestanda och dess riskstratifiering behöver monitoreras vid användning. En metod för monitorering behöver utvecklas, vilket kan göras i PACS.

I de verksamheter där cancerkontroller och hereditetskontroller hanteras via samma arbetsflöde som screeningen bör det också beaktas att även dessa undersökningar kommer att granskas med stöd av AI i händelse av ett sådant införande.

Vid AI-stödd granskning, där AI används som en av två granskare, kommer en stor del av screeningundersökningarna att enbart granskas av en radiolog. Särskilda kompetenskrav kommer därför att krävas för enkelgranskande radiolog. Ett riktmärke kan vara en granskningsvolym mellan 3 500 och 11 000 screeningundersökningar per år eller resultat från kvalitetsutvärdering av screeninggranskare. Ett exempel på en sådan screeninggranskare är Personal Performance in Mammographic Screening (PERFORMS). Det är en träningsdatabas, med en population som anrikats med cancerfall och används för individuell träning och utvärdering. Enkelgranskande läkare måste vara uppmärksam på symtom som screeningdeltagare har rapporterat vid undersökningen. Screeningenheten måste ombesörja så att mindre erfarna bröstradiologer kan börja screeninggranska på ett säkert sätt, exempelvis att en del av screeningundersökningarna dubbelgranskas oavsett riskstratifiering. Screeningenheterna måste utforma arbetslistor i PACS för att möta dessa behov. Före införande av AI måste läkargruppen utbildas i hur AI-granskning fungerar och hur den ska monitoreras. I dagsläget är granskning av bilder från bröstcancerscreening med mammografi en viktig del i utbildningen till bröstradiolog och en integrerad del av den kliniska verksamheten. Med färre bilder att granska och färre radiologer som arbetar med detta riskerar denna del av utbildningen att försvåras.

Liksom dagens screeningprogram måste AI-stödd bröstcancerscreening med mammografi utvärderas med sammanställning och uppföljning av screeningindikatorer med tillägg av falskt negativa AI-bedömningar. Informationsmaterial om hur AI-granskning går till måste finnas tillgänglig för målpopulationen för screeningen.

7 Ekonomiska aspekter

Ett införande av AI i bröstcancerscreening med mammografi kan påverka resursanvändning, kostnader, hälsa och livskvalitet på såväl kort som lång sikt. Detta avsnitt presenterar en analys av genomsnittskostnader per screenad kvinna på kort sikt för användning av AI i bröstcancerscreening med mammografi med hjälp av fyra alternativa scenarier vid ett möjligt införande. Scenarierna använder uppgifter från de två svenska studierna i sammanställningen, underlag från aktörer i Södra sjukvårdsregionen, en dansk studie samt antaganden för att undersöka hur förväntade effekter som rapporterats påverkar kostnader för bröstcancerscreening med mammografi. Analysen omfattar endast kostnader där användning av AI kan medföra förändringar, och rapporterade kostnader är därför lägre än bröstcancerscreeningsprogrammets totala kostnader.

Tabell 9 nedan beskriver kortfattat analysscenarier 1 till 4, som beskriver olika sätt att använda AI inom bröstcancerscreening med mammografi samt sedvanlig vård, där två radiologer granskar bilderna såsom enligt det svenska vårdprogrammet. Scenario 1 använder resultat från den randomiserade kliniska studien från Sverige (Lång 2023) medan scenario 2 följer en prospektiv, non-inferiority studie, också den från Sverige (Dembrower 2023). Scenarierna 3 och 4 är uppbyggda kring antaganden med syfte att illustrera vikten av enskilda egenskaper. Scenario 3 är en utveckling av scenario 1 och visar påverkan av att även fortsättningsvis använda två radiologer utöver AI. I scenario 3 granskas 38,6 % av bilderna av två radiologer i stället för att det bara är 7,2 % i scenario 1.

Scenario 4 illustrerar en betydligt större förändring jämfört med sedvanlig vård och kallas därför disruptiv användning av AI. Med utgångspunkt i en dansk studie (Lauritzen 2022) använder scenario 4 information från AI också för att sälla vilka bilder som endast AI bedömer och vilka bilder som även ska granskas av en radiolog. Detta scenario är disruptivt eftersom det dels minskar antalet granskare för bilder som AI bedömer har låg risk, dels endast använder en radiolog vilket har stor betydelse för måttet radiologbelastning.

Appendix F sammanställer i två tabeller detaljerade uppgifter om enhetskostnader och procentandelar samt vilka källor som dessa uppgifter hämtats från. Beräkningarna utgår från punktskattningar från publicerade studier och använder även skillnader som inte är statistiskt signifikanta. Analyserna redovisas som genomsnittlig kostnad per kvinna som gör mammografi, uppdelat på kostnader för bildgranskning, konsensusbedömning och återkallande för ytterligare undersökning. Det är 133 775

kvinnor som genomgår mammografi varje år i Region Skåne och i texten kommenteras kostnadsskillnader mellan de olika scenarierna på regional nivå.

Tabell 9. Illustrativa exempel för användning av AI i bröstcancerscreening med mammografi

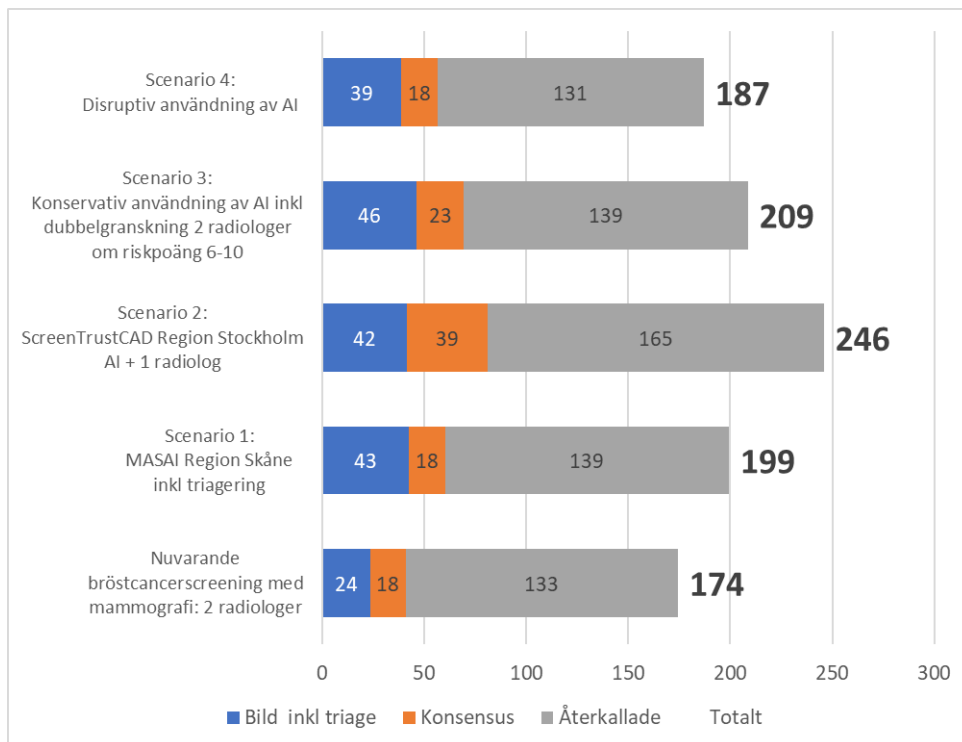
Jämförelselternativ	Innehåll
Scenario 1	<u>MASAI Region Skåne inkl. triagering och granskningsstöd</u> Interventionsgruppen i Lång (2023). AI ersätter radiolog 1 och radiolog 2 granskar alla bilder. Om AI ger riskpoäng 10 granskas bilderna av två radiologer utöver AI. Ökat behov av konsensusmöten och ökad andel återkallade. ^{a)}
Scenario 2	<u>ScreenTrustCAD Region Stockholm</u> Studiegrupp "AI + 1 radiolog" i Dembrower (2023). AI ersätter radiolog 1 och radiolog 2 granskar alla bilder. Ökat behov av konsensusmöten och högre andel återkallade jämfört med nuvarande bröstcancerscreening med mammografi i Skåne. ^{b)}
Scenario 3	<u>Konservativ användning av AI som granskningsstöd</u> Som i scenario 1 men om AI ger riskpoäng 6-10 ^{b)} granskas bilderna av två radiologer utöver AI. Ökat behov av konsensusmöten och ökad andel återkallade. ^{c)}
Scenario 4	<u>Disruptiv och större förändring i handläggning vid introduktion av AI</u> AI ersätter radiolog 1 och genomför triage för radiolog 2 som ej granskar bilder med låg riskpoäng. Oförändrat behov av konsensusmöten men ökad andel återkallade. Andel enbart granskade av AI som i Lauritzen (2022). Andel i konsensus och återkallade som i Lång (2023).
Nuvarande bröstcancerscreening med mammografi-	Två radiologer granskar bilderna. Om olika uppfattning görs bedömning med konsensus. Andel i konsensus och återkallade som i kontrollgrupp i Lång (2023).

^{a)} Andel konsensusmöten och andel återkallade ej statistiskt signifikant skilda mellan studiegrupper. Beräkningarna använder redovisade punktskattningar.

^{b)} Andel konsensusmöten och andel återkallade statistiskt signifikant skilda mellan studiegrupper i Dembrower (2023). Beräkningarna använder redovisade punktskattningar för AI + 1 radiolog där både intervention och kontroll är större än i Lång (2023) som utgör underlag med relevans för nuvarande praxis i Region Skåne.

^{c)} Scenariot inspirerat av Region Hallands planerade införande av användning av AI i bröstcancerscreening med mammografi och underlag om fördelning av riskpoäng från svensk studie (Lång 2023).

Figur 8 och figur 9 nedan jämför genomsnittlig kostnad per kvinna med två olika priser för användning av AI. Dessa prisnivåer används i analysen för att illustrera betydelsen av denna enhetskostnad som finns för varje granskad kvinna vid ett införande av användning av AI och som jämförelse med dagens lönekostnader för radiologtid. Figur 8 använder ett högt pris, 30 kr per granskad kvinna. Detta pris är förhållandevis högt i jämförelse med en beräknad lönekostnad inklusive påslag (11,75 kronor) för en radiolog, som använder genomsnittlig granskningstid (48 sekunder) enligt verksamhetsdata från Unilabs som sköter bröstcancerscreening med mammografi i Region Skåne. Efterföljande figur 9 redovisar därför kostnader per kvinna under samma scenarier, men där priset per kvinna för AI sattes till 10 kronor i stället, det vill säga en enhetskostnad som är lägre än genomsnittlig lönekostnad för en radiolog vid genomsnittlig granskningstid.



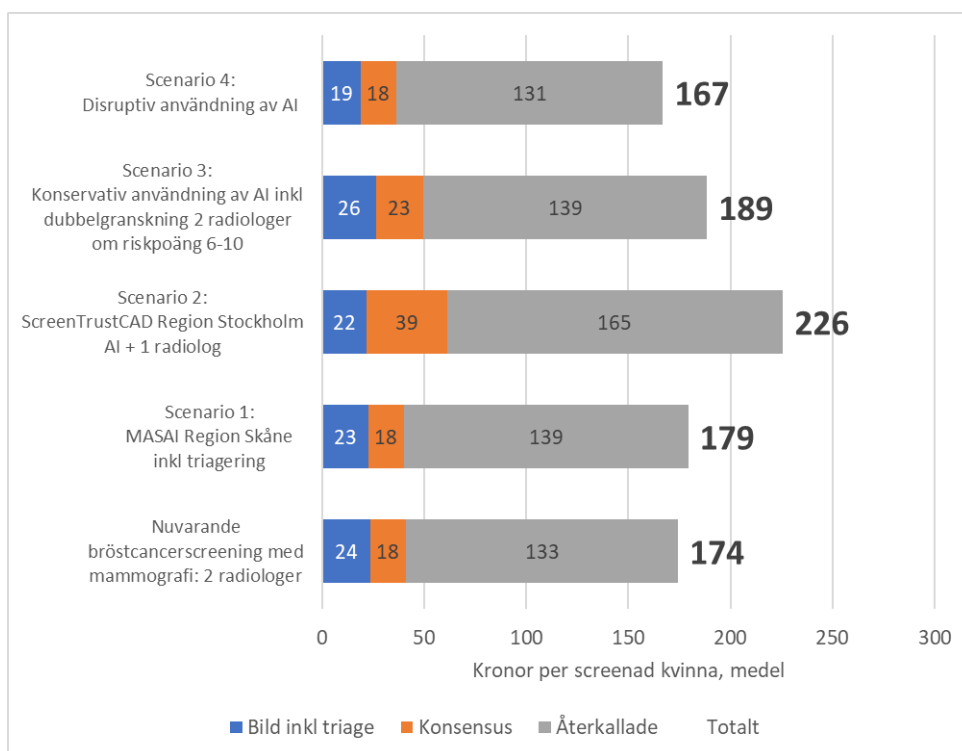
Figur 8. Kostnader som påverkas på kort sikt av användning av AI i bröstcancerscreening med mammografi (bildgranskning, konsensusbedömningar och återkallade). Nuvarande bröstcancerscreening med mammografi och 4 scenarier med användning av AI beskrivs i tabell 9 ovan. Pris per kvinna för AI granskning 30 kronor enligt antagande.

Alla fyra scenarier i figur 8 som använder AI på något sätt kostar mer än nuvarande granskning med två radiologer. Kostnadsskillnaden drivs framför allt av avgiften för AI, men kan också påverkas i hög grad om andelen återkallade blir högre än vid nuvarande mammografiscreening. Exempelvis bygger scenario 1 på den nivå av återkallande som redovisas från MASAI-studien i Region Skåne (2,2 procent i interventionsgruppen), och scenario 2 på en högre nivå såsom observerats i ScreenTrustCAD-studien i Region Stockholm (2,8 procent i gruppen som granskats av en radiolog tillsammans med AI). Dessa andelar är båda högre än kontrollgruppen i MASAI-studien som används som referens för nuvarande mammografiscreening i Region Skåne. ScreenTrustCAD studiens nivå på återkallade ligger i alla alternativ som involverar radiologer över nivån för båda grupperna i MASAI-studien. Resultaten för granskning av två radiologer (utan AI) från ScreenTrustCAD-studien pekar alltså på att återkallandenivån kan skilja mellan regioner. ScreenTrustCAD-studien redovisade också en högre andel konsensusbedömningar i sina studiegrupper än i MASAI-studien vilket bidrar till den sammanlagt högre kostnaden för scenario 3.

Analysen i figur 8 pekar på att användning av AI kan leda till kostnadshöjningar, när priset per granskad kvinna för att använda AI är så högt som 30 kronor. För 133 775 kvinnor per år i Region Skåne motsvarar det en kostnadsskillnad på 3,3 miljoner kronor mellan nuvarande mammografi och

mammografi som använder AI och scenario 1. Även de tre övriga scenarierna innebär kostnadsökningar.

Dessa skillnader blir mindre vid ett lägre pris för AI per mammografiscreenad kvinna såsom visas i figur 9, där motsvarande kostnadsökning minskas till knappt 700 000 kronor för 133 775 kvinnor i Region Skåne med scenario 1. Beräkningarna utgår från att priset för AI per screenad kvinna är 10 kronor. Även vid denna lägre kostnad, som ligger under lönekostnaden för granskning per radiolog, skulle inte användning av AI inom bröstcancerscreening med mammografi med nödvändighet innebära kostnadsbesparingar på kort sikt. Detta eftersom även små ökningar i andelen konsensusbeslut eller andelen återkallade ökar kostnaderna. Det är därför viktigt att fortsätta utvärdera om ett införande av AI som ett verktyg i bröstcancerscreening med mammografi också ger viktiga hälso- och livskvalitetsvinster, som i så fall kan motivera ökade kostnader. Av exemplen i figur 8 och 9 är det endast vid ett pris för AI på 10 kronor per granskning och i scenario 4, där AI också används som ensam bedömare av bilder och för att sälla vilka bilder som en radiolog ska granska, som den genomsnittliga kostnaden per kvinna för kostnaden blir lägre än i nuläget.



Figur 9. Kostnader som påverkas på kort sikt av användning av AI i bröstcancerscreening med mammografi (bildgranskning, konsensusbedömningar och återkallade). Nuvarande bröstcancerscreening med mammografi och scenarier beskrivs i tabell 9 ovan. Kostnad per kvinna för AI granskning hypotetiskt antagande 10 kronor.

Resultaten i figur 8 och figur 9 utgår ifrån att granskning genomförs av personer inom verksamheten och har utgått från lönekostnader för specialistläkare i Skåne. Verksamheten använder också externa

radiologkonsulter för att granska mammografibilder. En vanlig ersättning för detta är 25 kronor per granskning, vilket då kan jämföras med lönekostnader för en genomsnittlig granskningstid (48 sekunder och 11,75 kronor). Med nuvarande rutin för bröstcancerscreening med mammografi med två radiologer är genomsnittskostnaden per granskad kvinna 174 kronor, där omkring 25 kronor är lönekostnad för bildgranskning och triage (konsensusbedömning ej inkluderat). Om den ena radiologen är en extern konsult blir den genomsnittliga kostnaden per granskad kvinna i stället 188 kronor i nuvarande bröstcancerscreening med mammografi.

Sammantaget visar scenarioanalyserna att priset per granskning som AI genomför har en avgörande betydelse för hur stor budgetpåverkan kan förväntas bli vid ett införande av AI. Förutom i det disruptiva scenario 4, där AI också helt ersätter radiologgranskning i en del av populationen och med den lägre ersättningen per granskning (figur 9), kommer AI att innebära kostnadsökningar för Region Skåne.

En brittisk studie från 2023 har utvärderat kostnadseffektiviteten på lång sikt för mammografi som använder AI jämfört med det nuvarande brittiska screeningprogrammet, där kvinnor bjuds in till screening vart tredje år från 50 års ålder (Vargas-Palacios 2023). Denna modellbaserade analys är inte relevant för svenska förhållanden eftersom den utgår från en betydligt högre kostnad per screenad kvinna, samtidigt som det finns flera oklarheter kring hur denna kostnad inkluderats i modellens långsiktiga perspektiv.

8 Etiska aspekter

Alla AI-system är hitintills upplärda och tränade på en avgränsad informationsmängd. Vid bildtolkning av till exempel mammografibilder ifrån bröstcancerscreening används vanligen ett mycket stort antal bilder (det kan röra sig om hundratusentals). Varje sådant bildmaterial definieras efter de undersökta kvinnornas ålder vid bildtagningen, eventuell hormonbehandling, bröstcancerförekomst, etnicitet med mera. Om algoritmen används på en population med annan sammansättning, är det inte säkert att gjorda antaganden är relevanta i samma grad. En aspekt av detta är att etnicitet som begrepp inte används i svensk sjukvård.

All medicinsk information är integritetskänslig, och det ställs höga krav på hur sekretess säkerställs vid användning av automatiserade system. Detta väcker frågor angående användning av olika AI-system: Sker bildtolkningen i ett lokalt slutet system? Kommunicerar AI-systemet med externa aktörer? Sänds bildinformation externt? Hur är systemet skyddat för extern åtkomst? Flera av dessa frågor gäller samtliga datoriserade system i sjukvården, men skillnaden vid användning av AI-system är bristen på transparens.

Transparens är även viktigt i beslutsfattande. Hur och varför kom systemet fram till det givna resultatet? AI-system är mycket som en ”svart låda” vilket kan göra det svårt, eller omöjligt, att förstå grunderna för ett givet beslut. Transparens är viktigt för att vårdprofessionen ska kunna lita på systemet och den data som genereras. Kopplat till detta är också ansvarsfrågan vid felaktigheter. Vid dagens bröstcancerscreening med mammografi med dubbelgranskning utförd av två radiologer är det för dessa individer alltid ett delat ansvar för missad bröstcancer. Vid införandet av AI i utbyte mot en radiolog vilar ansvaret tyngre på den radiolog som granskar bilder ensam. Vid användning av AI som ensam granskare och vid så kallad triagefunktion för att selektera ut endast screeningbilder med högre bröstcancerriks för granskning, är det högst oklart vem som kan anses ansvarig för en missad bröstcancerdiagnos.

En annan tänkbar risk vid användning av AI-system, om dessa visar sig hålla en hög kvalitet, är att vårdprofessionen börjar lita på systemen i alltför hög grad och nonchalerar eller får minskad tilltro till den egna förmågan. Det måste ständigt betonas att AI-systemen bygger på statistiska sannolikheter och ska vara ett komplement och inte en ersättning för den professionella kompetensen.

Många nya metoder införs kontinuerligt inom sjukvården, och används utan patienternas kännedom. Skiljer sig AI på något sätt som ny metod så att det är rimligt eller viktigt att kvinnan i förväg informeras om att AI kommer att användas vid bildtolkningen, för att bibehålla transparens och förtroende? Kan en ”felbedömning” av AI leda till minskat förtroende för användning? Eller finns det redan idag en övertro på tekniska lösningar generellt i populationen?

Under 2023 genomförde Bröstcancerförbundet en attitydundersökning av kvinnor (både med och utan erfarenhet av egen bröstcancer) angående inställningen till AI inom bröstcancerscreening med mammografi (Bröstcancerförbundet 2023). Vid den undersökningen, som inkluderade diskussioner i fokusgrupper och en webbaserad undersökning genomförd av Novus, framkom att majoriteten av kvinnor ställer sig positiva till undersökning med AI, om stor säkerhet av systemet kan garanteras.

Eftersom det kan vara svårt som individ att förhålla sig till balansen mellan sensitivitet och specificitet, är det inte orimligt att tänka sig att enskilda kvinnor gör andra bedömningar än vården gällande den ”rätta” avvägningen. Skulle förtroendet för AI exempelvis tjäna på en högre återkallandefrekvens?

Flera av studierna som ingår i rapporten visar att AI möjligen kan öka cancerdetektionsfrekvensen, vilket naturligtvis kan anses som en fördel vid användningen av dessa system. Dock får man även beakta risken för överdiagnostik och överbehandling. Det är svårt att efter införandet av bröstcancerscreening med mammografi få en uppfattning om överdiagnostik. En tidigare studie genomförd i Malmö visar på en frekvens på 10 % i åldersgruppen 55–69 år hos de kvinnor som kallats till screening (Zackrisson 2006). Det går inte att vid diagnos avgöra om en bröstcancer är

överdiagnostiserad, men risken är större vid en lågradigt invasiv cancer eller en in situ-cancer (Grimm 2022; van Seijen 2019). En bröstcancerdiagnos innebär, utöver den behandling kvinnan utsätts för, en stor psykologisk belastning som påverkar kvinnan och eventuellt hennes anhöriga under en längre period. Det kan således antas att om cancerdetektionen ökar, ökar även antalet kvinnor som överdiagnostiseras, och med dem antalet individer som påverkas av en bröstcancerdiagnos, där själva canceren aldrig skulle ha påverkat kvinnans hälsa om den ej upptäckts.

Långsiktig användning av AI kan leda till ändrade arbetssätt och roller för den kliniska professionen och konsekvenserna av detta bör adresseras.

9 Identifierade kunskapsluckor

En stor del av kunskapsluckorna om effekterna av AI vid bröstcancerscreening med mammografi rör vilka cancrar som detekteras med stöd av AI jämfört med utan stöd av AI. Det finns otillräckligt med data rörande den histopatologiska karaktären på de cancrar som detekteras med AI jämfört med utan. Potentiellt finns andra effekter att utvärdera oavsett om det är fler eller färre invasiva cancrar som upptäcks med AI vid bröstcancerscreening med mammografi. Till exempel kan det vara relevant att veta om AI hittar cancrar som inte har någon klinisk signifikans eller hittar fler farliga cancrar tidigt. Det saknas också studier om effekterna på antalet intervallcancerar.

Vad gäller radiologbelastning talar studierna för att denna kan minska med användning av AI vid bröstcancerscreening med mammografi. Befintliga studier är baserade på mätningar av antal granskningar snarare än den faktiska tiden som granskningen tar. Det är därför svårt att värdera den faktiska tidsvinsten, som även kan påverkas av antalet konsensusmöten.

Andra kunskapsluckor är frånvaron av studier som mäter livskvalitet och bröstcancermortalitet vid användning av AI vid bröstcancerscreening med mammografi.

10 Diskussion

Denna rapport visar att användning av AI i kombination med radiolog inom bröstcancerscreening med mammografi ger likvärdig eller högre cancerdetektion som dubbelgranskning utförd av radiologer (måttlig tillförlitlighet). Återkallandefrekvensen ökar inte vid användning av AI i kombination med radiolog vid bröstcancerscreening med mammografi, men de resultaten har låg tillförlitlighet. Resultaten visar även att användning av AI vid bröstcancerscreening med mammografi minskar radiologbelastningen (låg tillförlitlighet).

Däremot visar resultaten i rapporten att AI som stand-alone inom bröstcancerscreening med mammografi ger lägre eller samma cancerdetektion som dubbelgranskning utförd av radiologer (låg tillförlitlighet) medan det inte går att dra någon slutsats om återkallandefrekvensen för AI stand-alone.

Rapporten visar också att det inte går att dra några slutsatser om effekterna på intervallcancerfrekvens, TNM-stadium, bröstcancer mortalitet eller livskvalitet. Således finns stora kunskapsluckor vad gäller patientnära utfallsmått.

Jämfört med de två systematiska översikter som identifierats under arbetet med denna HTA-rapport är resultaten inte helt samstämmiga. Den ena (Freeman 2021) visar att AI som enskild granskare var bättre än en radiolog för att korrekt detektera cancer, men sämre än två granskande radiologer, och den andra (Hickman 2022) visar att AI som enskild granskare samt radiologbedömning var likvärdiga för bröstcancerdetektion. Anledning till våra divergerande resultat kan vara att dessa systematiska översikter inkluderar ett flertal artiklar med hög risk för bias. Dessa systematiska översikter publicerades också före 2023, varför de saknar ett flertal av de studier som inkluderats i denna rapport, inklusive den randomiserade studie som publicerades av Lång 2023.

Det är tydligt i studierna att cancerdetektionen är beroende av hur man kalibrerar sina system och använder sig av AI utifrån den sensitivitet man önskar. Risken med detta är att man idag inte vet vilken typ av cancer man hittar och de långsiktiga konsekvenserna för patienterna och vården i stort.

Den återkallandefrekvens som visas i de två prospektiva studierna överensstämmer med återkallandefrekvensen i nuvarande rutin. Om man ökar sensitiviteten på AI-systemen kommer återkallandefrekvensen att öka, men konsekvenserna av detta är inte studerade. Dock visar resultaten från de studier som presenterar frekvensen av falskt återkallade i denna rapport att detta kan variera mycket beroende på hur AI används. Det är därför av vikt att följa upp detta vid eventuellt införande.

Avseende detektion av intervallcancer så anses inte retrospektiva studier kunna påvisa korrekta samband med utfallsmåttet. Detta då AI-produkten används i en tidig del av utredningsgången och man inte vet vad resultatet efter återkallelse för utvidgad bildtagning och eventuell provtagning skulle ha visat. Det är således möjligt att även om en intervallcancer noterats vid en tidigare screeningkontroll, behöver den inte ha diagnostiserats. De prospektiva studierna i denna rapport har inte med detta utfallsmått eftersom de saknar tillräckligt lång uppföljningstid.

Då bröstcancerscreening med mammografi tar en stor andel radiologresurser i anspråk var en primär utgångspunkt för den här HTA-rapporten att utreda om man kunde spara radiologtid med hjälp av AI. Rapporten stöder att det finns potentiella fördelar med AI-stödd granskning av screeningbilder i form av färre granskningstillfällen för radiologer med bibehållen cancerdetektion. Detta skulle i sin tur

kunna bidra till kortare väntetider för andra patienter och undersökningar. Detta är dock inte studerat i någon av de inkluderade studierna.

Forskningsområdet präglas av snabba framsteg och hög forskningsaktivitet samtidigt som det kliniska intresset för införandet är stort. Därför tillför den här HTA-rapporten en viktig sammanställning av det aktuella forskningsläget inför beslut om eventuellt införande av AI inom bröstcancerscreening med mammografi.

Under arbetet med denna HTA-rapport har ny litteratur kontinuerligt eftersökts, men det har inte identifierats någon, för rapportens frågeställning, ytterligare intressant publikation. Dock kan nya publikationer med viktiga resultat förväntas inom snar framtid.

En svaghet med HTA-rapporten är att det vetenskapliga underlaget inte möjliggör någon tydlig anvisning om hur AI ska användas. Det framgår dock att AI som enda granskare verkar fungera sämre än de olika kombinationer av AI och radiolog som förekommer i studierna.

En annan möjlig svaghet med HTA-rapporten är den breda ansatsen. Frågan som ställdes gäller effekterna av en diagnostisk metod snarare än en behandling. Det är då inte lika självklart som i många andra HTA-rapporter om PICO eller PIRO (Population, Indextest, Referenstest, Utfall) ska användas för att strukturera frågan eller vilka mallar som bäst fångar risken för bias. Samtidigt är detta kliniskt mer relevant än att bara efterfråga diagnostisk korrekthet. Vi har gjort en bedömning och valt den modell vi finner bäst efter att ha analyserat för- och nackdelar för just denna frågeställning, men ingen modell skulle varit klanderfri.

Gällande kostnader för AI beräknade utifrån de vetenskapliga studierna ses att användning av AI inte ger någon kostnadsbesparing, utan snarare kan vara dyrare på kort sikt jämfört med nuvarande rutin med dubbelgranskning. Detta utifrån alternativa scenarier som bygger på hur AI har använts i två svenska studier (Lång 2023; Dembrower 2023). Detta gäller även vid en större förändring av scenariot, då AI helt ersätter den ena radiologen och även sorterar bort kvinnor med låg bedömd risk, så att det endast återstår 75 procent av fallen att granska för den kvarvarande radiologen, såsom i den danska studien av Lauritzen (2022).

I den ekonomiska analysen har två prisnivåer använts, en högre och en lägre, i samråd med projektgruppen. De två prisnivåerna syftar till att illustrera den stora betydelse priset per granskning har. Kostnadsanalyserna utgår från att all granskning av bilder inom nuvarande bröstcancerscreening med mammografi kan hanteras inom anställda radiologers ordinarie tjänst. I de fall kompletterande kapacitet behöver köpas in från externa konsulter kan kostnaderna bli högre. Det är därför viktigt att inför ett införande av AI ta reda på i vilken utsträckning nuvarande verksamhet har tillräcklig kapacitet

för uppdraget och klargöra vem som har kostnadsansvar för eventuella merkostnader för externkonsulter i bildgranskningen.

I nuläget kan hälsovinster för screeningpopulationen i sig inte påvisas med användning av AI vid bröstcancerscreening med mammografi, men det kan ändå spekuleras i eventuella hälsovinster i andra delar av populationen, om frigjord radiologtid kan användas till annat. Detta förutsatt att AI tillsammans med en radiolog presterar minst lika bra som dubbelgranskning utförd av radiologer för cancerdetektion. Exempelvis skulle det vara en stor vinst för många patienter om väntetider för olika radiologiska undersökningar kunde kortas ner eller om man med fler radiologresurser kan underlätta eller snabba på diagnostiken av andra tillstånd. Dock bör här noteras att granskning av screeningbilder utgör en mycket liten del av en bröstradiologs arbetstid.

Vid ett eventuellt införande av AI inom bröstcancerscreening-programmet i Sverige är det angeläget att bidra till att fylla de många kunskapsluckor som kvarstår. Detta görs lämpligen genom fortsatta prospektiva studier. Hittills finns två publicerade utvärderingar av användning av AI vid bröstcancerscreening i Sverige inom ramen för större studier. Det är flera regioner som planerar någon form av användning av AI inom bröstcancerscreening med mammografi. Mot bakgrund av de kunskapsluckor som denna HTA-rapport identifierat är det fortsatt angeläget att systematiskt följa upp och utvärdera såväl kliniska effekter som påverkan på resursanvändning och kostnader. Detta för att lägga grunden för långsiktig uppföljning av screeningmetoders påverkan på cancerincidens och behandlingsmöjligheter samt övrig patientnytta.

Det är viktigt att arbetet med AI inom bröstcancerscreening med mammografi i de olika regionerna i Sverige registreras och följs upp inom ramen för kliniska studier för att kunna bidra till och säkerställa jämlik vård.

11 Referenser

Arbetsgivarverket. Arbetsgivarguiden: Arbetstid [Internet]. Stockholm: Arbetsgivarverket; [uppdaterad 2023-12-22; citerad 2024-05-16]. Tillgänglig via: <https://arbetsgivarverket.se/arbetsgivarguiden/arbetstid>

Brady AP, Allen B, Chong J, Kotter E, Kottler N, Mongan J, et al. Developing, purchasing, implementing and monitoring AI tools in radiology: practical considerations. A multi-society statement from the ACR, CAR, ESR, RANZCR & RSNA. *Insights Imaging*. 2024;15(1):16. doi: 10.1186/s13244-023-01541-3

Bramer WM, Giustini D, de Jonge GB, Holland L & Bekhuis T. De-duplication of database search results for systematic reviews in EndNote. *J Med Libr Assoc*. 2016;104(3):240-243. doi: 10.3163/1536-5050.104.3.014

Bröstcancerförbundet. Ny teknik i bröstcancervården välkomnas: en attitydundersökning till AI inom mammografi [Internet]. Stockholm: Bröstcancerförbundet; 2023. Tillgänglig via: <https://brostcancerforbundet.se/wt/documents/1352/attitydundersokning-VAIB.pdf>

Cancerfonden. Cancer i siffror 2023 [Internet]. Stockholm: Cancerfonden/Socialstyrelsen; 2023. Tillgänglig via: <https://www.socialstyrelsen.se/globalassets/sharepoint-dokument/dokument-webb/statistik/cancer-i-siffror-2023.pdf>

Dahlblom V, Andersson I, Lång K, Tingberg A, Zackrisson F & Dustler M. Artificial intelligence detection of missed cancers at digital mammography that were detected at digital breast tomosynthesis. *Radiol Artif Intel*. 2021;3(6):e200299. doi: 10.1148/ryai2021200299

Dembrower K, Crippa A, Colón E, Eklund M, Strand F; ScreenTrustCAD Trial Consortium. Artificial intelligence for breast cancer detection in screening mammography in Sweden: a prospective, population-based, paired-reader, non-inferiority study. *Lancet Digit Health*. 2023;10:e703-e711. doi: 10.1016/S2589-7500(23)00153-X

ECIBC. European guidelines on breast cancer screening and diagnosis [Internet]. Ispra: European Commission Initiative on Breast Cancer (ECIBC); 2024. Tillgänglig via: <https://cancer-screening-and-care.jrc.ec.europa.eu/en/ecibc/european-breast-cancer-guidelines>

Elhakim MT, Stougaard SW, Graumann O, Nielsen M, Lång K, Gerke O, et al. Breast cancer detection accuracy of AI in an entire screening population: a retrospective, multicentre study. *Cancer Imaging*. 2023;23(1):127. doi: 10.1186/s40644-023-00643-x

Freeman K, Geppert J, Stinton C, Todkill D, Johnson S, Clarke A, et al. Use of artificial intelligence for image analysis in breast cancer screening programmes: systematic review of test accuracy. *BMJ*. 2021;374:n1872. doi. 10.1136/bmj.n1872

Fridhammar A, Andersson E & Steen Carlsson K. Hälsoekonomisk analys av höjd åldersgräns från 74 till 79 år vid mammografiscreening i Region Stockholm [Internet]. Lund: Institutet för Hälso- och Sjukvårdsekonomi (IHE); 2022. IHE Rapport; 2022:11. Tillgänglig via: https://ihe.se/app/uploads/2022/11/IHE-Rapport-2022_11_.pdf

Grimm LJ, Rahbar H, Abdelmalak M, Hall AH & Ryser MD. Ductal carcinoma in situ: state-of-the-art review. *Radiology*. 2022;302(2):246-255. doi: 10.1148/radiol.211839

Heywang-Köbrunner SH, Hacker A, Jansch A, Hertlein M, Mieskes C, Elsner S, et al. Use of novel artificial intelligence computer-assisted detection (CAD-AI) for screening mammography: an analysis of 17,884 consecutive two-view full-field digital mammography screening exams. *Acta Radiol*. 2023;64(10):2697-2703. doi: 10.1177/02841851231187382

Hickman SE, Woitek R, Phuong Le EPV, Im YR, Mouritsen Luxhøj C, Aviles-Rimero AI, et al. Machine learning for workflow applications in screening mammography: systematic review and meta-analysis. *Radiology*. 2022;302(1):88-104. doi: 10.1148/radiol.2021210391

Kühl J, Elhakim MT, Stougaard SW, Rasmussen BSB, Nielsen M, Gerke O, et al. Population-wide evaluation of artificial intelligence and radiologist assessment of screening mammograms. *Eur Radiol*. 2023. doi: 10.1007/s00330-023-10423-7. Online ahead of print

Larsen M, Aglen CF, Hoff SR, Lund-Hanssen H & Hofvind S. Possible strategies for use of artificial intelligence in screen-reading of mammograms, based on retrospective data from 122,969 screening examinations. *Eur Radiol*. 2022;32(12):8238-8246. doi: 10.1007/s00330-022-08909-x

Lauritzen AD, Rodríguez-Ruiz A, von Euler-Chelpin MC, Lynge E, Vejborg I, Nielsen M, et al. An artificial intelligence-based mammography screening protocol for breast cancer: outcome and radiologist workload. *Radiology*. 2022;304(1):41-49. doi. 10.1148/radiol.210948

Lång K, Josefsson V, Larsson A-M, Larsson S, Högberg C, Sartor H, et al. Artificial intelligence-supported screen reading versus standard double-reading in the Mammography Screening with Artificial Intelligence trial (MASAI): a clinical trial safety analysis of a randomised, controlled, non-inferiority single-blinded screening accuracy study. *Lancet Oncol*. 2023;24(8):936-944. doi: 10.1016/S1470-2045(23)00298-X

Moher D, Liberati A, Tetzlaff J, Altman DG; PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS Med. 2009;6(7):e1000097. doi: 10.1371/journal.pmed.1000097

NCT04778670. Artificial intelligence in large-scale breast cancer screening. Tillgänglig via: <https://clinicaltrials.gov/study/NCT04778670>

NCT04838756. Mammography screening with artificial intelligence (MASAI). Tillgänglig via: <https://clinicaltrials.gov/study/NCT04838756>

NCT06032390. Artificial Intelligence in mammography screening in Norway. Tillgänglig via: <https://clinicaltrials.gov/study/NCT06032390>

NKBC. Årsrapport 2023 [Internet]. Stockholm: Nationellt kvalitetsregister för bröstcancer (NKBC); [uppdaterad 2024-05-02; citerad 2024-05-29]. Tillgänglig via: <https://statistik.incanet.se/brostcancer/>

Perry N, Broeders M, de Wolf C, Törnberg S, Holland R, von Karsa L, et al, redaktörer. European guidelines for quality assurance in breast cancer screening and diagnosis [Internet]. 4. uppl. Luxembourg: Office for Official Publications of the European Communities; 2006. Tillgänglig via: https://screening.iarc.fr/doc-ND7306954ENC_002-pdf

Raya-Povedano JL, Romero-Martín S, Elías-Cabot E, Gubern-Mérida A, Rodríguez-Ruiz A & Álvarez-Benito M. AI-based strategies to reduce workload in breast cancer screening with mammography and tomosynthesis: a retrospective evaluation. Radiology. 2021;300(1):57-65. doi: 10.1148/radiol.2021203555

Regionala cancercentrum i samverkan. Remissrunda: nationellt vårdprogram för bröstcancerscreening [Internet]. Stockholm: Regionala cancercentrum i samverkan (RCC); 2024. Tillgänglig via: <https://cancercentrum.se/globalassets/vara-uppdrag/kunskapsstyrning/varprogram/kommande-varprogram/2024/240215/nvp-brostcancerscreening-remissversion-240215.pdf>

Romero-Martín S, Elías-Cabot E, Raya-Povedano JL, Gubern-Mérida A, Rodríguez-Ruiz A & Álvarez-Benito M. Stand-alone use of artificial intelligence for digital mammography and digital breast tomosynthesis screening: a retrospective evaluation. Radiology. 2022;302(3):535-542. doi: 10.1148/radiol.211590

SBU. Etiska aspekter på insatser inom hälso- och sjukvården: en vägledning för att identifiera relevanta etiska aspekter [Internet]. Stockholm: Statens beredning för medicinsk och social utvärdering (SBU); 2021. Tillgänglig via: https://sbu.se/globalassets/ebm/etiska_aspekter_halso_sjukvarden.pdf

SBU. Utvärdering av metoder i hälso- och sjukvården och insatser i socialtjänsten: en metodbok [Internet]. Stockholm: Statens beredning för medicinsk och social utvärdering (SBU); 2023. Tillgänglig via: <https://sbu.se/metodbok>

Socialstyrelsen. Screening för bröstcancer: Socialstyrelsens rekommendation. Slutversion [Internet]. Stockholm: Socialstyrelsen; 2023. Tillgänglig via: <https://www.socialstyrelsen.se/globalassets/sharepoint-dokument/artikelkatalog/nationella-screeningprogram/2023-5-8564.pdf>

Statistikdatabasen SCB. Genomsnittlig grund- och månadslön samt kvinnors lön i procent av mäns lön efter region, sektor, yrke (SSYK 2012) och kön. År 2014 – 2022 [Internet]. Stockholm: Statistikmyndigheten (SCB) [uppdaterad 2024-03-12; citerad 2024-05-16] Tillgänglig via: https://www.statistikdatabasen.scb.se/pxweb/sv/ssd/START_AM_AM0110_AM0110A/LonYrkeRegion4A/?loadedQueryId=126559&timeType=top&timeValue=1

Sveriges Kommuner och Regioner. Arbetsgivaravgifter, PO-pålägg [Internet]. Stockholm: Sveriges Kommuner och Regioner (SKR) [uppdaterad 2024-05-29; citerad 2024-05-16] Tillgänglig via: <https://skr.se/skr/ekonomijuridik/ekonomi/arbetsgivaravgifterpopalaggpension/arbetsgivaravgifterpopalagg.1290.html>

Taylor P & Potts HW. Computer aids and human second reading as interventions in screening mammography: two systematic reviews to compare effects on cancer detection and recall rate. *Eur J Cancer*. 2008;44(6):798-807. doi: 10.1016/j.ejca.2008.02.016

van Seijen M, Lips EH, Thompson AM, Nik-Zainal S, Futreal A, Hwang ES, et al. Ductal carcinoma in situ: to treat or not to treat, that is the question. *Br J Cancer*. 2019;121(4):285-292. doi: 10.1038/s41416-019-0478-6

Vargas-Palacios A, Sharma N & Sagoo GS. Cost-effectiveness requirements for implementing artificial intelligence technology in the Women's UK Breast Cancer Screening service. *Nat Commun*. 2023;14(1):6110. doi: 10.1038/s41467-023-41754-0

Zackrisson S, Andersson I, Janzon L, Manjer J & Garne JP. Rate of over-diagnosis of breast cancer 15 years after end of Malmö mammographic screening trial: follow-up study. *BMJ*. 2006;332(7543):689-692. doi: 10.1136/bmj.38764.572569.7C

12 Appendix

Appendix A: Projektorganisation

Frågeställare

- Catharina Behmer, överläkare, medicinskt ansvarig, mammografi Unilabs, Skånes universitetssjukhus

Sakkunniggrupp

- Maria Erngrund, överläkare, bröstkirurgi, Kirurgkliniken, Blekingesjukhuset
- Kristina Lång, överläkare, docent, bröstradiologi, Unilabs, Skånes universitetssjukhus
- Maria Madestam, överläkare, bröstradiologi, Unilabs, Helsingborgs lasarett
- Martin Malmberg, överläkare, PhD, VO hema onk-strål, VE onkologi och strålningsfysik, Skånes universitetssjukhus

HTA syd

- Beata Borgström Bolmsjö, specialistläkare, docent, projektledare
- Sophia Frantz, överläkare, med dr, projektledare
- Linnea Huss, specialistläkare, med dr, projektledare
- Anna Saxne Jöud, epidemiolog, docent, enhetschef, projektledare
- Katarina Steen Carlsson, hälsoekonom, docent
- Ylva Sundin, informationspecialist, fil mag
- Erik Wikström, informationspecialist

Externa granskare[†]

- Maria Edegran, överläkare, radiolog, Mammografienheten NU-sjukvården, Uddevalla sjukhus
- Monica Hulterantz, med dr, enhetschef, HTA Region Stockholm, Centrum för hälsoekonomi, informatik och sjukvårdsforskning (CHIS)

Metodstöd

- Anders Castor, sektionschef, med dr, överläkare, VO barnmedicin, VE barnmedicin, Skånes universitetssjukhus, ledamot Etiska rådet, ledamot Region Skånes Metod- och prioriteringsråd

Intressekonflikter och jäv

Kristina Lång deltog och arvoderades under 2022 som expert i en workshop arrangerad av Siemens Healthineers Women's Health. Jävsdeklarationer för samtliga projektdeltagare finns tillgängliga på HTA syd.

Arbetsprocess

HTA syd har utfört samtliga delar av evidenssammanställningen inom HTA-rapporten. Det innebär att medarbetare på HTA syd har utfört litteratursökningarna, abstraktsällning, fulltextgranskning, bedömt risk of bias och utfört evidensgradering (GRADE). Sakkunniga i detta projekt har gett synpunkter på PICO och sökord och på av HTA syd framtagna bedömningar gällande relevanta artiklar, risk of bias och evidensgradering. Vidare har sakkunniga bidragit i rapportskrivningen vad gäller underlag till riktlinjer och rekommendationer, praxis och organisatoriska aspekter samt etik.

Inför arbetet med denna HTA-analys har intervju med patientföreträdare från Bröstcancerförbundet, ordförande Susanne Dieroff Hay, genomförts med fokus på PICO.

[†] HTA syd anlitar, i likhet med SBU, externa granskare av sina rapporter. De externa granskarna ger värdefulla kommentarer och bidrar i hög grad till att förbättra rapporten. Det är dock inte säkert att alla ändrings- eller tilläggförslag kan tillgodoses. I rapporten görs en sammanvägning av synpunkterna och HTA syd ansvarar för slutresultatet. Det är därför inte säkert att de externa granskarna står bakom samtliga formuleringar eller slutsatser i rapporten.

Projektid

	Datum
Projektnominering	2023-11-28
Projektstart:	2023-12-20
Avslutande litteratursökning	2024-03-05
Publiceringsdatum	2024-06-05

Appendix B: Sökstrategier och databaser

MEDLINE via Ovid

Ovid MEDLINE(R) ALL <1946 to March 04, 2024>

Datum: **2024-03-05**

Antal träffar: **985**

#	Query	Results from 5 Mar 2024
1	(mammograph* adj2 screen*).ti,ab,kf.	9682
2	((breast and (cancer* or neoplas* or tumo*)) adj2 screen*).ti,ab,kf.	23135
3	or/1-2	27863
4	mammograph*.ti,ab,kf.	34353
5	breast cancer*.ti,ab,kf.	357093
6	breast neoplas*.ti,ab,kf.	13028
7	breast tumo*.ti,ab,kf.	28414
8	exp Mammography/	33920
9	exp Breast Neoplasms/di, dg, pc [Diagnosis, Diagnostic Imaging, Prevention & Control]	79399
10	or/4-9	409343
11	screen*.ti,ab,kf.	1024074
12	exp Mass Screening/	145365
13	exp "Early Detection of Cancer"/	39643
14	or/11-13	1076944
15	10 and 14	45107
16	3 or 15	46498
17	(artific* adj3 intelligen*).ti,ab,kf.	48216
18	AI.ti,ab,kf.	53484
19	exp Artificial Intelligence/	190752
20	or/17-19	247125
21	16 and 20	1080
22	limit 21 to (danish or english or norwegian or swedish)	1062
23	limit 22 to (address or autobiography or bibliography or biography or clinical conference or comment or congress or consensus development conference or consensus development conference, nih or dictionary or directory or editorial or "expression of concern" or government publication or interview or legal case or legislation or letter or news or newspaper article or patient education handout or personal narrative or preprint or video-audio media or webcast)	45
24	22 not 23	1017
25	limit 24 to ("in data review" or in process or publisher)	32
26	24 not 25	985

Embase via Ovid

Embase <1974 to 2023 March 04> via Ovid

Datum: 2024-03-05

Antal träffar: 753

#	Query	Results
1	(mammograph* adj2 screen*).ti,ab,kf.	12,402
2	((breast and (cancer* or neoplas* or tumo*)) adj2 screen*).ti,ab,kf.	33,089
3	1 or 2	38,909
4	mammograph*.ti,ab,kf.	45,374
5	breast cancer*.ti,ab,kf.	522,445
6	breast neoplas*.ti,ab,kf.	15,010
7	breast tumo*.ti,ab,kf.	40,121
8	exp mammography/	69,496
9	exp breast cancer/di, pc [Diagnosis, Prevention]	73,349
10	or/4-9	585,861
11	screen*.ti,ab,kf.	1,435,093
12	exp screening/	837,659
13	exp cancer screening/	100,932
14	exp early cancer diagnosis/	14,700
15	or/11-14	1,789,002
16	10 and 15	74,961
17	3 or 16	77,443
18	(artific* adj3 intelligen*).ti,ab,kf.	55,580
19	AI.ti,ab,kf.	70,635
20	exp artificial intelligence/	97,309
21	or/18-20	155,103
22	17 and 21	1,190
23	limit 22 to (danish or english or norwegian or swedish)	1,167
24	limit 23 to (books or chapter or conference abstract or conference paper or "conference review" or editorial or letter or note or "preprint (unpublished, non-peer reviewed)" or short survey or tombstone)	414
25	23 not 24	753

Cochrane Library

Cochrane Central Register of Controlled Trials (CENTRAL)

Datum: 2024-02-29

Antal träffar: 90

ID	Search	Hits
#1	mammograph* NEAR/2 screen*:ti,ab,kw	921
#2	breast and (cancer* or neoplas* or tumo*) NEXT/2 screen*:ti,ab,kw	1391
#3	#1 or #2	1931
#4	mammograph*:ti,ab,kw	2747
#5	(breast NEXT cancer*):ti,ab,kw	42345
#6	breast NEXT neoplas*:ti,ab,kw	20614
#7	breast NEXT tumo*:ti,ab,kw	1878

#8	MeSH descriptor: [Mammography] explode all trees	1240
#9	MeSH descriptor: [Breast Neoplasms] explode all trees and with qualifier(s): [prevention & control - PC, diagnostic imaging - DG, diagnosis - DI]	2645
#10	{OR #4-#9}	45392
#11	screen*:ti,ab,kw	101240
#12	MeSH descriptor: [Mass Screening] explode all trees	5974
#13	MeSH descriptor: [Early Detection of Cancer] explode all trees	2527
#14	{OR #11-#13}	101927
#15	#10 and #14	3856
#16	#3 or #15	3955
#17	artific* NEAR/3 intelligen*:ti,ab,kw	1971
#18	MeSH descriptor: [Artificial Intelligence] explode all trees	3144
#19	AI:ti,ab,kw	5807
#20	{OR #17-#19}	9396
#21	#16 and #20	90

CINAHL via EBSCO

CINAHL with Full Text

Datum: **2024-03-05**

Antal träffar: 310

#	Query	Results
S25	S23 NOT S24	310
S24	S23 AND PT (Anecdote OR Bibliography OR Biography OR Book OR Book Chapter OR Book Review OR Brief Item OR Care Plan OR Cartoon OR Code of Ethics OR Commentary OR Computer Program OR Consumer/Patient Teaching Materials OR Directories OR Doctoral Dissertation OR Editorial OR Exam Questions OR Forms OR Games OR Glossary OR Individual Testimonial Website OR Interview OR Legal Case OR Letter OR Masters Thesis OR Nurse Practice Acts OR Nursing Diagnoses OR Nursing Interventions OR Obituary OR Poetry OR Practice Acts OR Proceedings OR Protocol OR Questionnaire/Scale OR Questions and Answers OR Responsens OR Teaching Materials OR Tracings OR Website)	38
S23	S22 AND LA (English OR Danish OR Norwegian OR Swedish)	348
S22	S17 AND S21	349
S21	S18 OR S19 OR S20	51,368
S20	MH Artificial Intelligence+	33,499
S19	AI	22,988
S18	artific* N3 intelligen*	15,327
S17	S3 OR S16	19,215
S16	S10 AND S15	18,655
S15	S11 OR S12 OR S13 OR S14	285,920
S14	MH Early Detection of Cancer	12,997
S13	MH Cancer Screening+	17,966
S12	MH Health Screening+	110,628
S11	screen*	267,550

S10	S4 OR S5 OR S6 OR S7 OR S8 OR S9	124,005
S9	MH Breast Neoplasms+/DI/PC	21,740
S8	MH Mammography+	13,594
S7	breast tumo*	3,110
S6	breast neoplas*	97,821
S5	breast cancer*	108,451
S4	mammograph*	16,564
S3	S1 OR S2	10,880
S2	(breast AND (cancer* OR neoplas* OR tumo*)) W2 screen*	9,649
S1	mammograph* W2 screen*	2,486

HTA-rapporter söktes på följande HTA-relaterade organisationers webbsajter:

AHRQ – Agency for Healthcare Research and Quality, USA

CADTH – Canada’s Drug and Health Technology Agency (numera Canada’s Drug Agency)

CMT – Centrum för utvärdering av medicinsk teknologi (Linköpings universitet)

Cochrane Library (Wiley, UK)

CRD – Centre for Reviews and Dissemination (University of York, UK)

DEFACTUM (Aarhus, Danmark)

Epistemonikos database (Epistemonikos Foundation, Chile)

EUnetHTA – European Network for Health Technology Assessment

FinCCHTA – Finnish Coordinating Center for Health Technology Assessment

FHI – Folkehelseinstituttet (Kunnskapssenteret – Nasjonalt kunnskapssenter for helsetjensten), Norge

HTA-centrum Västra Götalandsregionen (VGR)

HTA-enheten CAMTÖ – Centre for Assessment of Medical Technology in Örebro

HTA Region Stockholm

IHE – Institutet för Hälso- och Sjukvårdsekonomi (Lund)

INAHTA – International HTA database (International Network of Agencies for Health Technology Assessment)

Regional samverkansgrupp HTA Sydöstra sjukvårdsregionen

SBU – Statens beredning för medicinsk och social utvärdering

Sida/organisation	Datum	Sökord	Träffar	Varav relevanta
SBU	2024-03-21	artificiell	1	0
		artificial	4	0
HTA-centrum VGR	2024-03-21	artifici*	0	–

HTA Region Stockholm	2024-03-21	artifici*	0	–
CAMTÖ	2024-03-21	artifici*	0	–
Regional samverkansgrupp HTA	2024-03-21	artifici*	0	–
CMT	2024-03-21	artifici*	0	–
IHE	2024-03-21	artificiell	1	0
		artificial	1	0
FHI	2024-03-21	kunstig	2	0
		artificial	1	0
DEFACTUM	2024-03-21	kunstig	1	0
FinCCHTA	2024-03-21	tekoäly	0	–
INAHTA	2024-03-21	((mammograph* OR (breast cancer*) OR (breast neoplas*) OR (breast tumo*)) AND screen*) AND (artific* AND intelligen*)	0	–
EUnetHTA	2024-03-21	artificial intelligence	0	–
CADTH	2024-03-21	artificial intelligence	8	0
CRD	2024-03-21	((mammograph* OR breast cancer* OR breast neoplas* OR breast tumo*) AND screen*) AND (artific* AND intelligen*)	0	–
AHRQ	2024-03-21	artificial	0	–
Epistemonikos	2024-03-21	Title/Abstract: ((mammograph* OR (breast cancer) OR (breast neoplas*) OR (breast tumo*)) AND screen*) AND (artific* AND intelligen*)	0	–
Cochrane Library	2024-03-21	Intervention "Artificial intelligence"	0	–
		((mammograph* OR (breast cancer*) OR (breast neoplas*) OR (breast tumo*)) AND screen*) AND (artific* AND intelligen*)):ti,ab,kw	0	–
Totalt			19	0

Kliniska riktlinjer söktes i följande databaser:

ClinicalKey – Elsevier, The Netherlands

GIN (Guidelines International) – EBSCOhost, USA

GuidelineCentral – National Library for Health Guideline Finder, USA

PubMed – U.S National Library of Medicine (NLM), National Institute of Health (NIH), USA

NICE – The National Institute for Health and Care Excellence, UK

Socialstyrelsen (National Board of Health and Welfare), Sweden

TRIP Database – UK

UpToDate – Wolters Kluwer, The Netherlands

Källa/databas	Datum	Sökord	Träffar	Varav relevanta
GIN	2024-02-29	Breast cancer screening Publication scope: prevention, screening. Guideline publication status: published. Languages: English. Publication year: 2018-	3	0
GuidelineCentral	2024-02-29	Breast cancer screening Filter: Within 6 years	22	0
TRIP	2024-02-29	Breast cancer screening artificial intelligence Filter: Guidelines	8	0
UpToDate	2024-03-21	Breast cancer screening Filter: Society guideline links	0	–
NICE	2024-03-21	Breast cancer screening	0	–
		Breast cancer mammography	1	0
ClinicalKey	2024-03-21	Breast cancer screening Filter: Guidelines	62	2
PubMed	2024-03-21	Breast cancer[tw] AND mammography[tw] OR (screening[tw]) Filter: Practice Guideline Begränsning: 5 år, engelska, danska, norska, svenska	20	1*
Socialstyrelsen	2024-03-21	Mammografi	1	1
	2024-04-02	Screening bröstcancer	1	1
Totalt			118	5(4⁺)

*Dubblett

⁺Utan dubletter

Pågående kliniska studier söktes i följande databaser:

ClinicalTrials.gov (U.S. National Library of Medicine, NLM, National Institute of Health (NIH))

ICTRP – International Clinical Trials Registry Platform (WHO)

ISRCTN Registry – BioMed Central (BMC), UK

Sida	Datum	Sökord	Träffar	Varav pågående	Därav intressanta
ClinicalTrials	2024-04-30	((mammography OR breast cancer OR breast neoplasms OR breast tumor) AND screening) AND (artificial AND Intelligence)	33	27	7
ISRCTN	2024-04-30	((mammography OR breast cancer OR breast neoplasms OR breast tumor) AND screening) AND (artificial AND Intelligence)	6	4	3

ICTRP	2024-04-30	((mammography screening) OR (breast cancer screening) OR (breast neoplasm screening) OR (breast tumor screening)) AND (artificial intelligence)	20	17	1+10 [#]
Totalt			105	88	21(11[§])

[#]Dubbletter

[§]Utan dubbletter

Appendix C: Inkluderade artiklar

Included studies (original articles)	Overall rating
<p>Dahlblom 2021 Dahlblom V, Andersson I, Lång K, Tingberg A, Zackrisson F & Dustler M. Artificial intelligence detection of missed cancers at digital mammography that were detected at digital breast tomosynthesis. Radiol Artif Intel. 2021;3(6):e200299. doi. 10.1148/ryai2021200299</p>	Moderate risk of bias
<p>Dembrower 2023 Dembrower K, Crippa A, Colón E, Eklund M, Strand F & ScreenTrustCAD Trial Consortium. Artificial intelligence for breast cancer detection in screening mammography in Sweden: a prospective, population-based, paired-reader, non-inferiority study. Lancet Digit Health. 2023a;10:e703-e711. doi: 10.1016/S2589-7500(23)00153-X</p>	Moderate risk of bias
<p>Elhakim 2023 Elhakim MT, Stougaard SW, Graumann O, Nielsen M, Lång K, Gerke O, et al. Breast cancer detection accuracy of AI in an entire screening population: a retrospective, multicentre study. Cancer Imaging. 2023;23(1):127. doi: 10.1186/s40644-023-00643-x</p>	Low risk of bias
<p>Heywang-Köbrunner 2023 Heywang-Köbrunner SH, Hacker A, Jansch A, Hertlein M, Mieskes C, Elsner S, et al. Use of novel artificial intelligence computer-assisted detection (CAD-AI) for screening mammography: an analysis of 17,884 consecutive two-view full-field digital mammography screening exams. Acta Radiol. 2023;64(10):2697-2703. doi: 10.1177/02841851231187382</p>	Moderate risk of bias
<p>Kühl 2023 Kühl J, Elhakim MT, Stougaard SW, Rasmussen BSB, Nielsen M, Gerke O, et al. Population-wide evaluation of artificial intelligence and radiologist assessment of screening mammograms. Eur Radiol. 2023. doi: 10.1007/s00330-023-10423-7. Online ahead of print</p>	Moderate risk of bias
<p>Larsen 2022 Larsen M, Aglen CF, Hoff SR, Lund-Hanssen H & Hofvind S. Possible strategies for use of artificial intelligence in screen-reading of mammograms, based on retrospective data from 122,969 screening examinations. Eur Radiol. 2022b;32(12):8238-8246. doi: 10.1007/s00330-022-08909-x</p>	Low risk of bias

<p>Lauritzen 2022</p> <p>Lauritzen AD, Rodríguez-Ruiz A, von Euler-Chelpin MC, Lyng E, Vejborg I, Nielsen M, et al. An artificial intelligence-based mammography screening protocol for breast cancer: outcome and radiologist workload. <i>Radiology</i>. 2022;304(1):41-49. doi. 10.1148/radiol.210948</p>	<p>Low risk of bias</p>
<p>Lång 2023</p> <p>Lång K, Josefsson V, Larsson A-M, Larsson S, Högberg C, Sartor H, et al. Artificial intelligence-supported screen reading versus standard double-reading in the Mammography Screening with Artificial Intelligence trial (MASAI): a clinical trial safety analysis of a randomised, controlled, non-inferiority single-blinded screening accuracy study. <i>Lancet Oncol</i>. 2023;24(8):936-944. doi: 10.1016/S1470-2045(23)00298-X</p>	<p>Moderate risk of bias</p>
<p>Raya-Povedano 2021</p> <p>Raya-Povedano JL, Romero-Martín S, Elías-Cabot E, Gubern-Mérida A, Rodríguez-Ruiz A & Álvarez-Benito M. AI-based strategies to reduce workload in breast cancer screening with mammography and tomosynthesis: a retrospective evaluation. <i>Radiology</i>. 2021;300(1):57-65. doi: 10.1148/radiol.2021203555</p>	<p>Low risk of bias</p>
<p>Romero-Martín 2022</p> <p>Romero-Martín S, Elías-Cabot E, Raya-Povedano JL, Gubern-Mérida A, Rodríguez-Ruiz A & Álvarez-Benito M. Stand-alone use of artificial intelligence for digital mammography and digital breast tomosynthesis screening: a retrospective evaluation. <i>Radiology</i>. 2022;302(3):535-542. doi: 10.1148/radiol.211590</p>	<p>Moderate risk of bias</p>

Included studies (systematic reviews)	Overall rating
<p>Freeman 2021</p> <p>Freeman K, Geppert J, Stinton C, Todkill D, Johnson S, Clarke A, et al. Use of artificial intelligence for image analysis in breast cancer screening programmes: systematic review of test accuracy. <i>BMJ</i>. 2021;374:n1872. doi. 10.1136/bmj.n1872</p>	<p>Moderate risk of bias</p>
<p>Hickman 2022</p> <p>Hickman SE, Woitek R, Phuong Le EPV, Im YR, Mouritsen Luxhøj C, Aviles-Rimero AI, et al. Machine learning for workflow applications in screening mammography: systematic review and meta-analysis. <i>Radiology</i>. 2022;302(1):88-104. doi: 10.1148/radiol.2021210391</p>	<p>Moderate risk of bias</p>

Appendix D: Exkluderade artiklar

Excluded studies (original articles)	Motif for exclusion
<p>Detaljer om bedömningar av risk för bias i alla studerade utfallsmått finns att få från rapportförfattarna på begäran.</p>	
<p>Al-Bazzaz H, Janicijevic M & Strand F. Reader bias in breast cancer screening related to cancer prevalence and artificial intelligence decision support: a reader study. Eur Radiol. 2024; doi: 10.1007/s00330-023-10515-5. Online ahead of print</p>	Wrong PICO
<p>Arzamasov K, Vasilev Y, Vladzimirskyy A, Omelyanskaya O, Shulkin I, Kozikhina D, et al. An international non-inferiority study for the benchmarking of AI for routine radiology cases: chest x-ray, fluorography and mammography. Healthcare (Basel). 2023;11(12):1684. doi: 10.3390/healthcare11121684</p>	Wrong PICO
<p>Bao C, Shen J, Zhang Y, Wei W, Wang Z, Ding J, et al. Evaluation of an artificial intelligence support system for breast cancer screening in Chinese people based on mammogram. Cancer Med. 2023;12(3):3718-3726. doi: 10.1002/cam4.5231</p>	Wrong PICO
<p>Bassi E, Russo A, Oliboni E, Zamboni F, De Santis C, Mansueto G, et al. The role of an artificial intelligence software in clinical senology: a mammography multi-reader study. Radiol Med. 2024;129(2):202-210. doi: 10.1007/s11547-023-01751-1</p>	Wrong PICO
<p>Becker AS, Marcon M, Ghafoor S, Wurnig MC, Frauenfelder T & Boss A. Deep learning in mammography: Diagnostic accuracy of a multipurpose image software analysis in the detection of breast cancer. Invest Radiol. 2017;52(7):434-440. doi: 10.1097/RLI0000000000000358</p>	Wrong PICO
<p>Byng D, Strauch B, Gnas L, Leibig C, Stephan O, Bunk S, et al. AI-based prevention of interval cancers in a national mammography screening program. Eur J Radiol. 2022;152:110321. doi: 10.1016/j.ejrad.2022.110321</p>	Wrong PICO
<p>Dang L-A, Chazard E, Poncelet E, Serb T, Rusu A, Pauwels S, et al. Impact of artificial intelligence in breast cancer screening with mammography. Breast Cancer. 2022;29(6):967-977. doi: 10.1007/s12282-022-01375-9</p>	Wrong PICO

Dembrower K, Lindholm P & Strand F. A multi-million mammography image dataset and population-based screening cohort for the training and evaluation of deep neural networks: the Cohort of Screen-Aged Women (CSAW). J Digit Imaging. 2020;33(2):408-413. doi: 10.1007/s10278-019-00278-0	Wrong PICO
Dembrower K, Salim M, Eklund M, Lindholm P & Strand F. Implications for downstream workload based on calibrating an artificial intelligence detection algorithm by standalone-reader or combined-reader sensitivity matching. J Med Imaging (Bellingham). 2023b;10(Suppl 2):S22405. doi: 10.1117/1.JMI.10.S2.S22405	Wrong PICO
Elías-Cabot E, Romero-Martín S, Raya-Povedano JL, Brehl A-K & Álvarez-Benito M. Impact of real-life use of artificial intelligence as support for human reading in a population-based breast cancer screening program with mammography and tomosynthesis. Eur Radiol. 2023. doi: 10.1007/s00330-023-10426-4. Online ahead of print	Wrong PICO
Ghosh A. Artificial intelligence using open source BI-RADS data exemplifying potential future use. J Am Coll Radiol. 2019;16(1):64-72. doi: 10.1016/j.jacr.2018.09.040	Wrong PICO
Graewingholt A & Rossi PG. Retrospective analysis of the effect on interval cancer rate of adding an artificial intelligence algorithm to the reading process for two-dimensional full-field digital mammography. J Med Screen. 2021;28(3):369-371. doi: 10.1177/0969141320988049	Wrong PICO
Hickman SE, Payne NR, Black RT, Huang Y, Priest AN, Hudson S, et al. Mammography Breast Cancer Screening Triage Using Deep Learning: A UK Retrospective Study. Radiology. 2023;309(2):e231173. doi: 10.1148/radiol.231173.	High risk of bias
Kerschke L, Weigel S, Rodriguez-Ruiz A, Karssemeijer N & Heindel W. Using deep learning to assist readers during the arbitration process: a lesion-based retrospective evaluation of breast cancer screening performance. Eur Radiol. 2022;32(2):842-852. doi: 10.1007/s-00330-021-08217-w	Wrong PICO
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-e148. doi: 10.1016/S2589-7500(20)30003-0	Wrong PICO

<p>Kizildag Yirgin I, Koyluoglu YO, Seker ME, Ozkan Gurdal S, Ozaydin AN, Ozcinar B, et al. Diagnostic performance of AI for cancers registered in a mammography screening program: a retrospective analysis. <i>Technol Cancer Res Treat.</i> 2022;21. doi: 10.1177/15330338221075172</p>	<p>Wrong PICO</p>
<p>Koch HW, Larsen M, Bartsch H, Daehli Kurz K & Hofvind S. Artificial intelligence in BreastScreen Norway: a retrospective analysis of a cancer-enriched sample including 1254 breast cancer cases. <i>Eur Radiol.</i> 2023;33(5):3735-3743. doi: 10.1007/s00330-023-09461-y</p>	<p>Wrong PICO</p>
<p>Kooi T, Litjens G, van Ginneken B, Gubern-Mérida A, Sánchez CI, Mann R, et al. Large scale deep learning for computed aided detection of mammographic lesions. <i>Med Image Anal.</i> 2017;35:303-312. doi: 10.1016/j.media.2016.07.007</p>	<p>Wrong study design</p>
<p>Larsen M, Aglen CF, Lee CI, Hoff SR, Lund-Hanssen H, Lång K, et al. Artificial intelligence evaluation of 122 969 mammography examinations from a population-based screening program. <i>Radiology.</i> 2022a;303(3):502-511. doi: 10.1148/radiol.212381</p>	<p>Wrong PICO</p>
<p>Lee JH, Kim KH, Lee EH, Ahn JS, Ryu JK, Park YM, et al. Improving the performance of radiologists using artificial intelligence-based detection support software for mammography: a multi-reader study. <i>Korean J Radiol.</i> 2022;23(5):505-516. doi: 10.3348/kjr.2021.0476</p>	<p>Wrong PICO</p>
<p>Leibig C, Brehmer M, Bunk S, Byng D, Pinker K & Umutlu L. Combining the strength of radiologists and AI for breast cancer screening: a retrospective analysis. <i>Lancet Digit Health.</i> 2022;4(7):e407-e419. doi: 10.1016/S2589-7500(22)00070-X</p>	<p>Wrong PICO</p>
<p>Lång K, Dustler M, Dahlblom V, Åkesson A, Andersson I & Zackrisson S. Identifying normal mammograms in a large screening population using artificial intelligence. <i>Eur Radiol.</i> 2021a;31(3):1687-1692. doi: 10.1007/s00330-020-07165-1</p>	<p>Wrong PICO</p>
<p>Lång K, Hofvind S, Rodríguez-Ruiz A & Andersson I. Can artificial intelligence reduce the interval cancer rate in mammography screening? <i>Eur Radiol.</i> 2021b;31(8):5940-5947. doi: 10.1007/s00330-021-07686-3</p>	<p>Wrong PICO</p>
<p>Marinovich ML, Wylie E, Lotter W, Lund H, Waddell A, Madeley C, et al. Artificial intelligence (AI) for breast cancer screening: BreastScreen population-based cohort study of cancer detection. <i>EBioMedicine.</i> 2023;90:104498. doi: 10.1016/j.ebiom.2023.104498.</p>	<p>High risk of bias</p>

McKinney SM, Sieniek M, Godbole V, Godwin J, Antropova N, Ashrafian H, et al. International evaluation of an AI system for breast cancer screening. <i>Nature</i> , 2020;577(7788):89-94. doi:10.1038/s41586-019-1799-6	Wrong PICO
Ng AY, Glocker B, Oberije C, Fox G, Sharma N, James JJ, et al. Artificial intelligence as supporting reader in breast screening: a novel workflow to preserve quality and reduce workload. <i>J Breast Imaging</i> . 2023a;5(3):267-276. doi: 10.1093/jbi/wbad010	Wrong PICO
Ng AY, Obrije CJG, Ambrózay É, Szabó E, Serfőző A, Karpati E, et al. Prospective implementation of AI-assisted screen reading to improve early detection of breast cancer. <i>Nat Med</i> . 2023b;29(12):3044-3049. doi. 10.1038/s41591-023-02625-9	Wrong PICO
Obrije CJG, Sharma N, James JJ, Ng AY, Nash J & Kecskemethy PD. Comparing prognostic factors of cancers identified by artificial intelligence (AI) and human readers in breast cancer screening. <i>Cancers</i> . 2023;15(12):3069. doi: 10.3390/cancers15123069	Wrong PICO
Rodríguez-Ruiz A, Krupinski E, Mordang J-J, Schilling K, Heywang-Köbrunner SH, Sechopoulos I, et al. Detection of breast cancer with mammography: effect of an artificial intelligence support system. <i>Radiology</i> . 2019a;290(2):305-314. doi: 10.1148/radiol.2018181371	Wrong PICO
Rodríguez-Ruiz A, Lång K, Gubern-Mérida A, Broeders M, Gennaro G, Clauser P, et al. Stand-alone artificial intelligence for breast cancer detection in mammography: comparison with 101 radiologists. <i>J Natl Cancer Inst</i> . 2019b;111(9):916-922. doi: 10.1093/jnci/djy222	Wrong PICO
Salim M, Wählin E, Dembrower K, Azavedo E, Foukakis T, Liu Y, et al. External evaluation of 3 commercial artificial intelligence algorithms for independent assessment of screening mammograms. <i>JAMA Oncol</i> . 2020;6(10):1581-1588. doi: 10.1001/jamaoncol.2020.3321	Wrong PICO
Salim M, Dembrower K, Eklund M, Smith K & Strand F. Differences and similarities in false interpretations by AI CAD and radiologists in screening mammography. <i>Br J Radiol</i> . 2023;96(1151):20230210. doi: 10.1259/bjr.20230210	Wrong PICO
Schaffter T, Buist DSM, Lee CI, Nikulin Y, Ribli D, Guan Y, et al. Evaluation of combined artificial intelligence and radiologist assessment to interpret screening mammograms. <i>JAMA Netw Open</i> . 2020;3(3):e200265. doi: 10.1001/jamanetworkopen.2020.0265	Wrong PICO

<p>Seker ME, Koyluoglu YO, Ozaydin AN, Gurdal SO, Ozcinar B, Cabioglu N, et al. Diagnostic capabilities of artificial intelligence as an additional reader in a breast cancer screening program. <i>Eur Radiol.</i> 2024. doi: 10.1007/s00330-024-10661-3. Online ahead of print</p>	<p>Wrong PICO</p>
<p>Sharma N, Ng AY, James JJ, Khara G, Ambrózay É, Austin CC, et al. Multi-vendor evaluation of artificial intelligence as an independent reader for double reading in breast cancer screening on 275,900 mammograms. <i>BMC Cancer.</i> 2023;23(1):460. doi: 10.1186/s12885-023-10890-7</p>	<p>Wrong PICO</p>
<p>Skaane P, Kshirsagar A, Stapleton S, Young K & Castellino RA. Effect of computer-aided detection on independent double reading of paired screen-film and full-field digital screening mammograms. <i>AJR Am J Roentgenol.</i> 2007;188(2):377-384. doi: 10.2214/AJR.05.2207</p>	<p>Wrong PICO</p>
<p>Vargas-Palacios A, Sharma N & Sagoo GS. Cost-effectiveness requirements for implementing artificial intelligence technology in the Women's UK Breast Cancer Screening service. <i>Nat Commun.</i> 2023;14(1):6110. doi: 10.1038/s41467-023-41754-0</p>	<p>Wrong PICO</p>
<p>Waugh J, Evans J, Miocevic M, Lockie D, Aminzadeh P, Lynch A, et al. Performance of artificial intelligence in 7533 consecutive prevalent screening mammograms from the BreastScreen Australia program. <i>Eur Radiol.</i> 2023. doi: 10.1007/s00330-023-10396-7. Online ahead of print</p>	<p>Wrong PICO</p>
<p>Wu N, Phang J, Park J, Shen Y, Huang Z, Zorin M, et al. Deep neural networks improve radiologists' performance in breast cancer screening. <i>IEEE Trans Med Imaging.</i> 2020;39(4):1184-1194. doi: 10.1109/TMI.2019.2945514</p>	<p>Wrong study design</p>
<p>Yala A, Schuster T, Miles R, Barzilay R & Lehman C. A deep learning model to triage screening mammograms: a simulation study. <i>Radiology.</i> 2019;293(1):38-46. doi: 10.1148/radiol.2019182908</p>	<p>Wrong PICO</p>
<p>Yoon JH, Han K, Suh HJ, Youk JH, Lee SE & Kim E-K. Artificial intelligence-based computer-assisted detection/diagnosis (AI-CAD) for screening mammography: outcomes of AI-CAD in the mammographic interpretation workflow. <i>Eur J Radiol Open.</i> 2023;11:100509. doi: 10.1016/j.ejro.2023.100509</p>	<p>Wrong PICO</p>

Excluded studies (review articles)	Motif for exclusion
Anderson AW, Marinovich ML, Houssami N, Lowry KP, Elmore JG, Buist DSM, et al. Independent external validation of artificial intelligence algorithms for automated interpretation of screening mammography: a systematic review. J Am Coll Radiol. 2022;19(2 Pt A):259-273. doi: 10.1016/j.jacr.2021.11.008.	High risk of bias
Bhalla D, Rangarajan K, Chandra T, Banerjee S & Arora C. Reproducibility and explainability of deep learning in mammography: a systematic review of literature. Indian J Radiol Imaging 2023. Article-in-Press. doi: 10.1055/s-0043-1775737	Wrong PICO
Liu J, Lei J, Ou Y, Zhao Y, Tuo X, Zhang B, et al. Mammography diagnosis of breast cancer screening through machine learning: a systematic review and meta-analysis. Clin Exp Med. 2023;23(6):2341-2356. doi: 10.1007/s10238-022-00895-0.	High risk of bias
Omega Boro L & Nandi G. Diagnostic performance of deep learning in screened mammogram: systematic review. Curr Med Imaging. 2024;20:e170423215871. doi: 10.2174/1573405620666230417085624	High risk of bias
Xavier D, Miyawaki I, Campello Jorge CA, Freitas Silva GB, Lloyd M, Moraes F, et al. Artificial intelligence for triaging of breast cancer screening mammograms and workload reduction: a meta-analysis of a deep learning software. J Med Screen. 2023;9691413231219952. doi: 10.1177/09691413231219952.	High risk of bias
Yoon JH, Strand F, Baltzer PAT, Conant EF, Gilbert FJ, Lehman CD, et al. Standalone AI for breast cancer detection at screening digital mammography and digital breast tomosynthesis: a systematic review and meta-analysis. Radiology. 2023;307(5):e222639. doi: 10.1148/radiol.222639.	High risk of bias

Appendix E: Pågående studier

Ongoing studies per 2024-04-30

Registration number	Study title	Locations of study	Estimated completion	Interventions	URL
NCT06032390	Artificial Intelligence in Mammography Screening in Norway	Norway	2033-06-01	AI assisted mammography screening interpretation Standard mammography screening interpretation	https://clinicaltrials.gov/study/NCT06032390
NCT04778670	Artificial Intelligence in Large-scale Breast Cancer Screening	Capio St Göran Hospital, Stockholm, Sweden Karolinska University Hospital, Stockholm, Sweden	2024-12-01	AI CAD Radiologist reading	https://clinicaltrials.gov/study/NCT04778670

NCT04949776	Artificial Intelligence in Breast Cancer Screening Programs in Córdoba (AITIC)	Hospital Universitario Reina Sofia, Córdoba, Spain	2024-02-01	Mammograms	https://clinicaltrials.gov/study/NCT04949776
NCT04838756	Mammography Screening With Artificial Intelligence (MASAI)	Mammography Unit, Unilabs / Skane University Hospital, Malmö, Skåne, Sweden	2025-04-12	AI screening modality Conventional screening modality	https://clinicaltrials.gov/study/NCT04838756
NCT06187350	The Use of AI to Safely Reduce the Workload in Breast Cancer Screening With Mammography in Region Östergötland	Region Östergötland, Linköping, Östergötland, Sweden	2027-12-31	AI cancer detection system	https://clinicaltrials.gov/study/NCT06187350
NCT05024591	Artificial Intelligence for breast cancer Screening in mammography (AI-STREAM)	Kyung Hee University Hospital at Gangdong, Seoul, Korea, Republic of	2024-12-31	Lunit INSIGHT MMG CADe/x for medical imaging	https://clinicaltrials.gov/study/NCT05024591
NCT05800132	Diagnostic Performance of Breast Cancer Screening Second Reading Process Assisted by AI	Centre Régional de Coordination des Dépistages des Cancers	2027-09-01	MammoScreen	https://clinicaltrials.gov/study/NCT05800132

		(CRCDC) Région Sud, Lille, France			
ISRCTN95571932	A prospective study to assess the impact and benefits of an AI system in double reading for breast cancer screening	Kheiron Medical Technologies, London, United Kingdom	2024-08-31	Mammography intelligent assessment (Mia)	https://www.isrctn.com/ISRCTN95571932
ISRCTN88754382	Can we use an artificial intelligence system to improve the quality and efficiency of breast cancer screening?	Imperial College London, United Kingdom	2024-05-21	AI system	https://www.isrctn.com/ISRCTN88754382
ISRCTN60839016	Artificial intelligence in mammography study	Imperial College London, United Kingdom	2024-05-21	AI system	https://www.isrctn.com/ISRCTN60839016
ChiCTR2300068614	A multicenter cohort study of breast Ultrasound Artificial Intelligence Diagnostic System for breast cancer screening	West China Hospital of Sichuan University, Chengdu City, Sichuan Province, China		Breast Ultrasound Artificial Intelligence Diagnostic System (DeepBC)	https://trialsearch.who.int/Trial.aspx?TrialID=ChiCTR2300068614

Appendix F: Ekonomiska aspekter

Underlag för beräkning av kostnader

Underlag från Unilabs

Unilabs genomför omkring 133 775 mammografibilder ifrån bröstcancerscreeningsundersökningar i Skåne per år. Nuvarande screeningmodell innebär att två radiologer granskar varje kvinnas bilder. Tiden som går åt för en radiolog att granska bilderna varierar beroende på radiologens erfarenhet och på vad bilden visar. Enligt Unilabs egen uppföljning av verksamheten granskar radiologer i genomsnitt bilder från 75 kvinnor per timme, men det kan variera mellan bilder från 60 till 120 kvinnor. I genomsnitt motsvarar det 75 kvinnor per timme och 48 sekunder per kvinna.

För närvarande arbetare 12 bröstradiologer hos Unilabs i Region Skåne och ytterligare 4 är under upplärning. Granskning av mammografibilder utgör del av verksamheten. Utifrån att det i genomsnitt tar 48 sekunder att granska en kvinnas bilder, motsvarar det i effektiv arbetstid cirka 2 heltidsekvivalenter. Granskning av bilder från bröstcancerscreening med mammografi utgör alltså en mindre del av bröstradiologernas arbetstid. För rutingranskning av mammografibilder kan Unilabs också kontraktera externa radiologkonsulter som då får 25 kronor per granskad kvinna.

Enhetskostnader och andra antaganden

Tabell F1. Enhetskostnader och relaterade uppgifter

Resurs	Värde i analysen	Källa
Granskningstid per kvinna, medel	48 s	Kommunikation med Unilabs
Helårsarbetstid	1787 h	Underlag från Arbetsgivarverket samt hänsyn till lagstadgad semester
Arbetskostnad radiolog, per timme inkl. påslag	882 kr	Statistikmyndigheten SCB, lönestatistik och helårsarbetstid
Personalomkostnadspålägg	57,21 %	Sveriges Kommuner och Regioner
Klinisk mammografi och ultraljud	4250 kr	Fridhammar (2022)
Ytterligare undersökning inkl. biopsi	6 603 kr	Fridhammar (2022)
Bildgranskning 2 radiologer per kvinna	23,51 kr	Beräkning utifrån arbetskostnad och genomsnittlig granskningstid
Bildgranskning 1 radiolog per kvinna	11,75 kr	Beräkning utifrån arbetskostnad och genomsnittlig granskningstid
Bildgranskning extern radiologkonsult, per kvinna	25 kr	Uppgift från projektgruppen
Hypotetiskt antagande om kostnad för AI, per kvinna	30 kr 10 kr	En högre och en lägre kostnad för att illustrera vilken roll som avgift för AI spelar. Antaganden i samråd med projektgruppen.

Tabell F2. Nyckeluppgifter från kliniska studier

Uppgift	Värde i analysen	Källa
Triage		
2 radiologer granskar efter AI vid riskpoäng = 10	7,20 %	Lång (2023) enligt studiedesign. Table 3
2 radiologer granskar efter AI vid riskpoäng =6–10	38,6 %	Inspirerat av planerad strategi vid införande i Region Halland. Underlag enligt Lång (2023), Table 3
AI ensam granskare vid bedömd låg risk. Andel kvinnor som också granskas av 1 radiolog	74,8 %	Lauritzen (2022). Beräknat från att radiologtid minskade med 62,6 %
Konsensusmöten		
MASAI Region Skåne	4 %	Lång (2023), table 2, ej signifikant skillnad mellan interventions- och kontrollgrupp
ScreenTrustCAD Region Stockholm: AI+1 radiolog	8,95%	Dembrower (2023) table 2
ScreenTrustCAD Region Stockholm: 2 radiologer	7,38%	Dembrower (2023) table 2
Tidsåtgång 2 radiologer	30 min	Antagande i samråd med projektgruppen
Återkallade		
MASAI Region Skåne Intervention	2,20 %	Lång (2023)
MASAI Region Skåne Kontroll	2,00 %	Lång (2023)
ScreenTrustCAD Region Stockholm AI + 1 radiolog	2,80 %	Dembrower (2023) table 2
ScreenTrustCAD Region Stockholm 2 radiologer	2,93 %	Dembrower (2023) table 2
Ytterligare undersökningar inkl. biopsi		
ScreenTrustCAD Region Stockholm AI + 1 radiolog	0,69 %	Dembrower (2023) table 2
ScreenTrustCAD Region Stockholm 2 radiologer	0,73 %	Dembrower (2023) table 2



Region Skåne
HTA syd

ISBN 978-91-989049-0-1