

Datorassisterad granskning inom mammografiscreening (CAD)

SBU ALERT-RAPPORT NR 2011-05 • 2011-05-18 • WWW.SBU.SE/ALERT



Sammanfattning och slutsatser

I Sverige erbjuds alla kvinnor mellan 50 och 69 år mammografiundersökning med regelbundna intervall. I de flesta landsting får även kvinnor i åldrarna 40–49 år och 70–74 år erbjudande om sådan undersökning. Svenska och europeiska riktlinjer rekommenderar att röntgenbilderna granskas av två specialutbildade radiologer (bröstradiologer). CAD ("computer-aided detection") är en metod för datoriserad analys av mammografibilder vid screening. Metoden har funnits i cirka tio år, men den kliniska användningen i Sverige är än så länge begränsad.

SBU:s bedömning av kunskapsläget

- Det vetenskapliga underlaget räcker inte för att bedöma om CAD plus granskning av endast en bröstradiolog ger ett minst lika bra resultat som vedertagen praxis, där röntgenbilden granskas av två bröstradiologer var för sig, så kallad dubbelgranskning.
- Eftersom de medicinska följderna är oklara, går det inte heller att bedöma kostnadseffektivitet eller samhällsekonomiska konsekvenser av att ersätta den ena granskaren med CAD i samband med mammografiscreening.
- Efter denna genomgång av litteraturen har CAD-metoden utvecklats genom förbättrade egenskaper hos dataprogrammet samt digitalisering av bilderna. Det är angeläget att undersöka den aktuella metodens nytta, risker och kostnader i ytterligare studier.

Metod och målgrupp

Den undersökta metoden är CAD ("computer-aided detection"), dvs datorstyrd analys av mammografibilder vid screening för bröstcancer. Målgruppen är kvinnor från 40 till 74 år som genomgår mammografi inom ramen för screening i befolkningen.

Programmet som används vid CAD identifierar och markerar områden som avviker från normal bröstvävnad.

CAD-programmet är inte avsett att användas som självständig del vid analys av mammografibilder. Det är istället utformat som ett stöd för att uppmärksamma radiologen på områden som avviker från normal bröstvävnad. Metoden kräver således tolkning av en bröstradiolog.

Det har föreslagits att CAD i samband med mammografiscreening skulle kunna ersätta den ena av de två bröstradiologiska bedömningarna som görs enligt europeiska och svenska riktlinjer. En förutsättning för detta är att den diagnostiska tillförlitligheten och patientnyttan blir lika god med bildgranskning av en bröstradiolog plus CAD som med granskning av två bröstradiologer. En annan viktig förutsättning är att inte alltför många kvinnor måste kallas tillbaka för kompletterande undersökning. I Europa är högsta rekommenderade återkallningsfrekvens 5 procent.

Hög medelålder bland verksamma bröstradiologer och dålig återväxt i denna grupp av specialister har ökat intresset för datoriserad analys av mammografibilder.

I denna rapport har vi undersökt om den diagnostiska tillförlitligheten är minst lika god, och återkallningsfrekvensen inte högre, då CAD plus granskning av en bröstradiolog används i samband med mammografiscreening, istället för granskning av två bröstradiologer.

Frågor

Är granskning av mammografibilder med en bröstradiolog plus CAD minst lika effektivt som granskning med två bröstradiologer (nuvarande praxis) avseende

- sensitivitet (sannolikheten för att sjuka har ett positivt testresultat)?
- specificitet (sannolikheten för att friska har ett negativt testresultat)?
- cancerdetektionsfrekvens (antal upptäckta cancerfall per 1 000 undersökta kvinnor)?
- återkallningsfrekvens (kvinnan kallas tillbaka för fortsatt utredning)?
- kostnadseffektivitet?

Avgränsningar

Rapporten omfattar enbart populationsbaserade screeningstudier. Populationen ska vara jämförbar med svenska förhållanden och omfatta minst 5 000 kvinnor. Studierna ska jämföra mammografigranskning med en bröststradiolog plus CAD med granskning med två bröststradiologer.

Prospektiva studier som grundar sig på digital mammo- grafi kunde inte identifieras, varför användning av scan- nade analoga bilder accepterades.

Patientnytta

Kraven på enkelgranskning plus CAD är att metoden ska upptäcka minst lika många cancerfall som dubbelgransk- ning, utan att återkallningsfrekvensen ökar, dvs metodens specificitet ska vara lika hög som vid dubbelgranskning. Förnyade undersökningar leder kortsiktigt till ökad oro hos kvinnorna och medför ökade kostnader.

Endast en studie med tillräcklig kvalitet uppfyllde inklu- sionskriterierna. Den jämförde enkelgranskning plus CAD med dubbelgranskning i samband med mammografi- screening, och fann ingen skillnad i andelen upptäckta cancerfall. Återkallningsfrekvensen var däremot högre för enkelgranskning plus CAD (3,9% jämfört med 3,4% för dubbelgranskning). Studiens generaliserbarhet mins- kas pga att samtliga bröststradiologer som medverkade i studien hade lång erfarenhet av att granska bilder vid mammografiscreening. Denna enda studie med brister

i studiekvalitet och generaliserbarhet kan inte ligga till grund för slutsatser (otillräckligt vetenskapligt underlag ⊕○○○).

SBU:s utvärdering visar således att det saknas tillräckligt vetenskapligt stöd för att kunna uttala sig om mammo- grafigranskning med en bröststradiolog plus CAD är jäm- förbar med nuvarande praxis med granskning med två bröststradiologer.

Ekonomiska aspekter

Eftersom de medicinska följderna är oklara, går det inte heller att bedöma kostnadseffektivitet eller samhällseko- nomiska konsekvenser av att ersätta den ena granskaren med CAD i samband med mammografiscreening.

Gradering av styrkan i det vetenskapliga underlag som en slutsats grundas på görs i fyra nivåer:

Starkt vetenskapligt underlag (⊕⊕⊕⊕). Bygger på studier med hög eller medelhög kvalitet utan försvagande faktorer vid en samlad bedömning.

Måttligt starkt vetenskapligt underlag (⊕⊕⊕○). Bygger på stu- dier med hög eller medelhög kvalitet med förekomst av enstaka försvagande faktorer vid en samlad bedömning.

Begränsat vetenskapligt underlag (⊕⊕○○). Bygger på studier med hög eller medelhög kvalitet med försvagande faktorer vid en samlad bedömning.

Otillräckligt vetenskapligt underlag (⊕○○○). När vetenskapligt underlag saknas, tillgängliga studier har låg kvalitet eller där stu- dier av likartad kvalitet är motsägande anges det vetenskapliga underlaget som otillräckligt.

Datorassisterad granskning inom mammografiscreening (CAD)

Projektgrupp

- **Edward Azavedo**, docent, överläkare, Röntgenavdelningen, Karolinska Universitetssjukhuset, Stockholm
- **Sophia Zackrisson**, med dr, specialistläkare, Diagnostiskt centrum för bild- och funktionsmedicin, Skånes universitetssjukhus, Malmö
- **Marianne Heibert Arnlind**, projektledare, hälsoekonom, SBU, heibert.arnlind@sbu.se
- **Ingegerd Mejäre**, bitr projektledare, SBU
- Övriga medverkande från SBU: Lars-Åke Marké, hälsoekonom, Karin Rydin, litteratursökare, Lena Wallgren, projektassistent

Granskare

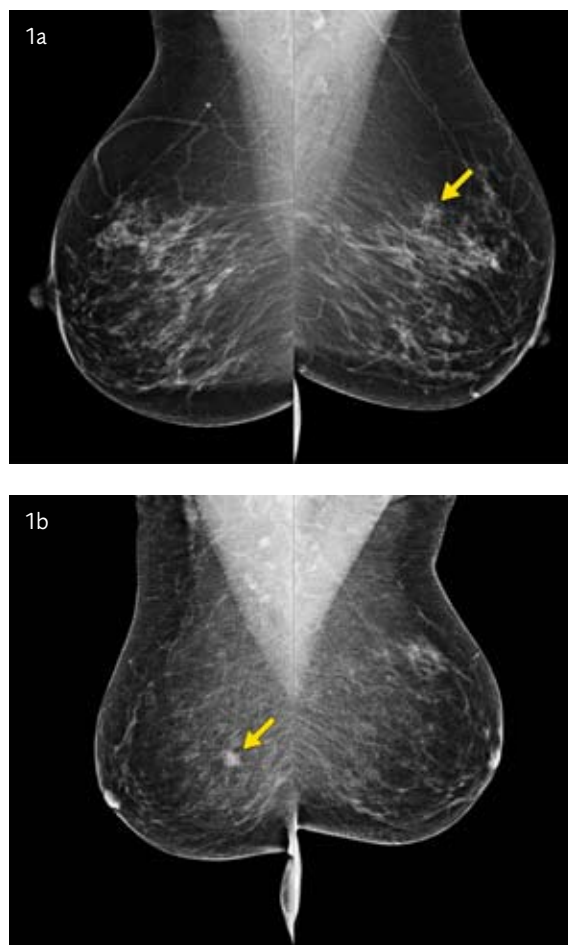
- **Olof Jarlman**, docent, ansvarig läkare, Sophiaklinikens bröstmottagning, Lund
- **Per Skaane**, professor, Department of Radiology, Oslo University Hospital, Oslo

Problembeskrivning

Drygt en fjärdedel av all cancer hos kvinnor är bröstcancer, vilket innebär att det är den vanligaste cancersjukdomen hos kvinnor i Sverige. Varje år ställs drygt 7 000 nya bröstcancerdiagnoser. Antal nya fall av bröstcancer i åldersgruppen 40–74 år var år 2009 2,76/1 000 kvinnor [1]. Sjukdomen drabbar framför allt medelålders och äldre kvinnor. Cirka 1 500 kvinnor avlider varje år av sjukdomen [2]. Med hjälp av mammografi kan bröstcancer upptäckas i ett tidigt stadium och minska dödligheten i sjukdomen för kvinnor i åldern 40–74 år med 24–33 procent [3].

Socialstyrelsen rekommenderar mammografiscreening för kvinnor mellan 40 och 74 år med intervall på 18 månader för kvinnor 40–54 år och 24 månader för kvinnor 55–74 år. Landstingen erbjuder alla kvinnor mellan 50 och 69 år möjlighet att genomgå mammografiundersökning. I de flesta landsting erbjuds även kvinnor 40–49 år samt 70–74 år undersökning, men det varierar när man kallas för första gången och hur länge undersökningarna fortsätter [4].

Varje år undersöks cirka 600 000 kvinnor med mammografi inom det nationella screeningprogrammet [5]. Uppgifterna är från 1995, och nyare saknas. Alla kvinnor i screeningåldrarna får en personlig inbjudan, och 70–92 procent deltar [6]. Vid undersökningen tas en eller två bilder av varje bröst. Bilderna granskas av två specialutbildade röntgenläkare (bröstradiologer) oberoende av varandra, och resultatet skickas till kvinnan inom två veckor.



Figur 1 Bild 1a visar en svårupptäckt cancer i vänster bröst (pil). Höger bröst är friskt. Bild 1b visar en lättupptäckt cancer i höger bröst (pil). Vänster bröst är friskt. Bilderna är spegelvända enligt radiologisk granskningspraxis.

Den radiologiska bilden av bröstvävnad är komplex. I Figur 1 visas exempel på röntgenbilder.

Ungefär 4–5 per 1 000 screenade kvinnor diagnostiseras med bröstcancer [6]. Förmågan att upptäcka bröstcancer (sensitivitet) vid mammografiscreening ligger mellan 70 och 85 procent, dvs upp till 30 procent av cancertumörer kan förbli oupptäckta vid screening [7]. Flera faktorer påverkar sensitiviteten i mammografiscreening, bl a

- förekomst av bröstcancer i den undersökta populationen
- åldersfördelning i den undersökta populationen
- mammografimetodens begränsningar att påvisa tumörer när bröstvävnaden är tät

- frekvensen tumörer med ett växtsätt som inte framträder tydligt på bilderna
- teknisk utrustning och bildkvalitet
- röntgensjuksköterskans kompetens vid bildtagning
- bröstradiologens kompetens vid granskning.

För att öka den diagnostiska tillförlitligheten bör, enligt svenska och europeiska riktlinjer, mammografibilderna granskas av två bröstradiologer oberoende av varandra [1,8]. Dubbelgranskning ökar sannolikheten för upptäckt av bröstcancer (cancerdetektionsfrekvensen) med 5–17 procent [9]. Att korrekt skilja ut de friska i populationen (specificitet) är minst lika viktigt som förmågan att hitta cancerfall, eftersom de flesta kvinnor som screenas faktiskt inte har bröstcancer. Specificiteten har rapporterats vara 82–98 procent [7].

Om bröstradiologen/-erna ser något misstänkt på bilderna, kallas kvinnan tillbaka för kompletterande bildtagning, i vissa fall även ultraljudsundersökning och eventuellt finnåls- eller mellannålspunktion för cytologi. I oklara fall kan operation göras. Om för få återkallas missas cancerfall, om för många återkallas, utreds och oroas många kvinnor i onödan. Enligt europeiska riktlinjer bör andelen kvinnor som återkallas vara mindre än 5 procent. Återkallningsfrekvensen i Sverige är cirka 3 procent. Av

Faktaruta 1 Definitioner.

Sensitivitet

Sannolikheten för att sjuka har ett positivt testresultat.

Specificitet

Sannolikheten för att friska har ett negativt testresultat.

Relativ sensitivitet

Antalet upptäckta cancerfall hos respektive granskare dividerat med det totala antalet upptäckta cancerfall.

Intervallcancer

Cancerfall som diagnostiseras mellan två screeningtillfällen.

Cancerdetektionsfrekvens

Antal upptäckta cancerfall per 1 000 undersökta.

Populationsbaserad screening

Samtliga kvinnor i vissa åldersgrupper får personlig inbjudan till mammografi med regelbundna intervall (1,5–2 år).

de återkallade har endast få kvinnor bröstcancer (cirka 20%), och majoriteten (cirka 80%) har alltså ingen bröstcancer (falskt positiva) [10,11].

Det finns inte några entydiga studier avseende psykosociala lång- och korttidseffekter hos kvinnor som fått besked om att fortsatt utredning behövs. Psykologiska korttidseffekter, såsom oro och ångest, har konstaterats i flera studier [12]. Resultat från studier om långtidseffekter är mer osäkra [13–15]. En högre återkallningsfrekvens innebär också högre kostnader för samhället [16].

Medelåldern hos radiologer som granskar mammografibilder är hög, och återväxten dålig. Att upprätthålla nuvarande praxis med dubbelgranskning kan därför tänkas bli svårt, även om detta förfarande ger en utmärkt utbildningsmöjlighet för nästa generation bröstradiologer. Intresset för att med datorns hjälp analysera mammografibilder har därför ökat. Utrustningen för så kallad CAD ("computer-aided detection"), som är datorstyrd bildanalys, bl a av mammografibilder, finns tillgänglig på marknaden sedan många år.

Frågor

I denna rapport har vi undersökt om den diagnostiska tillförlitligheten är minst lika god och återkallningsfrekvensen minst lika låg vid mammografiscreening vid användning av CAD + granskning av en bröstradiolog som vid granskning av två bröstradiologer.

Frågeställning enligt PICO [17]:

- P – Population/deltagare: kvinnor, 40–74 år, som deltar i mammografiscreening
- I – Intervention/metod: CAD + en bröstradiolog (enkelgranskning)
- C – Jämförelsemetod/kontroll: två bröstradiologer (dubbelgranskning)
- O – Effektmått: sensitivitet, specificitet, cancerdetektionsfrekvens och återkallningsfrekvens

Utvärderingen avser även att jämföra de två strategierna med avseende på kostnader och kostnadseffektivitet.

Avgränsningar

Inklusionskriterier

Studierna ska vara genomförda i en screeningpopulation och omfatta minst 5 000 kvinnor, motsvara svenska förhållanden (framför allt återkallningsfrekvens) och helst vara prospektiva. Studierna ska jämföra enkelgranskning + CAD med dubbelgranskning. Eftersom inga studier baserade på digital mammografi (DM) kunde identifieras, accepterades studier med scannade analoga bilder.

Beskrivning av den utvärderade metoden

CAD för mammografiscreening utvecklades för att kunna användas som en "andra granskare" i analogi med dubbelgranskning och med förhoppningen att öka framför allt den diagnostiska sensitiviteten [9]. CAD-programmet markerar områden som uppfattas som avvikande, och som kan misstänkas vara cancer. Många program ger möjlighet att balansera sensitivitet och specificitet med tyngdpunkt på det förstnämnda, dvs ökad känslighet för cancerupptäckt utan att samtidigt få för hög andel falskt positiva markeringar. Mjukvaran utvecklas kontinuerligt, och nya studier kommer således att baseras på förbättrade versioner. Det finns ett flertal CAD-program på marknaden med olika mjukvara och således med olika förmåga. Hittills finns CAD installerat på enstaka kliniker i Sverige, men används inte rutinmässigt.

En screeningundersökning består i allmänhet av fyra mammografibilder, där CAD ger ett antal markeringar per undersökning, varav de flesta är falskt positiva fynd [9]. Khoo och medarbetare redovisar 1,6 falskt positiva markeringar per screeningundersökning [18]. En översiktsartikel anger 2 falskt positiva markeringar per screeningundersökning och beräknar att CAD ger 400 falskt positiva markeringar för varje sant positiv markering [19]. En bröststradiolog måste alltså utvärdera ett stort antal markeringar, och CAD kan inte användas som "ensam granskare". Bröststradiologen granskar CAD-markeringar för att ta ett beslut; antingen att återkalla kvinnan för utredning, eller att avfärda markeringen som falskt positivt fynd. En bröststradiolog kan upptäcka sidoskillnad (höger och vänster bröst) och se eventuella nytillkomna förändringar jämfört med föregående mammografibilder, vilket ett CAD-program idag inte kan.

CAD-programmet kan användas på såväl digitaliserade analoga röntgenbilder (scannade bilder) som på digitala bilder. Merparten av de hittills publicerade studierna har använt digitaliserade analoga bilder. När man scannar/digitaliserar en analog mammografibild kan viss information gå förlorad. Flera studier, bl a en norsk studie [20], kunde dock inte visa någon skillnad på CAD-prestanda mellan scannade och direktdigitala mammografibilder.

Enligt en nyligen publicerad översikt visar CAD hög sensitivitet för malignitetsmisstänkta mikroförkalkningar (98–99%). Dessa utgör 15–20 procent av alla bröstcancerfall i populationen [10]. CAD visar lägre sensitivitet för tumörer som syns som andra typer av vävnadsförändringar (75–89%), och för riktigt svårupptäckta fall är sensitiviteten endast cirka 50 procent [19].

Patientnytta

Det saknas studier som jämför överlevnad mellan de två strategierna, och våra utfallsmått (cancerdetektions-

frekvens och återkallningsfrekvens) kan betraktas som surrogatmått. Endast en studie med medelhög kvalitet inkluderades i den sammanfattande resultattabellen [21]. Den är en prospektiv multicenterstudie i det nationella screeningprogrammet i Storbritannien med 28 204 kvinnor i åldrarna 50–70 år. Resultaten visade ingen statistiskt signifikant skillnad mellan enkelgranskning + CAD och dubbelgranskning avseende cancerdetektionsfrekvens (7,02/1 000 respektive 7,06/1 000) med en överensstämmelse på 74,9 procent (170/227). Enkelgranskning + CAD gav dock en signifikant högre återkallningsfrekvens (3,9% respektive 3,4%, $p=0,001$). Enkelgranskning + CAD jämfört med dubbelgranskning gav något lägre sensitivitet (87,2% respektive 87,7%) och specificitet (96,9% respektive 97,4%), men skillnaderna var inte statistiskt signifikanta. Ett observandum är att studien rapporterar en så kallad uppskattad sensitivitet, eftersom uppföljning inte hade gjorts av alla kvinnor, varför sensitiviteten sannolikt är överskattad. Generellt sett fanns ingen statistiskt signifikant skillnad mellan de två granskningsstrategierna avseende patologiska karakteristika hos de 57 upptäckta cancerumörerna. Resultaten från denna studie redovisas i Tabell 1 och 2.

Ytterligare tre studier inkluderades, men bedömdes inte ha tillräcklig kvalitet för att ligga till grund för slutsatser [18,22,23]. Två är genomförda i USA [22,23], där man inte har några populationsbaserade screeningprogram. Populationerna är inte väl beskrivna i artiklarna, och det är oklart om kvinnorna fått en personlig inbjudan, eller om endast kvinnor som själva sökt för mammografi hade inkluderats. Dessutom var återkallningsfrekvensen 8–12 procent, vilket är avsevärt högre än det i Sverige och Europa rekommenderade värdet (<5%). Den största studien är retrospektiv med 231 221 kvinnor som genomgick mammografiscreening [23]. Den andra är en prospektiv studie med 6 381 konsekutiva screeningundersökningar [22]. Ingen av dessa studier visade några statistiskt signifikanta skillnader i cancerdetektionsfrekvens mellan enkelgranskning + CAD och dubbelgranskning. Effekten på återkallningsfrekvensen var divergerande: lägre för enkelgranskning + CAD i den ena [23] och ingen statistisk skillnad i den andra [22].

Den tredje studien genomfördes inom ramen för the UK National Health Service Screening Programme [18]. Det är en prospektiv studie med 6 111 screeningundersökningar med en relativt hög cancerdetektionsfrekvens; 10/1 000 med enkelgranskning + CAD jämfört med 8,4/1 000 med dubbelgranskning. Skillnaden var dock inte statistiskt signifikant. Även kvinnor som var äldre än 64 år (övre gränsen för screening) hade inkluderats, vilket möjligen kan förklara den relativt höga andelen cancerfall. En annan bidragande förklaring kan vara att intervallet mellan screeningtillfällena var 3 år (vanligtvis 1,5–2 år).

I studien beräknades en så kallad relativ sensitivitet, där enkelgranskning + CAD gav lägre, dock inte statistiskt signifikant, sensitivitet på 91,5 procent jämfört med 98,4 procent med dubbelgranskning. Enkelgranskning + CAD gav en signifikant högre återkallningsfrekvens (6,1%) jämfört med dubbelgranskning (5,0%).

Dessa tre studier visar delvis motstridiga resultat. Cancerdetektionsfrekvensen förefaller dock påverkas mindre än återkallningsfrekvensen. Att dra några säkra slutsatser låter sig dock inte göras. Visserligen upptäckte man lika många cancerfall med båda metoderna men på bekostnad av avsevärt högre återkallningsfrekvens, som orsakar onödiga extra undersökningar. Återkallningsfrekvensen i amerikanska studier ligger generellt sett två till tre gånger högre än i Sverige, varför det är osäkert om resultaten kan överföras till svenska förhållanden [22,23].

Sammanfattningsvis fann vi endast en studie med medelhög kvalitet och som uppfyllde inklusionskraven [21], vilket ger otillräckligt underlag för att kunna dra några slutsatser om enkelgranskning + CAD kan ersätta dubbelgranskning i ett svenskt screeningprogram.

SBU tillämpar det internationellt utarbetade evidensgraderingssystemet GRADE [24], se Faktaruta 2.

Komplikationer och biverkningar

Högre återkallningsfrekvens innebär att kvinnor måste komma tillbaka för kompletterande undersökningar med ytterligare mammografibilder och ofta också ultraljudsundersökning. En stor andel genomgår också finnåls- eller vävnadsbiopsi och i vissa fall även kirurgiska ingrepp. Detta innebär också extra läkarbesök. Sammantaget tas extra resurser i anspråk och kvinnor oroas i onödan.

Tabell 1 Redovisning av studien av Gilbert och medarbetare.

Författare År, referens Land	Studiedesign Tid Population Bildgranskare	Indextest (I)	Referenstest	Resultat	Studiekvalitet Kommentarer
Gilbert et al 2008 [21] Storbritannien	<p>Prospektiv multicenterstudie</p> <p>År 2006–2007</p> <p><u>Population</u> 28 204 kvinnor 50–70 år (1% >70 år)</p> <p><u>Bildgranskare</u> Radiolog (n=17) + övrig specialutbildad personal (n=10)</p> <p>Samtliga granskare hade minst 6 års erfarenhet och över 5 000 granskningar per år</p>	<p>Ia: Enkelgranskning + CAD (n=28 204)</p> <p>Ib: Dubbelgranskning (n=28 204)</p>	Biopsi på misstänkta fall eller uppföljning (dock ej på samtliga, antal ej rapporterat)	<p><u>Cancerdetektionsfrekvens</u> <i>Enkelgranskning + CAD</i> 7,02/1 000</p> <p><i>Dubbelgranskning</i> 7,06/1 000</p> <p><i>Differens</i> 0,04/1 000 Ej statistiskt signifikant (p=0,89)</p> <p><u>Återkallningsfrekvens</u> <i>Enkelgranskning + CAD</i> 3,9%</p> <p><i>Dubbelgranskning</i> 3,4% (p=0,001)</p> <p><i>Differens</i> 0,5% (95% KI, 0,3; 0,8)</p> <p><u>Sensitivitet, specificitet</u> <i>Enkelgranskning + CAD</i> Sens: 87,2% Spec: 96,9%</p> <p><i>Dubbelgranskning</i> Sens: 87,7% Spec: 97,4%</p> <p><i>Differens sensitivitet</i> 0,5% (95% KI, -7,4; 6,6)</p>	<p>Medelhög studiekvalitet</p> <p>Begränsad generaliserbarhet</p> <p>Resultaten baseras på att enkelgranskning + CAD görs av erfaren bröstadiolog</p> <p>Ofullständig uppföljning, vilket påverkar skattningen av framför allt sensitivitet</p> <p>Scannade analoga bilder</p>

CAD = "Computer-aided detection"; KI = Konfidensintervall; Sens = Sensitivitet; Spec = Specificitet

Faktaruta 2 Evidensstyrka/GRADE.

Evidensstyrkan är en bedömning av hur starkt det sammanlagda vetenskapliga underlaget är för att besvara en viss fråga på ett tillförlitligt sätt. SBU tillämpar det internationellt utarbetade evidensgraderingssystemet GRADE [24]. Vid GRADE utgår man från effektmått. Styrkan på det vetenskapliga underlaget (evidensen) grundas på hur mycket man bedömer att man kan lita på nuvarande skattningar av effekt (resultat), dvs sannolikheten för att nya studier kommer att förändra tilliten till nuvarande resultat. Man utgår från fyra plus (⊕⊕⊕⊕), som är den starkaste evidensen. Fyra plus

innebär att man bedömer att nya studier med all sannolikhet inte kommer att förändra vår tillit till nuvarande resultat. Tre plus (⊕⊕⊕○) innebär att det är troligt att nya studier kommer att påverka och kan förändra vår tillit till nuvarande resultat. Två plus (⊕⊕○○) innebär att det är mycket troligt att ytterligare studier kommer att påverka och förändra vår tillit till nuvarande resultat, och ett plus (⊕○○○) innebär att resultaten är mycket osäkra. Avdrag görs för brister i studiekvalitet, t ex risk för bias, direktet (engelska "directness"), skillnader i patientpopulation, intervention, jämförelser eller utfall, precision (materialstorlek) och risk för publikationsbias.

Tabell 2 Sammanfattande resultattabell (GRADE). Skillnaden mellan enkelgranskning (en radiolog + CAD) och dubbelgranskning (två radiologer) avseende cancerdetektionsfrekvens respektive återkallningsfrekvens vid mammografiscreening. En studie [21].

Effektmått	Antal patienter (antal studier)	Sant positiva: Enkelgranskning + CAD (95% KI)	Sant positiva: Dubbelgranskning (95% KI)	Absolut skillnad (95% KI)	Vetenskapligt underlag	Kommentarer*
Cancer-detektionsfrekvens	28 204 (1)	0,702% (0,6; 0,8)	0,706% (0,6; 0,8)	0,004% (Ej statistiskt signifikant skillnad)	⊕○○○ Otillräckligt	Brister i studiekvalitet –1 Begränsad överförbarhet –1 Opresisa data –1
Effektmått	Antal patienter (antal studier)	Andel återkallade: Enkelgranskning + CAD (95% KI)	Andel återkallade: Dubbelgranskning (95% KI)	Absolut skillnad (95% KI)	Vetenskapligt underlag	Kommentarer*
Återkallningsfrekvens	28 204 (1)	3,9% (3,7; 4,1)	3,4% (3,2; 3,6)	0,5% (0,3; 0,8)	⊕○○○ Otillräckligt	Brister i studiekvalitet –1 Begränsad överförbarhet, endast en studie –2

CAD = "Computer-aided detection"; KI = Konfidensintervall

* Brister i studiekvalitet = Risk för bias, dvs sensitivitet troligtvis övervärderad (ofullständig uppföljning av kvinnor med negativt testresultat).
Begränsad överförbarhet = Endast bröstradiologer med lång klinisk erfarenhet ingick i studien.
Opresisa data = Vida konfidensintervall för skillnaden i sensitivitet mellan dubbelgranskning och enkelgranskning + CAD.

Användning av metoden i Sverige

För närvarande finns CAD installerat på ett fåtal screeningcentra, men används inte i klinisk praxis.

Ekonomiska aspekter

Det medicinska underlaget är otillräckligt för att bedöma kostnadseffektivitet eller samhällsekonomiska konsekvenser av att ersätta en bröstradiolog (andra granskaren) med CAD vid mammografiscreening. Med hänsyn till nedan redovisade resultat kan konstateras att införande av CAD för närvarande inte är kostnadseffektivt.

Kostnadseffektivitet

Litteraturgenomgången av hälsoekonomiska studier resulterade i att en HTA-rapport från Storbritannien inkluderades [25]. Syftet var att undersöka nyttan av CAD vid bröstcancerscreening, och för att besvara frågan genomfördes två studier inom ramen för rapporten. Dessa studier användes som underlag för beräkningar av kostnadseffektivitet. Två olika kostnadseffektivitetsutfall beräknades: kostnad per ytterligare upptäckt cancerfall och kostnad per kvalitetsjusterat levnadsår (QALY). Det senare beräknades med hjälp av en Markovmodell. Inga

statistiskt signifikanta skillnader erhöles mellan dubbelgranskning jämfört med enkelgranskning + CAD. Kostnadsberäkningarna visade att CAD ökade kostnaderna och att kostnadseffektiviteten var osäker. Det medicinska underlaget var dock bristfälligt, och studien bedömdes ha låg kvalitet.

Sjukvårdens struktur och organisation

Hälsundersökning med mammografi erbjuds i alla landsting i Sverige, men åldersgrupperna som bjuds in kan skilja något beroende på tillgängliga resurser. Populationsbaserad screening innebär att alla kvinnor erhåller en personlig inbjudan med vissa intervall, vanligen 18–24 månader. Det finns inget nationellt register för screeningverksamheten utan varje landsting eller screeningområde administrerar sitt eget register baserat på folkbokföringsregister. Antal screeningområden varierar, nyligen har 35 rapporterats [26].

Deltagarfrekvensen varierar över tid och mellan åldersgrupper, men ligger nationellt mellan 80 och 90 procent. Vanligtvis är deltagandet lägre i storstadsregionerna jämfört med övriga landet [5]. Återkallningsfrekvensen är 2–4 procent [26,27].

Vanligen finns flera screeningenheter i ett landsting, där enbart själva undersökningen utförs. I vissa glesbygdsområden används mobila enheter. Bilderna tas av röntgensjuksköterskor och själva undersökningen tar cirka fem minuter. Bilderna granskas därefter av bröstradiologer, och ett friskbrev skickas hem om allt är normalt. Om kvinnan behöver återkallas sker detta per telefon eller via brev. Återbesöket sker vanligtvis inom 14 dagar vid en mammografiklinik där det finns tillgång till klinisk mammografi, ultraljud, punktion etc för utredning.

Nuvarande rutin med dubbelgranskning innebär att en yngre kollega kan utbildas av en äldre och mer rutinerad bröstradiolog. Om enkelgranskning + CAD införs, måste frågan med utbildning av nya bröstradiologer lösas. Idag bygger den på att den mer erfarne utbildar den nyare och mindre erfarne. Detta är en viktig aspekt, eftersom bröstradiologens erfarenhet är avgörande för kvaliteten på diagnostiken vid enkelgranskning + CAD [28]. I framtiden borde dock datorstyrd utbildning med hjälp av CAD bli ett alternativ, förutsatt att metoden uppnår godtagbar sensitivitet och specificitet.

Idag görs mammografiscreening med direktdigital teknik (FFDM) istället för analog teknik i så gott som hela Sverige. De få ställen som inte har FFDM är på väg att införa sådan.

Etiska aspekter

Det grundläggande etiska problemet med all screening är att personer utan kliniska symtom inbjuds att göra en utredning för sjukdom. Socialstyrelsen rekommenderar att regelbundet inbjuda kvinnor i 40–74 års ålder till mammografiundersökning för att tidigt upptäcka bröstcancer. Målsättningen med att använda CAD skiljer sig inte från målsättningen med den nu använda metoden. Den vetenskapliga dokumentationen är bristfällig men CAD-metoden kan öka återkallningsfrekvensen, vilket ger ökade kostnader och att fler kvinnor oroas i onödan. Det finns även en teoretisk risk för att cancer utan klinisk betydelse upptäcks oftare om CAD-metoden används, med de negativa effekter det kan ha, framför allt för den drabbade kvinnan. Att otillförlitliga metoder används, särskilt i samband med screening, kan också skada allmänhetens tillit till sjukvården.

Identifierade kunskapsluckor

Det saknas tillräcklig kunskap om CAD som ersättning för en radiolog avseende såväl diagnostiska prestanda som kostnadseffektivitet. Det behövs populationsbaserade screeningstudier, och dessa bör ha en prospektiv och helst randomiserad studiedesign. Studierna bör vara utformade så att resultaten kan appliceras på europeiska (och svenska) förhållanden. Sedan tekniken introducerades för drygt tio år sedan, har CAD utvecklats och förbättrats. Därför behövs också nya studier som specifikt undersöker metodens detektionsförmåga och diagnostiska prestanda.

Samlad bedömning

Från ett patientperspektiv finns idag inte tillräckligt stöd i litteraturen för att kunna bedöma om enkelgranskning + CAD är likvärdigt med nuvarande praxis med dubbelgranskning.

Det medicinska vetenskapliga underlaget är också otillräckligt för att bedöma kostnadseffektivitet eller samhällsekonomiska konsekvenser av att ersätta en bröstradiolog (andra granskaren) med CAD vid mammografiscreening.

Det kommer inom överskådlig framtid alltid att krävas minst en radiolog som granskar mammografibilder, dvs tolkningen av bilderna kan inte automatiseras via datoranalys. Enkelgranskning + CAD för mammografiscreening kan dock komma att bli ett alternativ till dubbelgranskning, särskilt om CAD-programmet förbättras så att specificiteten blir högre (andelen falskt positiva markeringar reduceras) och sensitiviteten för svårdiagnostiserade cancerfall ökar. Oavsett detta krävs att bröstradiologen har tillräcklig kompetens, eftersom CAD inte ger någon vägledning för det kliniska beslutsfattandet om att fria eller återkalla. Om enkelgranskning + CAD införs, måste

frågan med utbildning av nya bröstradiologer lösas. Idag bygger den på att den mer erfarne utbildar den nyare och mindre erfarne. I framtiden borde dock utbildning med hjälp av CAD kunna ingå som en del av utbildningen, förutsatt att metoden uppnår godtagbar sensitivitet och specificitet.

Fördelen med enkelgranskning + CAD är att man sparar radiologresurser som kan användas inom andra delar av vården. Denna fördel begränsas dock delvis av behovet av utbildningsplatser (se avsnittet "Sjukvårdens struktur och organisation").

Nackdelen med enkelgranskning + CAD är (med reservation för osäkerheten i studieresultaten) att återkallningsfrekvensen är högre. Det motsvarar cirka 3 000 (0,5%* 600 000) kvinnor per år i Sverige.

Resultaten avseende sensitivitet visade ingen statistiskt signifikant skillnad mellan de två granskningsstrategierna (0,5%, se Tabell 1). Differensen visar dock vida konfidensintervall (-7,4% till 6,6%) vilket gör resultatet osäkert. Skillnaden i sensitivitet kan alltså vara 7 procentenheter lägre alternativt 6 procentenheter högre (Tabell 1) för enkelgranskning + CAD jämfört med dubbelgranskning. Den begränsade uppföljningen av kvinnorna i studien ökar ytterligare osäkerheten, eftersom man inte vet hur många så kallade intervallcancerfall som tillkom i respektive grupp.

Den teknik som använts i studien baseras på scannade analoga bilder. Med en helt digitaliserad teknik ökar möjligen kvaliteten hos CAD-metoden. I vilken mån detta skulle påverka en jämförelse mellan enkelgranskning + CAD och dubbelgranskning går inte att avgöra med nuvarande kunskap.

Pågående studier

Inga pågående studier har identifierats i databasen Clinical Trials.

Metodik för den systematiska litteraturgenomgången

Litteratursökning

Litteratursökning har utförts i databaserna Pubmed, Cochrane Library och Embase t o m september 2010. För en mer detaljerad beskrivning av vilka söktermer och begränsningar som använts, se Bilaga 1 på www.sbu.se/201105. Förutom sökningar i databaser har referenslistor granskats i relevanta arbeten.

Kvalitetsgranskning

De abstraktlistor som genererades vid litteratursökningen granskades av de två sakkunniga oberoende av varandra.

De studier som minst en av de sakkunniga bedömde som relevanta för frågeställningarna rekvirerades i fulltext. Studier som vid granskning i fulltext inte uppfyllde inklusionskriterierna exkluderades med angivande av huvudsakliga skäl för exklusion (Figur 2 och 3 nedan samt Bilaga 2 och 3 på www.sbu.se/201105). Som stöd för bedömningen användes en granskningsmall baserad på kriterier enligt QUADAS [29] och studierna graderades med måtten hög, medelhög eller låg studiekvalitet (Faktaruta 3). Data från studierna infördes i en tabell tillsammans med bedömd studiekvalitet samt eventuella kommentarer. Enbart studier med minst medelhög studiekvalitet tabellerades.

Faktaruta 3 Studiekvalitet.

Studiekvalitet avser den vetenskapliga kvaliteten hos en enskild studie och dess förmåga att besvara en viss fråga på ett tillförlitligt sätt. Expertgruppens bedömningskriterier av studiekvalitet har varit:

Hög kvalitet

Randomiserad kontrollerad studie (RCT), uppföljning av samtliga kvinnor mellan två screeningtillfällen, studiepopulation minst 10 000 (motsvarande minsta screeningenhet för svenska förhållanden), oselekterad population, jävsituation deklarerad.

Medelhög kvalitet

RCT alternativt prospektiv studie, studiepopulation minst 6 000, oselekterad population, jävsituation deklarerad.

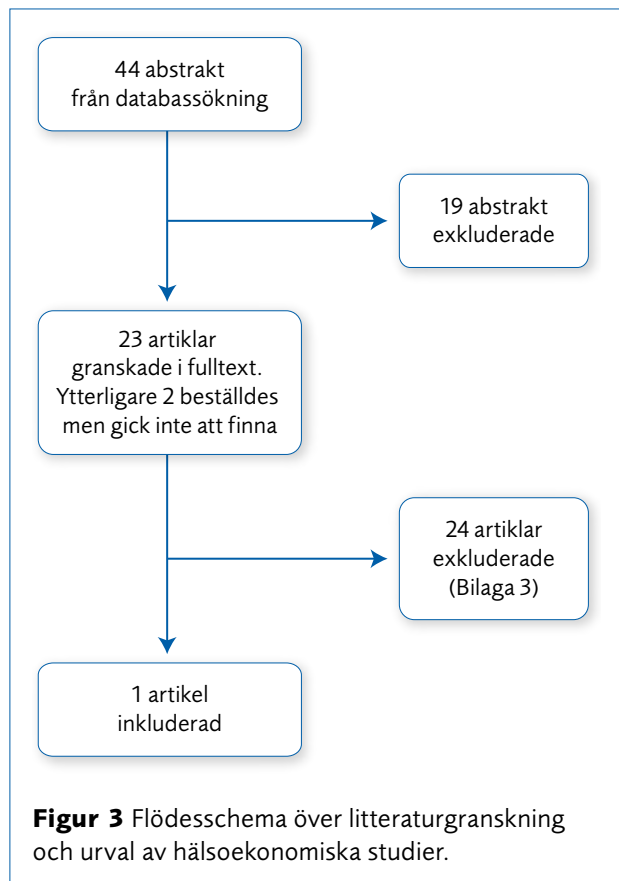
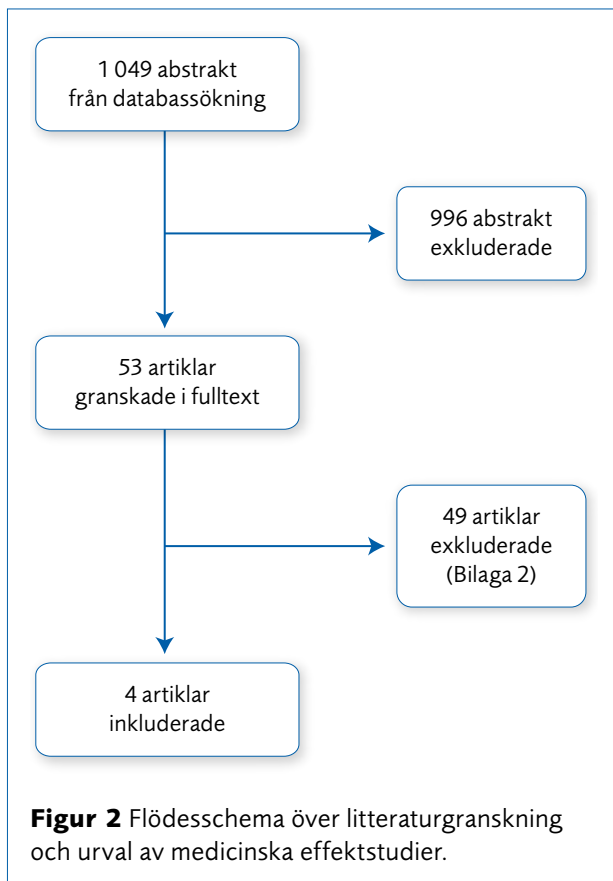
Låg kvalitet

Studier som inte uppfyller kraven för medelhög kvalitet.

Medicinska effektstudier

Femtiofyra artiklar granskades i fulltext. Nio stycken var översiktsartiklar [9,19,30–36]. Många var inte gjorda på ett screeningmaterial utan hade selekterade eller berikade populationer, ibland i kombination med att dubbelgranskning inte hade genomförts som jämförelse till enkelgranskning + CAD [20,25,28,37–63]. Nio av studierna hade stora material, men hade endast enkelgranskning som jämförelse [64–72]. En studie som beskrev cancertyperna exkluderades [73].

Initialt inkluderades fyra studier [18,21–23], vilka granskades med hjälp av en mall baserad på kriterier enligt QUADAS [29]. Studiernas metodologiska kvalitet och



kliniska relevans graderades som hög, medelhög eller låg (Faktaruta 3). Studien av Gilbert och medarbetare bedömdes ha medelhög kvalitet [21]. De övriga tre fick graderingen låg kvalitet [18,22,23], och därför inkluderades endast studien av Gilbert och medarbetare [21] i den sammanfattande resultattabellen. De tre studierna med låg kvalitet [18,22,23] kommenteras i texten.

Hälsoekonomiska studier

Två av de 25 hälsoekonomiska artiklarna som beställdes gick inte att finna [74,75]. Två var endast kommentarer [76,77], en rapport [78], en intervju med författarna [79] och en artikel var en "state of the art" [80]. Inga kostnader fanns med i sju av artiklarna [81–87] och tre artiklar var endast kommenterande ledartiklar [88–90]. Sex artiklar beskrev kostnader för att driva ett mammograficentrum, ersättningar och/eller "break-even" för ett center [33,91–95]. Artikeln av Lindfors och medarbetare exkluderades pga att alternativet, dvs den ordinarie mammografiverksamheten, inte beskrivs [96]. Studien av Taylor och medarbetare [25] svarade till viss del upp till inklusionskriterierna och redogörs för i kostnadsavsnittet.

Bindningar och jäv

Sakkunniga och granskare har i enlighet med SBU:s krav inlämnat deklARATION rörande bindningar och jäv. Dessa dokument finns tillgängliga på SBU:s kansli och kan rekvi-

reras från SBU (Box 3657, 103 59 Stockholm, eller e-post: registrator@sbu.se). SBU har på detta underlag bedömt att jäv inte föreligger.

Referenser

1. Socialstyrelsen. www.socialstyrelsen.se
2. Sveriges Kommuner och Landsting, Socialstyrelsen. Öppna jämförelser av hälso- och sjukvårdens kvalitet och effektivitet: jämförelser mellan landsting 2009. Stockholm: Sveriges Kommuner och Landsting; 2009.
3. Nyström L, Andersson I, Bjurstram N, Frisell J, Nordenskjöld B, Rutqvist LE. Long-term effects of mammography screening: updated overview of the Swedish randomised trials. *Lancet* 2002;359(9310):909-19.
4. 1177. Mammografi. www.1177.se
5. Olsson S, Andersson I, Bjurstram N, Frodis E, Håkansson S, Lithander E, et al. 600 000 kvinnor per år undersöks med mammografi. Var femte inbjuden avstår från screening. *Läkartidningen* 1995;92(6):552-6.
6. Swedish Organised Service Screening Evaluation Group. Reduction in breast cancer mortality from organized service screening with mammography: 1. Further confirmation with extended data. *Cancer Epidemiol Biomarkers Prev* 2006;15(1): 45-51.
7. Vainio H, Bianchini F, editors. IARC Handbooks of Cancer Prevention – Breast Cancer Screening. Lyon: IARC Press; 2002.
8. Perry N, Broeders M, de Wolf C, Törnberg S, Holland R, von Karsa L. European guidelines for quality assurance in breast cancer screening and diagnosis. Fourth edition – summary document. *Ann Oncol* 2008;19(4):614-22.
9. Helvie M. Improving mammographic interpretation: double reading and computer-aided diagnosis. *Radiol Clin North Am* 2007;45(5):801-11, vi.

10. Azavedo E, Svane G. Radiologic aspects of breast cancers detected through a breast cancer screening program. *Eur J Radiol* 1991;13(2):88-90.
11. Azavedo E, Svane G. Radial scars detected mammographically in a breast cancer screening programme. *Eur J Radiol* 1992;15(1):18-21.
12. Brett J, Bankhead C, Henderson B, Watson E, Austoker J. The psychological impact of mammographic screening. A systematic review. *Psychooncology* 2005;14(11):917-38.
13. Brett J, Austoker J. Women who are recalled for further investigation for breast screening: psychological consequences 3 years after recall and factors affecting re-attendance. *J Public Health Med* 2001;23(4):292-300.
14. Brewer NT, Salz T, Lillie SE. Systematic review: the long-term effects of false-positive mammograms. *Ann Intern Med* 2007;146(7):502-10.
15. Brodersen J, Thorsen H, Cockburn J. The adequacy of measurement of short and long-term consequences of false-positive screening mammography. *J Med Screen* 2004;11(1):39-44.
16. Lidbrink E, Elfving J, Frisell J, Jonsson E. Neglected aspects of false positive findings of mammography in breast cancer screening: analysis of false positive cases from the Stockholm trial. *BMJ* 1996;312(7026): 273-6.
17. Boudin F, Nie JY, Bartlett JC, Grad R, Pluye P, Dawes M. Combining classifiers for robust PICO element detection. *BMC Med Inform Decis Mak* 2010;10:29.
18. Khoo LA, Taylor P, Given-Wilson RM. Computer-aided detection in the United Kingdom National Breast Screening Programme: prospective study. *Radiology* 2005;237(2):444-9.
19. Houssami N, Given-Wilson R, Ciatto S. Early detection of breast cancer: overview of the evidence on computer-aided detection in mammography screening. *J Med Imaging Radiat Oncol* 2009;53(2):171-6.
20. 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. *AJR Am J Roentgenol* 2007;188(2):377-84.
21. Gilbert FJ, Astley SM, Gillan MG, Agbaje OF, Wallis MG, James J, et al. Single reading with computer-aided detection for screening mammography. *N Engl J Med* 2008;359(16):1675-84.
22. Georgian-Smith D, Moore RH, Halpern E, Yeh ED, Rafferty EA, D'Alessandro HA, et al. Blinded comparison of computer-aided detection with human second reading in screening mammography. *AJR Am J Roentgenol* 2007;189(5):1135-41.
23. Gromet M. Comparison of computer-aided detection to double reading of screening mammograms: review of 231,221 mammograms. *AJR Am J Roentgenol* 2008;190(4):854-9.
24. Guyatt GH, Oxman AD, Vist GE, Kunz R, Falck-Ytter Y, Alonso-Coello P, et al. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *BMJ* 2008;336(7650):924-6.
25. Taylor P, Champness J, Given-Wilson R, Johnston K, Potts H. Impact of computer-aided detection prompts on the sensitivity and specificity of screening mammography. *Health Technol Assess* 2005;9(6):iii, 1-58.
26. Hellquist BN, Duffy SW, Abdsaleh S, Björnelid L, Bordás P, Tabár L, et al. Effectiveness of population-based service screening with mammography for women ages 40 to 49 years: evaluation of the Swedish Mammography Screening in Young Women (SCRY) cohort. *Cancer* 2011;117(4):714-22.
27. Olsson S, Andersson I, Karlberg I, Bjurstam N, Frodis E, Håkansson S. Implementation of service screening with mammography in Sweden: from pilot study to nationwide programme. *J Med Screen* 2000;7(1):14-8.
28. Moberg K, Bjurstam N, Wilczek B, Rostgård L, Egge E, Muren C. Computed assisted detection of interval breast cancers. *Eur J Radiol* 2001;39(2):104-10.
29. Whiting P, Rutjes AW, Reitsma JB, Bossuyt PM, Kleijnen J. The development of QUADAS: a tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. *BMC Med Res Methodol* 2003;3:25.

Exkluderade studier

Se Bilaga 2 och 3 på www.sbu.se/201105.

SBU utvärderar sjukvårdens metoder

SBU, Statens beredning för medicinsk utvärdering, är en statlig myndighet som utvärderar hälso- och sjukvårdens metoder. SBU analyserar metodernas nytta, risker och kostnader och jämför vetenskapliga fakta med svensk vårdpraxis. Målet är att ge ett bättre beslutsunderlag för alla som avgör hur vården ska utformas.

SBU Alert-rapporterna tas fram i samarbete med sakkunniga inom respektive ämnesområde, Socialstyrelsen, Läkemedelsverket och Sveriges Kommuner och Landsting samt med en särskild rådsgrupp (Alerträdet).

Denna utvärdering publicerades år 2011. Resultat som bygger på ett starkt vetenskapligt underlag fortsätter vanligen att gälla under en lång tid framåt. Andra resultat kan ha hunnit bli inaktuella. Det gäller främst områden där det vetenskapliga underlaget är otillräckligt, begränsat eller motstridigt.

SBU Alert-rapport 2011-05 • ISSN 1652-7151 (webb)
Rapporten kan beställas från SBU:
Internet: www.sbu.se • Telefon: 08-412 32 00

Alerträdet

Jan-Erik Johansson, Ordförande, Professor, Urologi
Christel Bahtsevani, Dr Med Vet, Omvårdnad
Lars Borgquist, Professor, Allmänmed, Hälsoekonomi
Bo Carlberg, Docent, Internmedicin
Jane Carlsson, Professor, Sjukgymnastik
Per Carlsson, Professor, Hälsoekonomi
Björn-Erik Erlandsson, Professor, Medicinsk teknik
Mårten Fernö, Professor, Experimentell onkologi
Stefan Jutterdal, Utvecklingsdirektör (repr SKL)
Viveca Odling, Professor, Gynekologi (repr LV)
Anders Rydh, Docent, Med radiologi, Nuklearmedicin
Anders Tegnell, Med dr, Infektionssjukdomar (repr SoS)
Jan Wahlström, Professor emeritus, Klinisk genetik
Anna Åberg Wistedt, Professor, Psykiatri

SBU:s nämnds arbetsutskott

Susanna Axelsson, David Bergqvist, Håkan Ceder,
Tove Livered, Jan Liliemark, Nina Rehnqvist,
Måns Rosén och Ewalotte Ränzlöv.

Ansvarig utgivare: Måns Rosén, Direktör SBU
Programchef: Sofia Tranæus, SBU
Grafisk produktion: Elin Rye-Danjensen, SBU