

Screening for brystkreft

En oppsummering av kunnskapsstatus per august 2002

**SINTEF Unimed**

Postadresse: Boks 124, Blindern
0314 Oslo

Besøksadresse: Forskningsveien 1

Telefon: 22 06 73 00

Telefaks: 22 06 79 09

Foretaksregisteret: NO 948 007 029 MVA

SINTEF RAPPORT

TITTEL

Screening for brystkreft

FORFATTER(E)

Mørland, Berit; Lund Håheim, Lise; Linnestad, Kristin

OPPDRAUGSGIVER(E)

Sosial- og helsedirektoratet

RAPPORTNR.

SFT 78A2404

GRADERING

Åpen

OPPDRAUGSGIVERS REF.

GRADER. DENNE SIDE

Åpen

ISBN

82-14-02788-8

PROSJEKTNR.

ANTALL SIDER OG BILAG

72

ELEKTRONISK ARKIVKODE

Document2

PROSJEKTLEDER (NAVN, SIGN.)

Berit Mørland

VERIFISERT AV (NAVN, SIGN.)

ARKIVKODE

DATO

2002-08-30

GODKJENT AV (NAVN, STILLING, SIGN.)

SAMMENDRAG

Senter for medisinsk metodevurdering (SMM) ble april 2002 anmodet av Sosial- og helsedirektoratet om å foreta en sammenstilling av kunnskapsgrunnlaget for nytten av mammografiscreening.

STIKKORD

NORSK

ENGELSK

GRUPPE 1

mammografiscreening**mammography screening**

GRUPPE 2

brystkreft**breast cancer**

EGENVALGTE

Forord

Senter for medisinsk metodevurdering (SMM) ble april 2002 anmodet av Sosial- og helsedirektoratet (Avdeling for prioritering, retningslinjer og kvalitetsindikatorer) om å utføre en vurdering av nytten av mammografiscreening. Oppdraget ble gitt etter drøftinger med arbeidsutvalget i Nasjonalt råd for prioriteringer i helsevesenet og Helsedepartementet.

SMM har ved en intern prosjektgruppe utarbeidet utkast til rapport. Utkastet er forelagt følgende norske og nordiske eksperter til vurdering og uttalelser:

Divisjonsdirektør Arild Bjørndal, SHDir (samfunnsmedisin)
Sjeflege Bjørn Erikstein, Det Norske Radiumhospital (onkologi)
Sjeflege Rolf Kåresen, Ullevål Universitetssykehus HF (kirurgi)
Professor Eiliv Lund, Universitetet i Tromsø (epidemiologi)
Professor Steinar Østerbø Thoresen, Kreftregisteret (epidemiologi)
Professor Jan Erik Varhaug, Haukeland sykehus (kirurgi)
Dr med., Cand scient. Per-Henrik Zahl, Nasjonalt Folkehelseinstitutt, (statistikk)

Professor Elsebeth Lynge, Danmark (epidemiologi)
Professor Marjukka Mäkelä, Finland (medisinsk metodevurdering og klinisk epidemiologi)

De eksterne eksperter har hatt meget kort tid til arbeidet, og SMM er svært takknemlig for at de har stilt opp så velvillig og gitt meget nyttige kommentarer til rapportutkastet. Det er en selvsagt forutsetning at ekspertene ikke fremstår som garantister for den endelige rapporten, innholdet er SMM sitt ansvar.

Kommentarene fra de eksterne ekspertene er arkivert i SMM og er tilgjengelig på forespørsel.

Rapporten er godkjent av SMMs styringsgruppe.

Oslo, august 2002

Berit Mørland
Direktør

Innhold

FORORD	3
1. OPPSUMMERING	7
1.1 INFORMASJON TIL PASIENTER – HVA BØR DET INFORMERES OM	9
1.2 ENGLISH SUMMARY	10
2. INNLEDNING- OPPDRAG OG ARBEIDSFORM	13
3. HVA ER BRYSTKREFT OG STATUS FOR MAMMOGRAFI	16
3.1 ANTALL OG INSIDENS RATER I NORGE	17
3.1.1 ANTALL OG INSIDENS RELATERT TIL ALDER	17
3.1.2 LIVSTIDSRISIKO FOR Å DØ AV BRYSTKREFT	18
3.2 STATUS FOR MAMMOGRAFISCREENING I NORGE	18
3.3 STATUS I EUROPA	19
4. OM SCREENING	20
4.1 OM SCREENING GENERELT	20
4.2 FORSKNINGSMETODOLOGISKE SPØRSMÅL	21
4.3 HVORDAN ER OPPLEGGET FOR MAMMOGRAFISCREENING	21
5. DET VITENSKAPELIGE GRUNNLAGET FOR SCREENING FOR BRYSTKREFT	23
5.1 SYSTEMATISKE OVERSIKTER	25
5.2 BEHANDLINGSRETNINGSLINJER	27
5.3 NYTTE-KOSTNADS ANALYSER	28
5.4 ANNEN SPESIELT RELEVANT DOKUMENTASJON	28
6. HVORDAN SKAL NYTTEVERDIEN BESKRIVES – HVILKE MÅL PÅ EFFEKTER LEGGES TIL GRUNN	30
6.1 ANTATT POSITIVE KONSEKVENSER	30
6.2 ANTATT NEGATIVE KONSEKVENSER	31
6.3 SPESIELLE PROBLEMER	32
6.4 SCREENINGKVALITET	34
7. OPPSUMMERING AV KUNNSKAPSSTATUS	35
7.1 NYTTE-REDUSERT DØDELIGHET	37
7.2 NEGATIVE KONSEKVENSER ("SKADE") ELLER SPESIELLE PROBLEMER AV SCREENING	43
7.3 NYTTE-KOSTNADSVURDERINGER	47
8. VURDERING AV PROBLEMOMRÅDER	48
8.1 METODOLOGISKE PROBLEMER, HVOR STORE OG ALVORLIGE ER DE	48
8.2 EFFEKT AV SCREENING-DØDELIGHET	50
8.2.1 DØDELIGHET - KREFTSPESIFIKK ELLER TOTALDØDELIGHET?	50
8.2.2 EFFEKT IFT. ALDER	52

8.2.3	ANDRE MÅTER Å ILLUSTRERE NYTTEVERDIEN AV MAMMOGRAFISCREENING	54
8.3	OVERDIAGNOSTIKK OG OVERBEHANDLING	55
8.4	INTERVALLCANCERE	58
8.5	ER EFFEKTEN AV SCREENING REELL ELLER ER DEN RESULTAT AV FORBEDRET, IKKE-KIRURGISK BEHANDLING?	58
8.6	NYTTE-KOSTNAD	59
9.	PÅGÅENDE METODEVURDERINGER OG SCREENINGFORSØK	61
10.	AVSLUTTENDE KOMMENTARER	62
	VEDLEGG 1-5	63

1. Oppsummering

Denne metodevurderingen er basert på systematiske oversikter og metaanalyser publisert fom. 1995 av resultatene fra 7 (8) randomiserte studier, samt resultater fra populasjonsbaserte screeningprogrammer.

I vurderingen er det tatt stilling til

POSITIVE HELSEEFFEKTER		NEGATIVE HELSEEFFEKTER	
<u>Reduksjon i dødelighet</u> • Brystkreft-spesifikk absolutt dødelighet (faktisk) antall • Brystkreft-spesifikk relativ dødelighet • Total dødelighet ift. aldersgruppene 50-69 og 40-49 år	Resultatene er gitt i pkt. 7.1. Dokumentasjonsgrunnlaget er systematiske litteraturoversikter / metaanalyser av 7 (8) randomiserte kliniske studier	Falskt positive mammogram Flere med brystkreftdiagnose (overdiagnostikk, flere DCIS). Mulig overbehandling (av DCIS). Intervallcancer-problematikken.	Resultatene er gitt i pkt. 7.2 Dokumentasjonsgrunnlaget er resultater fra populasjonsbaserte screeningprogram

For **aldersgruppen 50-69** viser denne vurderingen at mammografiscreening reduserer den relative dødelighetsrisiko av brystkreft. Effektestimatene (relativ risiko; RR) er konsistent <1 i alle metaanalysene som er utført (med de unntak som er beskrevet i Cochrane-rapporten; se tabell s. 39), og med en dødelighetsreduksjon fra 6-27%. Det er Cochrane-rapporten som har den laveste risikoreduksjon (6%; forsøk med medium kvalitet), de øvrige rapporterer om en risikoreduksjon rundt 25%

Den relative dødsrisikoreduksjon må ses i forhold til det faktiske (absolutte) antall døde av brystkreft i hhv. mammografi- og kontrollgruppe i de utførte 7 (8) randomiserte forsøkene som er utført. (s. 37-38)

Mammografiscreening påvirker ikke totaldødeligheten for kvinner som screenes.

For denne aldersgruppen (50-69) er det lite sannsynlig at det vil komme ny primærforskning i overskuelig framtid som vil endre dette bildet vesentlig. Resultatene fra studien fra Singapore (se kap. 9) må selvsagt vurderes i sammenheng med de tidlige RCT.

Forøvrig kan frigivelse av rådata og reanalyser av allerede foreliggende data være en mulig kilde til ytterligere avklaring.

For **aldersgruppen 40-49 år** er det sannsynlighetsovervekt for at mammografiscreening ikke reduserer dødeligheten av brystkreft.

For denne aldersgruppen vil resultater fra det engelske forsøket (beskrevet i kapittel 9) kunne endre denne konklusjonen. Resultater vil tidligst foreligge i 2003.

Mammografiscreening øker antall brystkrefttilfeller ved påvisning av forstadier til kreft (ductal carcinoma in situ (DCIS)). DCIS utgjør 11-20 % av alle brystkrefttilfeller i første og senere screeningrunder. Ikke alle DCIS vil ha utviklet seg til klinisk sykdom og dermed ha behov for behandling. Med dagens kunnskap vet vi ikke hvilke saktevoksende kreftsvulster som vil ha utviklet seg til symptomgivende cancer, og hvilke som ikke vil føre til klinisk sykdom.

Falskt positive undersøkelser er et reelt problem. Basert på tall fra de populasjonsbaserte mammografiscreeningsprosjektene i England og København vil 10-14% av kvinnene ha en falsk positiv test (mammogram) etter 3 screeningrunder. Denne andelen vil øke for hver screeningrunde.

Minst 1/5-del av alle brystkrefttilfeller i et screeningprogram diagnostiseres utenom screening (dvs. mellom screeningrundene (intervallcancerfenomenet). Dersom en uttrykker intervallcancer ift. antall brystkrefttilfeller uten screening (proportional interval cancer rate), er denne andelen høyere.

Den vitenskapelige litteraturen om mammografiscreening har så langt hatt som hovedfokus å vise endringer i dødelighet (dødelighetsreduksjon). Data som fremstiller de negative sidene av screening er ikke lett tilgjengelige. De populasjonsbaserte mammografiscreeningsprosjektene er viktige informasjonskilder for slik informasjon.

Det er delvis tatt stilling til kostnads-nytte forholdet ved mammografiscreening i Norge, basert på 2 norske primærstudier. Kostnadene per vunnet leveår ved screening i aldersgruppen 50-69 år varierer. Isolert sett er det vanskelig å se at disse estimatene kan brukes i prioriteringssammenheng. Et kostnads-effekt forhold som ikke er sett i forhold til konkurrerende alternativ, vil ikke uten videre kunne legges til grunn for beslutninger. En problematisering av hva som inngår på nytte og kostnadssiden vil likevel være nyttig i en prioriteringssammenheng. Her må en også ta med økningen, forbigående eller permanent, av brystkrefttilfelle som følge av screening.

Det er i SMM-rapporten ikke konkludert ift.:

- Livskvalitet og kvinners reaksjon på screening
- Påstått overbehandling i form av flere mastektomier og flere strålebehandlede
- Påstått mer skånsom behandling etter mammografi
- Antatt stråleindusert sykdom pga. ekstern strålebehandling
- Påstått stråleindusert brystkreft forårsaket av gjentatte mammografier
- Effektene av endret screening intervall

Alle disse problemstillingene vil kreve en egen metodevurdering som det ikke har vært mulig å gjennomføre innen den frist som er gitt.

1.1 Informasjon til pasienter – hva bør det informeres om?

Kvinnene i de aktuelle aldersgruppene må informeres om alle sider ved mammografiscreening som et helsetilbud. Informasjon må dermed angi nytteverdien og også de negative konsekvensene av screening. Kunnskapen oppsummert i denne rapporten inneholder bl.a. følgende elementer:

- Mammografiscreening i aldersgruppen 50-69 år reduserer dødeligheten av brystkreft. Den vitenskapelige dokumentasjonen for dette har noen svakheter og beregningene av den relative dødelighetsreduksjonen varierer mellom 6%¹ og 27%². De øvrige rapporter om en risikoreduksjon rundt 20-25 %.
- Effekten for aldersgruppen 40-49 år er mindre og ikke sikker med et enda svakere dokumentasjonsgrunnlag.
- Mammografiscreening beskytter ikke mot utvikling av brystkreft mellom screeningrundene. Intervallcancer er en realitet og mer enn 1/5 av kvinnene som får diagnostisert brystkreft i et screeningprogram, får sin kreftdiagnose mellom screeningrundene. Kvinnene må ha kjennskap til dette fenomenet.
- Det foreligger en kumulativ risiko for falskt positive mammogram for hver screeningrunde, og dermed behov for nye undersøkelser og andre utredninger (inkl. vevsprøver /kirurgi).
- Det er usikkert om alle med DCIS trenger behandling. Problemet er at en ikke vet hvilke DCIS-tilfeller som trenger behandling, og hvilke som ikke gjør det.

Et viktig aspekt her er hvor stor del av de inviterte kvinnene som møter til screening. Dersom mammografiscreening skal være effektivt på befolkningsnivå, antas at flertallet av kvinnene møter frem til undersøkelse når de bli invitert.

Her er det antagelig kryssende hensyn. Gir en kvinnene "sann" og balansert informasjon om nytte versus skade, vil den usikkerheten som her skapes kunne redusere fremmøte. Dermed vil effekten av screening på populasjonsnivå reduseres.

En drøfting av disse spørsmål finnes i en nylig lederartikkel i JAMA³. I lederen heter det bl.a.

"Because screening is a population-based strategy, the responsibility to inform the choice falls on the larger system: managed care organizations, insurers, governmental agencies, and professional medical societies. This education may be an ideal part of the curriculum for group visits. However, the educational approach needs to shift. Most educational efforts to date have focused on persuading people to be tested, not informing them of the trade-offs involved".

Hva informasjonene forøvrig og hva det informerte samtykke ift. informasjonsplikten må inneholde, er ikke bare et juridisk problem ift. lovverket, men også et spørsmål som vil kreve klargjøringer av de rent medisinsk-faglige spørsmål reist i denne rapporten, og diskusjoner om etiske forhold rundt screening.

¹ Cochrane-rapporten, "medium quality" –studier : RR =0.94, 95% konfidensintervall 0,77-1,15. (ref. 10)

² Kerlikowske 1997 (ref. 6)

³ JAMA 2001; 258:2776-8.

1.2 English summary

The Norwegian Centre for Health Technology Assessment (SMM) was in April 2002 asked by the Norwegian Directorate for Health and Social Services to make an evaluation of the benefit of mammography screening. The task was given after consultation with the National Council on Priorities in Health Care and The Department of Health.

A working group at SMM made a draft report. The draft report was peer reviewed by Norwegian and Nordic experts.

This technology assessment on mammography screening is based on systematic reviews and meta-analyses of 7 (8) randomised controlled trials (RCT) and including results of population based screening programs published from 1995 and later.

The following is evaluated:

POSITIVE HEALTH EFFECTS		NEGATIVE HEALTH EFFECTS	
<u>Reduction in mortality</u> - Breast cancer mortality in absolute numbers - Relative breast cancer mortality - Total mortality Included are the age groups 50-69 and 40-49 years.	Results are given in section 7.1. Documentation from systematic reviews and meta-analyses of 7 (8) randomised controlled trials.	False positive mammograms More women with diagnoses of breast cancer (overdiagnosis, more Ductal Carcinoma in Situ (DCIS)) Possible overtreatment (of DCIS)	Results are given in section 7.2. Documentation from results of population based screening programs
		The issue on interval cancers	

For *the age group of 50-69 years* this assessment shows a reduction of the relative risk of breast cancer mortality. The effect estimates (relative risk; RR) are consistently under 1 in all the meta-analyses that have been carried out (with the exception of the Cochrane-report; se table **page 39**) with a reduction in breast cancer mortality of 6-27%. It is the Cochrane report that give the lowest risk reduction (6 % in trials of medium quality), the others report of a risk reduction of about 25 %.

The relative risk of mortality must be seen in relation to the absolute numbers in the randomised groups of the 7(8) trials. (pages 37-38)

Mammography screening does not influence the total mortality of the women who have undergone screening.

For this age group (50-69 years) results of new trials are not expected to change this picture to a great extent. The results of the ongoing trial in Singapore (chapter 9) must of course be assessed with the previous RCTs. Reanalyses of current data may be a source for clarification.

For *the age groups 40-49 years* it is not probable that mammography screening will reduce breast cancer mortality.

Results of the English trial on this age group (described in Chapter 9) may change this conclusion. The results are expected in 2003.

Mammography screening increases the incidence of breast cancer by also detecting early stages (Ductal Carcinoma in Situ (DCIS)). DCISs constitute 11-20 % of all cases of breast cancer diagnosed at the first and following screenings. Not all DCISs will cause clinical symptoms and require treatment. With to-days knowledge we do not know which of the slow growing tumours will give symptoms of cancer and which will not.

False positive mammograms are a real problem and based on figures from the population based mammography screening programs of England and Copenhagen, Denmark, 10-14 % of women will have a false positive test (mammography) after 3 rounds of screening. This proportion will increase with each screening round.

At least one fifth of all cases of breast cancer in a screening program are diagnosed outside the program (that is between the planned screening rounds (called interval cancers). If interval cancer is expressed relative to the number of breast cancer in a population without screening program (proportional interval cancer rate), then this proportion is higher.

The scientific literature on mammography screening has so far had its main focus on the changes in mortality (reduction in mortality). Data on negative effects of screening are not easily available. The population based screening projects are important sources of such information.

The cost-utility of mammography screening in Norway is partly considered, based on two Norwegian studies. Cost per life saved in the age group 50-69 varies. In isolation, it is hard to see how these figures can be used for making priorities. A cost-benefit analysis without making comparisons to other alternatives can not be used in making decisions without further considerations.

A discussion on what items are to be regarded as cost and benefit will be useful in conjunction with prioritisation. The increase, whether temporarily or permanent, in the number of cases of breast cancer due to the screening must also be considered.

In SMMs report there is no conclusion on:

- Quality of life and the reaction of women to screening ("distress")
- Claims of over treatment of more mastectomies and radiation therapy
- Claims of increased use of breast-conserving treatment after mammography
- Assumed increased risk of radiation induced disease due to external radiotherapy
- Claimed cases of radiation induced breast cancer due to repeated mammograms
- The effects of different lengths of screening interval

All these points need separate assessments that were not feasible to do within the time limit.

Information to patients – what information should be given?

The women in the relevant age groups must be informed about all aspects of mammography screening as a health offer. The information must include the benefit and also the negative consequences of screening. The knowledge as summed in this report includes the following elements:

- Mammography screening in the age group of 50-69 years reduces breast cancer mortality. The scientific documentation has some weaknesses and the estimated relative risk of reduction in mortality is in the order of 6 %¹ to 27 %². Other documentation reports a risk reduction of 20-25 %.
- The effect in the age group 40-49 years is less and uncertain with an even weaker evidence.
- Mammography screening does not protect against development of breast cancer between the screening rounds. Interval cancer is a reality and more than on fifth of the women who get the diagnosis breast cancer, get their cancer diagnosis between the screening rounds. The women must have knowledge of this phenomenon.
- There is a cumulative risk of false positive mammograms for each screening round, and thereby a need for new mammograms and other examinations (included biopsies or surgery).
- It is uncertain if all DCISs need treatment. The problem is that one does not know which DCISs needs treatment and which ones do not.

An important aspect here is what proportion of the women invited to screening does attend. If the mammography screening is to be effective at a population level it is assumed that the majority of the women attend when invited.

Here may be crossing considerations. If women are given “true” and balanced information about benefit versus harm, the insecurity that may arise could reduce attendance. Then the effect of screening on population level will be reduced.

From a discussion on these questions in a new Editorial in the JAMA³ is quoted:

“Because screening is a population-based strategy, the responsibility to inform the choice falls on the larger system: managed care organizations, insurers, governmental agencies, and professional medical societies. This education may be an ideal part of the curriculum for group visits. However, the educational approach needs to shift. Most educational efforts to date have focused on persuading people to be tested, not informing them of the trade-offs involved”.

What additional information and what the informed consent in accordance with the information duty must include, is not only a judicial problem according to the law, but also a question that will demand clarification of the purely medically related questions raised in this report and discussions on ethical issues around screening.

¹ Cochrane-rapporten, ”medium quality”-studier : RR = 0.94, 95% konfidensintervall 0,77-1,15 (ref.10).

² Kerlikowske 1997 (ref.6).

³ JAMA 2001; 258:2776-8.

2. Innledning og arbeidsform

Bakgrunn og oppdrag

Mammografiscreening er omdiskutert, framfor alt om nytten av tiltaket i form av antatt redusert dødelighet står i forhold til eventuelle negative konsekvenser av screening.

Diskusjonene har fått ny næring etter en nylig publisert systematisk litteraturoversikt om mammografiscreening utgått fra Nordic Cochrane Centre i København (publisert i Lancet) og senere også fra Cochrane Collaboration⁴.

Med mammografiscreening menes her undersøkelse av kvinners bryst med radiologisk diagnostikk (mammografi) i et organisert system der vurderingen av screeningen innbefatter alle trinn i dette systemet (valg av populasjon (aldersgrupper), resultatet av selve undersøkelsen, etterkontroll og evt. behandling av positive funn). Mammografi brukt som et supplement til klinisk undersøkelse på gitt indikasjon, eller mammografi som tas på legers eller kvinnens eget initiativ uten klinisk mistanke om sykdom (opportunistisk screening, "villscreening"), faller utenfor denne definisjonen og vil ikke bli omtalt.

I Norge har bl.a. brev til helseminister Dagfinn Høibråten fra Den nasjonale forskningsetiske komite for medisin av 17.12.01 (www.etikkom.no/NEM/01/mammografi.htm) ført til diskusjoner ikke bare i fagmiljøene, men også i Nasjonalt råd for prioriteringer i helsevesenet. Diskusjonen dreier seg om nytteverdien av screening, om mammografiscreening er å anse som et etablert og veldokumentert helsetiltak, og om informasjonen som kvinnene som inviteres til screening får er i tråd med best mulig kunnskap.

Dette er bakgrunnen for at Sosial- og helsedirektoratet (SHDir) i brev av 22.04.02 (Ref. 02/4256 tsl/trp; Vedlegg 1) ba Senter for medisinsk metodevurdering (SMM) om å vurdere nytten av mammografiscreening.

Tilsvarende problemstillinger er reist i andre land. Bl.a. har den hollandske minister for helse, velferd og sport bedt Gezondheidsraad (Health Council of the Netherlands)⁵, som er en søsterorganisasjon til SMM, om å foreta en vurdering av mammografi. Også i Holland er bakgrunnen diskusjonene som oppstod ifm. publiseringen av den omtalte Cochrane-oversikten.

Socialstyrelsen i Sverige har også på bakgrunn av diskusjonene etter Cochrane-rapporten ble publisert, avholdt et møte mellom "kritikere" og "talspersoner for"

⁴ SMM går ikke inn på uenighetene som her foreligger, delvis innad i Cochrane Collaboration (Breast Cancer Group), og delvis ift. publikasjonsetiske problemstillinger bl.a. reist av Lancets sjefsredaktør Richard Horton.

⁵ www.gr.nl

(”førespråkare”). Uttalelsen fra Socialstyrelsen er publisert (www.sos.se/SOS/PUBL/MEDBLAD/mb0203.htm).

Utenlandske organisasjoner eller institusjoner som U.S. Preventive Services Task Force (USPTTF) og National Cancer Institute (National Institute of Health) PDQ^{6A} har også drøftet mammografiscreening-problematikken og noen har revidert sine retningslinjer (eng. guidelines) for og informasjon om mammografiscreening etter Cochrane-rapporten ble publisert.

Arbeidsform

SMMs normale arbeidsform er på en systematisk måte å identifisere og innhente relevant vitenskapelig informasjon, vanligvis primærstudier, vurdere denne kritisk og oppsummere resultatene. Arbeidsformen følger beskrevne og aksepterte prinsipper for slikt arbeid⁶. SMM vil vanligvis opprette en egen ekspertgruppe som under ledelse av SMM forestår arbeidet etter fastlagte prosedyrer.

I dette tilfellet vil denne arbeidsmetoden fravikes. Som grunnlag for denne utredningen, vil SMM bruke *systematiske litteraturoversikter* (sekundær-litteratur). Dette er kunnskapsoppsummeringer basert på en systematisk litteraturidentifikasjon og kritisk vurdering av enkeltstudier.

Begrunnelsen er at SMM har fått kort tid på seg til å ferdigstille arbeidet, og at det synes som om det er liten uenighet om hvilke vitenskapelige arbeider om mammografi som legges til grunn. Uenigheten går mer på hvordan de vitenskapelige arbeidene vurderes og tolkes og hvilke konklusjoner som skal trekkes.

Inndelingen av vitenskapelig litteratur og bruk av andre kilder fremgår av kapittel 5. Som hovedregel vil kun litteratur publisert f.o.m. 1995 bli brukt. Grunnen til dette er å sikre lengre oppfølging av kvinnene som har deltatt i de randomiserte kliniske studiene.

SMM etablerte en egen intern prosjektgruppe⁷ som utarbeidet et utkast til rapport. Utkastet ble deretter sendt til norske og nordiske eksperter innen brystkreft og brystkreftbehandling, screening, epidemiologi, metodevurdering og økonomi for vurdering og kritikk. De skriftlige uttalelsene ble deretter vurdert av SMM og endringer foretatt på punkter der det etter vurdering av kommentarene var nødvendig. Uttalelsene fra de eksterne ekspertene er tilgjengelig i SMM. SMMs respons og kommentarer gis i Vedlegg 2

Rapporten er til slutt godkjent av SMMs styringsgruppe.

^{6A} www.nci.nih.gov/cancer_information/pdq/. PDQ = Physician Data Query. PDQ er definert slik: "An NCI Database that contain the latest information about cancer treatment, screening, prevention, genetic and supportive care, plus clinical trials". PDQ er imidlertid ikke NCI's "official policy voice", men er et uavhengig faglig panel. PDQ om mammografiscreening er oppdatert juni 2002.

⁶ Se f.eks. "Undertaking Systematic Reviews of Research on Effectiveness. CRD's Guidance for those Carrying Out or Commissioning Reviews". CRD Report Number 4 (2nd Edition, March 2001). Centre for Review and Dissemination, University of York.

⁷ Ledet av SMMs direktør, Berit Mørland, og med dr.phil. Lise Lund Håheim, M. Sci. Elisabeth Buntz og cand. oecon. Kristin Linnestad som faglige medarbeidere. SMMs øvrige ansatte har deltatt i diskusjonene om utformingen og innhold.

Rapporten vil i svært liten grad omtale og drøfte mammografi som diagnostisk test, dvs. hvor god eller dårlig testen mammografi er. Det understrekes likevel at dette er et viktig punkt. Kvaliteten på mammografiundersøkelsen og tolkning av bildene diskuteres nå også i media⁸. Effekt-kostnads-analysene vil bl.a. påvirkes av egenskapene til mammografi som test eller diagnostisk metode (se pkt. 8.6).

Mandatet som er gitt er svært omfattende og ville ideelt sett krevet flere separate metodevurderinger som det innenfor den fristen som er gitt ikke har vært mulig å gjennomføre.

⁸ Se Moss M. Spotting Breast Cancer Doctors are Weak Link. New York Times June 27, 2002.

3. Hva er brystkreft og status for mammografi

Brystkreft er den hyppigste kreftformen hos kvinner, og utgjør omlag 23% av alle krefttilfeller hos kvinner. Hver 12. kvinne vil utvikle brystkreft. I 1999 fikk 2380 kvinner denne diagnosen⁹.

Brystkreft rammer først og fremst kvinner over 60 år. Over 34 % av de nye tilfellene forekommer i denne aldersgruppen. Omlag 4,5% nye tilfeller inntraff hos kvinner under 40 år. Det var 477 krefttilfeller hos kvinner under 50 år i 1999 (20%).

Årlig dør rundt 700-750 kvinner av brystkreft i Norge. Dødeligheten har vært relativt konstant i mange år. Brystkreft er en viktig årsak til tapte leveår hos kvinner. Til enhver tid lever vel 20 000 kvinner med diagnosen brystkreft i Norge. Nesten alle er operert.

Brystkreftsykdommen deles inn i ulike stadier (dvs. etter hvor langt kommen kreftsykdommen er). Systemet for inndeling baseres på internasjonale konvensjoner (TNM-systemet)¹⁰ og baseres på informasjon om svulststørrelse (T-stadium), lymfeknutestatus (kreftspredning eller ikke; N-stadium) og fjernmetastaser (M-stadium). Brystkreftsykdommen kan på bakgrunn av slike data deles inn i sykdomsgrupper (4 stadier; stadium I-IV). Overlevelsessannsynlighetene på diagnosetidspunktet bestemmes i stor grad av sykdommens stadium.

I screeningsammenheng skilles mellom *invasiv cancer* (som vil ha varierende størrelse, dvs. diameter og utbredelse) og *in situ cancer* (internasjonalt kalt CIS).

In situ cancer (CIS) er forstadier til kreft og finnes som *lobulær carcinoma in situ (LCIS)* og *ductal carcinoma in situ (DCIS)*. LCIS har ikke noe korrelat til de radiologiske funn ved mammografi (dvs. forklarer ikke unormale mammogram) og er derfor i screeningsammenheng lite relevant.

Entiteten *ductal carcinoma in situ (DCIS)* oppdages ved mammografi. Deteksjonsraten for brystcancer er summen av ratene for invasiv kreft og DCIS. Det er fortsatt ikke avklart at alle DSIC-tilfeller må behandles (se diskusjonene under 8.4).

Kreftsykdom i stadium III og IV regnes vanligvis som ikke-operabel, dvs. at kreften ikke kan fjernes med operasjon alene.

Mammografiscreening har potensiell betydning ved å identifisere brystkreft som forstadier til kreft (DCIS; også kalt Stadium 0) eller i stadium I-II (tidlige sykdomsstadier).

⁹ Kreft i Norge 1999. Kreftregisteret 2002.

¹⁰ "TNM Classification of Malignant Tumours", 5th Edition, 2nd Revision, Springer Verlag 1998. Nasjonale tilpasninger kan finnes

3.1 Antall og insidensrater i Norge

3.1.1 Antall og insidens relatert til alder

Tabellene viser antall brystkrefttilfeller i Norge og antall brystkrefttilfeller per 100.000 kvinner for aldersgruppene 0-49 år, 50-69 år, 70-80 år og over 80 år¹¹.

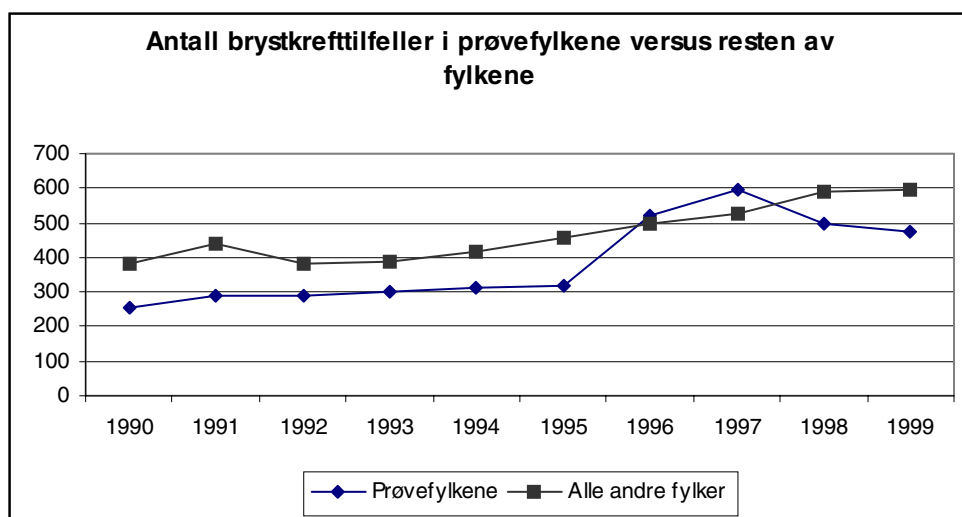
Antall brystkrefttilfeller blant kvinner i Norge

	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999
0-49	423	464	476	522	481	511	515	518	479	477
50-69	633	732	674	690	727	775	1015	1119	1088	1070
70-80	523	516	532	503	519	520	518	487	522	504
81+	279	274	289	243	284	306	300	288	304	306
Sum:	1858	1986	1971	1958	2011	2112	2348	2412	2393	2357

Antall brystkrefttilfeller per 100 000 kvinner i Norge (insidensrater)

	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999
0-49	29.2	31.7	32.3	35.2	32.4	34.3	34.6	34.7	32.0	31.7
50-69	156.7	183.5	169.6	173.3	180.5	189.1	241.8	260.4	248.1	239.1
70-80	259.0	252.3	257.8	242.5	249.6	251.1	250.8	238.4	259.0	253.8
81+	302.8	290.2	298.4	245.8	281.6	297.3	284.7	267.9	275.7	272.7
Alle	86.5	91.9	90.7	89.6	91.5	95.6	105.7	108.0	106.6	104.2
Justert rate ¹²	53.0	58.3	56.6	57.9	58.0	60.6	68.9	72.0	69.3	67.5

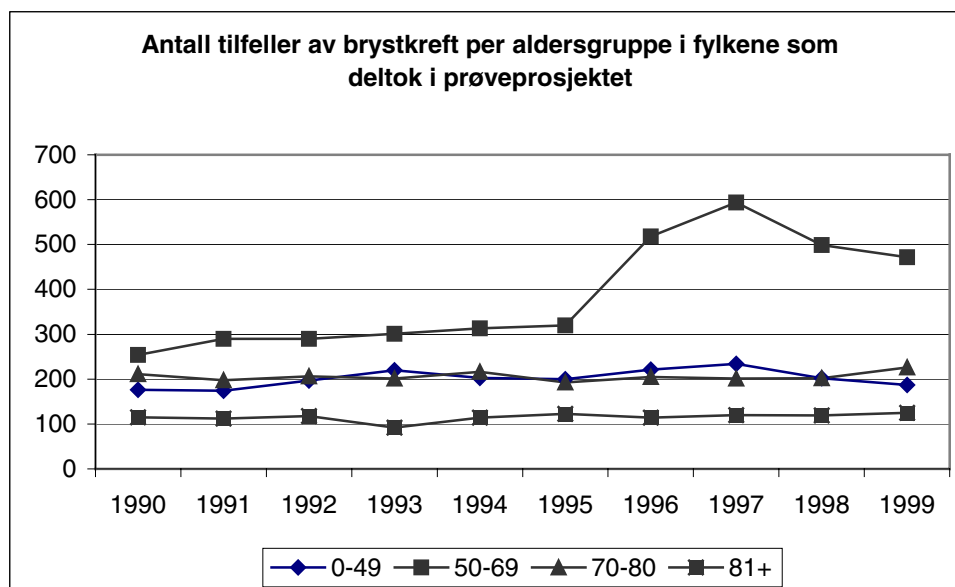
Figuren nedenfor viser endring i totalt antall brystkrefttilfeller fra 1990 til 1999 i aldersgruppen 50-69 år for prøvefylkene versus resten av fylkene. Prøveprosjektet med mammografiscreening i 4 fylker i Norge startet i 1995/96.



¹¹ Data oversendt fra S. Thoresen, Kreftregisteret mai 2002.

¹² Aldersjustert etter verdens standard befolkning per 100 000.

Under vises forekomst av brystkrefttilfeller før og etter at mammografiscreeningen startet i prøvefylkene for aldersgruppene 0-49 år, 50-69 år, 70-80 år og over 80 år.



3.1.2 Livstidsrisiko for å dø av brystkreft

En annen måte å fremstille brystkreftproblemet på er å angi risikoen for å utvikle brystkreft for kvinner med en gitt alder og med en antatt forventet livslengde. Livstidsrisikoen kan ta utgangspunktet i risiko ved fødsel, eller ved en gitt alder.

Tabellen viser ett eksempel på hvordan slik livstidsrisiko kan fremstilles basert på amerikanske data tatt fra informasjonssidene til National Cancer Institute, USA (PDQ).

Table 1: Probability of developing invasive breast cancer among women who are free from invasive breast cancer at their current age, shown as risk over specific intervals of time (based on an analysis of data from the Surveillance, Epidemiology, and End Results (SEER) registry for 1987-1988) *

Current age (in years)	Risk interval (in years)				
	10	20	30	40	Eventually
30	0.4%	2.0%	4.3%	7.5%	12.8% (1 in 8)
40	1.6%	3.9%	7.1%	10.3%	12.5% (1 in 8)
50	2.4%	5.7%	9.0%	10.8%	11.3% (1 in 9)
60	3.6%	7.1%	9.1%		9.6% (1 in 10)
70	4.1%	6.5%			7.1% (1 in 14)

* Tilsvarende tall fra Norge kan beregnes på bakgrunn av data fra Kreftregisteret (ikke prioritert i denne utredningen).

3.2 Status for mammografiscreening i Norge

Fra januar 1996 startet man i Norge organisert screening i Akershus, Hordaland, Oslo og Rogaland. Kvinner i alderen 50-69 år ble invitert til undersøkelse hvert annet år.

Det er vedtatt landsdekkende utbygning av mammografiscreening for kvinner i aldersgruppen 50-69 så snart personellsituasjonen tillater det (Stortingsvedtak, juni 1998). Per 2002 er det innført mammografiscreening i alle norske fylker med unntak av Vestfold, Hedmark og Sogn og Fjordane som starter i 2003.

I Norge er mammografiscreening (mammografiprogrammet) for kvinner i aldersgruppen 50-69 år organisert som et prosjekt som har ansvar for organisering (innkalling og praktisk gjennomføring), innhenting og sammenstilling av data og kvalitetskontroll. Prosjektledelsen er i Kreftregisteret. Prosjektet representerer en betydelig informasjonskilde for de som ønsker innsikt i konsekvensene av organisert mammografiscreening på populasjonsnivå.

3.3 Status i Europa

Tabellen¹³ gir en oversikt over de største *organiserte* screeningprogrammene i 2000. Mammografiscreening foretas på nasjonalt plan i 7 land, og med regionale programmer (dvs. ikke landsdekkende) i 8 (9)¹⁴ land.

Sted	Startår (2)	Inviteret aldersgruppe	Screenings interval 50+ år
Nasjonalt program			
Sverige (1)	1986	40/50-64/74	2
Finland	1986	50-59	2
Island	1987	40-69	2
UK	1988	50-64	3
Holland	1989	50-69	2
Luxembourg	1992	50-65	2
Israel	1997	50-74	2
Regionale programmer			
Canada	1988	50-69	2
Frankrig	1989	50-65/69	2-3
Italien	1990	50-69	2
Portugal	1990	40 ingen grænse	2
Spanien	1990	45-64	2
Australien	1991	40-69	2
Danmark	1991	50-69	2
Norge	1995	50-69	2

(1) organisert etter region; (2) år for første regionale program

Det pågår pilotprosjekter i Tyskland, Hellas, Ungarn, Japan og Uruguay¹⁵. USA og Belgia anbefaler screening, men har ikke innført organiserte screeningprogram.

¹³ Tabellen (modifisert) er tatt fra "Kvalitet af service-screening med mammografi". For fullstendig referanse, se kapittel 5.

¹⁴ Sveits har også startet innføring av mammografiscreening. Per 2002 er screening innført i 3 kantoner, med en planlagt trinnvis utvidelse til 26 kantoner innen 2007 (opplysninger på ISTAHC-konferansen, Berlin juni 2002).

¹⁵ Se Int J Technol Assess Health Care 2001;528-41

4. Om screening

4.1 Om screening generelt

Den engelske National Screening Committee¹⁶ som er et rådgivende utvalg for myndighetene i England, har definert screening slik:

”Screening is a public health service in which members of a defined population, who do not necessarily perceive they are at risk of, or are already affected by a disease or its complications, are asked a question or offered a test, to identify those individuals who are more likely to be helped than harmed by further tests or treatment to reduce the risk of a disease or its complications”.

Kriteriene eller forutsetningene for å starte opp et screeningtiltak i en befolkning er formulert i WHO Report (1968), kjent under betegnelsen ”Wilson & Jugner-kriteriene”. National Screening Committee har revidert disse.

Kort oppsummert er de viktigste kriteriene:

1. Sykdomsforløpet må være kjent og utvikles over tid fra en latent fase til symptomer opptrer og sykdommen manifesterer seg.
2. Det må finnes en enkel, presis og valid screeningstest som er akseptabel for befolkningen.
3. Det må finnes en effektiv behandling eller intervensjon, og tidlig behandling må føre til bedre helse for pasienten.
4. Det må foreligge en plan for organisering og gjennomføring, og screeningprogrammet må kvalitetssikres etter gitte kriterier.

Stortinget har indirekte sluttet seg til kriteriene for masseundersøkelser (screening) ved behandling av Stortingsmelding 26 (1999-2000)¹⁷ *Om verdier for den norske helsetenesta*. Stortingsmeldingene beskriver bl.a. de 10 kriteriene som må oppfylles før en setter i gang screeningsopplegg.

National Screening Committee har også publisert anbefalinger om hvordan screeningprogram bør gjennomføres og rapporteres for å sikre en best mulig objektiv vurdering. Den samme komiteen har videre publisert retningslinjer for hvordan et screeningprogram skal gjennomføres og vurderes (eng: assess)¹⁸.

Det etiske råd i Danmark har på oppdrag av helseministeren publisert en rapport i 2001 om de etiske problemene ved screening der det også diskuteres forutsetninger for screening av befolkningen (”Screening – a report”¹⁹).

¹⁶ www.doh.gov.uk/nsc/index.htm

¹⁷ <http://odin.dep.no/hd/norsk/publ/stmeld/030005-044001/index-ind001-b-n-a.html>

¹⁸ NHSBSP Publication No 49, April 2001. Clinical Guidelines for Breast Cancer Screening Assessment. Tilgjengelig på www.cancerscreening.nhs.uk.

¹⁹ www.etiskraad.dk

4.2 Forskningsmetodologiske spørsmål

Diskusjonene om hvor nyttig eller skadelig mammografiscreening er reiser en rekke forskningsmetodologiske problemer om mammografi som diagnostisk metode, design av studier (inkl. randomiseringsmåte) og til analyse av resultater, inkludert hvilke regler som er etablert mht. å følge opp kvinner med et positivt mammogram. Forskningsmetodologiske problemstillinger reises bl.a. i den informasjonsdatabase som National Cancer Institute (USA) har på sine hjemmesider om mammografiscreening.

Gjennomføring av et screeningtiltak reiser spørsmål om ulike former for bias²⁰. NCI-PDQ omtaler slike problemstillinger når det gjelder mammografiscreening slik:

”Numerous uncontrolled trials and retrospective series have documented the capacity of mammography to diagnose small, early stage breast cancers, including those that have a favorable clinical course.¹⁷ These trials also show that cancer-related survival is better in screened compared to nonscreened women. These comparisons are susceptible, however, to a number of important biases: 1) Lead-time bias: Survival time for a cancer found mammographically includes the time between detection and when it would have become detected because of clinical symptoms, but this time is not included in the survival time of those found because of symptoms; 2) Length bias: Mammography detects a cancer while it is preclinical and preclinical durations vary; cancers with longer preclinical durations are more likely to be detected by screening and these cancers tend to be slow growing and to have good prognoses; 3) Overdiagnosis bias, which is an extreme form of length bias: Screening may find cancers that are very slow growing and that would never have become manifest clinically; and, 4) Healthy volunteer bias: The screened population may be healthier or more health conscious than the general population. Since the extent of these biases is never clear in a particular study, one must rely on randomized controlled trials to assess the benefits of screening. (Refer to the Effect of Screening on Breast Cancer Mortality section.)”.

Ytterligere et problemområde er om resultatene fra eksperimentelle studier (som en randomisert studie er) kan overføres til og reproduseres i rutinevirksomhet (screeningprogram). Randomiserte kliniske forsøk vil ofte være gjennomført under mer optimale betingelser.

SMM har ingen muligheter innen den tidsfristen som er gitt å vurdere alle forskningsmetodologiske spørsmål (design, analyser) som reises i NCI-PDQ-rapporten, eller i ulike vitenskapelige publikasjoner, andre rapporter, kommentarer, etc. SMM vil i denne rapporten imidlertid trekke frem de relevante problemstillingene der det er naturlig.

4.3 Hvordan er opplegget for mammografiscreening?

Hvordan det praktiske rundt screening blir organisert og gjennomført vil ikke bli diskutert. Derimot er det relevant for forståelsen å fremstille mulige forløp for en kvinne som blir invitert til å delta (kalt flytdiagram).

²⁰ Bias = systematiske feil. Bias representerer et grunnleggende problem i tolkningen av forskningsdata. Faren for systematiske feil rokker ved sannhetsverdien av forskningsfunn.

Figuren²¹ viser mulige forløp i en forenklet form.



Figuren beskriver de rent tekniske sidene ved gjennomføring av mammografiscreening. Screening har imidlertid også opplevelse- og følelsesmessige dimensjoner for kvinnene. Ett eksempel kan være en kvinne som gjennomgår mammografi med negativt resultat (mammografi negativ) og som opplever dette som en forsikring om at hun er “frisk”, men som før neste screeningrunde får diagnostisert kreft (intervallcancer).

Kvinner som gjennomfører screening vil iht. denne figuren ha 4 mulige utfall (resultat) etter mammografiundersøkelsen:

- Mammografiundersøkelsen viser noe galt, og kvinnen viser seg etter videre undersøkelse å ha kreft (*sann positiv undersøkelse*);
- Mammografiundersøkelsen viser noe galt, men kvinnen viser seg etter utredning ikke å ha kreft (*falsk positiv undersøkelse*);
- Mammografiundersøkelsen er normal, men kvinnen viser seg likevel å ha kreft (*falsk negativ undersøkelse*);
- Mammografiundersøkelsen er normal, og kvinnen er frisk (*sann negativ undersøkelse*).

Se ytterligere diskusjon kap. 6 og 8.

²¹ Tatt fra CEMTV (Lyng og Olsen) 2002.

5. Det vitenskapelige grunnlaget for screening for brystkreft ("best evidence")

Som grunnlag for denne utredningen, vil SMM bruke *systematiske litteraturoversikter* (sekundær-litteratur, inkl. metaanalyser). Dette er kunnskapsoppsummeringer basert på en systematisk litteraturidentifikasjon og kritisk vurdering av enkeltstudier, i dette tilfelle de randomiserte kliniske studiene som er utført, og med sammenstilling av resultatene i form av kvalitative eller kvantitative metaanalyser.

Bakgrunnen for en slik tilnærming finnes i generelt aksepterte normer for vurdering av vitenskapelige arbeiders utsagnskraft ("level and strength of evidence") der metaanalyser og andre systematiske litteraturoversikter, spesielt av randomiserte kliniske forsøk, rangeres høyest²².

Et ytterligere argument for denne tilnærmingen er å finne i den kjennsgjerning at metaanalysene publisert av Gøtzsche & Olsen (ref. nr 8; se senere) og Cochrane Collaboration (ref. nr 10; se senere) har ført til debatt. Det beste sammenligningsgrunnlag for disse to publikasjonene er derfor andre metaanalyser, og ikke data fra individuelle studier.

Dette innebærer at vi i denne vurderingen ikke legger til grunn enkeltstudier (primærstudier, primærlitteratur) for å vurdere effekten av mammografiscreening på dødelighet. Det vil i vurderingen henvises til enkeltstudier i noen tilfeller, men det er for å belyse enkelte problemområder, og ikke for å studere effekter.

Søkene etter vitenskapelig litteratur f.o.m. 1995 (sekundærlitteratur) er foretatt i databasene:

- Medline
- Embase
- Cinahl
- PsycInfo
- Cochrane Library
- Database of Abstracts of Reviews of Effectiveness (DARE)
- International Network of Agencies for Health Technology Assessment (INAHTA) databasen
- National Guideline Clearinghouse.

Ved generelt søk ble disse søkeordene brukt:

Mammography screening AND breast neoplasms AND meta-analysis OR practice guidelines OR systematic review

Ved søk om intervall cancer:

Mammography AND mass screening AND breast neoplasms AND interval (ti) AND cancer (ti).

²² SMM bruker graderingssystemet utarbeidet av SIGN (Scottish Intercollegiate Guidelines Network)

Ved søk om insidens:

Mammography AND mass screening AND breast neoplasms (incidence or epidemiology(sh)) OR risk and cohort studies OR review OR randomized controlled trials.

Litteratur om helseøkonomi (nytte-kostnader) er identifisert i:

Sekundær litteratur:

Medline

Søkeord:

1. *Mammography screening AND systematic review AND economics*
2. *Mammography screening AND meta analysis AND economics*

OHE Health Economic Evaluation Database og NHS Economic Evaluation Database.

Søkeord:

1. *Mammography screening AND systematic review*
2. *Mammography screening AND meta analysis*

Primær litteratur:

Medline, OHE Health Economic Evaluation Database og NHS Economic Evaluation

Søkeord:

Cost-effectiveness AND mammography, avgrenset til norske studier (se pkt. 5.3).

Andre kilder:

WHO: Status i andre sammenlignbare land

Regional Office for Europe

fap@who.dk

Den nasjonale forskningsetiske komité for medisin (NEM)

www.etikkom.no/index.htm

National Screening Committee Program UK

www.cancerscreening.nhs.uk/breastscreen/index.html

The Internationale Agency for Research on Cancer (IARC)

www.iarc.fr

National Guideline Clearinghouse (NGC)

www.guideline.gov/index.asp

National Cancer Institute (NCI):

<http://www.nci.nih.gov/>

For tema (eller spørsmål som reises i bestillingen fra SHDir) der SMM ikke har kunnet finne systematiske litteraturoversikter, vil det ikke refereres funn eller resultater (unntaket er norske publikasjoner om helseøkonomi; se senere).

SMM har ikke søkt etter sekundærlitteratur (systematiske litteraturoversikter) som gir informasjon om kvinners reaksjon på organisert mammografiscreening pga. begrenset tidsramme for utredningen.

Når det gjelder spørsmålet om hvilke *effekter (nytte)* mammografiscreening har, vil SMM i denne rapporten legge til grunn resultatene av systematiske litteraturoversikter (SR) (Cochrane oversikter, andre systematiske litteraturoversikter (SR) inkl. formelle metaanalyser, metodevurderinger utført av SMMs søsterorganisasjoner²³ og behandlingsretningslinjer (guidelines) utformet etter prinsippene i SR.

Når det gjelder spørsmål om *hvilke mulige negative konsekvenser eller andre aspekter* ved bruk av mammografi (deteksjonsrater, behov for nye undersøkelser, fremmøte, etc) vil data fra land som har innført mammografiscreening som et organisert helsetiltak på populasjonsnivå (fylke, region, land) og som har oversikt over og kvalitetskontroll med denne, bli brukt. De populasjonsbaserte screeningprogrammene er her svært viktige og vil, med enhetlig og god rapportering, kunne gi ytterligere kunnskapsgrunnlag i fremtiden. SMM vil her peke på at en hittil har hatt en begrenset omtale og dokumentasjon av negative konsekvenser ut over 1. screeningrunde.

Det finnes internasjonale standarder for hvordan metaanalyser av randomiserte kliniske studier bør rapporteres (QUORUM statement)²⁴. Mange av metaanalysene om mammografiscreening er publisert før 1999, og SMM har derfor ikke vurdert analysene opp mot QUORUM-retningslinjene. Det er også publisert slike standarder i Norge²⁵. Likedan har INAHTA publisert en sjekkliste for metodevurderingsrapporter²⁶.

En alternativ tilnærming kunne ha vært å kritisk vurdere den seneste oppdateringen av de 7 (8) randomiserte kliniske forsøkene som er foretatt, og sammenstille resultatene i en ny metaanalyse. Dette er en svært arbeidskrevende prosess, og ville ikke ha vært mulig å gjennomføre i løpet av den tiden som er stilt til disposisjon.

5.1 Systematiske litteraturoversikter – kliniske effekter

Fra det systematiske litteratursøket i databasene Medline, Cinahl, Embase og PsycInfo ble 136 artikler identifisert, hvorav 11 ble inkludert og danner sammen med noen kliniske retningslinjer (guidelines; se pkt. 5.2) grunnlaget for vurdering av positive helseeffekter ved screening. Utvalget av arbeidene baseres på at det er metaanalyser

²³ Alle medlemmer i organisasjonen INAHTA = International Network of Agencies of Health Technology Assessment.

²⁴ Lancet 1999;354:1896-1900 ("Improving the quality of reports of meta-analyses of randomised controlled trials: the QUORUM statement").

²⁵ A. Bjørndal, S Flottorp og A. Klovning. Medisinsk kunnskapshåndtering. Gyldendal akademisk 2000. (Kapittel 9)

²⁶ INAHTA HTA checklist

eller systematiske litteraturoversikter (sekundærlitteratur), at de gir opplysninger om populasjon (kvinner som screenes vs. kontrollgruppe), om intervensjon (screening) og om helseeffekt (endepunkt).

1. Kerlikowske K, Grady D, Rubin SM, Sandrock C, Ernster VL. Efficacy of Screening Mammography. A Meta-analysis. JAMA 1995; 273(2):149-154.
2. Glasziou PP, Woodward AJ, Mahon CM. Mammographic screening trials for women aged under 50. A quality assessment and meta-analysis. Med J Aust 1995; 162:625-629.
3. Smart CR, Hendrick RE, Rutledge JH 3rd, Smith RA. Benefit of mammography screening in women ages 40 to 49 years. Current evidence from randomized controlled trials. Cancer 1995;75:1619-26
4. Cox B. Variation in the effectiveness of breast screening by year of follow-up. J Natl Cancer Inst Mongr 1997;22:69-72.
5. Hendrick RE, Smith RA, Rutledge JH III, Smart CR. Benefit of Screening Mammography in Women Aged 40-49: A New Meta-Analysis of Randomized Controlled Trials. J Natl Cancer Inst Mongr 1997; 22:87-92.
6. Kerlikowske K. Efficacy of Screening Mammography Among Women Aged 40 to 49 Years and 50 to 69 Years: Comparison of Relative and Absolute Benefit. J Natl Cancer Inst Mongr 1997; 22:79-86.
7. Larsson LG, Andersson I, Burstam N et al. Updated overview of the Swedish randomized trials on breast cancer screening with mammography: age group 40-49 at randomization. J Natl Cancer Inst Mongr 1997; 22:57-61.
8. Gøtzsche PC, Olsen O. Is screening for breast cancer with mammography justifiable? Lancet 2000;355:129-134.
9. Ringash J, with the Canadian Task Force on Preventive Health Care. Preventive health care, 2001 update: screening mammography among women aged 40-49 years at average risk of breast cancer. CMAJ 2001;164:469-476.
10. Olsen O, Gøtzsche PC. Screening for breast cancer with mammography (Cochrane review). The Cochrane Library, Issue 2, 2002. Oxford: Update Software.
11. Nystrom L, Andersson I, Bjurstam N, Frisell J, Nordenskjöld B, Rutqvist LE. Long-term effects og mammography screening: updated overview of the Swedish randomised trials. Lancet 2002;359(9310):909-19. Denne tas med fordi det i prinsippet er en metaanalyse og oppdatering av de svenske randomiserte kliniske studiene (unntatt fra Kopparberg).

Metodevurderinger (HTA-vurderinger)

Dette er utredninger utført av nasjonale / regionale organisasjoner med oppgave å utføre metodevurderinger. Utredningene samles / publiseres i National Health Service Centre for Reviews and Dissemination, University of York.

12. Mammography screening for early detection of breast cancer –systematic review. Koln: German Agency of Health Technology Assessment at German Institute for Medical Documentation and Information, 1998. Language: German
13. The early detection and diagnosis of breast cancer: a literature review – an update. New Zealand Health Technology Assessment. Christchurch: New Zealand Health Technology Assessment, 1999. (NZHTA Report Volume 2 number 2) ISBN: 1877235091

14. Dalton SO, Olsen J. Oversikt over ikke-randomiserte studier af mammografiscreeningprogrammer. Medisinsk Teknologivurdering 2002;4(1). CEMTV: Center for evaluering og Medisinsk Teknologivurdering. www.cemtv.dk.
15. Mammography screening: mortality rate reduction and screening interval. Harstall C. Edmonton: Alberta Heritage Foundation for Medical Research (AHFMR), 2000. (HTA 21) ISBN: 189695629716.
16. Diagnosis of mammary lesions detected through mammography screening. Gonzalez-Enriquez J, Amate-Blanco J M, Castellote-Olivito J. Madrid: Agencia de Evaluacion de Tecnologias Sanitarias (AETS), 1999. ISBN: 8493037990. Language: Spanish. (Ikke relevant iht. mandat og ikke vurdert p.g.a. språk)
17. Approach to the utilization of screening mammography in two health regions. Almazan C, Borrás JM, de la Puente ML, Gimenez A, Granados A. Barcelona: Catalan Agency for Health Technology Assessment, 1995. Language: Catalan and Spanish (Ikke relevant iht. mandat og ikke vurdert p.g.a. språk)

5.2 Behandlingsretningslinjer (eng. guidelines)

Det er her referert til behandlingsretningslinjer som er samlet og omtalt på hjemmesidene til National Guideline Clearinghouse.

Det er alment akseptert at retningslinjer bør utarbeides på bakgrunn av en systematisk litteraturvurdering. Det krever en systematikk i arbeidet som må dokumenteres i retningslinjene. Retningslinjer som er utarbeidet på denne måten vil bli brukt i denne utredningen. Andre retningslinjer der grunnlaget for utforming ikke er entydig gitt, vil bli oppgitt, men ikke brukt.

Utarbeidet etter prinsippene for systematiske litteraturvurderinger

18. Preventive health care, 2001 update: Screening mammography among women aged 40-49 years at average risk of breast cancer. Canadian Task Force on Preventive Health Care. 2001.
19. Screening for breast cancer: recommendations and rationale. United States Preventive Services Task Force. 2002.

Andre guidelines – grunnlaget for utforming ikke nærmere spesifisert²⁷

20. Breast cancer in women. A national clinical guideline. Scottish Intercollegiate Guidelines Network. 1998 Oct. 64 pages.
21. Early detection of breast cancer. Royal New Zealand College of General Practitioners. 1999. 61 pages.
22. Breast cancer screening for women ages 40-49. National Institutes of Health (NIH). Consensus Development Panel of Breast Cancer Screening for Women Ages 40-49. 1997 Jan. 36 pages.
23. American College of Radiology guideline for breast cancer screening. American College of Radiology. 1997. 5 pages.
24. ACS guidelines for the early detection of breast cancer: update 1997. American Cancer Society. 1997. 4 pages.

²⁷ Dette innebærer at resultater fra disse guidelines ikke brukes i denne rapporten

25. Recommended breast cancer surveillance guidelines. American Society of Clinical Oncology. 1997 May (updated 1998 Nov 13). 8 pages (1997 guideline); 3 pages (1998 update).
26. Recommendations regarding selected conditions affecting women's health. Centers for Disease Control and Prevention. 2000 Mar 31. 82 pages.
27. Diagnosis of breast disease. Institute for Clinical Systems Improvement. 1994 Jan (revised 2001 Nov). 42 pages.
28. Breast cancer screening. Kaiser Permanente-Southern California. 1993 (revised 2001 Apr). 4 pages.

5.3 Nytte-kostnads analyser

Systematiske oversikter

Søkene ga tilsammen 12 treff. Av disse ble 11 ekskludert fordi de ikke var systematiske oversikter, fordi de var studier skrevet før 1995, fordi de ikke hadde relevans i forhold til mandatet eller fordi noen artikler inngikk i flere databaser.

29. Barrat, A.L., Les M Irwig, Paul P Glasziou, Glenn P Salkeld and Nehmat Houssami. Benefits, harms and costs of screening mammography in women 70 years and over: a systematic review. Med J Australia 2002 Mar 18; 176(6):266-71.

Norske primærstudier

Søkene ga 3 treff. Dessuten kjente vi til en upublisert rapport (Ekerhovd).

30. Ekerhovd, K.M. Mammografiscreening i Norge-en kostnads-/effektivitetsanalyse. Hovedoppgave i sosialøkonomi, Universitetet i Oslo, 1996.
31. Norum, J. Breast cancer screening by mammography in Norway. Is it cost-effective? Annals of Oncology 1999; 10: 197-203.
32. Wang, H., Rolf Kåresen, Arild Hervik og Steinar Østerbø Thoresen. Mammography screening in Norway: Results from the first screening round in four counties and cost-effectiveness of a modeled nationwide screening. Cancer Causes and Control 2001; 12: 39-45.
33. Kåresen, R, Bø, JK, Haustveit, S, Hervik, A, Thoresen, S. Samfunnsøkonomisk lønnsomhet av mammografiscreening i Norge. Tidsskr Nor Lægeforen 1999;119:3553-9.

5.4 Annen spesielt relevant dokumentasjon

Dette er publikasjoner som ikke entydig lar seg ordne i de ovenfornevnte grupper, men som er brukt i denne utredningen. Her vil SMM spesielt peke på arbeid 37 fra Holland som gir vurderinger om endringer (reduksjon) i dødelighet uten at arbeidet isolert sett fyller kravene til en systematisk litteraturoversikt. Data fra Holland vil bli referert og brukt.

34. Mammografiprogrammet i Norge. Evaluering av prøveprosjektet 1996-2000. Kreftregisterets forskningsrapport nr. 2-2000.
35. Wang H. Epidemiological studies of breast cancer in Norway – with focus on implementation of organized mammography screening. Doktorgradsavhandling. Universitetet i Oslo 2002.

36. Lyng E, Olsen AH. Kvalitet af service-screening med mammografi. Medisinsk Teknologivurdering 2002; CEMTV: Center for evaluering og Medisinsk Teknologivurdering. www.cemtv.dk.
37. The benefit of population screening for breast cancer with mammograph. Gesundheitsraad (GR), Health Council of the Netherlands. 2002. En kortversjon er publisert som korrespondanseinnlegg i Lancet 2002;359 (11 May).

I det etterfølgende vil det bli henvist til den vitenskapelige litteraturen listet opp her med nummer.

I vedlegg 5 gis en oversikt over artikler og lignende som kan gi utfyllende informasjon. I de tilfellene der disse referansene brukes i denne utredningen vil det fremgå av fotnoter.

6. Hvordan skal nytteverdien beskrives - hvilke mål på effekter legges til grunn

Nytten eller effekten av bruk av medisinske metoder eller tiltak (her mammografiscreening) kan måles eller uttrykkes på ulike måter. *I denne rapporten brukes mål på effekt (endepunkter) som har direkte betydning for pasientene ved at det gir en målbar helsegevinst.*

Alle tiltak som brukes eller iverksettes i medisinen vil ha en *positiv* effekt (dvs. nytte), men samtidig også en *negativ* (skade eller utilsiktet virkning). Om et tiltak skal anbefales eller brukes vil derfor avhenge av om tiltaket gir *mer nytte enn skade* ("doing more good than harm").

Effekten av et medisinsk tiltak kan måles ved bruk av endepunkter som man antar uttrykker noe om effekt (intermediære eller surrogate endepunkter). Reduksjon i tumor diameter i invasive cancertilfeller som oppdages ved mammografi ift. diameter før mammografi eller endringer (reduksjon) over tid, er f.eks. et slikt mål som ofte brukes. Andre slike mål er andelen (%) DCIS av invasive cancertilfeller pluss DCIS, tilfeller av invasiv cancer påvist ved screening med spredning til de axillære lymfeknuter, malignitetsgraden ved invasive cancertilfeller påvist ved screening, og endring i stadiefordeling mot mindre avanserte kreftformer.

I denne rapporten vil intermediære eller surrogate endepunkt ikke tillegges vekt. Selv om det er generell enighet om at reduksjon i tumorstørrelse (diameter eller forskyvning mot tidligere stadier av sykdommen) er en forutsetning for bedre prognose og leveutsikter hos pasientene, er det bare i randomiserte kliniske undersøkelser at effekten av screening i form av redusert dødelighet kan fastslås.

Screening vil heller ikke hjelpe alle med brystkreft. F.eks. vil kvinnene som på screeningtidspunktet har sykdom i stadium 3 og 4 ikke ha tilsvarende nytte av screening fordi sykdommen er generalisert på diagnosetidspunktet.

Effekten av mammografiscreening kan også avhengige av hvilke aldersgrupper som inviteres for screening, screeningintervall, hvordan mammografiundersøkelsen utføres (antall "views"), og kvaliteten på arbeidet fra granskeren av bildet / bildene. Slike spørsmål vil ikke bli diskutert her, men diskuteres nå også i media²⁸.

I denne rapporten legges følgende mål (endepunkter) på effekt (negativ og positiv) til grunn:

6.1 Antatt positive konsekvenser

Redusert dødelighet

Dødeligheten i en gruppe som deltar i mammografiscreening kan i denne sammenheng beskrives som:

²⁸ M. Moss. Spotting Breast Cancer: Doctors are Weak Link. New York Times June 27, 2002.

- død av brystkreft (brystkreftspesifikk dødelighet)
- endring i livstidsrisiko for å dø av brystkreft
- total kreftdødelighet (alle kreftformer)
- alle dødsfall uansett årsak (totaldødelighet, "over all mortality")

Endringer i dødeligheten kvantifiseres, enten i *absolutte* tall (hvor mange som faktisk døde i screeninggruppe versus kontrollgruppe), som *faktiske (absolutte) endringer* i dødelighet, eller som en *relativ* størrelse (hvor store er endringene ift. et utgangspunkt vanligvis uttrykt i prosent). Dødelighetsbegrepet diskuteres videre i pkt. 8.3.

Mer skånsom behandling

Det argumenteres med at mammografiscreening fører til mer skånsom kirurgisk behandling. Dette tema er komplisert og omfattende og vil ikke bli vurdert i denne utredningen.

6.2 Antatte negative konsekvenser

Nye undersøkelser / intervensjoner (spes. ift. antall screeningrunder).

Det er en forutsetning for mammografiscreening at noen av kvinnene som undersøkes med mammografi, har funn ved undersøkelsen som leder til videre undersøkelser. I mammografiscreening-programmene rapporteres derfor om:

- gjeninnkallelse til ny vurdering eller videre undersøkelse (diagnostiske prosedyrer)²⁹
- gjenninnkallelse til behandling (kirurgi³⁰).

Dette er et trinnvis system og ingen programmer opererer med direkte innkalling til kirurgi uten forutgående vurdering.

"Falskt positive" omtales som andelen av de screenede som gjeninnkalles til ny vurdering eller behandling med mistanke om at de har brystkreft, men som likevel viser seg å være friske.

Med "falskt negative" menes brystkrefttilfeller som ikke oppdages ved mammografi. Falskt negative inngår alltid i intervallcancere (se 6.3).

"Overdiagnostikk" og "overbehandling".

Med *overdiagnostikk* menes her identifisering av krefttilfeller (DCIS og sakte-voksende infiltrerende brystkreft) som ikke vil føre til klinisk sykdom. Dette er så langt en teoretisk størrelse fordi vi med dagens kunnskap ikke vet hvilke saktevoksende krefttilfeller som vil utvikle seg til symptomgivende cancer, og hvilke som ikke vil føre til klinisk sykdom.

Overbehandling dekker dermed behandling av krefttilfeller som ikke utviklet seg til symptomgivende cancer eller behandling som ikke står i forhold til sykdommens

²⁹ Innebærer vanligvis palpasjon (undersøkelse av brystet), "three view" mammografi, evt. ultralyd og finåls aspirasjons cytologisk undersøkelse (FNAC) "og merkebiopsier".

³⁰ Omfatter exsisijonsbiopsi, brystbevarende tumorectomi eller mastectomi (fjerning av hele brystet)

stadium. Det er f.eks. usikkert om alle DCIS trenger både kirurgi og strålebehandling (brukes til dels i USA)³¹.

Påstått overbehandling i form av økt bruk av mastektomi som en konsekvens av screening er et tema som ikke vurderes i denne rapporten. SMM vil heller ikke vurdere om screening isolert sett kan føre til mindre omfattende behandling (gitt at alle brystkrefttilfeller (DCIS inkludert) som oppdages ved screening må behandles).

Begrepet overbehandling endres over tid og må derfor vurderes ift. den kunnskap og behandlingspraksis som er etablert på vurderingstidspunktet. Det henvises til pkt. 8.3 for videre diskusjoner.

6.3 Spesielle problemer

Intervallcancer. Med intervallcancer menes brystkrefttilfeller (både invasiv kreft og DCIS) som oppdages mellom 2 screeningrunder, der den første screeningrunden viste et normalt funn ved mammografiundersøkelsen.

Frekvensen (hyppigheten) av intervallcancer i vitenskapelige studier oppgis ikke bare som antall og rater (%) men også som "*proporsjonell intervall cancer rate*" som er observerte antall intervallcancer dividert med forventet antall tilfeller i gruppen av kvinner som er testet dersom ingen screening hadde funnet sted.

Stråleindusert brystcancer

Eventuelt som konsekvens av repeterte mammografier.

Annen stråleindusert sykdom

Eventuelt som konsekvens av ekstern strålebehandling som en del av behandlingsopplegget.

Mammografi er ioniserende stråling og kan i gjentatte (akumulerte) doser i seg selv muligens indusere kreftutvikling^{32,33}. Det hevdes i Cochrane-rapporten (10) at ekstern strålebehandling som inngår som en del av primærbehandlingen for brystkreft, kan føre til overdødelighet (RR=1.02, 0.96-1.06) bl.a. i "karsykdommer" (eng. cardiovascular disease) beregnet fra de svenske studiene.

SMM vil ikke omtale disse to aspektene nærmere. Begge hendelsene er likevel sjeldne og vil kreve en omfattende utredning, ikke minst ift. de endringer i ekstern strålebehandling som har foregått over tid. SMM har heller ikke hatt mulighet til å vurdere det mulige problemområdet med ny brystkreft i et strålebehandlet bryst.

³¹ I Norge er det utformet veiledninger for behandling som er tilgjengelig på www.kreft.no ("Brystkreft. Diagnostikk og behandling. En veiledning. Norsk Bryst Cancer Gruppe (NBCG). 6. utgave 2000.

³² se f.eks. Feig SA et al.: Radiation risk from screening mammography of women aged 40-49 years. J Natl Cancer Inst Monogr 1997;22:119-24.

³³ "Favourable and unfavourable effects on long-term survival of radiotherapy for early breast cancer". Lancet 2000;355:1757-70.

Livskvalitet

Med livskvalitet menes et uttrykk for en kvinnes livskvalitet, enten målt ved generelle (globale) eller sykdomsspesifikke måleinstrumenter.

Et screeningtiltak kan ha en positiv dimensjon for kvinnene ved at det skapes en trygghetsfølelse ift. et alvorlig helseproblem ved at undersøkelsen ikke viser sykdom eller at en påviser ”sykdom” i et tidlig stadium der sannsynligheten for varig helbredelse ved behandling er stor. Denne trygghetsfølelsen må imidlertid ses ift. intervallvancer-raten. Dersom intervallcancerraten er høy vil tryggheten for kvinnene som gruppe være tvilsom.

Screeningtiltaket kan også medføre en belastning for kvinnen ved uro og engstelse skapt av undersøkelsen og evt. etterfølgende undersøkelser av ”positive” funn.

Både trygghetsfølelsen og belastningen ved mammografi må sees ift. hvilken informasjon kvinnene får om mammografiscreening (nytte versus ”skade”).

Det er også reist spørsmål om slike masseundersøkelser endrer helse- og sykdomsadferd. Slike spørsmål faller utenfor rammen av denne forespørselen og diskuteres derfor ikke.

Endring i insidens

Teoretisk forventes det at første screeningrunde fører til en økning i sykdomshyppighet når man innfører et screeningprogram. Dette fordi en ved første screeningrunde detekterer ikke bare ”usynlig” cancer som kun kan oppdages med screening på det aktuelle tidspunktet, men også brystkrefttilfeller som kan oppdages ved vanlig kliniske undersøkelse (inkl. selvundersøkelse). Cancertilfeller oppdaget i første screeningrunde betegnes ofte *prevalente* krefttilfeller.

Ved senere screeningrunder skal en teoretisk stå igjen med bare ”nye” krefttilfeller (*insidente* krefttilfeller).

Som en konsekvens skal en etter første screeningrunde få en økning i antall brystkrefttilfeller. Deretter skal insidensen teoretisk sett falle til nivået før screening startet (forutsetningen er at >1 screeningrunde gjennomføres).

I diskusjonene om mammografiscreening argumenteres med at en får en vedvarende økning i insidens etter 2. screeningrunde som skyldes at screeningen detekterer kreftsvulster (DCIS spesielt) som ellers ikke ville ha utviklet seg til klinisk sykdom.

Spesifikke analyser fra populasjonsbaserte programmer, f.eks. fra Sverige, Holland, og fra de fylkene i Norge der screening er innført, vil kunne utdype dette spørsmålet. Dette diskuteres ytterligere i kapittel 8.3.

Problemstillingen henger videre nært sammen med spørsmålet om mammografiscreening fører til overdiagnostikk og overbehandling.

6.4 Screeningkvalitet

Med dette menes hvor godt eller dårlig screeningprogrammet fungerer; hvor mange som diagnostiseres, falskt positive, etc. EU har satt standarder for screeningkvalitet. Disse gjengis i egen ramme³⁴.

EU standarder for korttidsindikatorer fra organiserte mammografiscreenings programmer.

Indikator	Første screening		Efterfølgende screening	
	Acceptabel	Ønskelig	Acceptabel	Ønskelig
Deltagelsesrate	60%	>75%	60%	>75%
Yderligere bilder:				
- ved screeningen	<5%	<1%	<3%	<1%
- totalt	NG	NG	NG	NG
Genindkaldelsesrate	<7%	<5%	<5%	<3%
Cytologi/biopsi uegnet resultat (SMM: "mangelfull prøve")	<25%	<15%	<25%	<15%
Total cancer deteksjonsrate	3xIR	3xIR	1.5xIR	NG
- invasive				
- in situ				
Invasive ≤10 mm	NG	NG	NG	25%
Benign åben biopsirate	<0.5%	<0.4%	<0.35%	<0.2%
Benign ift malign biopsirate	<2 to 1	<1 to 1	<1 to 1	<0.5 to 1
NG Ikke angivet				
IR Incidensrate i fravær af screening				

³⁴ Tatt fra Lyng og Olsen (www.cemtv.dk). EU har publisert standardene i: European Commission. Europe against cancer programme. Evaluation and monitoring of screening programmes. Luxembourg 2001 (ISBN 92-894-0253-9).

7. Oppsummering av kunnskaps-status

Det er foretatt 7 randomiserte kliniske studier (RCT) der mammografiscreening som intervensjon er sammenlignet med en kontrollgruppe uten mammografi i en definert populasjon. Førsøkene ble igangsatt i perioden 1963-1982 og er:

- The HIP trial 1963 (Health Insurance Plan of New York trial)
- The Malmö trial 1976
- The Two-county trial 1977: Kopparberg 1977 og Östergötland 1978
- The Edinburgh trial 1978
- The Canadian trial 1980; Canada 1980(50-59 år) og Canada 1980 (40-49 år)
- The Stockholm trial 1981
- The Göteborg trial 1982

I noen vitenskapelige artikler refereres det til 8 RCTer. I disse arbeidene har en da delt Two-county trial iht. det geografiske området forsøket er utført. I litteraturen angis også de randomiserte studiene iht. antall randomiserte populasjoner (som vil være 9).

I denne rapporten legger vi til grunn resultater fra *allerede publiserte systematiske litteraturoppsummeringer (SR) eller meta-analyser (= sekundærlitteratur)*. Også slike studier har varierende kvalitet. I tabellen gis en vurdering og oppsummering av den sekundærlitteratur som brukes i denne rapporten.

SR, meta-analyse, HTA	Basis for studien	Er det foretatt en kritisk vurdering av enkeltstudiene?	Kommentarer
Kerlikowske, (1) 1995	De 7 randomiserte kontrollerte studier (RCT) og 4 kasus-kontroll-studier	Ja. Alle studier evaluert etter HTA-prinsipp før inklusjon.	Alder 40-74; 50-74 og 40-49 år. Meta-analyse for RCT og kasus-kontroll studier separat og samlet.
Glasziou, (2) 1995	Alle 7 RCT	Ja. Det er gjort blindet vurdering av studiene for å lage en kvalitetsscore	Alder 40-49 år. Sensitivitetsanalyse. Laget kumulativ meta-analyse ved addisjon av enkelt-studier
Smart, (3) 1995	Alle 7 RCT	Nei.	Alder 40-49 år. 7-18 års oppfølging.
Cox, (4) 1997	Alle 7 RCT	Studienes forskjeller er fremlagt og diskutert.	Alder 40-49. Laget kumulativ meta-analyse per år over 18 års oppfølging.
Hendrick, (5) 1997	Alle 7 RCT	Oppdatering av meta-analyse av Smart, 1995.	Alder 40-49 år. 12,7 års oppfølging i gjennomsnitt.
Kerlikowske, (6) 1997	Alle 7 RCT	Diskuterer forskjellene. Alle inkluderes for å minimere seleksjons bias.	Alder 40-49 og 50-69 år.
Larsson, (7) 1997	Alle svenske RCT	Nei	Alder 40-49 år.
German Agency of HTA, (12) 1998	7 RCT og 5 meta-analyser	Ja. Alle studier evaluert etter HTA-prinsipp før inklusjon.	Alder 40-74 år Grundig litteraturgjennomgang.

SR, meta-analyse, HTA	Basis for studien	Er det foretatt en kritisk vurdering av enkeltstudiene?	Kommentarer
New Zealand Agency of HTA, (13) 1999	I alt 20 studier med forskjellig studiedesign som dekker forskjellige aspekter ved mammografi screening er inkludert.	Ja. Alle studier evaluert etter HTA-prinsipp før inklusjon.	Alder 40-74 år. Omhandler også risiko faktorer, symptomer, diagnose, mammografi screening og økonomisk evaluering.
Gøtzsche, (8) 2000	Alle 7 RCT	Ja. Metodevurdering av studiene.	Alder 40-74 år
Ringash with Canadian Task Force on Preventive Health Care (9), 2001	Alle 7 RCT	Ja. Alle studier evaluert etter HTA-prinsipp før inklusjon.	Alder 40-49 år. Sammenlignet effektestimater fra tidligere meta-analyser.
Cochrane, (10) 2002	Alle RCT som graderes: 1. Medium kvalitet: Canada, Malmø 2. Lav kvalitet: Two-county, Gøteborg, Stockholm, 3. Dårlig kvalitet: HIP, Edinburgh	Ja, detaljert gjort rede for.	Alder 40-74 år Meta-analyser etter 7 og 13 års oppfølging for brystkreft død, kreft død og total død,
USPSTF 2002 (19)	Alle 7 RCT. Mener en RCT er av dårlig kvalitet.	Vurdert, men ikke gjort rede for.	Alder 40-74 år Vurdert konsekvensene /bias pga. svakhetene i studiene. Gjort sammenlignende metaanalyser av alle aksepterte RCT, ekskludert svakeste RCT og de som gjaldt mammografi-screening alene.
Gezondheidsraad, Nederland (37) 2002	Kun RCT: Canada, Malmø, Two-county, Gøteborg, Stockholm, HIP. Edinburgh utelates fra analysene.	Ja, detaljert gjort rede for.	Diskuterer Cochrane-rapporten grundig
Nystrøm, (11) 2002	De svenske RCT unntatt Kopparberg trial av Two-county trial	Basert på rådata. Metode og oppfølging godt redegjort for	Alder 40-74 år Siste oppdatering mht. dødelighet t.o.m. 1996

7.1 Nytte – redusert dødelighet

Effekt på brystkreftspesifikk dødelighet – absolutte tall

Her beskrives det absolutte, observerte antallet dødsfall av brystkreft i screeninggruppen og i kontrollgruppen. Disse tallene må vanligvis trekkes ut fra primærstudiene, og er ofte ikke spesifikt nevnt.

I den svenske oppfølgingsstudien (11) angis absolutt antall dødsfall. Tabellen oppsummerer tallene for studiene Malmø, Stockholm, Gøteborg, Östergötland (ikke Kopparberg) og for alle aldre:

Oppfølg- ingstid År	Antall døde av brystkreft		Antall i studien		Antall observasjonsår (i 1 000) ³⁵	
	Inviterte	Kontroller	Inviterte	Kontroller	Inviterte	Kontroller
Median 15,8 (5,8-20,2)	511	584	129 750	117 260	1 865	1 688

Alderen for de inviterte til screening og kontrollgruppen varierte, men de fleste kvinnene var i alderen 45-70 år.

Hendrick og medarbeidere (5) har oppsummert absolutt antall døde av brystkreft for **kvinner 40-49 år** slik:

Studie	Antall døde av brystkreft		Antall i studien		Antall observa- sjonsår (i 1 000)		Oppfølg- ingstid
	Inviterte	Kontroller	Inviterte	Kontroller	Inviterte	Kontroller	År
New York /							
HIP studien	49	65	14 432	14 701	248	253	18
Edinburgh	46	52	11 755	10 641	146	135	12,6
Kopparberg	23	18	9 650	5 009	144	75	15,2
Östergötland	27	27	10 240	10 411	143	147	14,2
Malmø	57	78	13 528	12 242	166	144	12,7
Stockholm	24	12	14 185	7 985	174	88	11,4
Gøteborg	18	39	11 724	14 217	138	168	12
Canada	82	72	25 214	25 216	283	283	10,5

³⁵ Antall observasjonsår er summen av alle årene kvinnene er fulgt fra screening og til studien avsluttes eller de får brystkreft.

Fra Cochrane-rapporten (10) er hentet absolutte tall for døde av brystkreft for *aldersgruppen 50 år eller eldre*:

Studie	Antall døde av brystkreft		Antall i studien		Oppfølgingstid
	Inviterte	Kontroller	Inviterte	Kontroller	År
New York / HIP	101	130	16 505	16 505	13
Edinburgh	129	134	17 149	15 748	13
Kopparberg	104	88	29 007	13 551	13
Östergötland	112	150	28 229	26 830	13
Malmö	79	92	17 730	17 426	13
Stockholm	42	33	25 476	12 840	13
Gøteborg	21	37	9 903	15 708	7
Canada	107	105	19 711	19 694	13

Det er de absolutte (faktiske) tallene over brystkreftdødsfall i gruppen av kvinner som screenes versus de som ikke screenes (kontrollgruppen) vist i disse 3 tabellene som er grunnlaget for beregningen av den relative endring i dødelighet som beskrives senere.

Effekt på brystkreftspesifikk dødelighet - relativ endring

I tabellen oppsummeres resultatene for kvinner i **aldersgruppen 50-69 år**. Litteraturen er oppført kronologisk.

Forfatter	Relativ risiko (RR) ³⁶	95% konfidens-intervall ³⁷	Oppfølgingstid
Kerlikowske, 1995 (1)	RCT: 0,77 Kasus.-kontr.: 0,45	0,69-0,87 0,29-0,70	7-12 7-12
Cox, 1997 (4)	0,74	0,62-0,87	7
Kerlikowske, 1997 (6)	0,73	0,67-0,94	7-9
Gøtzsche & Olsen, 2000; 40-74 år (8) Malmø og Canada.			
• Subtotal	1,04	0,84-1,27	
Gøteborg, Edinburgh, Stockholm, Two-County og New York.			
• Subtotal	0,75	0,67-0,83	
Cochrane 2002 (10) <u>Medium kvalitet</u> : Canada 1980 og Malmø			13
• Subtotal	0,94	0,77-1,15	
<u>Lav kvalitet</u> : Gøteborg(kun 7 års oppfølging; ikke med i subtotal).			
Kopparberg Stockholm Östergötland			
• Subtotal	0,64	0,54-0,76	
<u>Dårlig (eng. "flawed")kvalitet</u> : New York Edinburgh Data ikke angitt for aldersgruppen			
USPSTF 2002 (19)	0,77	0,67-0,89	11-20
Gezondheidsraad, Nederland 2002 (37) (Edinburgh ikke inkludert)	0,76	0,67-0,85	13
Nystrøm 2002 (11) Malmø Östergötland Stockholm Göteborg			15,8 (5,8-20,2)
<u>Etter alder totalt</u> :			
50-59	0,84	0,70-1,01	
55-64	0,73	0,60-0,89	
60-69	0,67	0,53-0,84	
40-74	0,79	0,70-0,89	

³⁶ Relativ risiko (RR) angir her forholdet mellom screeninggruppens dødsrate pga brystkreft og kontrollgruppens. Hvis verdien er lik 0,75, er risikoen 25% lavere i screeninggruppen.

³⁷ Hvis konfidensintervallet omfatter 1,00, er ikke gruppene signifikant forskjellige

Gøtzsche & Olsens arbeid (8) og Cochrane-rapporten (10) gir ikke relativ risikoendring for alle studiene samlet (uavhengig av studiekvalitet).

Resultatene for kvinner i **aldersgruppen 40-49 år** er gitt i neste tabell. Åtte av 10 metaanalyser/ systematiske oversikter viser ingen statistisk signifikant effekt av screening i denne aldersgruppen, mens 2 viser en effekt (som er mindre enn for aldersgruppen 50-69 år).

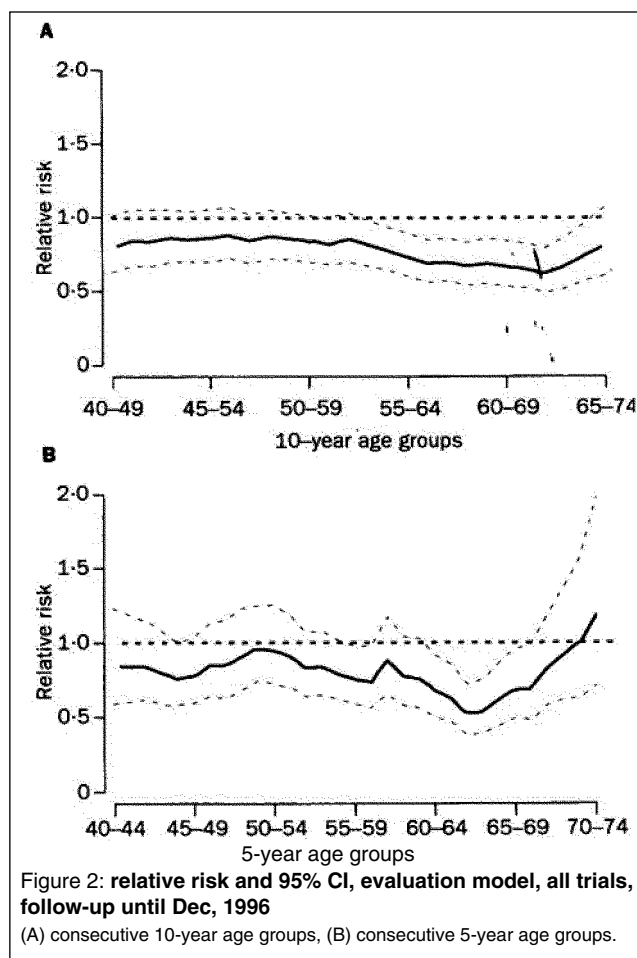
Forfatter	Relativ risiko (RR)	95% konfidens-intervall	Oppfølgingstid
Kerlikowske, 1995 (1)	0,84	0,69-1,02	7-9
Glasziou, 1995 (2)	0,95	0,77-1,18	7-9
Smart, 1995 (3)	0,84	0,69-1,02	10,4
Larsson, 1997 (7)	0,77	0,59-1,01	12,8 (9,7-16,8)
Cox, 1997 (4)	0,93	0,77-1,11	10
Hendrick, 1997 (5)	0,82	0,71-0,95	12,7
Kerlikowske, 1997 (6)	0,84	0,71-0,99	>10-12
Cochrane 2002 (10)			
<u>Medium kvalitet:</u>			13
Canada 1980 og Malmø			
• Subtotal	1,03	0,77-1,38	
<u>Lav kvalitet:</u>			
Gøteborg			
Kopparberg			
Stockholm			
Östergötland			
• Subtotal	0,77	0,57-1,04	
<u>Dårlig kvalitet:</u>			
Data ikke gitt for aldersgruppen.			
USPSTF 2002 (19)	0,83	0,64-1,04	11-20
Nystrøm 2002 (11)			
Malmø I 45-54			15,8 (5,8-20,2)
Malmø II 43-49			
Östergötland 40-49			
Stockholm 40-49			
Gøteborg 40-49			
• Totalt	0,80	0,63-1,01	

Neste tabell gir mer detaljert informasjon fra den hollandske vurderingen (37). Utgangspunktet for vurderingen er Gøtzsche & Olsens Lancet-artikkel (8) og Cochrane-rapporten (10). I tabellen gis også den relative dødelighetsendring for alle kreftsykdommer, samt totaldødelighet. Grunnen til at vi ikke har inkludert tallene her i de 2 foregående tabellene er at det ikke fremgår av rapporten hvordan man er kommet frem til disse resultatene.

Studie	Gj.snitt oppfølgingstid, år	Relativ dødelighet av brystkreft 50 år eller over*	Relativ dødelighet for all cancer. Alle 40-74 år ved screening	Relativ total dødelighet: 50 år eller over
GR rapporten (37): Vurdering av alle RCT	7 år	0,72 (0,61-0,85) Alle RCT unntatt Edinburgh		0,97 (0,93-1,00) Alle svenske RCT
	13 år	0,76 (0,67-0,85) Alle RCT unntatt Edinburgh og Gøteborg	1,00 (0,96-1,05) Alle RCT unntatt Edinburgh og Gøteborg	0,99 (0,97-1,02) Malmø, Canada, Two County

* Canada 40-49 år er per definisjon ekskludert.

For ytterligere å illustrere den relative effekten av screening vises figur fra den svenske oppfølgningsstudien som fremstiller den relative dødelighetsreduksjonen uttrykt som relativ risiko (RR; tykk strek) i ulike aldersgrupper. 95% konfidensintervall er påført (stiplede linjer).



Alle metaanalysene for aldersgruppen 50-69 år viser en relativ risiko (dvs. en relativ dødelighetsreduksjon) som konsistent er <1 og med et øvre (95%) konfidensintervall i de fleste studiene <1 . Også for aldersgruppen 40-49 år er relativ risiko <1 (med ett unntak; se tabell), men med et øvre konfidensintervall >1 i 8 av 10 studier.

Det kan virke noe rart at metaanalyser basert på samme primærstudier gir en relativ dødelighetsreduksjon som varierer fra studie til studie. Årsaken til dette er først og fremst å finne i noe forskjellige forutsetninger som legges til grunn, som forskjellig oppfølgingstid ved analysetidspunktet, og bruk av forskjellige matematiske analysemetoder.

Effekt på totaldødelighet

Informasjon om totaldødelighet (alle dødsårsaker medtatt) gis i tabellen.

Resultater fra Cochrane-rapporten (10):

Forfatter (referanse)	Relativ risiko (RR)	95% konfidensintervall	Oppfølgingstid
40-49 år			
Medium kvalitet: Canada 1980 og Malmö			13 år
• Subtotal	1,02	0,91-1,15	
<u>Lav (poor) kvalitet:</u> Stockholm (kun 7 års oppfølging; ikke med i subtotal). Gøteborg Kopparberg Östergötland			
• Subtotal	1,00	0,92-1,10	
50 år eller over			
<u>Medium kvalitet:</u> Canada 1980 og Malmö			13 år
• Subtotal	1,00	0,95-1,04	
<u>Lav (poor) kvalitet:</u> Stockholm (kun 7 års oppfølging; ikke med i subtotal). Gøteborg (kun 7 års oppfølging; ikke med i subtotal). Kopparberg Östergötland			
• Subtotal	0,99	0,97-1,02	
<u>Dårlig kvalitet:</u> Edinburgh (kun 7 års oppfølging, alle deltagere) New York (alle deltagere) Subtotal data ikke angitt.			

Den neste tabellen viser resultater fra den svenske oppsummeringen (11). Oppfølgingstiden er 15,8 år (median).

Studie	Aldersgruppe	Relativ risiko (RR)	95 % konfidens-intervall
Malmö I	45-70	0,99	0,97-1,01
Malmö II	43-49	1,03	0,89-1,20
(oppfølging av Malmö I)			
Östergötland	40-74	0,98	0,95-1,01
Stockholm	40-65	0,99	0,95-1,03
Göteborg	40-59	0,94	0,88-1,00
Total (aldersjustert)	40-74	0,98	0,96-1,00

I vurderingen fra GR (37); også referert i leserbrevet i Lancet) er den relative risiko for totaldødelighet hos kvinner over 50 år 0,99 (95 % CI 0,97-1,02) etter 13 års oppfølging.

7.2 Negative konsekvenser ("skade") eller spesielle problemer av screening

Intervall-cancer

Fremstilling av antall kvinner med intervallcancer krever kopling til et populasjons-basert kreftregister. Man må også kunne skille mellom intervallcancer som opptrer etter første screeningrunde og senere.

Tabellen viser "proportionell intervall cancer rate", kfr definisjon av begrepet i kap. 6.3, slik den er beskrevet i rapporten fra CEMTV³⁸ dvs "antall observerte intervall cancer delt på forventet antall cancer i gruppen dersom det ikke var screening". I tabellen er det i egen rekke lagt inn data fra Norge³⁹.

	Alder (1)	Proportional interval cancer rate		
		0-11 måneder	12-23 måneder	0-23 måneder
Første screening				
Sverige WE-undersøgelsen	50-69	[0.14]	[0.27]	[0.24]
England East Anglia	50-64	0.24	0.59	0.42
England North-West	50-64	0.31	0.52	0.42
København, Danmark	50-69	0.21	0.48	0.34
Holland	50-69	0.27	0.52	[0.40]
Uppsala, Sverige	40-69	[0.30]	[0.64]	[0.47]
Norge ³⁹		0.25	0.74	0,49
Efterfølgende screeninger				
Sverige WE-undersøgelsen	50-69	[0.14]	[0.33]	[0.24]
Holland	50-69	0.26	0.55	[0.41]

Noter:

1. alder ved invitation i WE-undersøgelsen (Two-county), alder ved screening i service programmer.

³⁸ "Kvalitet av service-screening med mammografi". CEMTV 2002. To kolonner i tabellen er tatt ut.

³⁹ Tatt fra doktorgradsavhandling (Hege Wang: "Epidemiological studies of breast cancer in Norway". Universitetet i Oslo 2002 (38). Av 247 intervallcancer var 17 DCIS.

Tolkningen av tallene i tabellen (første screening) er at fra 34-49% av forventede antall brystkrefttilfeller i perioden før screening ble innført, blir diagnostisert mellom 1. screeningrunde og i de neste 2 år (der en evt. ny screeningrunde følger)⁴⁰.

Absolutt antall intervallcancere fra det norske mammografiscreeningprosjektet oppsummeres slik:

- Antall screenede 1. screeningrunde (4 prøvetylker) : 127.064
- Antall brystkrefttilfeller diagnostisert 1. runde : 854
- Antall brystkrefttilfeller per 10.000 undersøkte : 68
- Antall intervallcancere (mellom 1. og 2. runde) : 249
- Antall intervallcancere per 10.000 screenede : 19,4

Oversikten viser at intervallcancere utgjorde > 1/5 av alle brystkrefttilfellene før 2. runde med screening. Dersom en legger til grunn at antall brystkrefttilfeller diagnostisert etter 1. runde er noe høyt fordi en diagnostiserer noen krefttilfeller (DCIS) som ikke vil utvikle seg til invasiv kreft, blir betydningen av intervallkreft større.

Konsekvenser av screening – screeningkvalitet

Konsekvenser av screening (screeningkvalitet) er oppsummert i rapporten fra CEMTV (Lyng og Olsen (36).

Korttidsindikatorer fra organiserte mammografiscreenings programmer (**første screeningrunde**) i de nordiske land med unntak av København, Danmark gis i tabellen s 45.

Tabellen viser følgende (uttrykt som % av alle deltakere i screeningprogrammet):

- Innkallelse til utredning : 2,1 – 5,7%
- Innkallelse til kirurgi : 0,65 – 1,33 %
- Cancer (invasiv og DCIS) : 0,48 – 0,98%
- Falskt positive : 1,8 – 5,0%

⁴⁰ En drøfting av de metodologiske problemene, spesielt sammenligninger mellom ulike mammografiscreeningsprogram, ligger utenfor denne utredningen

Indikator	Sverige Bohuslän 1987-8 50-74 år	Sverige Öster-götaland 1987-8 (1) 40-74 år	Sverige Uppsala 1988-9 40-69 år	Sverige Stockholm 1989-91 50-69 år	Sverige Nordbotten Väster NL 1991-3 40-74 år	Danmark Fyn 1993-5 50-69 år	Norge 4 amter 1996-7 50-69 år	Finland (1) 1987-97 50-59 år
Antal deltagere	[30040]	21914	37468	71085	[97435]	41596	127064	1493744
Deltagere, % af inviterede (1)	[81.7%]	84.4% (2)	87%	69%	89%	88%	79.5%	89%
Genindkaldelse til udredning	[1453]	[942]	[1712]	2634	[2046]	1138	[7157]	49020
Udredning, % af deltagere	[4.8%]	4.3%	4.57%	[3.66%]	2.1%	2.74%	[5.7%]	3.28%
Genindkaldelse til kirurgi	NR	[164]	[337]	[742]	NR	551	NR	9689
Kirurgi, % af deltagere	NR	0.75%	0.9%	[1.0%]	NR	1.33%	NR	0.65%
Opdagede tilfælde, Invasive+DCIS	[171]	[140]	178	[526]	NR	406	854	5595
Invasive+DCIS, % af deltagere	[0.57%]	0.64%	0.48%	0.74%	NR	0.98%	0.67%	0.37%
DCIS, % af Invasive+DCIS	NR	NR	11%	13%	NR	[16%]	19.8%	11%
Tumor diameter <10mm, % af invasive	NR	NR	11%	NR	NR	37%	53.1% (1)	[33%]
Lymfeknude status kendt, % af invasive	NR	NR	[94%]	NR	NR	NR	NR	NR
Positive, % af invasive lymfeknude status kendt	NR	NR	21%	20%	NR	33%	21.7%	NR
Falsk positive	[1282]	[802]	[1534]	[2108]	NR	[732]	[6303]	[43425]
Falsk positive, % af deltagere	[4.3%]	[3.7%]	4.1%	[3.0%]	NR	[1.8%]	[5.0%]	[2.9%]
Benign kirurgi	NR	[24]	[159]	216	NR	[145]	NR	[4094]
Benign kirurgi, % af deltagere	NR	[0.11%]	[0.42%]	[0.30%]	NR	[0.34%]	NR	[0.27%]

Noter:

1. kun første invitation

Östergötland noter:

1. Data fra service screenings program startet efter afslutning af WE-undersøgelsen. Arnesson LG et al, 1995 all except (2), 2. Vitak B et al, 1997

Norge noter:

1. <15mm

Finland noter: 1. total national aktivitet 1987-97, invitationer og screenings nummer 1 til 6

Resultater eller konsekvenser for **etterfølgende** screeninger for de organiserte mammografiscreenings programmer i England, Holland, og København, Danmark gis i neste tabell (ramme).

	England	England	England	Holland	København
	50-64 år	50-64 år	50-64 år	50-69 år	50-69 år
Indikator	1997/98	1998/99	1999/00	1990-7	1991-7
Antal deltagere	[744187]	[780902]	[855261]	1598075	40730
Deltagere, % af inviterede (1)	[76.2%]	[77.1%]	[76.3%]	78.5%	63.4%
Genindkaldelse til utredning	[32220] (1)	[33593] (1)	[36190] (1)	11014	1307
Utredning, % af deltagere	[4.3%]	[4.3%]	[4.2%]	0.69%	3.2%
Genindkaldelse til kirurgi	[1863] (2)	[1594] (2)	[1435] (2)	7213	311
Kirurgi, % af deltagere	[0.25%]	[0.20%]	[0.17%]	0.45%	0.8%
Opdagede tilfælde,					
Invasive+DCIS	NR	[4237]	[4817]	5656	242
Invasive+DCIS, % af deltagere	NR	[0.54%]	[0.56%]	0.35%	0.59%
DCIS, % af Invasive+DCIS	NR	[20%] (3)	[20%] (4)	14%	11%
Tumor diameter <10 mm,					
% af invasive	NR	NR	[24%] (5)	[29%] (1)	NR
Lymfeknude status kendt,					
% af invasive	NR	NR	NR	[94%]	NR
Positive, % af invasive					
lymfeknude status kendt	NR	NR	NR	[24%]	NR
Falsk positive	NR	[29356]	[31373]	[5358]	1065
Falsk positive, % af deltagere	NR	[3.8%]	[3.7%]	[0.3%]	2.6%
Benign kirurgi	[1863] (2)	[1594] (2)	[1435] (2)	[1557]	69
Bening kirurgi, % af deltagere	[0.27%] (2)	[0.20%] (2)	[0.17%] (2)	[0.10%]	0.17%

Noter: 1. kun etterfølgende screeninger

England noter:

1. eksklusiv kvinner med tidlig genindkaldelse efter initial screening, 2. kun åben biopsi. Omfatter formodentlig heller ikke kvinner opereret basert på cytologi eller fund fra kernebiopsi, da antallet af kvinner med åben biopsi er mindre end antallet af kvinner med opdaget kræft. Antallet formodes at representere benign kirurgi, 3. ikke invasive + micro-invasive i procent af alle, 4. ikke invasive og muligt micro-invasive i procent af alle, 5. procent af invasive eksklusiv micro-invasive

Holland noter:

1. 1990-1994, 1994 data fra syv af ni regioner, første+etterfølgende screeninger, 2. 1990-1994, 1994 data fra syv af ni regioner

Det bør her legges til at resultatene fra mammografiscreeningprogrammet i Holland avviker fra det som oppgis fra tilsvarende programmer i andre land. Holland har bevisst valgt å holde en lav gjeninnkallelsestrate. Dette gjenspeiles i andelen som etterinnkalles til utredning, andelen falskt positive mammografiundersøkelser, og andelen brystkrefttilfeller (invasiv + DCIS).

Overdiagnostikk og overbehandling

Andelen med DCIS er i de nordiske land 11-19% (se tabell under pkt. 7.2.). I Norge var andelen DCIS 19,7% i første screeningrunde (4 prøvefylker), og 19,0% i andre screeningrunde.

Det er ikke mulig, basert på foreliggende tall å si noe om hvor stor del av disse som representerer ”overdiagnostikk”.

SMM har ikke identifisert systematiske litteraturstudier som spesifikt tar opp problemstillingen med overdiagnostikk. Cochrane-rapporten (og den foregående Lancet-studien) er heller ikke spesielt ufyllende. Forøvrig vises til diskusjon under pkt. 8 .3.

7.3 Nytte-kostnadsvurderinger

I publiserte økonomiske analyser av mammografi-screening har nytte i forhold til kostnader stort sett vært målt som kostnader per vunnet leveår. Slike forholdstall basert på utenlandske studier har sannsynligvis moderat relevans for norske forhold. Selv om vi primært bare har hatt til hensikt å basere oss på systematiske oversiktsartikler, har vi derfor trukket inn norske primærstudier nettopp for å få kostnadsdata tilpasset norske forhold.

Av de 4 norske studiene er den ene mindre relevant for helseøkonomiske vurderinger av metodemessige grunner (Norum et al, 31.). Kåresen og medarbeideres artikkel (33) er i prinsippet lik Wang og medarbeideres (selv om stoffet er organisert og analysert på litt forskjellig måte). Siden Wangs artikkel (32) er mer utfyllende og senest publisert (2001) legges den til grunn.

I denne rapporten vil derfor studiene fra Ekerhovd (30) og Wang et al. (32) legges til grunn.

Ekerhovd i 1996 (30) beregnet at kostnaden per vunnet leveår var NOK 345.551, mens Wangs i 2001 (32) estimat var 3750 USD (ca. NOK 30.000). Wang beregnet leveårsgevinsten med basis i 30% dødelighetsreduksjon, mens Ekerhovd beregnet vunne leveår på basis av bedre stadiefordeling av brystkreft ved mammografi.

Sensitivitetsanalyser har vært gjort hos begge forfatterne. Spesielt kan det være relevant å påpeke at Wang i sin sensitivitetsanalyse ser på hvordan en variasjon i mortalitetsreduksjon fra 20% til 40% påvirker kostnads-effektforholdet. Denne sensitivitetsanalyse ble i hovedsak utført før Gøtzsche og Olsens meta-analyse i 2000 (8).

Variasjonen i kostnad-effekt-forholdene er stor og skyldes ulikheter i metode som har vært lagt til grunn for analysene. Ulikhetene gjelder så vel beregning av leveårsgevinster som beregning av samfunnets kostnader ved mammografi.

Et annet problem er hvorvidt disse forholdstallene kan brukes i prioriteringssammenheng. Rett nok sammenligner forfatterne screening med ikke-screening, men de har ikke drøftet tydelig nok kostnaden per vunnet leveår i andre anvendelser og tolkningen av deres egne kostnad-effekt-forhold. Tallene kan dermed vanskelig legges til grunn når beslutninger skal tas. Likevel vil en problematisering av hva som inngår på nytte og kostnadssiden kunne være til hjelp i beslutningsprosessen.

Kommentarer til de norske primærstudiene er lagt til Vedlegg 3. Den systematiske litteraturoversikten (Barrat et al. 2002, 29) tar opp et spesialproblem om hvorvidt det er kostnadseffektivt å utvide screeningprogrammet i Australia til også å omfatte kvinner over 70 år. SMMs oppsummering og vurdering av dette arbeidet vedlegges i Vedlegg 4.

8. Vurdering av problemområder

8.1 Metodologiske problemer, hvor store og alvorlige er de?

Mye av diskusjonen om mammografiscreening har dreiet seg om hvor gode eller dårlige de vitenskapelige arbeidene som inngår i Cochrane-oversikten (10) er. Denne diskusjonen dreier seg i stor grad om vurderingen av den vitenskapelige kvaliteten i de randomiserte studiene. Dette ble først belyst av Cochrane-rapportens forfattere i en metaanalyse publisert i Lancet (8).

Det er ikke nytt at mulige svakheter i de vitenskapelige arbeidene om mammografiscreening diskuteres Bl.a. har Cox (4) foretatt vurderinger av slike spørsmål. Det er likevel ikke tvil om at intensiteten i disse diskusjonene har økt etter at Lancet-artikkelen (8) og Cochrane-rapporten (10) kom.

Forfatterene av Cochrane-vurderingen delte de vitenskapelige arbeidene inn i:

- Høy kvalitets forsøk ("*high quality trials*"); ingen
- Medium kvalitets forsøk ("*medium quality trials*"); Canada og Malmø
- Forsøk med dårlig kvalitet ("*poor quality trials*"); Göteborg, Stockholm og Two-County;
- Forsøk med alvorlige feil ("*flawed trials*"); Edinburgh, New York

Bakgrunnen for denne klassifiseringen er primært knyttet til spørsmål om randomisering og om studiene er konsistent når det gjelder rapportering av antall pasienter. Tabellen oppsummerer hvordan Gøtzsche og Olsen (8) har vurdert slike spørsmål⁴¹

	Randomisation produced similar groups	Account of number of patients consistent
Malmö	Yes	Yes
Canada	Yes	Yes
Göteborg	No	Yes
Stockholm	No	No
Kopparberg	No	No
Östergötland	No	No
New York	No	No
Edinburgh	No	Yes

Table 1: Mammography screening trials according to methodological quality

Gøtzsche og Olsen konkluderer med at bare 2 studier hadde medium kvalitet, resten var av dårlig kvalitet eller hadde alvorlige feil. Dette synspunktet er utdypet og gjentatt i Bandolier (www.jr2.ox.ac.uk/bandolier/band72/b72-3.html) (43). Dersom man ser bort fra de dårlige studiene og bare legger de med "medium" kvalitet til grunn, vil effekten (dvs. reduksjonen i brystkreftspesifikk dødelighet) av screening falle bort.

⁴¹ tabellen er tatt fra Lancet 2000;355:129-134

Denne problemstillingen er sentral i den rapporten som Gesondheidsraad (GR) i Holland har utformet på oppdrag av den hollandske helseminister (37). Den vitenskapelige komiteen, med deltagelse bl.a. fra det hollandske Cochranesenteret, har også vurdert dette spørsmålet og avslutter sin vurdering slik:

”The committee rejects the contention that the inconsistency between the results of the ‘medium-quality’ RCTs and those of the ‘poor-quality’ RCTs is attributable solely to the methodological shortcomings of the latter. It is nevertheless accepted that such shortcomings might have influenced the estimation of breast-cancer mortality effects to a limited extent”.

International Agency for Research on Cancer (IARC) som er en del av WHO, har også drøftet de metodologiske problemene under et møte i Lyon 5.-12. mars 2002. 24 eksperter fra 11 land deltok. Diskusjonene og oppsummeringen er ennå ikke tilgjengelig men vil bli publisert i serien IARC Monographs høsten 2002. Som en oppsummering⁴² etter møtet ble det sendt ut en pressemelding der det heter bl.a.

”The working group found that many of the earlier criticisms were unsubstantiated, and the remaining deficiencies were judged not to invalidate the trials’ findings”.

I pressemeldingen er det også oppgitt en størrelse (35%) på dødelighetsreduksjonen av brystkreft. Denne kan ikke brukes her fordi grunnlaget for beregningene ikke er gitt. Det registreres at effektestimatet gir et bedre resultat enn det som fremgår av de andre metaanalysene som er referert i pkt. 7.1.

The U.S. Preventive Services Task Force har diskutert og revidert sine behandlings-retningslinjer etter at Lancet artikkelen (8) og Cochrane-rapporten (10) forelå. I de nye anbefalinger om brystkreft screening (19) heter det:

”Although the USPSTF was concerned about many (but not all) of the flaws identified in this review (Cochrane oversikten; SMMs tilføyelse), it did not consider the presence of flaws sufficient reason in itself for rejecting trial results. Instead, it examined whether observed mortality reductions in the trials were likely to be explained by the biases potentially introduced by such flaws. Studies rated to be of “fair” quality by the USPSTF contained flaws that were considered unlikely to account for observed benefits (or lack of benefits)”.

SMM kan ikke avgjøre den faglige uenigheten som her kommer til syne mellom forfatterene av Cochrane-oversikten og de organene eller organisasjonene som her siteres. Det vil alltid være element av skjønnmessige vurderingene ift. hva som er ”optimalt”, ”godt nok” eller ”ikke godt nok” når det gjelder forskningsdesign. Usikkerheten dette medfører må en leve med.

⁴² Det er ikke mulig å vurdere hvordan IARC har kommet frem til sine konklusjoner. Den relative dødsrisikoreduksjon som pressemeldingen anfører kan derfor ikke brukes. Publikasjonen fra IARC vil foreligge i september 2002 (IARC Handbooks of Cancer Prevention Series. Vol. 7. Breast Cancer Screening. IARC Press 2002 (ISBN 92-832-2412-4)). I en brosjyre for publikasjonen finnes følgende: ”The group concluded that there is sufficient evidence from randomised trials that mammographic breast cancer screening of women aged 50-69 years reduces mortality from breast cancer by 25%. There is only limited evidence for an effect in women aged 40-49 years”. Dette er resultater som ikke avviker fra de som gis i kapittel 7.

SMM kan heller ikke se at etablering av en egen norsk ekspertgruppe vil løse slik usikkerhet eller komme med avgjørende bidrag til oppklaring.

Rapporten fra GR og USPSTF og pressemeldingen fra IARC peker på noen svakheter i de vitenskapelige studiene. Konsekvensene av dette er, som påpekt av GR, ikke å avvise studiene, men å anta at estimatene av nytteverdi (dvs. effekt på kreftspesifikk dødelighet) i noen grad har vært overvurdert tidligere.

At det er forskningsmetodologiske svakheter i de 7 studiene som inngår i Cochrane-rapporten fremkommer bl.a. ved at USPSTF graderer det vitenskapelige dokumentasjonsgrunnlaget for mammografiscreening som FAIR⁴³. Fair defineres slik:

"Fair: Evidence is sufficient to determine effects on health outcomes, but the strength of the evidence is limited by the number, quality, or consistency of the individual studies; generalizability to routine practice; or indirect nature of the evidence on health outcomes".

European Institute of Oncology i Milano har avholdt et ekspertseminar 5.-7. juni 2002 for å diskutere liknende spørsmål ("Global Summit on Mammographic Screening"). Programmet er tilgjengelig (www.ieo.it/inglese/didattica/provprog.htm). SMM forventer ikke at dette seminaret vil endre kunnskapsgrunnlaget eller komme med bidrag som rydder av veien de uenighetene som foreligger. Det er bare publiserte vitenskapelige artikler der alle aspekter ved undersøkelsen kan vurderes kritisk som kan legges til grunn for vurdering av effekt.

8.2 Effekt av screening - dødelighet

8.2.1 Dødelighet – kreftspesifikk eller totaldødelighet?

Formålet med mammografiscreening er å redusere dødeligheten for kvinner med brystkreft. I Cochrane-rapporten (10) konkluderes det at brystkreftspesifikk dødelighet er et upålitelig (unreliable) mål⁴⁴ og biased i favør av kvinnene som er screenet, og at man bør bruke totaldødelighet. Begge måtene å uttrykke dødelighet på er relevante. Kreftspesifikk dødelighet gir et uttrykk for dødeligheten av en spesifikk kreftsykdom, i dette tilfelle brystkreft. Mammografiscreening kan bare påvirke dødeligheten av brystkreft.

Total dødelighet er viktig å registrere for å være sikker på at screeningtiltaket og de konsekvenser det innebærer, ikke *øker* den samlede dødelighet ved å øke dødeligheten av andre årsaker (f.eks. selvmord). Økt dødelighet i en screeninggruppe er bl.a. beskrevet i et pilotforsøk i Norge ved bruk av sigmoidoskopi⁴⁵ som screeninginstrument for å oppdage kolorektal cancer⁴⁶. Her er det også relevant å minne om at ekstern strålebehandling (som inngår som en del av behandlingen) er antatt å kunne øke dødeligheten av andre sykdommer⁴⁷.

⁴³ USPSTF bruker et tredelt system for gradering; good, fair and poor

⁴⁴ Den fullstendige argumentasjon gjengis ikke her, det vises til rapporten.

⁴⁵ Se SMM rapport nr. 3/2001 Screening for kolorektal kreft

⁴⁶ Kolorektal cancer = kreft i tykk- og endetarm

⁴⁷ Lancet 2000;355:1757-70. SMM har ikke foretatt en metodevurdering av denne problemstillingen. Argumentet fremføres for å synliggjøre at totaldødelighet også er et relevant endepunkt.

I et folkehelseperspektiv må en derfor unngå at en evt. positiv effekt av screening (reduisert dødelighet pga. brystkreft) ikke forsvinner ved at dødeligheten av andre årsaker øker etter innføring av screening.

Diskusjonene om dødelighet reiser spørsmålet om hvor sikkert dødsårsakene kan fastslås i de aktuelle vitenskapelige studiene som er utført, og om de som fastsetter dødsårsaken har kjennskap til hvilken gruppe (screening eller observasjon) den avdøde tilhørte⁴⁸.

”Gullstandard” for å fastslå dødsårsak er obduksjon (autopsi). I de studiene som inngår i Cochrane-rapporten var autopsi-raten fra 0-76%. Diskusjonene om sikkerheten i fastsettelse av dødsårsak er viktige når man ser på forskjellene i absolutt antall brystkreftdødsfall i screening- versus kontrollgruppe (se pkt. 7.1). I den svenske oppdateringen (11) er dødelighetstallene hentet fra offisiell svensk dødsårsakstatistikk. Denne fremgangsmåten er validert i en tidligere svensk studie⁴⁹. Unntaket fra dette er ”Two County Study” der dødsårsakene ble fastslått av ”a local team of physicians” (opplysninger tatt fra ref. 11)

Den ”beste” måten å uttrykke dødelighet på (kreftspesifikk eller total) er derfor ikke entydig gitt. Spørsmålet kan også stilles om det er rimelig å anta at screening for brystkreft skal kunne gi redusert totaldødeligheten.

Grunnen til at dette spørsmålet reises er at brystkreft-dødsfall i Norge utgjør 3,1 % av alle dødsfall hos kvinner (705 brystkreftdødsfall av 22698 dødsfall; kilde: www.ssb.no/dodsarsak/tab-2001-12-21-02.html).

I den svenske oppfølgingstudien (11) reduseres totaldødeligheten i de inviterte kohortene med 2% (RR 0,98; 95%CI 0,96-1,00). I artikkelen konkluderes med

“In our overview, there was a 2% lower total mortality in the invited cohorts. This estimate was not changed by adjustment for age. Mortality from breast cancer in the age groups 50–59, 60–69, 70–79, and 80–84 years in Sweden in 1990 constituted 13.2%, 7.2%, 3.2%, and 2.1%, respectively, of the total female mortality. Because the median age at death in women 40–49, 50–59, 60–69, and 70–74 years at entry was 55, 65, 75, and 82 years, respectively, and the relative effect according to the follow-up model was 9%, 12%, 17%, and –12%, respectively, the expected effect of the intervention on the total mortality was 2.3%, which is in accordance with our results”.

Ett problem er at det er ikke mulig å få tilgang til rådata (originaldata) til hele eller deler av studier. National Cancer Institute/PDQ (41) skriver om dette slik:

”The accessibility of the data for international audit and verification also varies, with the authors of several trials refusing to submit their raw data to independent review”.

⁴⁸ Kjenner en til hvilken gruppe den avdøde tilhørte, kan en bevisst eller ubevisst påvirkes ift. til det å fastsette dødsårsaken, mao. om døden skyldtes brystkreft eller andre årsaker.

⁴⁹ se Nyström L et al.: Determination of cause of death among breast cancer cases in the Swedish randomised mammography screening trials: a comparison between official statistics and validation by an endpoint committee. Acta Oncol 1995;34:145-52.

Problemet med manglende åpenhet blir ikke bare diskutert i fagtidsskrifter og andre faglige informasjonskilder, men også i mer almenne publikasjoner beregnet på et bredt publikum. Tidsskriftet TIME Magazine⁵⁰ henviser således til dette med følgende:

”Will we ever know the truth about mammograms? There’s push to make public the raw data from some of the original studies. Without new studies, that’s probably the closest we are likely to come in the near future to a scientific answer.”⁵¹

Fra et forskningsetisk synspunkt må mangel på åpenhet om rådata, spesielt i en situasjon der effekten av mammografiscreening diskuteres, være problematisk.

8.2.2 Effekt ift. alder.

Som påpekt av USPSTF (19) er ingen av de vitenskapelige forsøk som er utført (med unntak av NBSS Trial, Canada), og som danner grunnlaget for Cochrane-oversikten og andre tilsvarende analyser, utformet (designet) slik at de kan teste ut om det å starte screening hos kvinner under 50 år, er effektivt (dvs. reduserer dødeligheten). Målene på effekt i aldersgruppen 40-49 år har derfor stort sett sitt vitenskapelige grunnlag i subgruppe-analyser. Subgruppeanalyser er forbundet med metodeproblemer hvis ikke studiens styrke er beregnet for slike analyser. I forsøket som var designet for å ta rede på effekten av screening i aldersgruppen 40-49, var det ingen effekt av mammografiscreening.

USPSTF summerer opp analysene de har utført og diskusjonene om effekten av mammografiscreening hos kvinner 40-49 år på denne måten (s. 6):

”Earlier subgroup analyses from mammography trials raised questions about whether screening is effective in women younger than 50. Seven trials enrolled women aged 40-49. Six of these were rated by the USPSTF to be of at least “fair” quality, but only one of these was designed to specifically address the benefits of screening in this age group: it reported no reduction in breast cancer mortality with annual mammography and CBE.¹⁷

Of the remaining five fair-quality trials that included women younger than 50, two trials have now reported significant mortality reductions with screening in this age group^{12,14}, two have reported non-significant mortality reductions^{11,15}, and 1 found no benefit.¹³

In a meta-analysis performed for the USPSTF pooling results for women aged 40-49 in these six trials, the relative risk of breast cancer mortality was 0.83 (95 percent CI 0.64-1.04) among screened women; inclusion of the seventh, poor-quality study did not change results.² These results are similar to prior meta-analyses based on older data”.

Dataene fra USPSTF viser at det ikke er en sikker reduksjon i dødelighet i aldersgruppen 40-49 år (øvre grense for konfidensintervall 1.04). De fleste metaanalysene viser heller ingen sikker effekt (se pkt. 7.1).

⁵⁰ TIME, June 10, 2002 (fra s.45) ’Rethinking Breast Cancer. New detection techniques and treatments are exciting- and confusing. A guide to saving lives.’

⁵¹ Dette reiser forskningsetiske problemstillinger som ikke diskuteres.

Usikkerheten om screening reduserer dødelighet i aldersgruppen 40-49 år vises også i følgende tatt fra USPSTF guidelines:

”Because these data represent a subgroup analysis of trials not designed to test the benefits of beginning screening at a specific age, questions remain about the additional benefits of beginning screening before age 50. On average, the time until mortality benefits begin to be observed in these trials is longer in women younger than 50 than in older women (8 years vs 3 to 4 years) and some of the observed benefits could be due to screening after age 50.^{3,22} Analyses of individual studies suggest that at least some of the mortality reduction is due to early detection of tumors before age 50, but definitive estimates of the proportion of benefits due to early screening cannot be made.^{3,23}”

Effekten (dvs. nytteverdien) av screening øker noe med alderen, og den avveining som foretas mellom ”nytte” og ”skade” (eng. harm) blir mer positiv med stigende alder. Dette kommer frem i anbefalingene og konklusjonene fra USPSTF.

”The USPSTF found fair evidence⁵² that mammography screening every 12-33 months significantly reduces mortality from breast cancer. Evidence is strongest for women aged 50-69, the age group generally included in screening trials. For women aged 40-49, the evidence that screening mammography reduces mortality from breast cancer is weaker, and the absolute benefit of mammography is smaller, than it is for older women. Most, but not all, studies indicate a mortality benefit for women undergoing mammography at ages 40-49, but the delay in observed benefit in women younger than 50 makes it difficult to determine the incremental benefit of beginning screening at age 40 rather than at age 50. The absolute benefit is smaller because the incidence of breast cancer is lower among women in their 40s than it is among older women.

The USPSTF concluded that the evidence is also generalizable to women aged 70 and older (who face a higher absolute risk of breast cancer) if their life expectancy is not compromised by comorbid disease. The absolute probability of benefits of regular mammography increase along a continuum with age, whereas the likelihood of harms from screening (false-positive results and unnecessary anxiety, biopsies, and cost) diminish from ages 40-70.

The balance of benefits and potential harms, therefore, grows more favorable as women age. The precise age at which the potential benefits of mammography justify the possible harms is a subjective choice. The USPSTF did not find sufficient evidence to specify the optimal screening interval for women aged 40-49 (see Clinical Considerations)”.

Videre påpeker National Cancer Institute (USA) informasjonssystem PDQ at forholdet mellom nytte (dvs. mulig reduksjon i dødelighet) og skade (dvs. falskt positive undersøkelser) påvirkes av alder. I omtalen heter det bl.a.

”Based on these trials, there is general agreement that screening mammography for women aged 50 to 69 is beneficial, at the population level. The controversy arises regarding recommendations for younger women. Women in their 40s have a lower cancer incidence, but their clinically diagnosed cancers tend to be more aggressive. Premenopausal breast tissue is more likely to be dense, making interpretation of mammograms more difficult. The only trial designed with a specific pre-hoc intention to study these women was the NBSS1. Data from the age 40 to 49 subgroups of the other trials have been combined with data from this trial in meta-analyses.¹⁶⁻¹⁸ Combining data from all 8 trials, there may be a small benefit in favor of

⁵² se tidligere definisjoner av ”fair”

screened women that emerges approximately 10 to 12 years after the initiation of screening. Confidence intervals and level of statistical significance vary, depending on the statistic methodology. A concern is the variable age at entry into the trials, so that in the 40 to 49 subgroups, most of the women underwent mammograms in their late 40s or early 50s. It is possible that the demonstrated benefit would also accrue if women waited until age 50 to begin mammographic screening”.

Den australske helseøkonomiske litteraturoversikten (29) viser at det er relativt kostnadseffektivt å utvide screeningtilbudet for kvinner i alderen 70-79 år og at det fremstår som like kostnadseffektivt som å utvide tilbudet for kvinner i gruppen 40-49 år.

Tolkningen av resultatene kompliseres av at eldre kvinner er mindre flinke til å møte opp enn yngre⁵³ og at dødeligheten av andre årsaker øker med stigende alder.

8.2.3 Andre måter å illustrere nytteverdien av mammografiscreening

Number needed to treat (NNS)-begrepet uttrykker hvor mange kvinner som må screenes over en definert tidsperiode for å unngå ett dødsfall av brystkreft. Tabellen gitt nedenfor viser et eksempel på slike analyser. Tabellen⁵⁴ viser hvor mange kvinner med en gitt alder som må screenes over gjenværende levetid (som er anslått med forskjellige tid) for å forhindre ett dødsfall av brystkreft gitt en relativ dødsrisikoreduksjon på 26%.

	Alder 50 år	Alder 70 år	Alder 75 år	Alder 80 år	Alder 85 år	Alder 90 år
Forventet levetid i år	40 33 24,5	21,3 15,7 9,5	17 11,9 6,8	13 8,6 4,6	9,6 5,9 2,9	6,8 3,9 1,8
NNS	95 133 226	142 242 642	176 330 1361	240 533 ...	417 2131 ...	1066

National Cancer Institute har i sin oppsummering (PDQ-databasen) illustrert effekten av screening på denne måten:

”Another way to view the benefit of breast cancer screening is to estimate the number of lives extended because of early breast cancer detection.^{17,19} Harris ²⁰ calculated the outcomes of 10,000 women age 50 to 70 who are screened annually. Mammograms will be normal (true negatives and false negatives) in 9,500 women. Of the 500 abnormal examinations, between 466 and 479 will be false positives, and 100 to 200 of these women will undergo invasive procedures. The remaining 21 to 34 abnormal examinations will be true positives, indicating breast cancer. Some of these women will be incurable in spite of mammographic detection, and some may have been cured even if the disease had been discovered later. Thus the extended lives attributable to mammographic detection is between 2 and 6. Another expression of this analysis is that one life is extended per 1,700 to 5,000 mammograms performed. The same analysis for 10,000 women age 40 to 49, assuming the same 500 abnormal examinations, results in an estimate that 488 of these will be false positives, and 12 will be breast cancer. Of these 12, there will probably be only 1 to 2 lives extended. Thus, for women aged 40 to 49, it is estimated that 1 to 2 lives will be extended per 5,000 to 10,000 mammograms”.

⁵³ se bl.a. Br J Cancer 1987;55:547-51.

⁵⁴ Tatt fra JAMA 2001;285:2750-2756. Grunnlaget for ”Forventet levetid i år” er kvartiler av forventet gjenlevende tid for kvinner.

8.3 Overdiagnostikk og overbehandling

Et av hovedpoengene i Gøtzsche og Olsens Lancet-artikkel (8) er at mammografiscreening fører til overdiagnostikk og overbehandling. I den hollandske rapporten fra GR (s. 42) oppsummeres kritikken slik :

“According to the Cochrane review, the estimated cumulative risk of a false positive result after 10 mammograms⁵⁵ is 50 per cent. Screening also leads to more women being diagnosed with breast cancer and to increases both in the overall frequency of breast surgery and in more aggressive breast surgery, it is claimed. According to the Cochrane review, the rise in frequency of breast cancer cannot be dismissed as a temporary effect resulting from earlier diagnosis. Rather, it is a structural phenomenon attributable to the discovery of slow-growing cancers that would otherwise not have led to ill health. It is said that screening results in ductal carcinoma in situ (a precursor of breast cancer) being detected six times more often.

The authors claim that the resulting overtreatment involves 20 per cent more mastectomies and a 30 per cent increase in the overall number of breast operations (including breast-conserving surgery). It is also suggested that screening leads to a 25 to 40 per cent increase in the use of radiotherapy in the treatment of breast cancer”.

Falskt positive undersøkelser

Lancet-artikkelen (8) og Cochrane rapporten (10) refererer til en amerikansk undersøkelse ved Elmore og medarbeidere⁵⁶ på 2.400 kvinner at det er en 49 % kumulativ risiko for å ha minst ett falsk positiv resultat etter 10 screeningrunder. GR-rapporten mener at dette anslaget bygger på feile forutsetninger, og viser til at anslaget ikke stemmer for Holland.

Spørsmålet er vurdert i rapporten fra det danske metodevurderingsorganet (36). I rapporten anføres:

”Undersøgelsen af testen fra Stockholm viste at vurderingen efter en positiv mammografi nogle gange tog hele måneder [48], og en falsk positiv test har været forbundet med midlertidig bekymring [49-50]. Andelen af deltagere med en falsk positiv test var 5,6% ved første screening i København og 2,6% ved efterfølgende screeninger. Under antagelse om uafhængighed forventedes 10% ($=1 - (1-0.056)(1-0.026)(1-0.026)$)⁵⁷ af de kvinder, der deltog i de screeninger, der blev tilbudt i 1991-1997 at få foretaget mindst én falsk positiv test. Rent faktisk oplevede 10% dette (1636/15898) [20].

De landsdækkende programmer i Storbritannien og Holland har fulgt meget forskjellig praksis hvad genindkaldelse angår, og antallet af kvinder med en falsk positiv test vil derfor også være meget forskjellig. I 1999/00 var der 7,5% af kvinderne i England, som hadde en falsk positiv test ved første screening, og 3,7% ved efterfølgende screeninger. Hvis vi antager uafhængighet, giver dette os 14% med mindst én falsk positiv test ud af de første 3 screeninger. For Holland var procentdelen af falsk positive tests hhv. 0,7% og 0,3%, og dette fører til en forventet procentdel på 1% med mindst én falsk positiv test ud af de første 3 screeninger”.

⁵⁵ SMM forstår dette slik at dersom en med 10 screeningrunder, og med screening hvert annet år, vil den teoretiske kumulative risiko være 50% over en 20-års periode.

⁵⁶ Elmore JG, Barton MB, Mocerri VM et al. Ten-year risk of false positive screening mammograms and clinical breast examinations. N Engl J Med 1998;338:1089-96.

⁵⁷ dvs. etter 3 screeningrunder (SMMs tilføyelse)

Dersom en bruker samme tilnærming som i den danske vurderingen referert her, vil en kunne forvente en andel med falskt positive mammogram som vist i tabellen:

Screeningprogram	Etter 5 screeningrunder	Etter 10 screeningrunder
København	15%	25%
England	20%	34%
Holland ⁵⁸	2%	3%

Det foreligger altså en kumulativ risiko for et falskt positivt resultat som følger av etableringen av et screeningprogram. Risikoen avhenger av antall screeningrunder og i hvilken alder screeningen og avhenger av rutiner og praksis for gjeninnkalling. Det er imidlertid bare en regelmessig oppdatering av de populasjonsbaserte mammografi-screening programmene, også i Norge, som kan gi svar på hva som faktisk er den reelle frekvensen.

Overdiagnostikk

Den hollandske rapporten (37) erkjenner at det vil skje en overdiagnostikk i et screeningprogram. Men det påpekes at dette er en logisk konsekvens eller følge av ethvert screeningprogram, og at en økning var, i alle fall i Holland, forventet før screeningprogrammet startet.

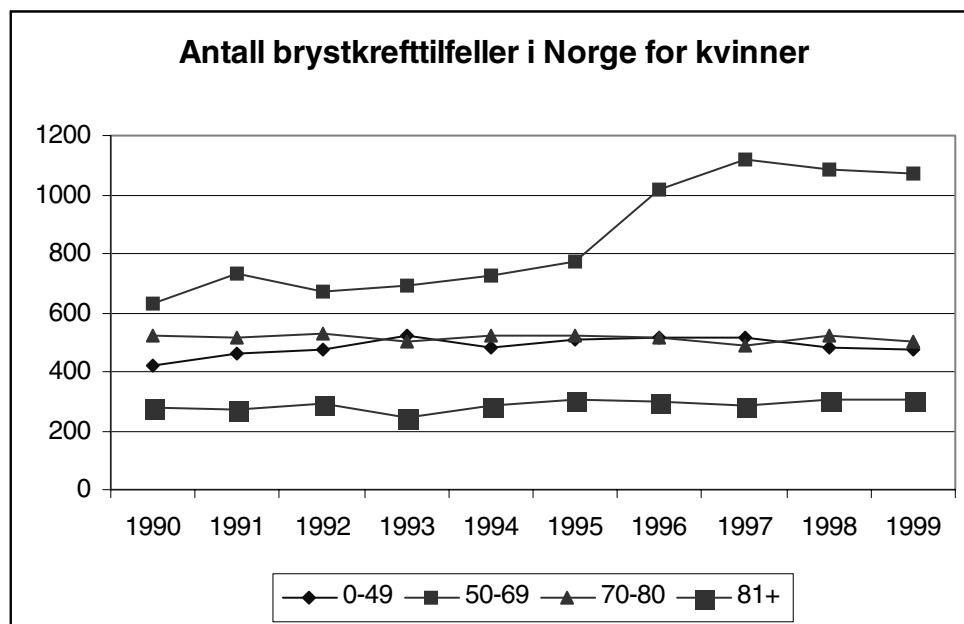
Den hollandske rapporten påpeker (s. 46) imidlertid at overdiagnostikken er mindre enn det Cochrane-rapporten angir, og anfører med bruk av tall fra det hollandske screeningprogrammet :

”According to cancer registry data, 114 new cases of (invasive) breast cancer per 100,000 women were found in the Netherlands in 1990, and 127 in 1996. In the same period, the frequency of the diagnosis of in-situ cancer of the breast rose from five to eleven per 100,000 women. This rise is attributable mainly to the screening programme. Some 13 per cent of women between the ages of fifty and seventy with breast cancer detected by screening have in-situ cancer. The corresponding figure for women in the same age group with breast cancer detected during interscreening intervals or outside the context of screening is 3 to 6 per cent, depending on age (Fra 01)”.

Figuren viser antall tilfeller av brystkreft i ulike aldersgrupper i Norge før og etter innføring av mammografiscreening (prøveprosjektet)⁵⁹. Det foreligger en økning i antall brystkrefttilfeller i Norge etter innføring av screening. Hvor stor den reelle økningen er kan ikke direkte leses fra figuren. Bl.a. fordi screeningen er organisert etter fødselskohorter og ikke etter løpende alder som i figuren.

⁵⁸ se tidligere kommentarer om screeningprogrammet i Holland

⁵⁹ Data fra S. Thoresen, Kreftregisteret



Det er en generell antagelse at det skjer noe overdiagnostikk i alle mammografi-screeningprogram og at dette er et fenomen som en må forvente⁶⁰.

Overbehandling

Dette dreier seg om 2 problemstillinger; om økningen i DCIS fører til behandling av "kreft" som ellers ikke ville utvikle seg til invasiv cancer, og om behandlingen av DCIS er for omfattende ift. den entitet DCIS er.

I den hollandske rapporten (s. 45) besvares dette spørsmål bl.a. slik :

"...If one looks only at the initial period following the introduction of screening, even if that is nearly ten years, it is inevitable that one will find an increase in the frequency of breast cancer and breast surgery. However, there is no evidence of any serious level of overdiagnosis in the longer term (Pee89).

Furthermore, the claims regarding overtreatment are based on incomplete data that do not distinguish between age groups under 50 and over 50 years or provide insight into the percentage distribution of the types of breast surgery. If such data had been presented, they would have shown that, particularly among screened women, there has been a relative decline in mastectomies and a relative increase in breast-saving surgery".

Rapporten konkluderer med:

"Screening enables breast cancer to be detected earlier, when the tumour is smaller and metastasis to the lymph nodes is less likely to have occurred. This means that 'aggressive' forms of diagnostic and therapeutic intervention (e.g. excision biopsy, axillary sampling or dissection and mastectomy) are necessary less often. In recent decades, screening has accelerated the development of less mutilating forms of intervention, thereby reducing

⁶⁰ Epidemiologen E. Lynge mener at det er mulig å lage et screeningprogram uten overdiagnostikk

suffering and the cosmetic implications of the disease. The Cochrane review ignores these facts, choosing only to look backwards”.

USPSTF (19) har diskutert overdiagnose og overbehandling på denne måten (s. 8):

”Some experts view the over-diagnosis and treatment of ductal carcinoma in situ (DCIS) as a potential adverse consequence of mammography. Although the natural history of DCIS is variable, many women in the United States are treated aggressively with mastectomy or lumpectomy and radiation.² Given the dramatic increase in the incidence of DCIS in the past two decades (750 percent) and autopsy series suggesting that there is a significant pool of DCIS among women who die of other causes³, screening may be increasing the number of women undergoing treatment for lesions that might not pose a threat to their health”.

8.4 Intervallcancere

Frekvensen av intervallcancer er gitt i 7.2. I Norge utgjør intervallcancer rundt 1/5 av alle brystkrefttilfeller (se pkt. 7.2). Den praktiske betydningen av funnene presentert i tabellen er at kvinnene må informeres om intervallcancer-fenomenet, spesielt at screening ikke beskytter mot kreft eller er en garanti for at brystkreft ikke kan forekomme etter screening og før neste screeningrunde, og at helsevesenet må kontaktes ved klinisk mistanke om brystkreft.

Intervallcancer-fenomenet reiser spørsmålet⁶¹ om forkortelse av screeningintervallet, dvs. hvor lang tid det skal gå mellom hver screeningrunde. USPSTF omtaler dette spørsmålet slik:

”In the trials that demonstrated the effectiveness of mammography in lowering breast cancer mortality, screening was performed every 12-33 months. For women aged 50 and older, there is little evidence to suggest that annual mammography is more effective than mammography done every other year. For women aged 40-49, available trials also have not reported a clear advantage of annual mammography over biennial mammography. Nevertheless, some experts recommend annual mammography based on the lower sensitivity of the test and on evidence that tumors grow more rapidly in this age group”.

En fullstendig vurdering av hvilken betydning screeningintervallenes lengde har, vil kreve en egen utredning og faller utenfor rammene av forspørselen.

8.5 Er effekten av screening reell eller er den resultat av forbedret, ikke-kirurgisk, behandling?

Dette spørsmålet reises, og definitive og autoritative svar finnes ikke. Mer kunnskap er nødvendig som også påpekt i en nylig leder i Lancet⁶².

“There has been a reduction in breast cancer mortality of nearly 30% since 1990. 21-23 How much of this is due to screening and early diagnosis and how much to improved therapy? During the era when most women were being accrued to the trials contributing to the latest analysis,

⁶¹ Mammografiprogrammet i Norge. Kreftregistrets forskningsrapport nr. 2-2000 (s. 43)

⁶² Lancet 2002;359:904-5

systemic therapy was just beginning to be adopted and certainly was not as commonly used as today, when most women with breast cancer receive some sort of adjuvant systemic therapy. It is not known whether the benefits of screening and adjuvant therapy are additive. Population-based analyses to tease out the relative effects of screening and treatment are needed”.

I Kreftregisterets oppsummering (42) av sitt syn på den aktuelle debatten (om mammografiscreening) påpekes også at

”.....effekten av screening kan være vanskelig å skille fra eventuelle effekter av adjuvant behandling” (s. 212).

8.6 Nytte-kostnad

Hensikten med en nytte-kostnadsvurdering er å belyse hvordan samfunnet kan bruke begrensede ressurser slik at nytten maksimeres for enkeltindivider og samfunnet som helhet. Siden ressursene er begrensede, vil man få størst nytte for samfunnet ved å prioritere tiltak med lavest kostnad i forhold til nytte. For å kunne foreta denne type prioriteringer kreves det at man sammenligner to eller flere alternativer. I en nytte-kostnadsvurdering av mammografi screening ønsker man derfor å kartlegge hva nytten og kostnaden ved tiltaket er samt hva alternativet til et slikt tiltak er.

Dersom samfunnets eneste mål med mammografi er å maksimere leveår, er det tilstrekkelig å måle nytten som vunne leveår eller redusert dødelighet. Reduksjon i dødelighet og vunne leveår har oftest vært brukt som nyttemål i de analyser som er publisert. Disse nyttemålene omfatter bare konsekvenser for kvinner som har brystkreft, er sant positive ved mammografi.

Dersom samfunnets mål er å maksimere nytte for alle kvinnene som deltar i screening, må analysene også omfatte nytte eller preferanser for alle typer resultater av screening (se pkt. 4.3). Etter mammografi kan kvinnene klassifiseres i fire grupper: sann eller falsk positiv eller negativ. I et slikt utvidet perspektiv er det derfor viktig å vite hvor stor andel av undersøkte kvinner som faller i de fire gruppene samt kvinnes verdsetting av å falle i hver av gruppene.

I så måte vil følgende nytteaspekter / konsekvenser av screening være: Sant positive som ikke lever lenger tross mammografi, får fremskyndet kreftdiagnosen med en slik diagnose. Blant de falskt positive vil testen skape engstelse for en kortere eller lengre periode og kvinnene må ha etterinnkalling, etterundersøkelse og evt. med kirurgi (inkludert de konsekvenser et kirurgisk inngrep medfører). Falsk negativ mammografi kan gi forsinket kreftdiagnose.

Hver av disse tre situasjonene innebærer et nyttetap for kvinnen i motsetning til de sanne negative. De sanne negative er de fleste og vil oppleve en lettelse ved å få bekreftet at de ikke har kreft. Det kan også tenkes at mammografi-screening gir økt oppmerksomhet om brystkreft blant kvinner og en mulig økt bekymring i befolkningen.

Selv om man kan estimere andelen sanne og falskt positive og negative med norske tall, finnes det ingen systematiske undersøkelser av de ulike positive og negative nytteaspektene ved mammografi, og i praksis er det bare evt. vunne leveår som blir brukt i økonomisk evaluering.

På kostnadssiden har målet vært å trekke inn kostnadene for samfunnet som helhet, inklusive individuelle omkostninger for kvinnen.

I litteraturen har alternativet til mammografi for å redusere brystkreftdødelighet vært ikke-screening. Likevel har man i liten grad drøftet hva ikke-screening alternativet innebærer økonomisk. Ved å prioritere screening legger man beslag på ressurser som har alternativ anvendelse. Kostnaden ved å bruke ressurser på screening er målt ved alternativ kostnaden, dvs ved den verdien ressursene ville hatt i beste alternative anvendelse. Hva denne beste alternative anvendelsen er har i liten grad vært drøftet i litteraturen (bare at det er ikke-screening). Et kostnad-effekt-forhold forbundet med mammografi-alternativet som ikke er satt i forhold til konkurrerende alternativer, vil ikke uten videre kunne legges til grunn når beslutninger skal tas.

9. Pågående metodevurderinger og screeningforsøk

I dette kapittelet gis informasjon om prosjekter som omhandler mammografiscreening som SMM har identifisert under arbeidet med utredningen.

Pågående prosjekter i regi av vurderingsorganer (refert i INAHTA-databasen)

- Mammography screening: screening interval for women aged 40 to 69 years; effectiveness in women aged 40 to 49 years (project). Alberta Heritage Foundation for Medical Research (AHFMR).
- Economic and psychosocial evaluation of mammography screening in the county of Funen (co-funded by DIHTA) - primary research, nonsystematic review (project). Danish Centre for Evaluation and Health Technology Assessment.
- Breast cancer mortality in Copenhagen following the introduction of mammography screening (co-funded by DACEHTA) - Primary Research (project). Danish Centre for Evaluation and Health Technology Assessment (DACEHTA) (formerly DIHTA).

Pågående screeningforsøk (populasjonsbaserte studier)

Singapore – aldersgruppe 50-64 år

Dette er en ny RCT i aldersgruppen 50-64 år

Det er identifisert 12 vitenskapelige artikler fra dette forsøket, men ingen om endringer i dødelighet. Forsøket og resultater – basert på surrogate endepunkter – er beskrevet i:

- Ng EH et al. Results of intermediate measures from a population-based, randomized trial of mammographic screening prevalence and detection of breast carcinoma among Asian women: the Singapore Breast Screening Project. *Cancer* 1998;82:1521-8.

England – aldersgruppe 40-41 år

Dette er et viktig pågående forsøk som spesifikt er utformet (designet) slik at studien skal kunne gi svar på om screening i aldersgruppen 40-49 år og med startalder for screening 40-41 år, reduserer dødeligheten av brystkreft.

Det er identifisert 4 vitenskapelige artikler fra dette forsøket. De første resultatene om evt. endringer i dødelighet forventes tidligst i 2003. Forsøket er beskrevet i:

- Moss S. A trial to study the effect on breast cancer mortality of annual mammographic screening in women starting at age 40. Trial Steering Group. *J Med Screen* 1999;6:144-8.

Det må også forventes at det med jevne mellomrom vil komme oppdateringer av de randomiserte studiene.

10. Avsluttende kommentarer

Mammografiscreening har vært omdiskutert i mange år. Debatten har tildels vært polariserende der ”troende” og ”tvilere” steilt har stått på sine standpunkter. Gøtzsches og Olsens Lancet-artikkel (8) og Cochrane-rapporten (10) har initiert nye diskusjoner om nytten ved mammografiscreening og også fått beslutningstagere til å be om nye faglige vurderinger. Begge metaanalysene har vært diskutert og kritisert slik alle vitenskapelige arbeider bør. Uten at man tar stilling til om disse 2 metaanalysene er ”rette” eller ”feilaktige”, er det i ettertid liten tvil om at de har initiert nye vurderinger av temaet, nye analyser og modifiserte resultater. Det er fortsatt muligheter for ytterligere klargjøringer, men dette vil bl.a. kreve tilgang til originaldata (rådata) som til nå ikke har vært tilgjengelige (se pkt. 8.2.1).

SMM har forsøkt å synliggjøre hva den internasjonale uenigheten om mammografiscreening består i, og noen positive og negative helseeffekter ved mammografiscreening. SMM har tillit til at debatten som må føres i Norge om mammografiscreening vil dreie seg om mammografiscreening gir *mer nytte enn skade* og om hvordan kvinnene skal informeres. Det forutsettes selvsagt at data fra det norske screeningprosjektet vil være tilgjengelig for ytterligere studier av ulike aspekter ved mammografiscreening.

Rapporten inneholder ikke en dokumentasjon av kvinners preferanser og valg, og om de følelsesmessige og livskvalitetsmessige aspektene ved mammografi. Dette er tema som det har vært relativt lite fokus på. Forutsetningene for at kvinnene skal kunne foreta selvstendige og begrunnede valg er at informasjonen som gis er nøktern ift. kunnskapen om de positive, men også de negative sidene ved mammografi.

Interessekonflikter (“Conflicts of interests”)

SMM har identifisert disse mulige interessekonfliktene:

Berit Mørland er nestleder i Det vitenskapelige råd, Den Norske Kreftforening (DNK), og er medlem av rådsforsamlingen (General Council) for IARC. Professor Odd Søreide ved SMM er styreleder i DNK.

Vedlegg 1



Sosial- og helsedirektoratet

Senter for Medisinsk Metodevurdering
Forskningsveien 1
Postboks 124 Blindern
0314 Oslo

Deres ref	Vår ref (oppgis ved henv.)	Arkivkode	Dato
	02/4256 tsl/trp		22.04.2002

Mammografiscreening

Sosial- og helsedirektoratet vil med dette be Senter for Medisinsk Metodevurdering om å vurdere nytten av mammografiscreening.

Mammografiscreening er en metode for tidlig å oppdage svulster i brystet. Verdien av screening med hensyn til lengre overlevelse har vært vurdert i et begrenset antall randomiserte forsøk. Disse forsøkene er bl. a. oppsummert i en Cochrane-oversikt. Denne oversikten er kritisk til det metodiske grunnlag studiene hviler på. I tillegg ble oversikten publisert i Lancet med tilleggsdata om økt forekomst av fjerning av brystet i intervensjonsgruppene, et funn som redaktørgruppen i Cochrane Collaboration mente burde underkastes ytterligere peer-review.

SMM bes spesielt vurdere:

- Hva betyr de metodiske svakheter for originalarbeidenes utsagnskraft?
- Er det grunn til å tro at screening også medfører skade? Herunder økt frekvens av aggressiv behandling og negative effekter av falske positive svar.
- Gitt de eksisterende studier, hva er det beste anslag for medisinsk nytteverdi i relative og absolutte tall?
- Hvor pålitelig er dette anslaget? Hvilke feilmarginer bør en regne med?
- Er en eventuell nytteverdi aldersavhengig?
- Såfremt det antas at screening har positiv nytteverdi, anslå nytte-kostnadsforholdet av screening for ulike aldersgrupper med tilhørende usikkerhetsmarginer.
- Kan det påvises at mammografiscreening gjennom tidlig oppdagelse av brystkreft gir økt livskvalitet i sykdomsperioden?
- Har informasjonen som er gitt til kvinner som har deltatt i mammografiscreening vært dekkende, gitt vurderingen av punktene ovenfor.

Av helt spesiell betydning er spørsmålet om brystkreftdiagnosen post mortem er pålitelig eller ikke.

Oppdraget gis etter drøftinger med arbeidsutvalget i Nasjonalt råd for prioriteringer i helsevesenet og Helsedepartementet. Sosial- og helsedirektoratet ønsker selv å tolke disse vurderingene inn i norsk virkelighet og gi helseministeren råd om screeningen bør fortsette eller ikke, og eventuelt i hvilken form.

Det er begrenset antall studier, som alle er vel kjent. Søkearbeidet vil derfor ikke ta lang tid. Vi er imidlertid på det rene med at vurderingene er vanskelige. Samtidig haster det for departementet og Sosial- og helsedirektoratet å få grunnlag for videre vurdering av tilbudet.

Uttalelsene fra SMM ønskes fremlagt som underlag for diskusjoner i Prioriteringsrådet i slutten av august 2002. Vi håper derfor at dere kan ha vurderingen ferdig innen 31.7.02.

Med vennlig hilsen


Frode Forland
Fung. avdelingsdirektør


Tefje Sletnes
Rådgiver

Vedlegg 2

SMMs kommentarer til ekspertvurderingene

SMM har vurdert uttalelsene fra de norske og nordiske ekspertene. Det har vært mange meget nyttige kommentarer, som SMM har benyttet i revisjon av utkastet. Forhåpentligvis er også direkte feil eller misforståelser rettet opp.

SMM vil gi noen ytterligere kommentarer til ekspertenes synspunkter:

- Det er ikke reist spørsmål ved SMMs metodologiske tilnærming til denne metodevurderingen: vurdering av allerede publisert sekundærlitteratur. Det er ikke påvist at vurderingen mangler sekundærlitteratur.
- E. Lynge mener at SMM burde ha referert resultatene fra de siste oppdateringene av de 7 (8) randomiserte undersøkelsene. SMM er enig i at dette kunne være en måte å besvare oppdraget på. SMM har imidlertid valgt en alternativ tilnærming slik det er gjort rede for.
- Flere har vist til primærlitteratur (enkeltstudier) som SMM ikke har vurdert eller lagt til grunn. SMMs utredning baseres på sekundærlitteratur (systematiske litteraturoversikter) slik det er gjort rede for i pkt. 5 (unntaket er de helseøkonomiske studiene; se begrunnelse i teksten). Det er spesielt etterlyst arbeider av Tabar (svensk, hovedansvarlig for "Two County Study" i Sverige). Tabar har ikke publisert sekundærlitteratur selv om konferanserapporten (Falun) gitt i vedlegg 5 ofte refereres til. Møtereferat tilfredstiller ikke de vitenskapelig krav som stilles til sekundærlitteratur. Det er alment kjent at rådata fra Kopparberg ikke er alment tilgjengelig.
- Det er reist spørsmål om søkestrategien som er brukt for søk i litteraturdatabaser. SMM er enig i at strategien kan virke avgrensende (dreier seg om bruken av ordet *and* og ikke *or*). Begrunnelsen for valget var å få et håndterbart søk med relevante vitenskapelige artikler. Metodevurderingen fra FINOHTA som ikke ble identifisert var kjent av SMM, men ble ikke brukt fordi den var skrevet på finsk og dermed vanskelig tilgjengelig. Vår finske referee finner den også lite relevant for vår problemstilling. Kåresen har etterlyst en artikkel om helseøkonomi (ref. 33). Den var identifisert i våre søk, men ikke anført i rapportutkastet som ble sendt ut til vurdering fordi den ble oppfattet som lik artikkelen fra Wang og medarbeidere (ref. 32; for øvrige kommentarer, se Vedlegg 4).
- Det er anført at Gøtzsches og Olsens Lancet-artikkel (8) og Cochrane-utredningen (10) er "uvitenskapelige". Begge arbeidene er etter SMMs mening gjennomført på en transparent måte og ved bruk av aksepterte metoder og analyser. Grunnlaget for klassifiseringen de har foretatt av den vitenskapelig litteratur kan (og bør) selvsagt diskuteres. Mht. effekt på dødelighet må derfor begge arbeidene tillegges vekt på linje med de andre systematiske litteraturoversiktene og metaanalysene gjengitt i pkt. 7.1. Mht. spørsmålet om overbehandling i form av mastektomi og strålebehandling er dette temaet ikke vurdert i denne utredningen (se pkt. 8).
- Flere etterlyser beskrivelse av trender i sykdomsutvikling og behandling over tid. SMM er enig i at dette er viktig informasjon, men slike analyser faller utenfor

rammen for denne vurderingen. Slike vurderinger vil representere primærforskning, og vil heller ikke kunne slå fast om screening er effektiv (reduserer dødelighet) eller ikke. SMM er ikke enig i at all økning i sykdomshyppighet (insidens) må betraktes som overdiagnostikk. Overdiagnostikk må vurderes ift. problemområdene (endepunkter) som er beskrevet i pkt. 6.

- SMM er ikke enig i at populasjonsbaserte screeningprogrammer ("servicescreening") vil kunne gi informasjon om hvor effektiv (nyttig) et screeningprogram er. Med utgangspunkt i endepunktet "brystkreftspesifikk dødelighet" kan sikker informasjon bare hentes fra kontrollerte studier. Men slike screeningprogram vil gi nyttig informasjon om de negative sidene ("harm") ved screening. Denne muligheten må utnyttes i Norge.
- IARC's nye publikasjon om mammografiscreening vil etter SMMs mening neppe være det avgjørende bidrag i diskusjonene om screening. Slik resultatene fremstilles i brosjyren som lanserer monografien som kommer i september 2002, er resultatene innenfor den variasjonsbredde som angis i denne rapporten (se for øvrig fotnote 41).
- SMM registrerer at flere peker på den gode kvaliteten og kvalitetssikringen i det norske mammografiscreeningprosjektet. SMM har ikke vurdert dette punktet, men viser til resultater gitt i pkt. 7.2. SMM er derimot enig i at prosjektet gir gode muligheter for forskning.
- Flere har pekt på at det finnes norsk vitenskapelig litteratur om kvinners reaksjon på screening. Disse artiklene må vurderes sammen med utenlandsk vitenskapelig litteratur om samme tema før en kan trekke sikre konklusjoner
- Flere etterlyser klare konklusjoner og anbefalinger og SMMs syn. Det vises her til SHDir's forepørsel (vedlegg 1) der SHDir selv ønsker å tolke vurderingene fra kunnskapsgrunnlaget inn i norsk virkelighet. For øvrig er det heller ikke SMMs arbeidsform å gi konkrete anbefalinger, råd, eller liknende. SMM skal derimot gi et beslutningsgrunnlag i form av best mulig kunnskap slik den fremkommer i systematiske litteraturvurderinger.

Vedlegg 3

En vurdering av 4 norske artikler om helseøkonomi ved mammografiscreening.

Ekerhovd (1996)

Dette er en hovedoppgave av en samfunnsøkonom utarbeidet med veiledning fra forskere ved Folkehelsa. Analysens hovedoppgave var å estimere kostnad per vunnet leveår ved mammografi screening for kvinner i alderen 50-69 år.

Forfatteren begrunner nytten av mammografi med en metaanalyse av Kerlikowske (JAMA 1995) som konkluderer med at relativ dødsrisiko er 0.74 ved screening av kvinner i aldersgruppen 50-74 år. Leveårsgevinsten er imidlertid basert på stadiespesifikk overlevelse (data fra Kreftregisteret) og antagelser om forbedring av stadiefordelingen ved diagnosetidspunktet etter innføring av mammografi.

Forfatteren tar utgangspunkt i sist registrerte kostnad for en screeningundersøkelse i Hordaland som var bare 38 kroner (marginalkostnad). Dette er langt lavere enn den gjennomsnittskostnaden som er beregnet i Oslo av Wang (et al.) (2001) (50 USD). Den lave kostnaden kan forklares bl.a. med at han har utelatt arbeidskostnader fra beregningen av kostnadene forbundet med selve screeningen. I tillegg har han beregnet reisekostnader (65 kroner per person) og tidskostnader som er verdsatt på basis av produksjonstap etter gjennomsnittlig timelønn. Det antas at tidstapet ved selve mammografien er 2 timer.

Bortsett fra at Ekerhovds analyse er basert på svært lave kostnader ved selve mammografien, har den ingen klare svakheter.

Kåresen et al (1999)

Arbeidet er en cost effectiveness analyse basert på første runde av mammografiscreeningprogrammet i Norge. Kostnadstallene er basert på data fra screeningvirksomheten i Oslo. Effektestimatene er basert på finske data⁶³, og beregner leveårsgevinsten utfra forutsetning om en 30% dødelighetsreduksjon. Med en diskonteringsfaktor på 4,5% beregnes kostnad per vunne leveår til NOK 30.000.

Norum (1999)

Forfatteren beregner kostnad per vunnet leveår basert på norske data. Effekt-estimatet bygger på det faktum at ca 800 kvinner dør årlig av brystcancer i Norge, noe som er ca 38% av insidensen av brystkreft og at screening antas å redusere dødeligheten med 30%.

⁶³ Hristova L. Effect of screening for cancer on mortality, costs and quality of life in Finland. Tammerfors:Acta Universaliensis Tamperensis 1995

Videre antas det at screening vil oppdage 337 cancere. Antallet unngåtte dødsfall er derfor $337 \cdot 0.38 \cdot 0.3$ som så omregnes til leveår på basis av aldersspesifikk levetid. Metoden for beregning av vunne leveår er ikke standard metode, og det er vanskelig å vurdere validiteten av resultatene.

Wang et al. (2001)

Målet har vært å vurdere kostnad per vunnet leveår av det norske screening programmet som startet i 1996 og som dekker kvinner i aldersgruppen 50-69 år.

Nytte eller effekt er målt som vunne leveår. Effekt er beregnet med utgangspunkt i et forventet antall brystkrefttilfeller for kvinnene dersom de ikke screenes (dvs risiko for nye tilfeller av cancer uten screening*antallet). Forventet antall "sparte dødsfall" er så beregnet med utgangspunkt i tall fra Finland (Hristova) hvor man antar at 50% av det forventete antallet brystkrefttilfeller uten screening vil dø og at 30% av disse igjen vil overleve etter 10 år som følge av screening (som tilsvare de sparte/unngåtte dødsfallene). For å beregne vunne leveår ble de sparte leveårene multiplisert med forventet gjennomsnittlig levetid for disse kvinnene basert på tall fra Statistisk sentralbyrå. Forfatterne har diskontert de første ti leveårene (diskonteringsrate 4.5%), men det er uklart om de påfølgende årene har vært diskontert. Metoden for å beregne leveårsgevinsten (Hristova) er heller ikke presentert.

Kostnader inkluderer kostnader ved innkallelse til screening og testing fram til endelig diagnose er bestemt, samt kostnader ved bruk av helsepersonell og kjøp av utstyr og materiale til screening. Likevel foreligger det ingen data i artikkelen som kostnadsanalysen bygger på. Kostnadene er basert på data fra screeningvirksomheten i Oslo. Gjennomsnittskostnad per screening er estimert til 50 USD. Det er rimelig å anta at det vil være stordriftsfordeler i produksjonen av screeningstjenester, slik at kostnadene vil kunne bli høyere i andre deler i landet. Kostnadene ved screening er derfor trolig underestimert. Dessuten er reisekostnader og tidskostnader (tid som går med for individet i forbindelse med behandling målt ved den verdien denne tiden ville hatt for individet i beste alternative anvendelse, f.eks. målt ved et produksjonstap ved at kvinnen ikke er på jobb) utelatt fra kostnadsanalysen.

Avsluttende kommentarer

Arbeidet fra Kåresen og medarbeidere og Wang og medarbeidere er i prinsippet like (med noen felles forfattere). Begge studiene vurderer screening ift. ikke-screening, og begge inkluderer kostnader fram til endelig diagnose er stilt. Reisekostnader og tidskostnader er utelatt i begge studiene. De bruker samme diskoteringsfaktor, men Kåresen et al. gir kostnadsestimatene i NOK, mens Wang et al. bruker USD. Kåresen et al. ser på kostnadseffektiviteten av det norske screeningprogrammet, mens Wang et al. også vurderer framtidig dødelighetsreduksjon. Det er også noen forskjeller i sensitivitetsanalysene.

SMMs vurdering er at Wang et al. er en videreføring av Kåresen og medarbeideres artikkel og legger derfor bare den til grunn for vurderingene.

Vedlegg 4

Resymé og vurdering av helseøkonomisk metaanalyse ved mammografiscreening - Barrat et al. (2002)

Målet har vært å vurdere kostnadseffektiviteten av å utvide screeningprogram for kvinner i alderen 50-69 år i Australia til også å gjelde kvinner over 70 år. Systematisk søk har vært gjort, hvor inklusjons- og eksklusjonskriteriene har vært definert. 5 artikler ble inkludert. Disse sammenligner 70-79 åringer med 50-69 åringer og 3 vurderer også 40-49 åringer mot basisgruppen (50-69 år).

Nytte for hver av gruppene 70-79 år og 40-49 år ble målt som den relative leveårsgevinsten (og også den kvalitetsjusterte leveårsgevinsten i to av studiene) for den aktuelle gruppen i forhold til basisgruppen. For de eldre (70-79 år) varierte dette tallet mellom 40%-72% og mellom 18%-62% når man inkluderte nytteaspekter som ikke knytter seg direkte til helseforbedring i form av vunne leveår. Tilsvarende tall for gruppen 40-49 år var henholdsvis 18%-62% og 92%-80%.

Kostnader inkluderer kostnader ved screening og påfølgende behandling av de kreftsyke.

Kostnad per vunnet leveår ble estimert (og også kostnad per kvalitetsjustert leveår i to av studiene). Resultatet av analysen viser at det er relativt kostnadseffektivt å utvide screeningtilbudet for kvinner i alderen 70-79 år (AUD 8.119-AUD 27.751 per QALY) og at det fremstår som minst like kostnadseffektivt som å utvide tilbudet for kvinner i gruppen 40-49 år (AUD 24.000-AUD 65.000 per vunnet leveår).

Det kan invendes her at dersom studiene hadde bygget på mer moderate antagelser om dødelighetsrisiko, ville dette ha trukket i retning av høyere kostnad-effekt-forhold. Dessuten kompliseres tolkningen av resultatene av at eldre kvinner er mindre flinke til å møte opp enn yngre.

Vedlegg 5. Supplerende litteratur

38. Report of the Organising Committee and Collaborators Falun Meeting. Breastcancer screening with mammography among women aged 40-49 years. Swedish Cancer Society and the Swedish National Board of Health and Welfare. Int J Cancer 1996;68(6):693-9. Dette er en mye sitert publikasjon, men kan ikke legges til grunn fordi arbeidet er et møtereferat).
39. Mushlin AI, Kouides RW, Shapiro DE. Estimating the accuracy of screening mammography: a meta-analysis. American Journal of Preventive Medicine 1998; 14(2):143-153.
40. NHS Cancer Screening Programmes. Informing choice in breast screening. Breast Screening Programme Annual Review 2001.
<http://www.cancerscreening.nhs.uk/breastscreen/index.html>
41. National Cancer Institute. Breast Cancer (PDQ): Screening (Health Professional). Oppdatert 06/2002. www.cancer.gov.
42. Wang H, Hofvind S, Tretli S, Thoresen S, Langmark F. Reduserer organisert mammografiscreening dødelighet av brystkreft? Kreftregisterets syn på den aktuelle debatten. Tidsskr Nor Laegeforen 2002;122(2):211-13.
43. Bandolier Library: Breast Cancer Screening. Feb 2000; 72-73.

Styringsorganene

SMM ble etablert i 1997 og er finansiert av Helsedepartementet og administrert under SINTEF Unimed. En styringsgruppe, oppnevnt av Helsedepartementet, gir faglige innspill og leder virksomheten. Styringsgruppen er bredt sammensatt, og består av 10 personer. I tillegg har SMM etablert et faglig nettverk bestående av omlag 50 eksperter innen de fleste medisinske fagmiljøer. I dette nettverket inngår også jurister, samfunnsvitere, etikere og økonomer.

Styringsgruppen består av følgende personer:

Professor Olav Helge Førde, RiTØ, Universitetet i Tromsø (leder)
Ass. helsedirektør Geir Sverre Braut
Fagdirektør Helge Bryne, Helse Vest RHF
Strategidirektør Finn Henry Hansen, Helse Nord RHF
Direktør Bjørn-Inge Larsen, Sosial- og helsedirektoratet
Allmennlege Kjeld Malde, Lægeforeningens Kvalitetsforbedringsutvalg
Professor Hans Olav Myhre, St.Olavs Hospital/Univ. i Trondheim
Professor Britt-Ingjerd Nesheim, Ullevål sykehus/UiO
Fagdirektør Hanne Thurmer, Helse Sør RHF
Forskningsjef Jon Magnussen, SINTEF Unimed

SMMs sekretariat

Dr. philos Berit Mørland, direktør	22 06 78 08
Prof. dr.med. Odd Søreide	22 06 76 78
Professor Ivar Sønbo Kristiansen	22 06 75 26
Cand. oecon Kristin K. Linnestad	22 06 79 39
Cand. scient Torbjørn Wisløff	22 06 79 69
Dr.scient Anita Lyngstadaas, med.kons	22 06 75 10
Cand.med. Kurt I. Myhre, med. kons	22 06 73 28
Dr. philos. Ellen M. Nilsen, med.kons	22 06 78 39
Dr.philos Inger Norderhaug, med.kons	22 06 78 65
Dr. philos Lise Lund Håheim, med.kons	22 06 74 07
Cand.pharm Krystyna Hviding (perm)	
Cand mag Dagny Fredheim, info.rådgiver	22 06 79 73
Cand.polit Berit Kolberg Rossiné, info.rådgiver (perm)	
Cand.mag Irene Waage, bibliotekar (perm)	
Elisabeth Buntz, bibliotekar	22 06 79 59
Kari Waitz, adm.sekretær	22 06 73 31
Wenche Nesting, sekretær	22 06 79 61
Kari Ballandby, sekretær (perm)	
 Forsker Randi Selmer, Nasjonalt folkehelseinstitutt, Rådgiver	 22 06 79 28

Kontakt SMM:

Senter for medisinsk metodevurdering
SINTEF Unimed
Postboks 124 Blindern
0314 Oslo

Telefon: 22 06 79 61
Telefaks: 22 06 79 79
E-post: smm@unimed.sintef.no
Internett: www.sintef.no/smm



Senter for medisinsk metodevurdering har kritisk vurdert kunnskapsgrunnlaget for
nytteten av mammografiscreening.

Senter for medisinsk metodevurdering
SINTEF Unimed
Postboks 124 Blindern
0314 Oslo

Telefon: 22 06 79 61
Telefaks: 22 06 79 79
E-post: smm@unimed.sintef.no
Internett: www.sintef.no/smm