

Reduserer organisert mammografiscreening dødelighet av brystkreft?

Kreftregisterets syn på den aktuelle debatten

Brystkreft er den klart hyppigste kreftform blant kvinner i Norge, og mer enn 2 300 kvinner fikk diagnostisert sykdommen i 1998 (1). Hvert år dør om lag 800 kvinner av brystkreft, og i underkant av halvparten av dødsfallene rammer kvinner under 70 år (2). Sykdommen utgjør et betydelig helseproblem og oppfyller dermed et av WHO's krav for å innføre et screeningprogram (3). Til tross for kjennskap til mange risikofaktorer, finnes det ingen praktisk anvendbar mulighet for å forebygge brystkreft. En tilnærmet stabil mortalitetsrate i Norge de siste 30 årene indikerer at det så langt ikke kan dokumenteres noe gjennombrudd på behandlingssiden (4).

Prognosen for brystkreftpasienter er sterkt relatert til stadium av sykdommen ved diagnose (1, 4). Ideen bak mammografiscreening er å diagnostisere og behandle flest mulig mens brystkreften er i et tidlig stadium, i håp om at dødeligheten av sykdommen reduseres. Randomiserte studier har vist at dødeligheten av brystkreft reduseres med om lag 30 % blant kvinner i alderen 50–69 år som inviteres til mammografiscreening (5, 6). Effekten blant dem som møter er enda større (7).

Det norske mammografiprogrammet startet som et fireårig prøveprosjekt i fire fylker (Akershus, Hordaland, Oslo og Rogaland) i 1995/96. Med bakgrunn i et stortingsvedtak fra 1998, vil alle landets fylker gi dette til-

budet til kvinner i alderen 50–69 år i løpet av 2002. Kvinnene blir invitert til mammografundersøkelse annethvert år. Mammografiprogrammet er et offentlig helsetilbud, ikke en randomisert studie med en kontrollgruppe. Hovedmålet i Mammografiprogrammet er å oppnå en reduksjon i dødeligheten av brystkreft blant de inviterte på 30 %, dvs. samme størrelsesorden som de randomiserte studiene indikerte. De første resultatene fra prøveprosjektet foreligger i rapporter og artikler (8–10). Tidligindikatorer (svulststørrelse, spredningsstatus osv.) som kan sammenliknes med tilsvarende mål fra de randomiserte studiene, tyder på at programmet vil nå sitt hovedmål.

Effekten av mammografi kan dokumenteres ved randomiserte studier eller ved analyser av dødelighetstrender for brystkreft. Nyelig er dokumentasjonen som finnes om effekt av mammografiscreening blitt trukket i tvil, særlig av Gøtzsche & Olsen (11–13). Hensikten med denne artikkelen er å gjennomgå noe av Gøtzsche & Olsens kritikk samt å klargjøre debatten om dødelighetsreduksjon. Vi vil først drøfte kritikken mot de eksisterende randomiserte studiene, deretter diskutere analyse av dødelighetstrender.

Randomiserte studier

Det eksisterer sju randomiserte studier av mammografiscreening, hvorav fire ble utført i Sverige (5, 6, 12, 14–16) (tab 1). Flere

av disse studiene har også publisert data fra delstudier, slik at de kan oppfattes som to (f.eks. WE-studien). Den første (HIP-studien fra New York) foregikk så tidlig som i midten av 1960-årene, mens den siste (Göteborg-studien) startet i 1982. Protokollene (f.eks. antall bilder per bryst og tid mellom hver undersøkelse) varierte mellom studiene. I de svenske studiene var mammografi eneste undersøkelse, de andre studiene benyttet også klinisk undersøkelse av brystet (5). Det har skjedd mye innen design, innsamling og behandling av data i randomiserte studier siden de første mammografistudier ble igangsatt. Man må derfor forvente å finne noen uklarheter i gamle studier. Å granske disse bringer etter vår oppfatning lite ny kunnskap.

Gøtzsche & Olsen har i flere omganger argumentert for at de fleste av de sju randomiserte studiene ikke holder god kvalitet (11–13). De grupperte studiene i fire kvalitetsklasser: høy, middels, dårlig og uakseptabel. Dette ble gjort med vekt på randomiseringsmetode, eventuelle forskjeller mellom studiegruppene i faktorer som kunne tenkes å påvirke resultatet (f.eks. alder), eksklusjon av brystkrefttilfeller etter randomisering og vurdering av dødsårsak. Deres hovedanalyse inneholdt kun data fra studiene i Canada og Malmö, som ble rangert som de beste studiene (middels kvalitet). Ifølge Greenland kan det «å ekskludere studier på

Tabell 1 Randomiserte studier med mammografi (5, 6, 12, 14–16)

| Studie | Oppstart | Randomiseringsmetode | Intervensjon | Antall i studiegruppen | Antall i kontrollgruppen | Alder (år) | RR (95 % KI) | |
|------------|----------|------------------------------------|-----------------------------------|------------------------|--------------------------|----------------|---------------------|---------------------|
| HIP | 1963 | Matchede par | Mammografi + klinisk undersøkelse | 31 000 | 31 000 | 40–64 | 0,71 (0,55–0,93) | |
| Malmö | 1976 | Individuell | Mammografi | 20 695 | 20 783 | 45–70 | 0,81 (0,62–1,07) | |
| WE-studien | 1977 | Cluster, geografisk | Mammografi | 77 080 | 55 985 | 40–74 | 0,68 (0,59–0,80) | |
| Edinburgh | 1978 | Cluster, geografisk | Mammografi + klinisk undersøkelse | 28 628 | 26 026 | 45–64 | 0,79 (0,6–1,02) | |
| Canada | 1980 | Individuell | Mammografi + klinisk undersøkelse | 44 925 | 44 910 | 40–49 50–59 | 1,36 (0,84–2,21) | 0,97 (0,62–1,52) |
| Stockholm | 1981 | Cluster, fødselsdag Individuell | Mammografi | 38 525 | 20 651 | 40–65 | 0,80 (0,53–1,22) | |
| Göteborg | 1982 | Cluster, fødselsdag | Mammografi | 20 724 | 28 809 | 40–59 | 0,86 (0,54–1,37) | |

bakgrunn av en kvalitetsklassifisering føre til seleksjon basert på fordømmer mot bestemte metoder og rene spekulasjoner om bias» (17). Dersom andre kvalitetskriterier var lagt til grunn, ville kanskje den kanadiske studien blitt ekskludert fordi den ikke var populasjonsbasert (14). Med et utgangspunkt på sju studier trenger man ikke å gjøre noen analyse for å vite hva resultatet blir. Dermed blir det vanskelig å forholde seg til objektiviteten i en kvalitetsbedømmelse som ekskluderer fem av sju studier. For å eksemplifisere kritikken vil vi ta for oss Gøtzsche & Olsens granskning av WE-studien som de klassifiserte til å være av dårlig kvalitet, blant annet på grunn av randomiseringsmetode (13). Det var resultatene fra denne studien som initierte spørsmålet om mammografiscreening i Norge i 1980-årene. Studien foregikk i lenene Kopperberg (W) og Östergötland (E) og inkluderte 133 065 kvinner mellom 40 og 74 år ved randomisering (15). Kvinnene ble randomisert etter geografisk område (cluster). Ifølge de Koning «kan det se ut som Gøtzsche og Olsen betrakter andre randomiseringsmetoder enn individuell, til å være synonymt med lavere kvalitet i utførelse og analyse av studier» (18). Selv om individuell randomisering er å foretrekke fremfor clusterrandomisering, er det ikke alltid praktisk gjennomførbart i store studier (19). Det er også utført en analyse på WE-studien hvor det ble korrigert for clusterrandomisering (20). Resultatene ble ikke nevneverdig endret.

Gøtzsche & Olsen har videre påpekt aldersforskjeller mellom studiegruppen og kontrollgruppen. De fant at studiegruppen i gjennomsnitt var 0,45 år eldre enn kontrollgruppen i Kopperberg, 0,27 år i Östergötland. Det ble ikke kommentert hvorledes denne ulikheten påvirket resultatene, men aldersforskjellen ble brukt som en indikator på at hele WE-studien var av dårlig kvalitet. I 1989 ble det publisert en analyse av WE-studien som tok hensyn til aldersforskjellene (21). Resultatene fra den analysen viste at resultatene ble bedre, dvs. de justerte estimatene var konservative. Aldersforskjellen bidro til en litt høyere totaldødelighet i studiegruppen, som ble borte etter justering (21). Også aldersforskjellen i WE-studien har sammenheng med randomiseringsmetoden som ble brukt.

Videre ble det påpekt at gruppestørrelsen varierte noe i tidlige og senere rapporter. Årsaken til avviket er forklart i artikkelen fra 1989 (21). Det skyldes at kobling av studiepopulasjonen med data fra det svenske krefregisteret ble mulig, slik at alle kvinner som hadde en brystkreftdiagnose før studiestart, ble ekskludert. Fra 1989 og videre er en «ren» populasjon brukt (opprinnelig 134 867, etter kobling 133 065). WE-studien inngikk i den svenske metaanalysen fra 1993 (6). Der ble det brukt eksakt alder ved randomisering, mens det i de individuelle studiene er brukt årskull. Metoden innebar at de kvinne-

ne som ikke hadde fylt 40 år samt de som hadde fylt 74 år før randomiseringsdatoen ble ekskludert. Endelig populasjon for WE-studien ble da 132 588. Det er verdt å merke seg at disse endringene ikke påvirket effektestimaten.

Vurdering av dødsårsak er meget viktig i slike studier. I WE-studien ble dødsårsaken fastlagt med kjennskap til screeningstatus, noe som kunne påvirke resultatene. Det ble derfor opprettet en komité for vurdering av dødsårsak som gjennomgikk dødsårsakene i alle de svenske studiene (6, 22). Resultatene i WE-studien ble ikke påvirket av denne fornyede granskningen. Likevel karakteriserer Gøtzsche & Olsens dødelighetsdataene i WE-studien som upålitelige (12).

Endepunktet i alle studiene var dødelighet av brystkreft. Effekten ble målt blant alle inviterte, ikke bare blant de møtte. Fordi død av brystkreft utgjør i underkant av 5% av alle dødsfall blant kvinner, kan det ikke forventes at total dødelighet påvirkes (2, 23).

Trender i brystkreftdødelighet

På grunn av brystkreftsykdommens natur, vil kvinnene som dør av brystkreft et gitt år, ha stor variasjon i året de fikk diagnosen. Derfor vil brystkreftdødsfallene i årene etter at man har startet opp en masseundersøkelse, bestå både av kvinner som har fått tilbud om mammografi og av kvinner som ikke har fått tilbud om mammografi. Tilgang til individuelle data, med opplysninger om kvinnene er screenet eller ikke, kan være nødvendig for å kunne vurdere effekten på en akseptabel måte (24).

I Norge ble Mammografiprogrammet startet i 1995/96, og effekten skal selvsagt evalueres. Fordi vi har et landsdekkende krefregister, har vi en unik mulighet til å studere effekten av organisert mammografiscreening. Ved Krefregisteret er det etablert en database med individuelle data for alle inviterte. I de randomiserte studiene skilte kurvene for dødelighet av brystkreft seg 5–7 år etter randomisering. Lengre oppfølgingstid enn det vi har i Norge i dag er nødvendig før effekten av en masseundersøkelse kan vurderes (25). Det vil være mulig å studere kontraster i dødelighetstrender mellom fylker som startet mammografiscreening på ulike tidspunkt og mellom aldersgrupper som har hatt ulikt tilbud om screening. Analyser av slike trender vil alltid ha svakheter. Et eksempel på dette er at effekten av screening kan være vanskelig å skille fra eventuelle effekter av adjuvant behandling.

I andre land er effekten av masseundersøkelser med mammografi vurdert med ulike metoder. I Nederland er geografisk kontroll benyttet (uten individuelle data). Der ble det funnet liten effekt (26). I Finland ble programmet innført trinnvis for ulike årskull, noe som gav mulighet for sammenlikning. Det ble funnet en ikke signifikant reduksjon i brystkreftdødelighet på 24% (27). I England har man sammenliknet observert med

predikert mortalitet (uten individuelle data). Resultatene viste 21% reduksjon i dødelighet av brystkreft (ikke oppgitt om signifikant). Kun 6% ble tilskrevet mammografi, resten adjuvant behandling (Tamoxifen) (28). I Sverige har også effekten av masseundersøkelse med mammografi blitt studert. Der ble masseundersøkelsene innført over en lang tidsperiode (1986–97) (29), og etter ulike protokoller. En studie viste ingen effekt (30), men er blitt kritisert for å ha oversett de ovennevnte faktorene (25, 31). En nylig publisert studie sammenlikner svenske len som startet masseundersøkelse til ulik tid (geografisk kontroll). Studien hadde ikke individuelle data. Man fant en ikke-signifikant effekt på 20% (32). En annen studie med individuelle data fra området der WE-studien ble gjennomført, sammenlikner tre perioder (historisk kontroll). Der ble det vist en signifikant effekt på 48% (7). I oktober 2001 offentliggjorde den svenske Socialstyrelsen at dødeligheten av brystkreft for kvinner i alderen 50–64 år var redusert med 21% i perioden 1987–99 (33).

Konklusjon

Vi har i denne artikkelen valgt ikke å drøfte falskt positive og falskt negative testresultater, intervallkreft og faren for overdiagnostisering. Vi innser at man kan finne forhold ved de randomiserte studiene som skaper usikkerhet. Imidlertid omfattet de svenske studiene alle kvinnene i bestemte områder og gode offentlige registre som folkeregister, krefregister og dødsårsaksregister ble benyttet. Vi vurderer derfor disse studienes validitet slik at de gir berettiget tro på en positiv effekt av mammografiscreening. Nye randomiserte studier kan av etiske grunner ikke gjennomføres, fordi indikasjonene av effekt er for klare.

Med dagens kunnskap er svulststørrelse og spredning til armhulens lymfeknuder de viktigste forhold av prognostisk verdi. Gjennom Mammografiprogrammet har andelen små svulster som oppdages økt betraktelig, og andelen pasienter med spredning til armhulens lymfeknuder er nesten halvert sammenliknet med tall fra før den organiserte mammografiscreeningen startet (8, 34). Både på bakgrunn av de fakta vi har presentert i denne artikkelen og de alternativer som kan tenkes, synes stortingsvedtaket fra 1998 om å tilby offentlig mammografiscreening til alle kvinner i alderen 50–69 år, å være god helsepolitikk.

Hege Wang

hege.wang@krefregisteret.no

Solveig Hofvind

Steinar Tretli

Steinar Thoresen

Froydis Langmark

Krefregisteret

Montebello

0310 Oslo

Litteratur →

Litteratur

1. Krefte i Norge 1998. Oslo: Krefteregisteret, 2001.
2. Dødsårsaker 1996. Oslo: Statistisk sentralbyrå, 2000.
3. Wilson JMG, Jungner G. Principles and practice of screening for disease. Genève: WHO, 1968.
4. Wang H, Thoresen SØ, Tretli S. Breast Cancer in Norway 1970–1993: a population based study on incidence, mortality and survival. *Br J Cancer* 1998; 77: 1519–24.
5. Kerlikowske K, Grady D, Rubin SM, Sandrock C, Ernster VL. Efficacy of screening mammography. A meta-analysis. *JAMA* 1995; 273: 149–54.
6. Nyström L, Rutqvist LE, Wall S, Lindgren A, Lindquist M, Ryden S et al. Breast cancer screening with mammography: overview of Swedish randomised trials. *Lancet* 1993; 341: 974–8.
7. Tabár L, Vitak B, Tony HH, Yen MF, Duffy SW, Smith RA. Beyond randomized controlled trials. *Cancer* 2001; 91: 724–31.
8. Ertzaas AK, Hofvind S, Thoresen SØ. Mammografi-programmet i Norge. Krefteregisterets forskningsrapport nr. 2/2000. Oslo: Krefteregisteret, 2001.
9. Wang H, Kåresen R, Hervik A, Thoresen SØ. Mammography screening in Norway: results from the first screening round in four counties and cost-effectiveness of a modelled nationwide screening. *Cancer Causes Control* 2001; 12: 39–45.
10. Wang H, Bjurstam N, Bjørndal H, Braaten A, Eriksen L, Skaane P et al. Interval cancers in the Norwegian Breast Cancer Screening Program: frequency, characteristics and use of HRT. *Int J Cancer* 2001; 94: 594–8.
11. Gøtzsche PC, Olsen O. Is screening for breast cancer with mammography justifiable? *Lancet* 2000; 355: 129–34.
12. Olsen O, Gøtzsche PC. Screening for breast cancer with mammography. *Cochrane Breast Cancer Group*. I: *Cochrane Library*, Issue 4, 2001. Oxford: Update Software, 2001.
13. Olsen O, Gøtzsche PC. Cochrane review on screening for breast cancer with mammography. *Lancet* 2001; 358: 1340–2.
14. Miller AB, Baines CJ, To T, Wall C. Canadian National Breast Screening Study: 1. Breast cancer detection and death rates among women aged 40 to 49 years. *Cancer Causes Control* 1992; 147: 1459–76.
15. Tabár L, Vitak B, Chen HH, Duffy SW, Yen MF, Chiang CF et al. The Swedish Two-County Trial twenty years later. Updated mortality results and new insights from long-term follow-up. *Radio Clin North Am* 2000; 38: 625–51.
16. Alexander FE, Anderson TJ, Forrest APM, Kirkpatrick AE, Prescott RJ, Smith A. 14 years of follow-up from the Edinburgh randomised trial of breast-cancer screening. *Lancet* 1999; 353: 1903–8.
17. Greenland S. Metaanalysis. I: Rothman K, Greenland S, red. *Modern epidemiology*. Philadelphia: Lippincott-Raven, 1998: 643–73.
18. Koning HD. Assessment of nationwide cancer-screening programmes. *Lancet* 2000; 355: 80–1.
19. Duffy SW. Interpretation of the breast screening trials: a commentary on the recent paper by Gøtzsche and Olsen. *Breast* 2000; 9: 1–4.
20. Tabár L, Fagerberg G, Duffy SW, Day N, Gad A, Grøntoft O. Update of the Swedish two-county program of mammographic screening for breast cancer. *Rad Clin North Am* 1992; 30: 187–209.
21. Tabár L, Fagerberg G, Duffy SW, Day N. The Swedish two county trial of mammography screening for breast cancer: recent results and calculation of benefit. *J Epidemiol Community Health* 1989; 43: 107–14.
22. Nyström L, Larsson LG, Rutqvist LE, Lindgren A, Lindqvist M, Ryden S et al. Determination of cause of death among breast cancer cases

- in the Swedish randomized mammography screening trials. A comparison between official statistics and validation by an endpoint committee. *Acta Oncol* 1995; 34: 145–52.
23. Nyström L, Larsson LG, Wall S, Rutqvist LE, Andersson I, Bjurstam N et al. An overview of the Swedish randomised mammography trials: total mortality pattern and the representivity of the study cohorts. *J Med Screen* 1996; 3: 85–7.
 24. Hakama M, Pukkala E, Soderman B, Day N. Implementation of screening as a public health policy: issues in design and evaluation. *J Med Screen* 1999; 6: 209–16.
 25. Rutqvist LE. Mammografiscreeningen minskar dødligheten. Naturalförløppet, grova metoder ledde till felkalkyl om brystcancer. *Läkartidningen* 1999; 96: 1211.
 26. Broeders MJ, Peer PG, Straatman H, Beex LV, Hendriks JH, Holland R et al. Diverging breast cancer mortality rates in relation to screening? A comparison of Nijmegen to Arnhem and the Netherlands, 1969–1997. *Int J Cancer* 2001; 92: 303–8.
 27. Hakama M, Pukkala E, Heikkilä M, Kallio M. Effectiveness of the public health policy for breast cancer screening in Finland: population based cohort study. *BMJ* 1997; 314: 864–7.
 28. Blanks RG, Moss S, McGahan CE, Quinn MJ, Babb PJ. Effects of NHS breast screening programme on mortality from breast cancer in England and Wales, 1990–8: comparison of observed with predicted mortality. *BMJ* 2000; 321: 665–9.
 29. Olsson S, Andersson I, Karlberg I, Bjurstam N, Frodis E, Håkansson S. Implementation of service screening with mammography in Sweden: from pilot study to nationwide programme. *J Med Screen* 2000; 7: 14–8.
 30. Sjønell G, Ståhle L. Hålsokontroller med mammografi minsker inte dødlighet i brystcancer. *Läkartidningen* 1999; 96: 904–13.
 31. Rosen M, Rehnqvist N. Breast screening. No need to reconsider breast screening programme on basis of results from defective study. *BMJ* 1999; 318: 809–10.
 32. Jonsson H, Nyström L, Törnberg S, Lenner P. Service screening with mammography of women aged 50–69 years in Sweden: effects on mortality from breast cancer. *J Med Screen* 2001; 8: 152–60.
 33. Minskad dødlighet i brystcancer for kvinner mellom 50 och 74 år. Pressemelding nr. 64. Stockholm: Sosialstyrelsen, 2001.
 34. Kåresen R, Jacobsen U, Sauer T, Skaane P. En prospektiv registrering av diagnostikk og behandling av tumor mammae ved Ullevål sykehus 1988–95. *Tidsskr Nor Lægeforen* 1999; 119: 3562–6.

Annonsse

○