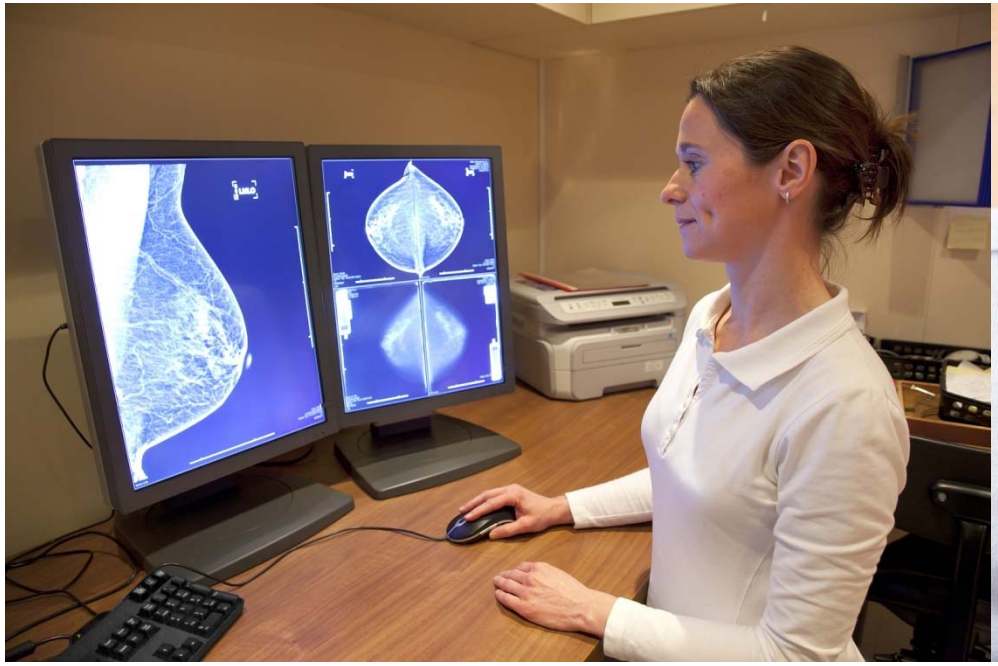
KCE REPORTS 172A

Federaal Kenniscentrum voor de Gezondheidszorg
Centre Fédéral d'Expertise des Soins de Santé
Belgian Health Care Knowledge Centre

BORSTKANKERSCREENING: HOE VROUWEN MET EEN VERHOOGD RISICO IDENTIFICEREN - WELKE BEELDVORMING GEBRUIKEN?



Het Federaal Kenniscentrum voor de Gezondheidszorg

Het Federaal Kenniscentrum voor de Gezondheidszorg is een parastatale, opgericht door de programmawet (1) van 24 december 2002 (artikelen 259 tot 281) die onder de bevoegdheid valt van de Minister van Volksgezondheid en Sociale Zaken. Het Centrum is belast met het realiseren van beleidsondersteunende studies binnen de sector van de gezondheidszorg en de ziekteverzekering.

Raad van Bestuur

	Effectieve Leden	Plaatsvervangende Leden
Voorzitter	Pierre Gillet	
Leidend ambtenaar RIZIV (vice-voorzitter)	Jo De Cock	Benoît Collin
Voorzitter FOD Volksgezondheid (vice-voorzitter)	Dirk Cuypers	Chris Decoster
Voorzitter FOD Sociale Zekerheid (vice-voorzitter)	Frank Van Massenhove	Jan Bertels
Administrateur-generaal FAGG	Xavier De Cuyper	Greet Musch
Vertegenwoordigers Minister van Volksgezondheid	Bernard Lange	François Perl
	Marco Schetgen	Annick Poncé
Vertegenwoordigers Minister van Sociale Zaken	Oliver de Stexhe	Karel Vermeyen
	Ri De Ridder	Lambert Stamatakis
Vertegenwoordigers Ministerraad	Jean-Noël Godin	Frédéric Lernoux
	Daniel Devos	Bart Ooghe
Intermutualistisch Agentschap	Michiel Callens	Anne Remacle
	Patrick Verertbruggen	Yolande Husden
	Xavier Brenez	Geert Messiaen
Beroepsverenigingen van de artsen	Marc Moens	Roland Lemye
	Jean-Pierre Baeyens	Rita Cuypers
Beroepsverenigingen van de verpleegkundigen	Michel Foulon	Ludo Meyers
	Myriam Hubinon	Olivier Thonon
Ziekenhuisfederaties	Johan Pauwels	Katrien Kesteloot
	Jean-Claude Praet	Pierre Smiets
Sociale partners	Rita Thys	Leo Neels
	Paul Palsterman	Celien Van Moerkerke
Kamer van Volksvertegenwoordigers	Maggie De Block	

Controle

Regeringscommissaris	Yves Roger

Directie

Algemeen Directeur	Raf Mertens
Adjunct Algemeen Directeur	Jean-Pierre Closon
Programmadirectie	Christian Léonard Kristel De Gauquier

Contact

Federaal Kenniscentrum voor de Gezondheidszorg (KCE)

Doorbuilding (10e verdieping)
Kruidtuinlaan 55
B-1000 Brussel
Belgium

T +32 [0]2 287 33 88

F +32 [0]2 287 33 85

info@kce.fgov.be

http://www.kce.fgov.be

KCE REPORTS 172A
GOOD CLINICAL PRACTICE

Federaal Kenniscentrum voor de Gezondheidszorg
Centre Fédéral d'Expertise des Soins de Santé
Belgian Health Care Knowledge Centre

BORSTKANKERSCREENING: HOE VROUWEN MET EEN VERHOOGD RISICO IDENTIFICEREN - WELKE BEELDVORMING GEBRUIKEN?

VERLEYE LEEN, DESOMER ANJA, GAILLY JEANNINE, ROBAYS JO

COLOFON

Titel:	Borstkankerscreening: hoe vrouwen met een verhoogd risico identificeren - welke beeldvorming gebruiken?
Auteurs:	Leen Verleye (KCE), Anja Desomer (KCE), Jeannine Gailly (KCE), Jo Robays (KCE)
Reviewers:	Dominique Paulus (KCE), Sabine Stordeur (KCE), Pascale Jonckheer (KCE), Joan Vlayen (KCE)
Externe experten:	Martine Berliere (UCLouvain), Joelle Desreux (CHR Citadelle), Valérie Fabri (Mutualités Socialistes), Eric Legius (KULeuven), Patrick Neven (UZLeuven), Anne Remacle (Mutualité Chrétienne), Mireille Van Goethem (UZA), Geert Villeirs (UGent), Joost Weyler (UA), Erwin De Clerck (Vlaamse Liga tegen Kanker), André-Robert Grivegnée (Institut Jules Bordet), Myriam Provost (Médibou)
Externe Validatoren:	Geert Page (RZH Jan Yperman), Chantal Van Ongeval (UZ gasthuisberg Leuven), Alicia Framarin (INESSS, Canada)
Belangenconflict:	Geert Page (Vergoeding voor spreekbeurten over gynaecologische en verloskundige onderwerpen alsook EBM. Betaling voor deelname aan symposia op vlak van perinatale geneeskunde), Patrick Neven (Financial support attending meeting/speaker fee)
Layout:	Ine Verhulst, Sophie Vaes
Disclaimer:	• **De externe experten werden geraadpleegd over een (preliminaire) versie van het wetenschappelijke rapport. Hun opmerkingen werden tijdens vergaderingen besproken. Zij zijn geen coauteur van het wetenschappelijke rapport en gingen niet noodzakelijk akkoord met de inhoud ervan.** • **Vervolgens werd een (finale) versie aan de validatoren voorgelegd. De validatie van het rapport volgt uit een consensus of een meerderheidsstem tussen de validatoren. Zij zijn geen coauteur van het wetenschappelijke rapport en gingen niet noodzakelijk alle drie akkoord met de inhoud ervan.** • **Tot slot werd dit rapport unaniem goedgekeurd door de Raad van Bestuur.** • **Alleen het KCE is verantwoordelijk voor de eventuele resterende vergissingen of onvolledigheden alsook voor de aanbevelingen aan de overheid.**
Publicatiedatum:	4 juni 2012 (2nd edition ;1st edition 10 januari 2012)
Domein:	Good Clinical Practice (GCP)
MeSH:	Breast Neoplasms ; Risk Factors ; Early Detection of Cancer ; Mammography ; Ultrasonography, Mammary ; Magnetic Resonance Imaging ; Risk Assessment ; Mass Screening
NLM classificatie:	WP870

Taal: Nederlands, Engels

Formaat: Adobe® PDF™ (A4)

Wettelijk depot: D/2011/10273/90

Hoe refereren naar dit document? Verleye L, Desomer A, Gailly J, Robays J. Borstkankerscreening: hoe vrouwen met een verhoogd risico identificeren - welke beeldvorming gebruiken? Good Clinical Practice (GCP). Brussel: Federaal Kenniscentrum voor de Gezondheidszorg (KCE). 2011. KCE Reports 172A. D/2011/10.273/90

Dit document is beschikbaar op de website van het Federaal Kenniscentrum voor de Gezondheidszorg.

VOORWOORD

De vroegtijdige opsporing van borstkanker blijft een belangrijke uitdaging voor ons gezondheidsbeleid. Maar het is evengoed een zeer controversieel onderwerp waarbij de discussies soms hoog oplaaien. Zelfs al kon het KCE in een vorig jaar gepubliceerde studie een systematische screening voor vrouwen beneden de 50 jaar niet aanbevelen, bleef de vraag of dit ook geldt voor vrouwen met een verhoogd risico. En indien deze risicogroep inderdaad wel baat heeft bij screening, hoe ga je deze groep dan aflijnen?

Een tweede heikel punt is de vraag of naast de screeningsmammografie met dubbele lezing er ook een plaats is voor de echografie of een MRI als opsporingsonderzoek. Het is, met andere woorden, een vraag naar de balans tussen een eventuele betere opsporing van tumoren en de risico's van vals positieve resultaten, overdiagnose en uiteindelijk overbehandeling die ermee gepaard gaan.

Het KCE hanteert zoals gebruikelijk zijn kritische benadering van de gepubliceerde onderzoeken en richtlijnen, aangevuld door onderzoek van de Belgische zorg-consumptie-data. Voor dit laatste aspect gaat onze oprechte dank naar de medewerkers van het IMA die ons met hun ruime expertise ter zake hebben bijgestaan.

Tot slot willen wij ook alle experten van 'het terrein' bedanken die dank zij hun kritische inbreng mee hebben geholpen om dit rapport gestalte te geven, en op die manier hebben bijgedragen, zo hopen wij, tot een meer evidence-based opsporingsbeleid.

Om deze reeks te vervolledigen, brengt het KCE in het voorjaar van 2012 een derde rapport in deze reeks uit, met name over borstkankerscreening bij vrouwen ouder dan 70 jaar.

Jean-Pierre CLOSON
Adjunct Algemeen Directeur

Raf MERTENS
Algemeen Directeur

SAMENVATTING

INTRODUCTIE

Borstkankerscreening heeft als doel borstkanker op te sporen in een vroeg, preklinisch stadium wanneer de kanker een goede prognose heeft. Dit heeft in principe een positieve impact op zowel de borstkankergerelateerde mortaliteit als de morbiditeit van de behandeling. Borstkankerscreening heeft echter ook een aantal negatieve effecten, zoals de morbiditeit veroorzaakt door 'overdiagnose' (gedefinieerd als het diagnosticeren van kankers die klinisch niet zouden evolueren en niet zouden leiden tot sterfte) en door 'vals positieve' resultaten (een screeningsonderzoek is vals positief als een letsel wordt gezien wanneer er geen kanker aanwezig is). Voor- en nadelen dienen dus zorgvuldig te worden afgewogen.

Het KCE publiceerde in het verleden twee rapporten over borstkankerscreening. Een rapport uit 2005 betrof borstkankerscreening in het algemeen bij vrouwen zonder risicofactoren. In 2010 werd een gedeeltelijke update gepubliceerd over borstkankerscreening bij vrouwen in de leeftijdsgroep 40-49 jaar zonder risicofactoren.

Cijfers van het intermutualistisch agentschap (IMA), gebaseerd op nomenclatuurgegevens, tonen dat er in België veel opportunistische screening plaatsvindt, gefactureerd als 'diagnostische mammografie'. We schatten dat tussen de 80 en 90% van de mammografieën aangerekend als 'diagnostisch' in werkelijkheid screeningsmammografieën buiten de georganiseerde screening betreffen. Het valt op dat 85% van deze mammografieën vergezeld wordt door een echografie die op dezelfde dag wordt uitgevoerd. Dit betekent dat echografie in België, voornamelijk in Wallonië en Brussel, vaak als screeningsmethode wordt ingezet. Dit kan mogelijk mede verklaren waarom het aantal puncties en biopsies per 100 000 vrouwen per jaar hoog is in België. Omgerekend worden 5.5 puncties of biopsies uitgevoerd per kankerdiagnose.

Verder blijkt dat ook voor 15% van de Belgische vrouwen tussen 35 en 39 jaar en 37% van de vrouwen tussen 40 en 49 jaar een diagnostische mammografie werd aangerekend. De precieze indicatie voor deze mammografieën kan niet uit de IMA-gegevens afgeleid worden. Het aandeel van klinische symptomen, screening van vrouwen met een

(vermeend) verhoogd risico op borstkanker en opportunistische screening bij vrouwen zonder verhoogd risico is bijgevolg niet duidelijk.

Deze gegevens roepen een aantal vragen op waarop dit rapport een antwoord wil formuleren. Het risico op borstkanker is niet homogeen verspreid over de ganse bevolking. Een aantal vrouwen maakt meer kans om in de loop van hun leven borstkanker te krijgen door familiale belasting en andere factoren. De vraag stelt zich bijgevolg welke factoren dit zijn, hoe hoog dit verhoogd risico is, op welke basis men vrouwen kan indelen volgens risico en wat de optimale screeningsstrategie is per risicogroep of profiel. Daarnaast onderzoekt dit rapport welke de optimale screeningsmethodes zijn voor vrouwen zonder en vrouwen met een verhoogd risico op borstkanker.

ONDERZOEKSVRAGEN

Welke vrouwen lopen extra risico op borstkanker en hoe kan dit risico bepaald worden?

Familiaal risico

Hoe kunnen vrouwen met een familiaal risico geïdentificeerd worden?

Welke instrumenten en modellen bestaan hiervoor en wat is hun validiteit en toepasbaarheid in de Belgische context?

Niet-familiaal risico

Welke niet-familiale risicofactoren dienen in overweging te worden genomen?

Wat is het relatief risico op borstkanker voor deze risicofactoren en welk absoluut risico is hiermee verbonden?

Hoe kunnen familiaal en niet- familiaal risico met elkaar gecombineerd worden?

Welke instrumenten en modellen bestaan hiervoor en wat is hun validiteit en toepasbaarheid in de Belgische context?

Wat is de waarde van de technische methoden die gebruikt worden tijdens de screening?

Wat is de accuraatheid en impact op morbiditeit en mortaliteit van:

- Mammografie met dubbele lezing en ‘computer aided detection’ (CAD)
- Digitale mammografie
- Echografie
- MRI scan

Wat is de optimale diagnostische strategie per risicogroep?

Wat is geweten over de voordelen en de nadelen?

METHODOLOGIE

Voor de beschrijving van de Belgische situatie werd samengewerkt met het IMA, dat gegevens leverde over alle leeftijdsgroepen in aanvulling van hun rapport uit 2010 over borstkankerscreening in België in de leeftijdsgroep 50-69 jaar.

Voor de aanbevelingen werd de ADAPTE-methodologie gebruikt, waarbij internationale praktijkrichtlijnen aan de Belgische context werden aangepast. Hiervoor werd gezocht in Medline, National Guideline Clearinghouse en websites van richtlijnorganisaties en oncologische organisaties. De gevonden richtlijnen werden door middel van het AGREE-instrument beoordeeld op hun kwaliteit door twee reviewers. Op basis hiervan werden 1 richtlijn en 1 HTA rapport over risicofactoren geselecteerd en voor de relevante klinische vragen geüpdatet door bijkomend bewijsmateriaal te zoeken in Medline, EMBASE en de Cochrane Database of Systematic Reviews.

Voor elke technische screeningsmethode werd ook een aparte zoekstrategie gebruikt. De meest recente zoekstrategieën werden in mei 2011 uitgevoerd.

Een niveau van bewijskracht werd toegekend aan elke aanbeveling door gebruik te maken van het GRADE-systeem. Op basis van het gevonden bewijsmateriaal stelde de multidisciplinaire richtlijnontwikkelingsgroep (i.e. de auteurs van deze richtlijn) de aanbevelingen op. Een review van deze

aanbevelingen werd uitgevoerd door externe experts. Belangenconflicten werden genoteerd.

RESULTATEN

Welke vrouwen hebben een verhoogd risico op borstkanker?

De belangrijkste risicofactor is het voorkomen van borstkanker in de familiale voorgeschiedenis. Op basis van familiale anamnese worden drie risicogroepen onderscheiden: gemiddeld, verhoogd en sterk verhoogd risico. Voor een overzicht van deze drie risicogroepen verwijzen we naar de aanbevelingen.

Personen die op jonge leeftijd radiotherapie met mantelveld ondergingen hebben een sterk verhoogd risico op borstkanker en ook vrouwen met zeer dens borstweefsel (BIRADS 4) kunnen tot de categorie met een verhoogd risico gerekend worden (levenslang risico +/- 17%).

Voorgeschiedenis van een precancereus letsel, zoals ductaal of lobulair carcinoma in situ gaat gepaard met een verhoogde kans op borstkanker. Opvolging en behandeling van deze letsels worden niet besproken in dit rapport, hiervoor verwijzen wij naar de nationale richtlijnen voor borstkanker in KCE rapport 143C.

Andere risicofactoren, zoals dens borstweefsel BIRADS 3, obesitas, alcoholgebruik, hormonale substitutietherapie, vroege menarche, nulliparitеit, hormonale contraceptie of gebruik van andere exogene hormonen (bv. diethylstilbestrol of DES) hebben slechts een beperkte invloed op risico op borstkanker. Op basis van deze risicofactoren hoeven vrouwen geen screeningsonderzoeken te ondergaan buiten het georganiseerde bevolkingsonderzoek. In de praktijk worden ze enkel gebruikt binnen veelomvattende risicomodellen die op basis van familiale en niet-familiale factoren een individueel risico berekenen.

Door de talrijke interacties en overlappingen tussen verschillende risicofactoren dient men voorzichtig te zijn met het combineren van factoren. Daarom zijn in de loop der tijd twee types modellen ontwikkeld die hiermee rekening houden. Een eerste type model voorspelt het individueel risico op borstkanker (hetzij het risico op borstkanker in de komende 5 of 10 jaar, hetzij het levenslang risico) en wordt gebruikt om een vrouw in een risicogroep in te delen. Een ander type model voorspelt het risico op een genetisch mutatie die een vrouw sterk voorbeschikt op borstkanker (vooral BRCA1 & 2 en TP53) en wordt gebruikt om te bepalen wie in aanmerking komt voor genetische tests. Dit is nodig omwille van de hoge kost van deze laatste. Sommige modellen kunnen voor beide doeleinden worden gebruikt.

Het Gail-model is het meest bestudeerde model, maar heeft een aantal nadelen, zoals het feit dat het risico in sommige gevallen te laag ingeschat wordt en dat slechts een beperkt aantal risicofactoren in rekening gebracht wordt. Het Tirer-Cuzick (IBIS) model brengt meer risicofactoren in rekening, maar niet de borstdensiteit. Er zijn ook aanwijzingen dat het wat accurater en beter gekalibreerd is dan het Gail-model maar dit dient nog bevestigd te worden. Er zijn recente pogingen ondernomen om borstdensiteit in de modellen op te nemen, zoals in het zogenaamde Tice model, maar er is nog geen onafhankelijke validatie van die modellen gebeurd. Het vermogen van modellen die het risico op een genetische mutatie voorspellen, is matig en vergelijkende studies tonen niet aan dat de accuraatheid of voorspellende waarde van het ene model beter is dan het andere. We kunnen dan ook geen uitspraken doen over welk model nu beter is dan het andere.

Er zijn geen studies die direct de impact aantonen van genetische screening op mortaliteit of aan borstkanker gerelateerde morbiditeit.

Screeningsmethoden

Dubbele lezing door twee onafhankelijke lezers en beslissing op basis van consensus of arbitrage verhoogt de gevoeligheid van borstkankerscreening: er is een stijging van de kanker detectie (stijging met 2.9-11.2 per 10.000 gescreende vrouwen) en een daling van het aantal teruggeroepen vrouwen (daling met 38-149 per 10 000 gescreende vrouwen).

Het interpreteren van mammografieën met behulp van computerdetectie, vergeleken met mammografie met een enkele lezing, geeft een beperkte verhoging van de gevoeligheid maar gaat gepaard met een toename van het aantal vals positieven. Er zijn geen studies die aantonen dat deze techniek een toegevoegde waarde heeft ten opzichte van dubbele lezing.

Analoge en digitale mammografie kunnen als evenwaardig beschouwd worden voor het detecteren van borstkanker. Het gebruik van digitale

mammografie kan voordelig zijn voor jonge vrouwen en vrouwen met borstweefsel met hoge densiteit.

Het gebruik van echografie in het bevolkingsonderzoek naar borstkanker bij een (op risico) ongeselecteerde populatie werd in geen enkele studie bestudeerd. Er bestaan alleen studies bij vrouwen met verhoogd risico. Er worden slechts weinig extra borstkankers opgespoord en het aantal extra onderzoeken en valse positieve resultaten is hoog.

MRI verhoogt in belangrijke mate de gevoeligheid bij vrouwen met hoog risico, met een gevoeligheid variërend van 68 tot 100 %. Het aantal doorverwijzingen voor verder onderzoek (recall rate) kan oplopen tot 24 %. De positief voorspellende waarde van een positieve MRI blijft echter hoog in die groep (39 – 58%) (ter vergelijking, in Vlaanderen lag de positief voorspellende waarde van een positieve mammografie in een vervolgscreening tussen de 14 en de 19 %).

Screeningsmethoden per risicogroep

Er zijn geen studies die rechtreeks meten welke impact het gebruik van de verschillende technologieën en het uitbreiden van screening naar jongere leeftijd bij (sterk) verhoogd risico heeft op morbiditeit en mortaliteit.

DISCUSSIE

We vonden alleen wetenschappelijke studies over risicofactoren, risicomodellen en diagnostische validaties van screeningstechnieken, maar niet over de rechtstreekse impact van screeningsstrategieën op mortaliteit of morbiditeit. De aanbevelingen zijn bijgevolg gebaseerd op indirect bewijs en expertopinie. Risicomodellen kunnen een toegevoegde waarde hebben, maar zijn nog in ontwikkeling. Het is daarom nog te vroeg om uitspraken te doen over welk model nu het beste is.

AANBEVELINGEN[ab]

Welke vrouwen moeten beschouwd worden als hebbende een verhoogd risico op borstkanker?

- **Een risicobepaling moet eerst het onderscheid maken tussen vrouwen die een risico hebben vergelijkbaar met de algemene populatie en zij die een verhoogd risico hebben. Dit moet in de eerste plaats gebeuren op basis van een eenvoudige familiale anamnese.**
- **Bij vrouwen met een verhoogd risico kan een grondigere risicobepaling volgen om individueel advies te kunnen geven over de screeningstrategie, genetische tests en profylactische maatregelen. Dergelijke individuele risicobepaling moet steeds besproken worden met de vrouw, rekening houdend met alle mogelijke maatregelen, limieten, onzekerheden en alternatieven.**

[a] Alleen het KCE is verantwoordelijk voor de aanbevelingen aan de overheid

[b] GRADE was gebruikt voor de aanbevelingen, zie samenvatting en appendix

HOE HET INDIVIDUELE RISICO BEPALEN? [c]

A. De belangrijkste risicofactor is familiale belasting.

1. Op basis van eenvoudige familiale anamnese kunnen vrouwen in drie risicogroepen ingedeeld worden, *(sterke aanbeveling, matig niveau van bewijs)*:

Gemiddeld risico:

- **Geen of één eerste- of tweedegraads familielid met borstkanker, waarbij de diagnose werd gesteld op een leeftijd ouder dan 40 jaar.**

Verhoogd risico (dit betekent een 10-jaar risico tussen 3 en 8 procent voor vrouwen van 40 tot 49 jaar of een levenslang risico op borstkanker tussen 17 en 29%)

- **Eén eerstegraads familielid met borstkanker gediagnosticeerd op een leeftijd jonger dan 40 jaar**

of

- **Twee eerste- of tweedegraads familieleden met de diagnose van borstkanker op een gemiddelde leeftijd boven de 50 jaar**

of

- **Drie eerste- of tweedegraads familieleden gediagnosticeerd met borstkanker op een gemiddelde leeftijd ouder dan 60 jaar**

Sterk verhoogd risico (dit betekent een 10-jaar risico hoger dan 8% procent voor vrouwen van 40 tot 49 jaar of een levenslang risico op borstkanker van 30% of hoger)

- **Twee eerste- of tweedegraads familieleden met de diagnose van borstkanker op een gemiddelde leeftijd jonger dan 50 jaar waarbij minstens 1 van de 2 familieleden eerstegraads verwant is**

of

- **Drie eerste- of tweedegraads familieleden met de diagnose van borstkanker op een gemiddelde leeftijd jonger dan 60 jaar waarbij minstens 1 van de 3 familieleden eerstegraads verwant is**

of

- **Vier familieleden met borstkanker onafhankelijk van de leeftijd bij diagnose waarbij minstens 1 van de 4 familieleden eerstegraads verwant is**

c Borstkanker bij de vrouw zelf als risicofactor valt onder opvolging van borstkanker na behandeling en word in dit rapport niet besproken

of

- **van Joodse afkomst zijn**

of

- **één van de volgende elementen in de familiale voorgeschiedenis:**
  - **bilateraal borstkanker**
  - **borstkanker bij een man**
  - **ovariumcarcinoom**
  - **sarcoma gediagnosticeerd op een leeftijd jonger dan 45 jaar**
  - **glioma of carcinoma van bijnierschors op kinderleeftijd**
  - **patroon van multipele carcinomen op jonge leeftijd**
  - **sterke familiale belasting in de vaderlijke stamboom (vier familieleden met diagnose van borstkanker op een leeftijd jonger dan 60 jaar aan vaderlijke zijde)**

2. Vrouwen met op basis van familiale anamneses een sterk verhoogd risico moeten een individuele risicobepaling aangeboden krijgen gevolgd door overleg over de screenings-strategie, eventuele genetische tests of profylactische maatregelen.

- **Het bepalen van het individueel risico omvat een grondige familiale anamnese en eventueel een gevalideerde, gecomputeriseerde schaal, zoals bijvoorbeeld het Gail model of het Tirer-Cuzick model. Andere modellen die ook rekening houden met de densiteit van het borstweefsel, zoals bijvoorbeeld het Tice model, zijn nog onvoldoende gevalideerd.**
- **Deze risicobepaling zou moeten uitgevoerd worden door professionelen met voldoende expertise terzake en gepaard gaan met uitgebreide counseling en aandacht voor de voorkeuren van de patiënt en voldoende ondersteuning**

(zwakke aanbeveling, zeer zwak niveau van bewijs).

B Andere risicofactoren

3. Personen die op jonge leeftijd radiotherapie met mantelveld ondergingen moeten ingedeeld worden in de groep van vrouwen met een sterk verhoogd risico op borstkanker *(sterke aanbeveling, matig niveau van bewijs)*

4. Vrouwen met zeer dens borstweefsel (BIRADS 4) kunnen tot de categorie met een verhoogd risico gerekend worden (levenslang risico +/- 17%) *(zwakke aanbeveling, zeer zwak niveau van bewijs)*

5. Ductale of lobulaire atypische hyperplasie zou moeten worden beschouwd als sterk verhoogd risico *(zwakke aanbeveling, zwak niveau van bewijs)***.**

6. Screeningsonderzoeken buiten het bevolkingsonderzoek worden niet aanbevolen op basis van risicofactoren zoals dens borstweefsel BIRADS 3, obesitas, alcoholgebruik, hormonale substitutietherapie, vroege menarche, nulliparitеit, hormonale contraceptie of andere exogene hormonen (bv. Diethylstilbestrol of DES). In de praktijk zijn deze risicofactoren enkel te gebruiken binnen een geïntegreerd risicomodel omdat hun invloed op het risico op borstkanker slechts beperkt is. *(sterke aanbeveling, zwak niveau van bewijs)*.

WELKE SCREENINGSTECHNIEKEN GEBRUIKEN?

7. Elke screeningsmammografie moet beantwoorden aan de Europese kwaliteitsvereisten en beoordeeld worden door twee onafhankelijke lezers. Bij verschillende interpretatie wordt de eindbeslissing genomen op basis van consensus of arbitratie *(sterke aanbeveling, sterk niveau van bewijs)***.**

8. Interpretatie van mammografieën met behulp van computerdetectie is niet aanbevolen en kan niet de dubbele lectuur zoals beschreven in aanbeveling 7 vervangen *(sterke aanbeveling, zeer zwak niveau van bewijs)***.**

9. Zowel analoge als digitale mammografie zijn aanbevolen technieken voor het vroegtijdig detecteren van borstkanker. Het gebruik van digitale mammografie kan voordelig zijn voor jonge vrouwen en vrouwen met dens borstweefsel *(zwakke aanbeveling, zeer zwak niveau van bewijs)***.**

10. Het gebruik van echografie is niet aanbevolen in het bevolkingsonderzoek naar borstkanker omdat er slechts weinig extra borstkankers opgespoord worden en het aantal extra onderzoeken en valse positieve resultaten te hoog is *(sterke aanbeveling, zwak niveau van bewijs)***.**

11. Op basis van de beschikbare gegevens is het eveneens niet aanbevolen om bij vrouwen met dens borstweefsel de echografie te gebruiken als screeningsonderzoek. Screening door middel van echografie bij vrouwen met zeer dens borstweefsel (BIRADS 4) wordt niet aanbevolen buiten het kader van kinische studies *(sterke aanbeveling, zwak niveau van bewijs)***.**

12. Bij vrouwen met een verhoogd risico op borstkanker is het aanbevolen een jaarlijkse mammografie aan te bieden vanaf 40 tot 49 jaar, uitgevoerd volgens de Europese richtlijnen en kwaliteitsvereisten. Vanaf de leeftijd van 50 – 69 jaar kunnen vrouwen met een verhoogd risico deelnemen aan het bevolkingsonderzoek met tweejaarlijks mammografie *(zwakke aanbeveling, zeer zwak niveau van bewijs)***.**

13. Voor vrouwen met een bewezen sterk verhoogd risico op borstkanker wordt jaarlijks MRI en mammografie aanbevolen vanaf de leeftijd van 30 jaar of 5 jaar vóór de leeftijd van het jongst gediagnosticeerde familielid *(sterke aanbeveling, zeer zwak niveau van bewijs)***. Het gebruik van echografie kan eveneens overwogen worden bij deze risicocategorie, bijvoorbeeld om het interval te verkorten of als aanvullend onderzoek bij positieve MRI of mammografie** *(zwakke aanbeveling, zwak niveau van bewijs)***.**

14. Alle vrouwen die deelnemen aan een screening moeten geïnformeerd worden over de mogelijkheid van vals positieve onderzoeken, het blijvend risico op intervalkanker en het ontbreken van data die een invloed op morbiditeit en mortaliteit aantonen voor screening buiten het algemeen bevolkingsonderzoek *(sterke aanbeveling, zeer zwak niveau van bewijs)*

SCIENTIFIC REPORT

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

ABBREVIATION	DEFINITION
BIRADS	Breast Imaging Reporting and Data System
BIRADS M	Breast Imaging Reporting and Data System for Mammography
BIRADS-US	Breast Imaging Reporting and Data System for Ultrasound
BRCA1-BRCA2	BReast CAncer (susceptibility genes)
CBE	Clinical breast examination
CI	Confidence Interval
CS	Cohort study
E/O	Expected-to-Observed rate
FFDM	Full-Field Digital Mammography
FNA	Fine needle aspiration
FU	Follow-up
GIN	Guidelines International Network
HTA	Health Technology Assessment
IMA	Intermutualistic Agency
INAMI/RIZIV	National Institute for Health and Disability Insurance
KCE	Belgian Healthcare Knowledge Centre
MA	Meta-analysis
MRI	Magnetic Resonance Imaging
MD	Diagnostic mammography
MT	Mammotest, screening mammography
Mx	Mammography
NCI	National Cancer Institute (United States)
NICE	National Institute for Health and Clinical Excellence (England and Wales)
NIS	National Institute for Statistics
NPV	Negative Predictive Value
NZ	New-Zealand

NZHTA	New Zealand Health Technology Assessment
OC	Oral contraceptive
pBSO	prophylactic bilateral salpingo-oophorectomy
PPV	Positive Predictive Value
QoL	Quality of Life
RCT	Randomized Controlled Trial
Sens	Sensitivity
Spec	Specificity
SR	Systematic Review
TP53	Tumor Protein 53
UK	United Kingdom
US	Ultrasound
USA	United States of America

.

INTRODUCTION

There are a lot of discussions between scientific experts in Belgium but also at an international level about breast cancer screening.. Clinical questions concern the necessity to screen younger or older women, the choice of the technical methods used for screening, the inclusion of women at higher risk of breast cancer in an organized screening program, the need of specific technical screening in case of women with high breast density.

To select the most important questions, the Belgian Healthcare Knowledge Centre (KCE) organized a stakeholder consultation.

1. STAKEHOLDERS' REPRESENTATIVES

Representatives of following stakeholders' organizations were invited to collaborate:

- Gynaecologists : Vlaamse Vereniging voor Obstetrie en Gynaecologie (VVOG) and Groupement des Gynécologues Obstétriciens de Langue Française de Belgique, (GGOLF)
- General practitioners: Société Scientifique de Médecine Générale (SSMG) and Domus Medica (Domus),
- Radiologists: Royal Belgian Society of Radiology (RBSR),
- Patients: Ligue des Usagers des Services de Santé (LUSS) and Vlaams Patiëntenplatform (VPP),
- Associations against cancer: Fondation contre le cancer/ Stichting tegen kanker and Vlaamse Liga tegen Kanker (VLK),
- National Institute for Health and Disability Insurance (INAMI/RIZIV),
- Belgian organized breast cancer screening programs: Brumammo (Bruxelles), Centre Communautaire de Référence pour le dépistage des cancers (CCRef) (Communauté Française) and BorstKankerOpsporing (BKO) (Vlaamse Gemeenschap).

The Vlaams Patiëntenplatform (VPP) chose to be represented by the VLK because they have no specific group dealing with breast cancer screening.

The KCE sought advice from the stakeholders at two moments: for the selection of questions before the literature search, and at the end of the process for the formulation of recommendations.

2. SELECTION OF CLINICAL QUESTIONS

First, KCE experts listed several clinical questions relative to breast cancer screening. Then, the stakeholders were invited to review the choice and the formulation of the questions and put forward priority questions to be investigated.

The selected questions were then split up in several KCE reports. A previously published KCE report focused on breast cancer screening with mammography for women in the age group of 40-49 years (KCE report 129) and another report is currently in progress about breast cancer screening with mammography for women in the age group over 70 years.[1]

3. SCOPE OF THIS REPORT

The Belgian federal and regional governments signed a protocol agreement in 2001 for an organized screening program for women aged 50-69 years, to be organized by the regional governments with appropriate financial resources supplied by the federal government. Since 2001, Flanders, the Walloon region and the Brussels capital region have each introduced an organized screening program at a different pace and within their specific context of already existing practices. A first chapter of this report gives an overview of the current breast cancer screening practices in Belgium. These data have generated questions that form the subject of this report: are there indications for the use of other techniques than routine mammography in breast cancer screening? In the general population or only for women with an increased breast cancer risk? Who is eligible for screening outside the program for the general population and how can these women be identified?

The specific clinical questions to be answered in this report are listed below.

3.1. Women at risk for breast cancer

3.1.1. Risk evaluation

Familial breast cancer risk:

- Between all women, how to select the women with a possible familial risk of breast cancer on the base of the family history?
  - What are the existing assessment tools?
  - What are their validity and their applicability in Belgian context?
- Between women with an identified possible familial raised risk of breast cancer, how to select the women eligible for a genetic test?
  - What are the existing assessment tools?
  - What are their validity and their applicability in Belgian context?

Non familial breast cancer risk:

- Which are the risk factors of breast cancer to be considered outside the familial risk?
- Which is the risk ratio or the life time risk for each of these risk factors?

Combination of familial (outside genetic) and non familial breast cancer risk

- Which are the existing models for individual risk assessment?
- What are their validity and their applicability in Belgian context?

3.2. Technical methods for breast cancer screening in women with average, raised and high risk

3.2.1. Advantages and burden

Mammography with double reading (including computer-aided detection)

- Accuracy if compared with single reading mammography?
- Accuracy of computer-aided detection compared with double reading mammography

Digital mammography

- Accuracy if compared with analogue mammography?

Ultrasound

- What are advantages and burden of a combination of mammography and ultrasound if compared with a screening by mammography alone?
- in asymptomatic women with an average risk
- in asymptomatic women with dense breast tissue on mammography
- in asymptomatic women with a high breast cancer risk

MRI (Magnetic Resonance Imaging)

- What are advantages and burden of MRI alone (or MRI plus mammography; or MRI plus mammography plus ultrasound) compared with mammography alone (or mammography plus ultrasound) in women with high breast cancer risk?

4. METHODS

For each clinical question, a systematic search of the literature is performed and discussed with the support of external experts chosen for their scientific competency in several fields: gynecology, radiology, clinical genetics or epidemiology. The methodology used and the results are described in each chapter.

Clinical recommendations are then written, based on the evidence available. The strength of the recommendations is estimated with the tool GRADE, with particular attention to the application of GRADE to diagnostic studies.[2, 3] Those recommendations are finally submitted to the stakeholders by e-mail and discussed during a meeting for adaptation to the Belgian context.

CHAPTER 1 DESCRIPTION OF THE BELGIAN CONTEXT

1. INTRODUCTION

Following section describes data compiled by the Intermutualistic Agency (IMA), a body that centralizes data coming from all Belgian sickness funds. IMA compiled and published several reports on the national screening program containing data on the target age groups as defined by the program (50 – 69 years). IMA complemented this with information on persons outside the target age-group, with a particular focus on the tests used, delays between screening tests and possible confirmation and treatments following testing.

2. METHODOLOGY, DESCRIPTION OF THE DATA

The methodology used is largely the same as in the IMA report n° 7 on breast cancer screening of 2010.[4]

The data concern the two year period from 1 January 2006 until 21 December 2007.

Two types of data are used:

- Female population by age-group (5 year age-bands) and province, determined using the NIS (National Institute for Statistics) code
- Billing data on health care reimbursed by RIZIV/INAMI.

Following billing codes were used:

- Diagnostic outpatient mammography (450096, 461090)
- Screening mammography (the so-called 'mammotest') first reading (450192-4502031).
- Screening mammography (mammotest) second reading (450214-450225)
- Breast ultrasound (460132-460143, 469394-469405)
- Breast MRI (459476-459480)
- Surgical biopsy of the breast (227091-227102)
- Breast puncture, +/- image guidance (355670-355681, 355913-355924)
- Axillary lymph node dissection (226936-226940)
- Ablation of a tumor or mammary gland cyst (227032-227043)
- Tumorectomy (227054-227065)
- Mastectomy (226951-226962, 226973-226984, 226995-227006, 227010-227021).

The term "mammotest", often used in the French speaking part of Belgium, thus always refers to a mammography performd within an organized screening program.

There is no separate billing code for non-surgical breast biopsies such as a core needle biopsy. This type of procedures is normally billed similar to puncture procedures for cytology. Both fine needle aspiration (FNA) and core needle biopsy are thus included in the codes 355670-355681, 355913-355924.

The term 'diagnostic mammography' simply refers to the fact that the mammography was billed with the codes 450096 or 461090. However, these billing codes are also used for mammographies with a screening purpose, often in combination with an ultrasound on the same day. This type of screening outside the organized screening program is further called opportunistic screening.

Since there is no code to invoice an opportunistic screening mammography and billing data contain no information on diagnosis or symptoms, it is impossible to distinguish directly outpatient mammographies done for screening purposes and those done for diagnostic purposes or as follow up after treatment. We will try to make an indirect distinction making use of some assumptions. However, when the billing codes for the screening mammographies are used, one can be sure that their purpose is screening. These codes can only be billed by certified mammographic screening centers for the screening of women in the eligible age-group 50-69 years. The second reading is compulsory but an ultrasound can never be performed on the same day as the screening mammography.

Coverage is calculated using the notion of eligible population. For the organized screening program, women aged 50 – 69 years are eligible. In general, women who died during the study period and women for whom information was incomplete were excluded. Only women who are in the compulsory insurance program are included. For the women who go from one age category to another in the study period, an approximation is used by allocating 50 % of the population to the lower age category and 50 % to the higher, assuming a continuous transition linear in time.

3. RESULTS

Study population and coverage with screening and diagnostic mammography (per region and per age-group) are displayed in Table 1. Supplementary tables breaking up the data per age-group in 5 years and per province, tables giving the absolute numbers on which the calculations are based as well as the eligible population per year per region and province are given in appendix 4 (Table 40- Table 47).

Three coverages are calculated:

- 'Covered by mammotest' implies that a woman, aged 50 to 69 years, got at least one screening mammography in the study period. Since the opportunistic screening is not included, this coverage is an underestimation of the real screening coverage.
- A woman is considered 'covered by diagnostic' if she got at least one diagnostic mammography in the study period. The category diagnostic mammography comprises 'real' diagnostic mammographies and opportunistic screening mammographies.
- A woman is considered 'covered' in the 'total coverage' column if she got at least one mammography during the study period, be it whatever the type. Follow up mammographies and symptomatic women are included, so total coverage is an overestimation of the coverage that is relevant for screening and prevention of mortality and morbidity.

In Flanders, screening mammographies dominate in the age-group 50-69. The coverage by diagnostic mammography drops compared to the age-group 40-49, which may indicate a partial switch to organized screening as soon as women are eligible. In the age-group 70-79 overall coverage drops, mainly due to the stopping of organized screening. The coverage with diagnostic mammography decreases also with 3%, indicating that substitution of screening mammography by opportunistic screening at the age of 70 is not important in Flanders.

In the Brussels and Walloon region, 'diagnostic' mammographies are dominating, even in the age-group 50-69. Coverage with diagnostic mammographies is comparable in the 40-49 group and the 50-69 group. The coverage of 9 % screening mammography seems to add, indicating that there is not really a switch from diagnostic to screening in the transition 40-49 to 50-69, although we cannot exclude a switch to screening accompanied by an increase in opportunistic screening in the 50 – 69 group.

These regional differences in the use of diagnostic or screening mammography result in a higher overall coverage in Flemish women aged 50-69 than in the Walloon and Brussels region among women of the same age. In the other age-groups, however, total coverage is higher in the Walloon and Brussels region.

As shown in Table 48 in appendix, overall for Belgium, of all the women that were examined in the period 2006-2007, 80 % were examined only once in that period, 15 % were examined once in 2006 and once in 2007 and 2.8 % got several mammographic examinations either in 2006 or 2007. In Flanders the % of women with one mammography is somewhat higher (85 %).

Table 1 Study population and coverage with screening mammography (mammotest) and diagnostic mammography per region and per age-band, IMA data - period 2006-2007

REGIONS	AGE	*study population*	coverage by screening mammography	coverage by diagnostic mammography	total coverage
Flemish region	35-39 years	*211.561*	0%	12%	**12%**
	40-49 years	*462.186*	0%	31%	**31%**
	50-69 years	*716.873*	45%	21%	**65%**
	70-74 years	*148.246*	0%	18%	**18%**
	75-79 years	*135.373*	0%	8,20%	**8,2%**
	Total	***1.674.239***	**19%**	**21%**	**40%**
Region Brussels Capital	35-39 years	*36.831*	0%	15%	**15%**
	40-49 years	*64.269*	0%	44%	**44%**
	50-69 years	*97.416*	9,50%	43%	**53%**
	70-74 years	*19.077*	0%	33%	**33%**
	75-79 years	*19.259*	0%	18%	**18%**
	Total	***236.852***	**3%**	**36%**	**40%**
Walloon region	35-39 years	*115.858*	0%	19%	**19%**
	40-49 years	*246.854*	0%	46%	**46%**
	50-69 years	*395.072*	9.1%	46%	**55%**
	70-74 years	*75.217*	0%	30%	**30%**
	75-79 years	*75.338*	0%	15%	**15%**
	Total	***908.339***	**4%**	**39%**	**43%**
Belgium	35-39 years	*364.250*	0%	15%	**15%**
	40-49 years	*773.309*	0%	37%	**37%**
	50-69 years	*1.209.361*	30%	31%	**61%**
	70-74 years	*242.540*	0%	23%	**23%**
	75-79 years	*229.970*	0%	11%	**11%**
	Total	***2.819.430***	**13%**	**28%**	**41%**

Table 2 shows the medical imaging following diagnostic mammography per age-band and per region.

Table 3 shows the medical imaging following screening mammography per age-band and per region. It shows that the majority (85%) of diagnostic mammographies is followed by an ultrasound in all 3 regions, this in contrast with screening mammographies of which only 4.3% is followed by ultrasound. This proportion drops somewhat to 70 % in the age-groups above 70 in Flanders and Region Brussels capital. The decrease is less in the Walloon region. Decreasing breast density may play a role but this is uncertain.

Proportion of diagnostic mammographies followed by MRI is twice the proportion for screening mammographies and more or less constant over the ages, Supplementary tables breaking up the data per age-group in 5 years and per province are given in appendix 4 as well as the eligible population per year per region and province (Table 49-Table 52). Note that women are eligible in the year that they become 50, so a small proportion of screening mammographies falls into the category 40-49.

Table 2 Medical imaging following diagnostic mammography per age-band and per region, IMA data - 2006

AGE	REGION	*N**	% followed by an echography	% followed by MRI
35-39 years	Flemish region	*10.037*	**88%**	1,9%
	Region Brussels capital	*2.230*	**90%**	0,9%
	Walloon region	*7.978*	**94%**	1,2%
	Belgium	***20.245***	**91%**	**1,5%**
40-49 years	Flemish region	*49.629*	**85%**	1,5%
	Region Brussels capital	*8.918*	**87%**	0,7%
	Walloon region	*34.802*	**93%**	0,8%
	Belgium	***93.349***	**88%**	**1,2%**
50-69 years	Flemish region	*42.242*	**81%**	1,4%
	Region Brussels capital	*11.734*	**79%**	0,9%
	Walloon region	*49.726*	**88%**	1,0%
	Belgium	***103.702***	**84%**	**1,2%**
70-74 years	Flemish region	*8.444*	**66%**	1,3%
	Region Brussels capital	*1.806*	**70%**	0,9%
	Walloon region	*6.201*	**82%**	1,0%
	Belgium	***16.451***	**72%**	**1,2%**
75-79 years	Flemish region	*3.329*	**70%**	1,7%
	Region Brussels capital	*962*	**70%**	0,7%
	Walloon region	*3.145*	**83%**	1,2%
	Belgium	***7.436***	**75%**	**1,4%**
Total	Flemish region	*113.681*	**82%**	1,5%
	Region Brussels capital	*25.650*	**82%**	0,8%
	Walloon region	*101.852*	**89%**	1,0%
	Belgium	***241.183***	**85%**	**1,2%**

Table 3 Medical imaging following screening mammography (mammotest) per age-band and per region, IMA data - 2006

AGE	REGION	*N**	% followed by a diagnostic mammography	% followed by an echography	% followed by MRI
40-49 years	Flemish region	13.141	3,1%	**6,3%**	0,4%
	Region Brussels capital	117	2,6%	**7,7%**	0,0%
	Walloon region	501	6,0%	**11%**	0,4%
	Belgium	**13.759**	**3,2%**	**6,5%**	**0,4%**
50-69 years	Flemish region	110.902	1,9%	**3,4%**	0,3%
	Region Brussels capital	3.191	2,1%	**6,0%**	0,1%
	Walloon region	10.209	6,4%	**9,9%**	0,4%
	Belgium	**124.302**	**2,3%**	**4,0%**	**0,3%**
Total	Flemish region	124.046	2,0%	**3,7%**	0,3%
	Region Brussels capital	3.308	2,1%	**6,0%**	0,1%
	Walloon region	10.710	6,3%	**10%**	0,4%
	Belgium	**138.064**	**2,3%**	**4,3%**	**0,3%**

Table 4 shows the number and proportions of biopsies, punctures and surgery after diagnostic mammography per age-band and per region. Table 5 shows the number and proportions of biopsies, punctures and surgery after screening mammography per age-band and per region. As for diagnostic imaging, the proportions biopsies, punctures and surgery after diagnostic mammography is between two and three times the proportion for screening mammographies in the age-group between 50 and 69 y. The proportion increases with age, this may reflect increasing incidence but also differences in the mix opportunistic screening-mammographies done for clinical reasons.

Table 4 Punctures, biopsies and surgery following diagnostic mammography, Belgium, 2007

			Punctions/ biopsies		Surgery after punctions/ biopsies		
AGE	**REGION**	*Nb ref*	Nbr [a]	**%**	Nbr [b]	**% [b/a]**	**% [b/Nb ref]**
35-40 years	Flemish region	*10.037*	379	3,8%	133	35%	1,3%
	Region Brussels capital	*2.230*	117	5,2%	6	5,1%	0,3%
	Walloon region	*7.978*	533	6,7%	74	14%	0,9%
	Belgium	***20.245***	**1.029**	**5,1%**	**213**	**21%**	**1,1%**
40-49 years	Flemish region	*49.629*	1.688	3,4%	579	34%	1,2%
	Region Brussels capital	*8.918*	346	3,9%	55	16%	0,6%
	Walloon region	*34.802*	1.798	5,2%	283	16%	0,8%
	Belgium	***93.349***	**3.832**	**4,1%**	**917**	**24%**	**1,0%**
50-69 years	Flemish region	*42.242*	1.374	3,3%	723	53%	1,7%
	Region Brussels capital	*11.734*	354	3,0%	129	36%	1,1%
	Walloon region	*49.726*	1.852	3,7%	545	29%	1,1%
	Belgium	***103.702***	**3.580**	**3,5%**	**1.397**	**39%**	**1,3%**
70-74 years	Flemish region	*8.444*	369	4,4%	266	72%	3,2%
	Region Brussels capital	*1.806*	69	3,8%	33	48%	1,8%
	Walloon region	*6.201*	277	4,5%	112	40%	1,8%
	Belgium	***16.451***	**715**	**4,3%**	**411**	**57%**	**2,5%**
75-79 years	Flemish region	*3.329*	252	7,6%	184	73%	5,5%
	Region Brussels capital	*962*	29	3,0%	15	52%	1,6%
	Walloon region	*3.145*	205	6,5%	106	52%	3,4%
	Belgium	***7.436***	**486**	**6,5%**	**305**	**63%**	**4,1%**
Total	Flemish region	*113.681*	4.062	3,6%	1.885	46%	1,7%
	Region Brussels capital	*25.650*	915	3,6%	238	26%	0,9%
	Walloon region	*101.852*	4.665	4,6%	1.120	24%	1,1%
	Belgium	***241.183***	**9.642**	**4,0%**	**3.243**	**34%**	**1,3%**

Table 5 Punctures, biopsies and surgery following screening mammography (mammotest), Belgium, 2007

			Punctions/ biopsies after exam ref.		Surgery after punctions/ biopsies		
AGE	**REGION**	*Nb ref*	Nbr [a]	**%**	Nbr [b]	**% [b/a]**	**% [b/Nb ref]**
40-49 years	Flemish region	*13.141*	152	1,2%	48	32%	0,4%
	Region Brussels capital	*117*	0	0,0%	0	/	0,0%
	Walloon region	*501*	11	2,2%	1	9,1%	0,2%
	Belgium	***13.759***	**163**	**1,2%**	**49**	**30%**	**0,4%**
50-69 years	Flemish region	*110.902*	887	0,8%	500	56%	0,5%
	Region Brussels capital	*3.191*	28	0,9%	13	46%	0,4%
	Walloon region	*10.209*	198	1,9%	39	20%	0,4%
	Belgium	***124.302***	**1.113**	**0,9%**	**552**	**50%**	**0,4%**
Total	Flemish region	*124.046*	1.039	0,8%	548	53%	0,4%
	Region Brussels capital	*3.308*	28	0,8%	13	46%	0,4%
	Walloon region	*10.710*	209	2,0%	40	19%	0,4%
	Belgium	***138.064***	**1.276**	**0,9%**	**601**	**47%**	**0,4%**

The fact that the proportion surgery is higher in the group undergoing diagnostic mammography but not as high as would be expected if it were not mixed with opportunistic screening can be used to give a very rough estimation of the proportion opportunistic screening. We assume here that the proportion women undergoing surgery amongst opportunistic screening is the same as in the organized screening in the group 50 - 69 (0.4%). We let the expected proportion of women undergoing surgery after a 'true' diagnostic mammography vary between 3 and 7 %, based on a study of Barlow et al[5], and use this for the estimation. We find that the proportion opportunistic screening varies between 80% and 90 % under those assumptions.

If proportion surgery is higher in the opportunistic screening group than in the organized screening group then the estimations of the proportion opportunistic screening are higher. This may be true as women at higher risk may preselect themselves and may have a higher tendency to seek or be offered opportunistic screening compared to organized screening, e.g. because of worries about family history or overweight.

Proportion biopsies could be used using the same reasoning but they seem to be more variable, as in the organized screening proportion in Walloon region are already twice the proportion in Flanders. More details on the estimation method are given in appendix 4.

IMA estimated that 3,58% of women undergoing a diagnostic mammography had a mammography in one breast and 5,55 % had a past history of a tumor, either benign or malignant, so they concluded that at least 10 % is done for clinical reasons. Their figures are comparable to ours. However, the estimations used by IMA and KCE are all very dependent on assumptions, hence they should be interpreted with (a lot of) caution. Figure 1 shows the evolution of screening mammographies per 100 000 women in the age-group 50 – 69 from 2002 to 2007 by region. It shows that the screening increased in Flanders and in a lesser degree in the region Brussels and stagnated and even dropped in the Walloon region reflecting different attitudes towards the organized screening program.

Figure 2 shows the evolution of diagnostic mammographies per 100 000 women in the age-group 70 – 74 from 2002 to 2007 by region.

Figure 3 shows the evolution of diagnostic mammographies per 100 000 women in the age-group 40 – 49 from 2002 to 2007 by region. Numbers are considerably lower for Flanders in both age-groups. More detailed data and the breakup in age-groups are provided in Table 53 of Appendix 4.

Figure 1 Evolution of screening mammographies per 100 000 women in the age-group 50 – 69 from 2002 to 2007 by region

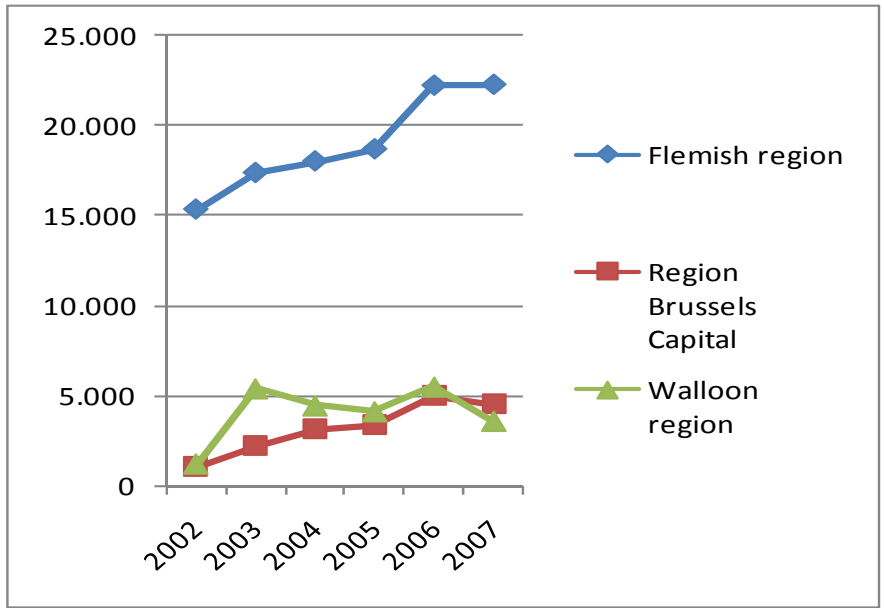


Figure 2 Evolution of diagnostic mammographies per 100 000 women in the age-group 70 – 74 from 2002 to 2007 by region

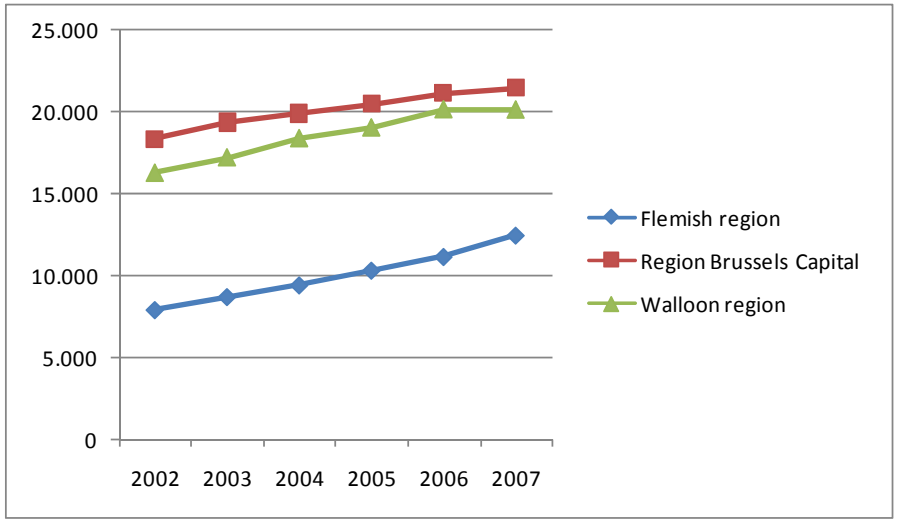


Figure 3 Evolution of diagnostic mammographies per 100 000 women in the age-group 40 – 49 from 2002 to 2007 by region

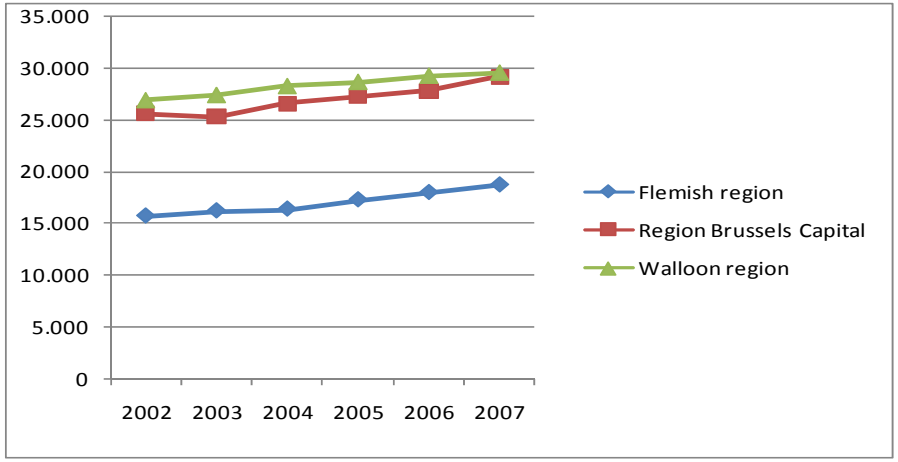


Table 6 (Table 54) shows the number of biopsies and punctures per 100 000 women per region, per year, period 2002 – 2007. The number of biopsies is considerably higher in Walloon region and region Brussels capital, in spite of a similar surgery rate. According the data from the cancer register, the incidence of invasive cancer in the age-group 50 – 59 was 364 per 100 000, than the rate puncture/biopsy to cancer is 5.5. If you assume that all declared cancers underwent at least either a puncture or a biopsy, then these figures suggest that the number of false positive screening or diagnostic examinations is too high, especially in Brussels and Wallonia, and this should be a point of attention. Possible explaining factors include the frequent use of ultrasound for screening purposes (see 3.4), different practices among physicians and variation in attitude versus the use of biopsies as seen between different countries[6]

Table 6 Number of biopsies and punctures per 100 000 women per year, period 2002 – 2007

		Biopsy						Punctures					
		2002	2003	2004	2005	2006	2007	2002	2003	2004	2005	2006	2007
Flemish region	35-39 years	16	11	15	18	20	18	328	323	334	349	352	396
	40-49 years	30	34	43	36	26	37	805	852	861	905	953	986
	50-69 years	42	35	42	38	31	42	877	939	948	972	990	1.024
	70-74 years	31	26	21	32	20	35	373	463	509	558	660	673
	75-79 years	20	19	26	18	19	35	311	362	423	493	523	551
	Total	**32**	**29**	**35**	**33**	**26**	**37**	**688**	**741**	**760**	**796**	**831**	**865**
Region Brussels Capital	35-39 years	66	68	89	94	47	83	540	624	561	599	526	543
	40-49 years	154	198	215	221	105	198	1.322	1.425	1.515	1.339	1.456	1.511
	50-69 years	170	179	233	172	135	214	1.474	1.532	1.630	1.302	1.560	1.483
	70-74 years	134	178	169	127	151	197	966	1.140	1.029	881	1.206	1.155
	75-79 years	114	102	134	131	64	146	689	739	823	797	859	893
	Total	**142**	**160**	**192**	**166**	**109**	**183**	**1.175**	**1.261**	**1.315**	**1.125**	**1.285**	**1.269**
Walloon region	35-39 years	37	31	37	20	21	34	813	830	849	781	895	866
	40-49 years	74	53	74	69	30	63	2.100	2.092	2.169	2.084	2.091	2.060
	50-69 years	93	77	69	62	35	71	2.021	2.154	2.214	2.028	2.046	1.932
	70-74 years	60	44	40	58	31	53	1.127	1.113	1.210	1.262	1.222	1.219
	75-79 years	30	45	31	49	21	41	659	791	850	921	922	983
	Total	**72**	**59**	**60**	**57**	**30**	**60**	**1.675**	**1.746**	**1.815**	**1.721**	**1.745**	**1.692**
Belgium	35-39 years	27	23	29	26	23	30	499	510	517	510	542	561
	40-49 years	54	54	67	62	34	59	1.269	1.300	1.336	1.319	1.358	1.373
	50-69 years	69	61	66	57	41	66	1.295	1.382	1.415	1.343	1.381	1.358
	70-74 years	50	44	39	48	34	53	672	731	776	807	880	879
	75-79 years	32	36	37	38	24	47	466	543	603	662	683	720
	Total	**55**	**50**	**57**	**52**	**34**	**56**	**1.049**	**1.110**	**1.148**	**1.123**	**1.164**	**1.166**

The evolution of the number of Halsted operations, mastectomies, partial mastectomies and tumorectomies is given in Table 55 and Table 56 of appendix 4. Numbers remain stable over this period and there are no marked differences between regions. A shift towards more breast sparing surgery after the introduction of the screening program cannot yet be seen.

Table 7 shows the delays (number of days) in percentiles between mammographies (diagnostic and screening) and different complementary tests for Belgium per age-group. A P10 of 21 days means that 10 % of women has a delay less than 21 days, a P90 of 58 days means that 90 % of women has a delay less than 58 days.

Table 7 Delays (days, in percentiles) between diagnostic or screening mammography and complementary tests, for Belgium, data from 2007

		Diagnostic mammographies *followed by complementary tests*						***Mammotests*** *followed by complementary tests*					
		N	**P 10**	**P 25**	**P 50**	**P 75**	**P 90**	*N*	**P 10**	**P 25**	**P 50**	**P 75**	**P 90**
Outpatient Diagnostic Mammography	35-39 years	*27.481*	0	0	0	0	0	/	/	/	/	/	/
	40-49 years	*118.636*	0	0	0	0	0	*521*	21	26	35	45	58
	50-69 years	*129.623*	0	0	0	0	0	*3.565*	18	24	33	45	60
	70-74 years	*20.869*	0	0	0	0	0	*1*	0	0	0	0	0
	75-79 years	*9.434*	0	0	0	0	0	/	/	/	/	/	/
	Total	***306.043***	**0**	**0**	**0**	**0**	**0**	***4.087***	**18**	**24**	**34**	**45**	**60**
Inpatient Diagnostic Mammography	35-39 years	*21*	12	18	28	35	42	/	/	/	/	/	/
	40-49 years	*135*	8	15	28	45	66	*14*	35	35	46	63	74
	50-69 years	*249*	13	20	28	46	64	*238*	28	36	47	62	75
	70-74 years	*73*	14	20	26	35	49	/	/	/	/	/	/
	75-79 years	*35*	13	16	22	37	52	/	/	/	/	/	/
	Total	***513***	**12**	**18**	**27**	**42**	**62**	***252***	**30**	**36**	**47**	**62**	**75**
Echography	35-39 years	*25.029*	0	0	0	0	0	/	/	/	/	/	/
	40-49 years	*104.680*	0	0	0	0	0	*1.030*	17	24	32	44	62
	50-69 years	*108.565*	0	0	0	0	0	*6.263*	15	22	32	45	60
	70-74 years	*15.154*	0	0	0	0	0	/	/	/	/	/	/
	75-79 years	*7.095*	0	0	0	0	0	/	/	/	/	/	/
	Total	***260.523***	**0**	**0**	**0**	**0**	**0**	***7.293***	**15**	**23**	**32**	**44**	**61**
MRI	35-39 years	*377*	4	9	16	31	49	/	/	/	/	/	/
	40-49 years	*1.360*	5	10	18	34	54	*69*	21	35	44	59	76
	50-69 years	*1.432*	5	9	18	32	50	*460*	24	33	44	61	76
	70-74 years	*236*	6	10	17	30	47	/	/	/	/	/	/
	75-79 years	*119*	6	11	17	28	50	/	/	/	/	/	/
	Total	***3.524***	**5**	**9**	**18**	**32**	**51**	***529***	**24**	**33**	**44**	**60**	**76**
Ponction ou biopsy	35-39 years	*1.370*	0	0	0	7	22	/	/	/	/	/	/
	40-49 years	*4.810*	0	0	0	9	27	*193*	20	26	39	55	70
	50-69 years	*4.456*	0	0	0	9	25	*1.397*	17	24	35	50	66
	70-74 years	*923*	0	0	2	10	23	/	/	/	/	/	/
	75-79 years	*626*	0	0	1	9	22	/	/	/	/	/	/
	Total	***12.185***	**0**	**0**	**0**	**9**	**25**	***1.590***	**17**	**24**	**35**	**51**	**67**

Since a diagnostic mammography is an outpatient diagnostic mammography, delays are of course 0 days. Delay between diagnostic mammography and ultrasound is also 0 days, reflecting the fact that ultrasound is usually done at the same time and is not the consequence of findings in the index mammography. Delays for MRI and biopsies are considerably longer for screening mammography. The same data for the three regions are given in Table 57,Table 58 and Table 59. There are no marked differences between regions and between age-groups.

Table 8 shows the delays between biopsy and surgery after diagnostic and screening mammography.

Delays are shorter for the Flemish region in general. For region Brussels capital and Walloon region, delays are grossly comparable between diagnostic and screening mammography. For Flanders the delays after screening mammography are somewhat shorter. More detailed data are displayed in appendix 4, Table 60, Table 61 and Table 62, with a breakup in age-groups.

Table 8. Delays between biopsy and surgery after diagnostic and screening mammography per region, 2007.

Diagnostic mammography

	Within the month		Between 1 and 3 month		Between 3 and 6 month		More then 6 months	
	Nbr	*Pct*	Nbr	*Pct*	Nbr	*Pct*		
Flemish region	3.572	*78%*	553	*12,0%*	246	*5,4%*	148	*3%*
Region Brussels Capital	452	*49%*	303	*33%*	109	*12%*	61	*7%*
Walloon region	2017	*52%*	1279	*33%*	290	*8%*	293	*8%*
Belgium	6.041	*65%*	2135	*23,0%*	645	*6,9%*	537	*6%*

Screening mammography

age 50-69 years	**Within the month**		**Between 1 and 3 month**		**Between 3 and 6 month**		**More then 6 months**	
	Nbr	*Pct*	Nbr	*Pct*	Nbr	*Pct*		
Flemish region	1.010	*87%*	115	*9,9%*	22	*1,9%*	18	*1,5%*
Region Brussels-Capital	21	*51%*	16	*39%*	2	*4,9%*	2	*5%*
Walloon region	112	*52%*	89	*41%*	6	*2,8%*	9	*4%*
Belgium	**1.143**	***80%***	**220**	***15%***	**30**	***2,1%***	**29**	***2%***

Finally, Table 9 shows the delays between diagnostic tests: Diagnostic mammography (DM), Screening mammography (MT), mean and percentile, for Belgium. Mean delays between DM-DM are a year, this can be an indication that a large part of the diagnostic mammographies are opportunistic screening mammographies. Delays MD-MT are similar, indicating a transition from opportunistic to organized screening. The shorter delays MT-DM probably partly reflect the fact that a part of these mammographies are supplementary mammographies after a suspected screening mammography. In Appendix 4 same data are displayed broken up by region, no marked differences between regions are noted.

Table 9. Delays (days) between diagnostic tests: Diagnostic mammography (DM), Screening mammography (MT), mean and percentile, for Belgium, 2007.

	DM-DM							DM-MT							MT-DM						
	N	Mean	P 10	P 25	P 50	P 75	P 90	*N*	Mean	P 10	P 25	P 50	P 75	P 90	*N*	Mean	P 10	P 25	P 50	P 75	P 90
35-40 year	*6.921*	364	179	304	370	435	539	/	/	/	/	/	/	/	/	/	/	/	/	/	/
40-49 year	*63.619*	381	218	342	377	440	533	*1.588*	398	235	322	389	476	582	*598*	57	21	28	38	56	139
50-69 year	*113.614*	373	214	341	371	420	517	*9.485*	376	204	288	367	458	573	*19.042*	228	24	36	172	398	518
70-74 year	*13.445*	362	199	334	368	405	493	/	/	/	/	/	/	/	*9*	201	8	45	130	363	498
75-79 year	*6.353*	358	197	329	366	403	486	/	/	/	/	/	/	/	/	/	/	/	/	/	/
Total	***203.952***	**374**	**210**	**340**	**371**	**425**	**521**	***11.073***	**379**	**210**	**294**	**370**	**462**	**573**	***19.649***	**223**	**24**	**36**	**147**	**393**	**513**

4. DISCUSSION

There are marked regional differences in coverage, with Flanders having the highest overall coverage and coverage with organized screening (mammotest) in the age-group 50-69 years. Coverage with what is labeled as diagnostic mammography among the other age-groups is consistently higher in the Flemish region and the region Brussels capital. We cannot possibly make out if these women are considered at risk in some way or another, it is to be noted that among women not at risk the balance risk-benefit of screening at younger ages is uncertain.[1]

For the age-group 50-69 we tried to estimate the proportion opportunistic screening with a number of assumptions using the proportion of screened women undergoing surgery in one form or another. We found that the likely proportion is above ninety percent under most assumptions. We did not use proportion biopsies nor proportion confirmatory diagnostic imaging for this estimation as there are already marked regional differences and estimations becomes even more unstable. Uncertainty around this estimation remains high though and needs to be interpreted with care. For other age-groups it is not possible to do similar estimations. Another indication for the high proportion of opportunistic screening is the fact that the delay between most diagnostic imaging is around a year.

Coverage by organized screening drops in Walloon region during the period 2002- 2007, number of diagnostic mammographies goes up in all regions during the same period. For all regions most diagnostic mammographies are accompanied by an ultrasound, data on delays show that this is done systematically at the same time. This is in contrast with common practice in other countries and the value of such an ultrasound is unclear at best (see next chapter).

Number of biopsies and punctures is considerably higher in the region of Brussels-capital and Walloon region. One possible explanation is the higher use of opportunistic screening, accompanied by an ultrasound, leading to more false positives and need for biopsy and puncture without a comparable increase in number of surgical interventions. Delays between mammographies and confirmatory test are higher for organized screening, delays between mammography and surgery with shorter for organized screening compared to diagnostic screening in Flanders but longer in the other regions.

Key points

- **In the age-group 50-69, overall coverage for mammography and coverage with organised screening for the period 2006-2007 is higher in Flanders (65%) than in the Brussels (53%) and Walloon region (55%).**
- **In the other age-groups coverage with diagnostic mammography is higher in Walloon and Brussels region. In Flanders, coverage of women aged 40-49 years is 31%. The percentage of coverage of Brussels and Walloon women of that age is respectively 44% and 46%. Older women, aged 70-74, are less covered in Flanders (18%) than in Brussels (33%) and Wallonia (30%).**
- **More than 80% of screening mammographies performed outside the organized screening program is accompanied by a breast ultrasound on the same day.**

- **In Belgium, the number of breast punctures and biopsies per 100 000 women per year is as high as 1222 per 100 000 women., Figures for Flanders, Region Brussels capital and Walloon region are 902, 1452 and 1752 per 100 000 women respectively.**
- **In the category diagnostic mammography, proportion opportunistic screening is estimated to be between 80 % and 90 % under most assumptions.**
- **It is unclear how many mammographies are done amongst women considered to be at high risk.**
- **Most diagnostic mammographies are accompanied by an ultrasound on the same day.**
- **Delays between mammographies and confirmatory tests are higher for organised screening. Delays between mammography and surgery, however, is shorter after a screening mammography compared to diagnostic mammography. The shortest delays for surgery are seen in Flanders.**

CHAPTER 2 WOMEN AT RISK FOR BREAST CANCER

1. INTRODUCTION

The assessment of breast cancer risk has a number of different aspects and reposes essentially on 3 pillars:

- the evaluation of family risk by asking information on relatives, first, second or third degree
- identification of genetic risk factors, mainly faulty *BRCA1*, *BRCA2* or *TP53* gene in the person or her family, but recently also single-nucleotide polymorphisms (SNP)
- individual non genetic risk factors, including dense breast tissue, benign breast disease, hormonal and dietary factors

Simply adding or multiplying measurements of risk factors however is not appropriate due to the multiple interactions, confounding and overlaps (e.g. dense breast is partially hereditary). Therefore a number of risk models have been developed, combining different familial and non familial risk factors.

We can distinguish two types or risk models:

models that assess the risk of developing breast cancer, either 5 year, 10 year or lifetime risk

models that assess the risk of carrying a germline mutation, such as BRCA1, BRCA2 and TP53 mutations

Validation studies usually asses 2 elements of the models: validation and discrimination.

Calibration concerns how accurately a model predicts the observed rate of breast cancer and is measured by the ratio of the expected-to-observed rate (E/0). For example, if a model predicts a 5-year breast cancer rate of 3% and a rate of 3.2% is observed in a population, then the model has an expected-to-observed ratio of 0.94.[7] Poor calibration is sometimes assessed with the Hosmer–Lemeshow test, a statistical test for goodness of fit for models, assessing whether or not the observed event rates match expected event rates in subgroups of the model population test, giving a p value derived from a chi square distribution.

Discrimination refers to how well the model differentiates between women who develop cancer and women who remain free of cancer. It is measured by the *c* –statistic, representing the area under the receiver operating characteristics (AUC) curve. The use of AUC to assess discriminative power has been criticized for not being clinically relevant. Other approaches assessing the fit of the predicted probabilities to the observed data have been suggested, such measuring the proportion of confirmed cases for whom the model assess the risk to be under a certain threshold (e.g. 10 %). These approaches may be clinically more relevant but have the disadvantage that no consensus exists on the thresholds that need to be chosen [8].

We focus mainly on risk assessment models and less on prediction of mutations, we consider the latter, as decisions on this need to be taken in specialized centers, out of scope. We discuss them however as there is some overlap between both and some are used for both, after some modifications.

2. METHODS

2.1. Literature search strategy

First a general search on breast cancer was perfomed to search for guidelines and HTA reports on risk assessment National guidelines Clearinghouse, Guidelines international Network (GIN), SBU, NICE, DACEHTA, MSAC, MAS, HAS, AHRQ, BCBS, AETSA, AATRM, CCOHTA,ECRI, DIMDI, IQWIG.

A search for HTA reports was performed in Center for review and dissemination databases CRD: DARE, NHS EED and HTA.

A NICE guideline and a report on individual non familial risk factors was identified and assessed as a valid and relevant.

Then 3 separate searches were performed in the Cochrane Database of Systematic Reviews (CDSR), Medline, and EMBASE to retrieve meta-analyses (MA), systematic reviews (SR), cohort studies (CS) and model validation studies on following topics:

- Family risk assessment
- Non familial risk factors
- Risk models
- An overview of the search strategy is given in appendix 1.

2.2. Selection criteria

A classic 'PICO' structure is not applicable to our research question.

Our final selection was limited to meta-analyses (MA), systematic reviews (SR), cohort studies (CS) and model validation studies. Only studies published in full were included.

Studies in English, German, Dutch, French and Portuguese were considered eligible.

2.3. Selection procedure

Our study selection started by looking at titles and abstracts to exclude any studies considered not relevant for our purposes. Articles that appeared relevant or for which we had doubts were assessed by reading the full text. A list of studies assessed in full text but excluded is given in Annex together with the reason for their exclusion.

The reference lists of the selected studies were checked for additional relevant studies that could be included in our review.

A hierarchical approach was followed by which:

Firstly, the analysis focused purely on MA and SR published up to the date of our search. Secondly, the selected evidence synthesis was updated by looking at all relevant original literature (RCTs) published after the search data of selected MA and SR and found via our search.

2.4. Critical appraisal

The reviewer critically appraised the SRs and MAs according to the checklist of the Dutch Cochrane collaboration (http://dcc.cochrane.org/dutch-cochrane-centre),

No specific checklist to evaluate studies that validate risk models exists. However, validation of risk prediction models is done on cohort studies, even if they are not in the first place set up for this purpose. Therefore we applied the checklist of the Dutch Cochrane centre to evaluate the quality of this kind of validation studies. Studies that assess the risk assessment models carrying germline mutations have a design that resembles more classic test validation studies, in the sense that they are transversal studies using the genetic test as gold standard and evaluate the model as was it a test. Therefore we applied the checklist for test validation studies for this type of study, although some items of the checklist were not relevant, mainly those concerning dealing with confounding (as the models themselves are partially developed for this purpose).

2.4.1. Data extraction

The reviewers synthesised the characteristics of the studies and the available results in evidence tables.

Results from the selected evidence synthesis were confronted with those from the original studies published after it. If conclusions were similar a descriptive analysis of the results from both the meta-analysis and the original studies was completed.

2.5. Research and selection

For family risk, the search strategy generated 1233 publications but only involving risk assessment models. All publications concerned refer to a collaborative reanalysis of individual data from 52 epidemiological studies including 58 209 women with breast cancer and 101 986 women without the disease published in the lancet in 2001.

The search for risk models generated 833 publications of which 17 were selected. The search for risk factors generated 1633 publications, of which 5 were selected. The flowcharts and detailed reasons for the exclusion are listed in appendix 2.

Evidence tables are available in appendix 3.

2.6. Findings

We first describe the findings and recommendations of the NICE guidelines of 2004 and the partial update of 2006 NICE, 2006 [9, 10] and the findings of the NZHTA systematic review by weir et al[11] on non familial risk factors. Then we describe the update on risk factors, starting from the search date of the NZHTA systematic review. Finally we describe the findings on the model validation studies.

2.6.1. *Risk assessment based on number of affected family members.*

The NICE guidelines base their risk classification in average, moderate an high risk on data from both Claus and co-workers (1994)[12] and the Collaborative Group on Hormonal Factors in Breast Cancer study, where a meta-analysis was performed using the primary data of 52 epidemiological studies (2001) [13] to guide the levels that are presented in the guideline.

Women are considered to be **at average risk** if the family history shows only one first-degree or second-degree relative diagnosed with breast cancer at an age older than 40 years.

Women are considered to be **at raised risk** (that is, a 10-year risk of 3–8% for women aged 40–49 or a lifetime risk of 17% or greater but less than 30%). Women at raised risk should be offered secondary care and do not require referral to tertiary care.

Women who meet the following criteria should be considered at raised risk:

- one first-degree relative diagnosed with breast cancer at younger than age 40 years, or
- two first-degree or second-degree relatives diagnosed with breast cancer at an average age of older than 50 years, or
- three first-degree or second-degree relatives diagnosed with breast cancer at an average age of older than 60 years, or
- a formal risk assessment (usually carried out in tertiary care) or a family history pattern is likely to give a 10-year risk of 3–8% for women aged 40–49 years, or a lifetime risk of 17% or greater but less than 30%

The conditions for being at average risk and raised risk are applied provided that none of the following are present in the family history:

- bilateral breast cancer
- male breast cancer
- ovarian cancer
- Jewish ancestry
- sarcoma in a relative younger than 45 years of age
- glioma or childhood adrenal cortical carcinomas
- complicated patterns of multiple cancers at a young age
- very strong paternal history (four relatives diagnosed at younger than 60 years of age on the father's side of the family).

Women are considered to be **at high risk** (that is, a 10-year risk at age 40–49 years of greater than 8% or a lifetime risk of 30% or greater, or a 20% or greater chance of a faulty BRCA1, BRCA2 or TP53 gene in the family) if

- At least the following female breast cancers in the family:
  - two first-degree or second-degree relatives diagnosed with breast cancer at younger than an average age of 50 years (at least one must be a first-degree relative), or
  - three first-degree or second-degree relatives diagnosed with breast cancer at younger than an average age of 60 years (at least one must be a first-degree relative), or
  - four relatives diagnosed with breast cancer at any age (at least one must be a first-degree relative).

or

- Families containing one relative with ovarian cancer at any age and, on the same side of the family:
  - one first-degree relative (including the relative with ovarian cancer) or second-degree relative diagnosed with breast cancer at younger than age 50 years, or
  - two first-degree or second-degree relatives diagnosed with breast cancer at younger than an average age of 60 years, or
  - another ovarian cancer at any age.

or

- Families containing bilateral cancer (each breast cancer has the same count value as one relative):
  - one first-degree relative with cancer diagnosed in both breasts at younger than an average age of 50 years, or

- one first-degree or second-degree relative diagnosed with bilateral breast cancer and one first-degree or second-degree relative diagnosed with breast cancer at younger than an average age of 60 years.

or

- Families containing male breast cancer at any age and on the same side of the family, at least:
  - one first-degree or second-degree relative diagnosed with breast cancer at younger than age 50 years, or
  - two first-degree or second-degree relatives diagnosed with breast cancer at younger than an average age of 60 years.

or

- A formal risk assessment has given risk estimates of:
  - a 20% or greater chance of a BRCA1, BRCA2 or TP53 mutation being harboured in the family, or
  - a greater than 8% chance of developing breast cancer age 40–49 years, or
  - a 30% or greater lifetime risk of developing breast cancer

The NICE guidelines recommend 3 referral levels. Primary care is considered appropriate for women at average risk (this is population risk; the term was modified in order to avoid the term 'low risk'). Women with a raised breast cancer risk should be referred to a secondary level, essentially a breast clinic, and only the high risk group should be referred to the tertiary level, this is a specialized genetic clinic.

They further recommend that women with a raised risk should be offered yearly mammographic surveillance from the age of 40 years on, however this is based on expert opinion as there is no proof that this approach is beneficial. The main argument is that these groups have a risk of developing breast cancer that is comparable of that of women above 50.

They did a literature search on risk models up to 2004 and concluded that existing computer models (Gail, Claus, BRCAPRO) underestimate in a family history setting in terms of breast cancer risk prediction, although the manual Claus tables produce risks close to those seen in a screened familial risk population. They identified one USA study that found that BRCAPRO predicted BRCA 1 & 2 mutation status better than genetic counsellors and also concluded that the degree of correlation between different risk models is relatively poor.

Based on these findings they consider that computerised risk-assessment models can be helpful aids to risk assessment, but can be misleading and should not yet totally replace careful clinical assessment of family trees with a manual approach.

2.6.2. Non familial risk factors:

2.6.2.1. Findings of the NZHTA report

The NZHTA report of Weir et al 2007 [11] reviewed non familial risk factors for breast cancer, based on systematic reviews of observational studies on the association between non familial risk factors and breast cancer.

They found 3 strong risk factors: a past history of (*in situ*) breast cancer, dense breast tissue alcohol intake.

A past history of breast cancer was a risk factor for a second primary breast cancer. Four primary research studies were identified, the relative risk estimates ranged between 2.8 and 7.4. The RR for a range of lesions associated with increased risk of breast estimated:

- ductal hyperplasia RR 1.5 - 2
- atypical ductal hyperplasia RR 4
- lobular carcinoma *in situ* RR 6-10
- ductal carcinoma *in situ* RR 8 - 10

The association between increased breast density and risk of breast cancer was considered in 12 primary research studies. The relative risk approximated four across these studies, when comparing the highest category (usually BIRADS 4) to the lowest.

One risk factor was considered moderate: alcohol intake. It was considered in three systematic reviews and 10 primary research studies. The increased risk was in the order of 10% for 10g alcohol/day, 25% for 25g alcohol/day and 55% for 50g alcohol/day.

Other risk factors were only modestly associated with breast cancer:

Nulliparity was considered in one secondary research study and 28 primary research studies and shown to be a risk factor for breast cancer.

Among the larger studies, the relative risk estimates appeared to decrease by approximately 0.09 for each additional birth.

Early menarche was associated with increased risk of breast cancer in the one secondary research study and 29 primary research studies but it is difficult to give an estimate due to variation in cut-points for categorisation of age at menarche and uncertainty due to potential biases but it is likely to be moderate.

Post menopausal obesity was considered in three systematic reviews and 14 primary research studies. The systematic review that compared BMI with risk of breast cancer estimated a relative risk of 1.12 the overweight category and 1.25 for the obese category.

Hormone replacement therapy was considered in eight systematic reviews. Most studies found an increased risk of 1.2-1.4.

Hormonal contraceptives were considered in 37 primary research studies. The results were consistent with the findings of the Collaborative Review (which re-analysed primary data from over 50 relevant studies). The results from this reanalysis were:

- current users: RR 1.24 (95% CI 1.15-1.33)
- 1-4 years after stopping: RR 1.16 (95% CI 1.08-1.23)
- 5-9 years after stopping: RR 1.07 (95% CI 1.02-1.13)
- >10 years after stopping: RR 1.01 (95% CI 0.96-1.05).

The role of exogenous hormones (stilboestrol, xenoestrogens and phytoestrogens) is unclear and the data in the literature conflicting, moreover, it is not feasible to use these risk factors for risk assessment of decision on screening modalities

The same is true for dietary fat, vegetable fat and polyunsaturated fat, where results were in any case conflicting or unclear.

2.6.2.2. Update of the NZHTA report (Weir et al.)

We did an update starting from the search date of the review. We included only risk factors that can be used for risk assessment and clinical decisions, excluding studies on e.g. diary use, use of soy beans, serum vitamin D. These studies are in the first place useful to give health and cancer prevention advice; this is not the scope of this report. The reasons for this type of exclusion are explained for each excluded study in the list of excluded studies.

5 supplementary systematic reviews were identified.

Vrieling et al [14] found that risks ratios for breast cancer were different according to oestrogen and progestogen receptor status (ER & PR), with a higher association between weight gain and risk for ER+PR+ and ER+ tumors combined (11 observational studies; RR = 2.03; 95% CI 1.62, 2.45). Clinical implications of these findings in our context are unclear.

Cummings et al 2009 [7] did an update of the meta-analysis of Mc Cormack et al [15] and, based on 47 prospective observational studies, found that breast density was strongly associated with breast cancer (RR = 4.03 [95% CI = 3.10 to 5.26] for BI-RADS category IV (extremely dense) vs category I (fatty); RR =4.20 [95% CI = 3.61 to 4.89] for >75% vs <5% dense area). These findings are grossly similar to the findings of the NZHTA review.

Kahlenborn et al, 2006 [16] did a meta-analysis of 34 studies and found that the use of OC was associated with an increased risk of premenopausal breast cancer in general (OR, 1.19; 95% CI, 1.09- 1.29) and across various patterns of OC use. Among studies that provided data on nulliparous and parous women separately, OC use was associated with breast cancer risk in both parous (OR, 1.29; 95% CI, 1.20-1.40) and nulliparous (OR, 1.24; 95% CI, 0.92-1.67) women. These findings are grossly similar to the findings of the NZHTA review.

Henderson et al, 2011 [17] reviewed 8 prospective studies and 3 case control studies and found that chest radiation and mantle irradiation for Hodgkin in particular was a strong risk factor with rate ratio's ranging from 13 to 55.

Zhou et al, 2011 [18] did a meta-analysis of nine studies, , including 2,340 cases and 4,422 controls and found that atypical ductal hyperplasie (ADH) increased risk (OR = 2.93, 95% CI 2.16-3.97) and that atypical lobular hyperplasia (ALH) increased the risk even more (OR = 5.14, 95% CI 3.52-7.52). Women with a first-degree family history and atypical hyperplasia (AH) were at highest risk (OR = 4.87, 95% CI 2.89–8.20).

Supplementary evidence on risk models

After the search date of NICE 2006 we did a literature search from 2006 to current date (search date june 2011).

All publications evaluating familial risk concerned different forms of risk models; no publications updating the Collaborative Group on Hormonal Factors in Breast Cancer paper of 2002 were identified.

No systematic review fulfilling minimum quality criteria was identified. A narrative review was identified and used for reference tracking.

2.6.2.3. Models assessing the risk of developing breast cancer:

Overview (description based on narrative review Amir 2010[19])

Before continuing the discussion on risk models, we will give first an overview of the different existing models and models that are currently under development. One of the difficulties assessing models is that some of the models are evolving themselves, mostly adaptations to recent trends in epidemiology.

The risk assessment model that is most used and studied is the Gail model. This model was initially designed in 1989 using data that were collected as part of the Breast Cancer Detection and Demonstration Project, a nested case–control study of almost 300 000 women who were undergoing breast screening between 1973 and 1980. It was modified and updated in 1999. Both the original and the modified versions of the Gail model use six breast cancer risk factors, namely age, hormonal or reproductive history (age at menarche and age at first live birth), previous history of breast disease (number of breast biopsies and history of atypical hyperplasia), and family history (number of first-degree relatives with breast cancer)[19]

The Claus Model uses data from the Cancer and Steroid Hormone Study, a nested population-based case–control study conducted between 1980 and 1982 using breast cancer patients registered in eight SEER (Surveillance, Epidemiology, and End Results database) regions. Unlike the Gail model, it only uses family history to estimate risk but incorporates a substantially more comprehensive history than the Gail model, including unaffected first- and second-degree relatives and the age at which cancers in those relatives were diagnosed.[19]

The BRCAPRO Model, originally developed to assess the likelihood of carrying a BRCA gene mutation, also includes an extension software package enabling to calculate overall breast cancer risk. The Jonker model is a combination of the Claus model and BRCAPRO. The IBIS model, also known as the Tyrer–Cuzick model, based in part on a dataset acquired from the International Breast Intervention Study and other epidemiological data includes the most comprehensive set of variables of all the models. [19] The BOADICEA model, just like the BRCAPRO model, was originally developed to predict BRCA carriage but has been extended to enable it to estimate cancer risk as well.[19]

The validation studies presented hereunder are comparisons of different models, or are attempts to improve the original Gail model by either adding information, such as breast density, or by recalibrating the model using data on breast cancer incidence amongst different (mostly non USA) populations.

Validation studies: main findings

Tice et al 2005 [20] estimated and compared the predictive accuracy of the Gail model and of the Gail model combined with a measure of the breast density (BIRADS) and found a concordance index (c-index) of 0.67; [95% CI 0.65–0.68] for the Gail model and 0.68 [95% CI .66–.70] when breast density was included, a small but statistically significant improvement of the Gail model alone, ($p < 0.01$). Also breast density alone had a similar discriminative power (c-index 0.67 [95% CI 0. 65–0.68]). Chen et al 2006[21] also developed a modified version of the Gail model but did not assess accuracy.

Tice et al. 2008 [22] developed and evaluated a new model (sometimes referred to as the Tice model) with the inclusion of breast density as a parameter. The breast density model was well calibrated with an overall expected–observed ratio of 1.03 [95% CI, 0.99 to 1.06] but with a modest discriminatory accuracy (concordance index, 0.66 [CI, 0.65 to 0.67]), being no improvement compared to the Gail model.

Barlow et al. 2006[23] developed and validated a model using logistic regression on cohort data. Logistic regression on a 'learning' subsample identified following risk factors among premenopausal women: age, breast density, family history of breast cancer, and a prior breast procedure. For postmenopausal women more risk factors were identified: age, breast density, race, ethnicity, family history of breast cancer, a prior breast procedure, body mass index, natural menopause, hormone therapy, and a prior false-positive mammogram.

They validated the resulting model on a validation subsample, giving a c-statistics of 0.631 [95% CI = 0.618 to 0.644] for premenopausal women and 0.624 [95% CI = 0.619 to 0.630] for postmenopausal women, accuracies similar to the Tice model. It must be noted that this validation was done on different subsamples of the cohort on which the development of the model was done, so accuracy may be overestimated.

Decarli et al. 2006 [24] modified the Gail model using data from an Italian case control study and found that the calibration was slightly improved with overall expected/observed (E/O) ratios of 0.96 [95% CI 0.84 to 1.11] and 0.93 [95% CI 0.81 to 1.08] for the modified Gail model and the 'classic' Gail model, respectively. The average age-specific concordance statistics were 58.6% [95% CI 54.4% to 62.8%] for the modified Gail model and 58.8% [95% CI = 54.6% to 63.1%] for the 'classic' Gail model, indicating that discriminative power was not improved.

Evans et al, 2006 did a validation of Gail, Claus, BRUCAPRO and IBIS (Cuzick-Tyrer) on a Family History Evaluation and Screening Program in Manchester, UK, amongst 1,933 women with a mean follow-up of 5.27 years, of which 52 developed cancer. They found that the Gail, Claus and BRCAPRO model were poorly calibrated with ratios of expected to observed numbers of breast cancers of 0.48 [95% CI 0.37–0.64] for the Gail model, 0.56 [95% CI 0.43–0.75] for the Claus model, 0.49 [95% CI 0.37–0.65] for the BRCAPRO model and that calibration was better for the Cuzick-Tyrer model, namely 0.81 [95% CI 0.62–1.08] although confidence intervals overlap somewhat. Accuracy was similar for all models with an AUC of 0.735 for the Gail model, 0.716 for the Claus model, 0.737 for the BRCAPRO model and 0.762 for the Cuzick–Tyrer model.

Chlebowski et al. 2007 [25] validated the Gail model in post-menopausal women and their ability to estimate prevalence of both estrogen receptor positive and estrogen receptor negative tumors and found that the Gail model was poorly calibrated and underestimated 5-year invasive breast cancer incidence by approximately 20% (p <.001), mostly among those with a low estimated risk.

Accuracy was similar to other studies, with an AUC for the Gail model of 0.58 [95% CI 0.56 to 0.60]. Discriminatory performance was better for the risk of ER-positive cancer (AUC = 0.60, 95% CI = 0.58 to 0.62) than for the risk of ER-negative cancer (AUC = 0.50, 95% CI = 0.45 to 0.54) but clinical meaning or importance of this finding is unclear.

Schonfeld et al, 2010 [26] calibrated the Gail model and compared the newly calibrated Gail model with the 'classic' model on two different cohort studies. The Gail model significantly underpredicted the number of invasive breast cancers in both cohorts, with an expected-to-observed ratio of 0.87 [95% CI, 0.85 to 0.89], and 0.86 [95% CI, 0.82 to 0.90]. The updated model had an expected-to-observed ratio of 1.03 [95% CI, 1.00 to 1.05] and 1.01 [95% CI: 0.97 to 1.06].

Vacek et al 2011 [27] compared 4 models (Gail model, the Tice modification of the Gail model, the Barlow model, and the Vermont model) amongst women of 70 years and older and found that accuracy in this group was poor. C-statistics were 0.54 [95% CI = 0.52–0.56] for the Gail model, 0.54 [95% CI = 0.51–0.56] for the Tice modification of the Gail model, 0.55 [95% CI = 0.53–0.58] for a model developed by Barlow and 0.55 [95% CI = 0.53–0.58] for a Vermont model, which is a modification of the Barlow model.

Crispo et al [28] tried to improve the Gail model by adding information on second degree relatives to the model, but discriminatory power did not improve much. The concordance for the 'classic' Gail model was 0.55 [95% CI 0.53–0.58], for model including second degree relatives 0.56, [95% CI 0.53–0.59] and a concordance statistic of 0.57 [95% CI 0.54–0.60] for the combination of the two models.

Several authors attempted to improve models with genetic data. Wacholder 2010[29] compared the Gail model with the Gail model modified using 10 common genetic variants associated with breast cancer and found that accuracy was only modestly improved, from an AUC of 0.580 to an AUC of 0.618.

Mealiffe 2010[30] found a similar modest improvement for a model adding single-nucleotide polymorphisms (SNP) with area under the curve of 0.594 compared with area under the curve of 0.557 for Gail risk alone ($P < .001$).

Two authors evaluated models amongst women with benign breast disease. Pankratz 2008[31] applied the Gail model on the Mayo Benign Breast Disease cohort and found a very poor performance with a concordance statistic of 0.50 [95% CI, 0.44 to 0.55]. Boughey 2010[32]

applied the Tyrer-Cusick model on the same cohort with a similar poor performance of the model with an observed-to-predicted ratio of 0.53 [95% CI 0.37 to 0.75] and a concordance statistic of 0.540.

2.6.2.4. Models that assess the risk of carrying a germline mutation, such as BRCA1, BRCA2 and TP53 mutations

A different branch of models assess the risk of carrying a germline mutation such as BRCA1, BRCA2 and TP53 mutations and not the risk of developing breast cancer. They are nearly all evaluated in specialized genetic clinics and aim at reducing the need for expensive genetic testing. As stated before, some of those models have extensions that enable them to assess or estimate the breast cancer risk. These models are not tested on cohorts but in transversal studies, where the model serves as 'test' and where the results of genetic testing are applied as 'gold standard'.

Kang et al. 2006 [33] evaluated the accuracy of the prediction algorithms BRCAPRO, Manchester, Penn and the Myriad-Frank and found that accuracy was moderate and similar for all models: BOADICEA Manchester 0.759 (CI 0.688-0.831), BRCAPRO 0.743 (CI 0.672-0.814), Myriad 0.753 (CI 0.680-0.827), Penn 0.757 (CI 0.686-0.827) and that all models have high false-negative and false-positive rates using 10 % probability thresholds.

Ruodgari et al 2007 [34] evaluated the accuracy of the probability estimation models COS, Manchester scoring system (MSS), BOADICEA and Tyrer–Cuzick (T–C). COS and MSS models demonstrated the greatest sensitivities and area under ROC curves for the majority of family structures. They also showed the highest sensitivities (91–92%) and AUCs (76–78%) for the entire dataset overall. However, BOADICEA and T–C had the highest specificities for the majority of the family structures. BOADICEA and T–C generated the best estimates for the prevalence of mutations in the population.

Parmigiani 2007 et al [35] evaluated the accuracy of BRCAPRO, family history assessment tool, Finnish, Myriad, Yale, NCI Penn. All models showed similar AUC: BRCAPRO 0.82 (0.81–0.84) Yale 0.71 (0.68–0.74) Myriad 0.77 (0.75–0.79) NCI Penn 0.76 (0.74–0.79) FHAT 0.77 (0.75–0.8) Finnish 0.78 (0.75–0.8) and all models have high false-negative and false-positive rates across a range of probability thresholds used to refer for mutation testing.

Antinou 2008 et al [36] evaluated the calibration and accuracy of the prediction algorithms BOADICEA, BRCAPRO, IBIS, the Manchester scoring system and Myriad tables and found that only BOADICEA was well calibrated (only for BOADICEA no statistically significant difference E/O), that all models underestimate probability in low risk population and that accuracy was moderate and similar for all models (BOADICEA=0.77, BRCAPRO=0.76, IBIS=0.74, Manchester=0.75, Myriad=0.72).

Panchal 2008 [37] evaluated the accuracy of the BRCAPRO, Manchester, Penn II, Myriad II, FHAT, IBIS and BOADICEA models. They found that BRCAPRO, Penn II, Myriad II, FHAT and BOADICEA models all have similar AUCs of approximately 0.75 for BRCA status and that the Manchester and IBIS models have lower AUCs (0.68 and 0.47 respectively).

At a 10 % testing threshold, the sensitivities and specificities for a BRCA mutation were, respectively, as follows: BRCAPRO (0.75, 0.62), Manchester (0.58, 0.71), Penn II (0.93, 0.31), Myriad II (0.71, 0.63), FHAT (0.70, 0.63), IBIS (0.20,0.74), BOADICEA (0.70, 0.65).

Lindor 2010 [8] evaluated the calibration and accuracy of LAMBDA, BRCAPRO, modified Couch tables and Myriad II tables and found that all models gave similar areas under the ROC curve of 0.71 to 0.76. All models except LAMBDA substantially under-predicted the numbers of carriers.

3. DISCUSSION

Risk assessment based on family history is largely based on the results of the Collaborative Group on Hormonal Factors in Breast Cancer study, where a meta-analysis was performed using the primary data of 50 epidemiological studies (2001) to guide the levels that are presented in the guideline. We didn't find similar more recent studies and guidelines of NICE and U.S. Preventive Services Task Force still use this study to establish the family risk categories.

The NZHTA report of Weir et al 2007 reviewed non familial risk factors for breast cancer and identified as strong risk factors history of breast cancer, a range of breast lesions including ductal carcinoma in situ as risk factors, but this is more related to follow up issues and not useful for our purpose. Breast density was identified as a strong risk factor; this was confirmed in latter systematic reviews, with a relative risk of about 4 when comparing the higher risk groups with the lower categories, be it measured with Wolfe grade, BI-RADS or % of breast area that is dense. This must be nuanced however, Carney et al [38] presents the frequency of the different BIRADS categories among women aged from 50 to 69 undergoing screening in 7 population-based mammography registries in the US:

Almost entirely fatty (BIRADS 1) 42 237 (9.1)

Scattered fibroglandular tissue (BIRADS 2) 218 129 (47.0)

Heterogeneously dense (BIRADS 3) 167 003 (36.0)

Extremely dense 36 303 (BIRADS 4) (7.8)

This implies that the lowest category is only present in a minority of the women and that RR with BIRADS 2, representing the majority of women, is only around 2..

Chest radiotherapy and mantel irradiation for Hodgkin lymphoma is a strong risk factor.

Some risk factors cannot be used for risk assessment in routine practice, such as intake of soy products. We excluded this kind of risk factors. Factors as alcohol use or body mass index could be measured more easily in routine practice but one can question feasibility of such an approach and they may be more useful for advice on prevention, however this is not the scope of this report.

Other risk factors that may be useful are related to hormonal status of the women, such as parity, age of menarche and use of oral contraceptives or hormone replacement therapy, although association with breast cancer is weaker than dense breasts.

All individual risk factors have numerous interactions amongst themselves and cannot be simply added. This is the main reason why there is an increasing interest in risk models, where the Gail model is the best known and the most studied.

A first class of models estimates the risk of developing breast cancer, either expressed as a 5 years, 10 years or lifetime risk. Calibration, which is a measure of the degree the risk % given by the model corresponds to the actual risk, may be the most important measure here. Validation studies find the 'classic' Gail model under-predicts risk, and attempts are done to 'recalibrate' the model, using more recent data or data of different populations, be it minority groups (such as Afro-American or Asian people in the US) or populations in the countries where the model needs to be used such as Decarli at al did on an Italian population. It may be useful to do the same for a Belgian population, but this would require databases that are currently not available in Belgium.

Another major disadvantage of the Gail or Claus model is that they only use a limited number of elements. Several studies attempted to improve the Gail model or to develop a new model using a more comprehensive set of risk elements. The Cuzick Tirer model, also known as the IBIS model includes, apart from elements from the family history, BMI index, a number of hormonal factors and antecedents of breast cancer and breast lesions. Studies indicated that they have a better calibration and accuracy then the Gail or Claus models, but this needs confirmation. They do not include however breast density. Several attempts were done to include breast density, such as the model by Tice et al. but there is more need for independent validation. Accuracy of this type of models is rather poor and the area under the curve rarely exceeds 0.6, seriously limiting their ability to target invasive prevention measures. Attempts to improve the model using more common single-nucleotide polymorphisms (SNP) only had a

limited impact on model performance, and do not seem to be useful at the moment given the considerable cost involved in the testing.

Models that assess the risk of carrying a germline mutation, such as BRCA1, BRCA2 and TP53 mutations seem to have a somewhat better discriminatory power; none of the models came out as being really superior to the other models.

Key points

- **Family risk**

Women can be categorised in 3 risk categories based on family history.

Average risk:

- **only one first-degree or second-degree relative diagnosed with breast cancer at older than age 40 years.**

<u>Raised risk</u> (that is, a 10-year risk of 3–8% for women aged 40–49 or a lifetime risk of 17% or greater but less than 30%):

- **one first-degree relative diagnosed with breast cancer at younger than age 40 years, or**
- **two first-degree or second-degree relatives diagnosed with breast cancer at an average age of older than 50 years, or**
- **three first-degree or second-degree relatives diagnosed with breast cancer at an average age of older than 60 years**

<u>High risk</u> (that is, a 10-year risk at age 40–49 years of greater than 8% or a lifetime risk of 30% or greater):

- **two first-degree or second-degree relatives diagnosed with breast cancer at younger than an average age of 50 years (at least one must be a first-degree relative), or**
- **three first-degree or second-degree relatives diagnosed with breast cancer at younger than an average age of 60 years (at least one must be a first-degree relative), or**
- **four relatives diagnosed with breast cancer at any age (at least one must be a first-degree relative).**

In case one of the following is present in the family history, women should always be considered at high risk:

- **bilateral breast cancer**
- **male breast cancer**
- **ovarian cancer**
- **Jewish ancestry**
- **sarcoma in a relative younger than 45 years of age**
- **glioma or childhood adrenal cortical carcinomas**
- **complicated patterns of multiple cancers at a young age**

• very strong paternal history (four relatives diagnosed at younger than 60 years of age on the father's side of the family).

- **Risk factors:**

Breast density was identified as a strong risk factor with a RR of around 4 when comparing the highest risk group with the lowest category, be it measured with Wolfe grade, BI-RADS or % of breast area that is dense. However, it must be noted that women with the lowest category are a minority of the women, and that the RR drops to 2 compared with women with a BIRADS 2.

Chest radiation and mantle irradiation for Hodgkin in particular is a strong risk factor with rate ratio's ranging from 13 to 55.

Atypical epithelial hyperplasia (lobular and ductal) is associated is a strong risk factor

Risk factors that related to hormonal status of the women, such as parity, age of menarche and use of oral contraceptives or hormone replacement therapy, are more weakly associated with breast cancer

- **Risk models:**

All individual risk factors have numerous interactions amongst themselves and cannot be simply added.

The Gail model is the most studied, but has a number of disadvantages such as underprediction of risk and the use of a limited number of factors.

The Cuzick Tirer model (IBIS) include a more comprehensive list of risk elements and studies indicated that they have a better calibration and accuracy than the Gail or Claus models, but this needs confirmation. They do not include however breast density.

Several attempts were done to include breast density in the models, such as the model by Tice et al. but there is more need for independent validation.

Models perform poorly in women with benign breast disease.

CHAPTER 3 TECHNICAL METHODS FOR BREAST CANCER SCREENING

1. INTRODUCTION

The ultimate goal of screening is the reduction of breast cancer related mortality by detecting the disease in an early and curable stage. Ideally, all tests considered for breast cancer screening should be evaluated for their effect on breast cancer related mortality, both in randomized controlled trials and after implementation in a population-based screening program.

This type of evaluation requires a large sample size and an observation period of minimal 7 years in a clinical trial and even longer to detect benefits outside a trial setting. Therefore, several short-term parameters to assess possible screening tests are used. A valuable early surrogate of mortality is the rate (not proportion) of advanced cancers[6].

To date, such extensive evaluation is only available for mammography. Most often, alternatives for or adjuncts to mammography are evaluated for their accuracy by measuring the sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV) in cross-sectional or cohort studies. This limited evaluation of tests holds several dangers.

Firstly, overdiagnosis and length bias are not taken into account. Overdiagnosis refers to the diagnosis of cancers that would never have become clinically apparent (and thus never have lead to treatment morbidity and mortality) if not detected by screening. Length bias refers to the undue proportion of cancers with a long sojourn time (defined as the detectable preclinical phase) and probably a good prognosis in the group of screening detected cancers. In other words, it is not sufficient to show an additional detection rate for a test to proof its beneficial effect on treatment decisions and mortality[6].

Secondly, results achieved in an ideal trial setting may not be applicable if applied in a decentralized population-based setting. This may especially be the case for techniques with a significant operator dependence and inter-observer variability.

Thirdly, problems arise when defining true and false positive and negative results of a test. The 'gold' standard to define a true positive result is cancer (invasive or *in situ*) proven on cytology or biopsy. As a fine needle aspiration (FNA) or a biopsy is only performed when the test result is considered positive, a work-up bias is inherent to all screening studies. To discriminate true negative from false negative test results another reference standard is thus needed. An acceptable definition of false negative test results is the group of women presenting with clinical disease during a follow-up period, e.g. one year. The question also arises if 'overdiagnosed' cancers (see above) can be considered true positives[6].. Furthermore, accuracy parameters will differ depending on diagnostic thresholds and whether the parameters are calculated for the test solely or for the complete screening episode. For example, if a 'positive' mammography is followed by a diagnostic ultrasound which is 'negative', this patient would be included differently in the calculations[6].

In this chapter, keeping in mind the considerations mentioned above, we attempt to answer the following questions:

- What are the possible benefits and limitations of double reading, including computer-aided detection, versus single reading mammography?
- What are the possible benefits and limitations of full-field digital mammography versus film-screen mammography?
- What is the current level of evidence supporting the use of ultrasound in breast cancer screening in the general population or in selected populations? What is the balance harm-benefit?
- What is the current level of evidence to promote MRI as a breast cancer screening tool in high risk populations? What is the balance harm-benefit?

2. METHODS

2.1. Literature search strategy

The search for the clinical literature about full-field digital mammography, computer-assisted interpretation/detection of mammography, ultrasound and magnetic resonance imaging (MRI) as screening tools for breast cancer included the consultation of electronic databases up to June-July 2011.

The search was done in 2 steps. Firstly, the following databases were searched to retrieve meta-analysis (MA), systematic reviews (SR), health technology assessments (HTA) and evidence based guidelines: Embase, Medline via Ovid, Center for review and dissemination databases (CRD, DARE, NHS, EED, HTA), Cochrane database of Systematic review (CDSR), National guidelines clearinghouse, guidelines international Network, CBO, Evidence based medicine guidelines, Guidelines finder UK, New Zealand guidelines group, HAS, NICE, SIGN.

Secondly, after selection and critical appraisal (see below), a search was performed to identify primary studies published after the most recent selected SR, meta-analysis or evidence-based guideline The following databases were searched to retrieve randomized controlled trials, cross-sectional studies and prospective cohort-studies: Embase, Medline via Ovid and the Cochrane Library for Clinical trials. Studies published between 2007 and search dates were included.

An overview of the search strategies is captured in appendix 1.

Reference lists of selected papers were checked for additional useful publications.

2.2. Selection criteria

Table 10. Selection criteria for SR, meta-analyses, HTA and evidence-based guide

Selection criteria	*Inclusion criteria*	*Exclusion criteria*
Population	All ages women without symptoms of breast cancer, with or without risk factors	Current breast cancer or breast diseases symptoms
Intervention	Screening with mammography (single or double reading) compared with digital mammography (computer aid?) and/or mammography + ultrasound and/or MRI (with or without mammography)	Other tests used for screening (clinical examination, doppler sonography,…) or for diagnosis (biopsy, scintimammography, PET-scan,…)
Outcome	Accuracy (sensitivity, specificity, PPV, PPN), mortality, morbidity, radiations risks	Physiological outcomes
Design	HTA, SR, MA or guidelines based on systematic review	Other design: primary studies, letters, editorial, narrative review, guidelines based on consensus, cost effectiveness studies

Table 11. Selection criteria for the primary studies

Selection criteria	*Inclusion criteria*	*Exclusion criteria*
Population	All ages women without symptoms of breast cancer, with or without risk factors	Current breast cancer or breast diseases symptoms
Intervention	Screening with mammography (single or double reading) compared with digital mammography (or computer assisted) and/or mammography + ultrasound and/or MRI (with or without mammography)	Other tests used for screening (clinical examination, doppler sonography,…) or for diagnosis (biopsy, scintimammography, PET-scan,…)
Outcome	Accuracy (sensitivity, specificity, PPV, PPN), mortality, morbidity, radiations risks, safety	Physiological outcomes, cost-effectiveness
Design	Primary studies: RCT, cross-sectional, prospective cohort studies	Other design: letters, editorial, narrative review, guidelines, cost effectiveness studies, SR, meta-analysis

No language restrictions were applied at this stage.

2.3. Selection procedure

Our study selection started by looking at titles and abstracts to exclude any studies considered not relevant for our purposes. Articles that appeared relevant or for which we had doubts were assessed by reading the full text.

Relevant titles and abstracts were selected in parallel by two reviewers. Any disagreements were discussed and a common decision and approach adopted. Following that, the full articles of studies found were evaluated by two reviewers.

A hierarchical approach was followed by which:

Firstly, the analysis focused purely on MA, SR, HTA and guidelines published up to the date of our search. Secondly, the selected evidence synthesis was updated by looking at all relevant original literature published after the search data of selected MA, SR, HTA and guidelines found via our search.

Finally, the reference lists of the selected studies were checked for additional relevant studies that could be included in our review.

2.4. Critical appraisal

The reviewers critically appraised the SRs and MAs according to the Checklist for systematic review of diagnostic research of the Dutch Cochrane Centre, guidelines were appraised using the AGREE II checklist. Primary studies were assessed following the QUADAS checklist for diagnostic accuracy studies for cross-sectional studies and using the checklist for randomized controlled trials from the Dutch Cochrane centre for randomized controlled trials.

3. RESULTS

3.1. Research and selection

The 2 steps of the literature search gave the following results:

Literature selection process for SR, MA, HTA, guidelines

After automated eliminating duplicates, searches on the previously mentioned databases listed 550 citations. Of those, 514 did not meet our inclusion criteria based on title or abstract or were duplicates. Of the 36 citations left, eleven were excluded from the analysis after exploring the full version of the study leaving us with a total of 25 relevant studies. Critical appraisal excluded a further 14 articles. The results of the critical appraisal are summarized in appendix 2, 0.

Two of the selected systematic reviews reported on double reading[39, 40], one review on computer-aided detection[41] and one review on digital screening[42].

Four of the selected systematic reviews reported on ultrasound[43-46] and also four on MRI[43-45, 47].

Literature selection process for primary studies: randomized controlled trials, cross-sectional studies and prospective cohort-studies.

After automated eliminating duplicates, searches on the previously mentioned databases listed 1160 citations. Of those, 1059 did not meet our inclusion criteria based on title or abstract or were duplicates. Of the 101 citations left, 52 were excluded from the analysis after exploring the full version of the study leaving us with a total of 45 relevant studies. Of these 45 studies, 8 report on MRI, 8 on ultrasound and 4 on both MRI and ultrasound, 15 on digital screening and 6 on computer-aided detection and 4 on double reading.

After consultation with the experts, an extra search in Medline and Embase is performed on double reading, resulting in a supplementary 10 studies. This additional search is performed to find out what the value in clinical outcomes is of double reading compared to single reading and computer-aided detection.

Searching of reference lists of selected studies resulted in 2 additional studies reporting on both MRI and ultrasound. No additional studies were selected for double reading and digital mammography.

Flowcharts of search results are listed in appendix 2

3.2. Double reading of mammography as screening tool

Computer-aided detection is nowadays not widely implemented in the Belgian screening units. Therefore, it is decided only to mention the computer-aided detection as a comparator to double reading and not as a widespread screening tool.

3.2.1. Systematic reviews, meta-analyses, health technology assessments and evidence based guidelines

3.2.1.1. Double reading

After selection and critical appraisal, two systematic reviews on the use of double reading in breast cancer screening, were selected. No good quality meta-analysis could be identified in the literature.

A summary of characteristics and results of the two reviews is presented in table 24.

Based on 10 cohort studies, Dinnes et al, 2001,[39] found an increase in cancer detection rate after double reading (overall increase ranging from +2.9 to +11.2 per 10.000 women screened).

The change in recall rate depended on the recall policy: double reading with unilateral recall increased the recall rate (by between 38 and 149 per 10,000 women screened), whereas double reading with arbitration or consensus decreased the recall rate (by between 61 and 269 per 10.000 women screened). We see a similar pattern for specificity: a decrease in specificity with unilateral recall and an increase in specificity with arbitration or consensus. The sensitivity increased with double reading, independent from the recall policy.

In addition, the cancer detection rate increased more in the studies with single-view mammograms compared to the studies with two-view mammograms (4.4-6.9 per 10,000 versus 3.0-4.4 per 10,000).

The review of Dinnes et al concludes that a screening protocol consisting of double reading with arbitration or consensus improves the sensitivity. The consensus or arbitration procedure after double reading of the mammograms can decrease the number of women recalled for unnecessary assessment.

The review of Taylor et al, 2008[40],[40], based on 17 studies, confirms the results of Dinnes et al[39]: an overall increase in the cancer detection rate and a decrease in recall rate after double reading combined with arbitration or consensus, in contrast to the increased recall rate after double reading combined with unilateral recall.

3.2.1.2. Computer-aided detection (CAD)

After critical appraisal of the reviews on computer-aided detection mammography (CAD), Two reviews (Noble et al, 2008[41] and Taylor et al, 2008 [40]) were maintained.

In the review of Noble et al[41] 7 studies were included and pooled results were calculated where possible. As mentioned in the methods section, main focus is put on sensitivity, specificity, recall rate and cancer detection rate. Next to these main variables, other results, such as biopsy rates, variables grouped per subgroups, etc will also be presented.

The pooled sensitivity in the review (based on 3 studies, n= 347 324 women) was 86.0% (95%CI 84.2-87.6%) and specificity was 88.2% (95% CI 88.1-88.3%). Despite the heterogeneity in the estimation of the sensitivity (I^2=87.2%) and specificity (I^2= 99.7%), the 95% confidence intervals were narrow and the sensitivity analysis was robust.

In comparison with single-read mammography the incremental cancer detection rate with CAD was 50 women per 100 000 women screened (95%CI 30-80 women).

The additional recall rate in healthy women was 1190 per 100 000 (95% CI 1090-1290). These women would not have been recalled based on single-read mammography only. Of these recalled women, 4.1% (95% CI 2.7-6.3%) were diagnosed with cancer and 96% (95% CI 93.9-97.3%) were healthy. Unexplained heterogeneity and lack of robustness affect credibility of these findings.

Next to recall rate, biopsy rate was calculated. Based on the CAD system, an additional of 80 per 100 000 biopsies (95% CI 60-110) were performed. Of these women, 65% (95% CI 52.3-76.0%) underwent biopsy but were healthy and 35.9% (95% CI 24.7-48.9%) were diagnosed with cancer. These rates are calculated with the data of five studies (n= 51 162 women).

The above-mentioned review of Taylor et al[40], 2008 gathered 10 studies which compare single reading with single-reading combined with CAD No statistically significant increase in cancer detection rate could be found. However an increased recall rate was seen, independent from the heterogeneity between the studies. The authors conclude that more evidence exists for the improvement in screening performance with double reading with arbitration compared to single-reading combined with CAD.

3.2.2. Primary Studies: randomized controlled trials, prospective cohort studies and cross-sectional studies published 2007-2011

The available evidence from the reviews was updated with primary studies published after the search date of the most recent systematic review. Findings of primary studies are discussed in the following paragraphs (3.2.2.1 and 3.2.2.2)

3.2.2.1. Single reading versus double reading

In the studies, published between 2007 and 2011, no primary studies could be found specifically on the comparison of double reading versus single reading. Next to the clinical outcomes after double reading of the mammograms, five primary studies assess more in detail the inter-observer variability. A summary of results of the studies can be found in appendix.

Hofvind et al[48], 2009, based on data from the Norwegian Breast Cancer Screening Program, found, after independent, blinded double reading of 1 033 870 screening mammographies, 54447 (5.3%) discordant interpretations and 21 928 (2.1%) positive concordant interpretations. Consensus was sought for discordant interpretations and concordant positive findings. 66.8% of the discordant findings and 17.9% of the concordant positive findings were found negative. The recall rate was 3.5%. Of the total detected cancers (5611 cancers), 23.6% were detected following discordant interpretations. There were some significant differences between discordant and concordant cancers:

- Proportion of micro-calcifications was higher in discordant cancers (24.9% versus 17.7%)(p<.001)
- Mass or density with micro-calcifications was lower in discordant cancers (11.1% versus 15.4%)(p<.001)
- Proportion of DCIS was higher in discordant cancers (23.9% versus 15.7%)(p<.001)
- Lobular cancers were less frequent in discordant cancers (7.3% versus 9.1%)(p=.035)

Of the total number of interval cancers (n=1791), 117 (6.5%) were found in dismissed discordant interpretations, revealing a substantially higher number of cancers compared to negative screenings. The authors conclude that the disagreements on microcalcifications are possibly due to a lack of competences of the readers in detecting microcalcifications on mammograms.

Caumo et al[49], 2010 examined the effect of a third reader at arbitration of discordant interpretations on the recall rate. In this study consisting of a consecutive series of 7.660 double readings of screening examinations, all discordant interpretations were redirected for further assessment, independent from the judgment of the third reader. Of the 49 detected cancers (43 concordant and 6 discordant cancers) 6 cancers are detected in the arbitrated cases (5 positive and 1 negative arbitrations). The one negative arbitration implies one missed cancer if only positive arbitrations would be redirected for assessment. Nevertheless the missed cancer (0.13% absolute or 2.0% relative reduction of cancer detection rate), the arbitration of discordant interpretations would spare out 216 assessments, resulting in a 2.8% absolute or 40.9% relative decrease in recall rate. The overall recall rate in this study was 528 (6.8%) of which 312 cases consisted of concordant interpretations and positive arbitrations. The arbitration of discordant interpretations is stated by the authors as the preferred practice in order to reduce the amount of (unnecessary) recalls. Ciatto et al[50], 2005 confirms these findings: arbitration reduced the referral rate from 3.82% to 2.59% and the number of cancers detected per 1000 women screened decreased from 4.58 to 4.50.

The review of discordant interpretations by a consensus panel is a possible alternative to arbitration. In the study of Shaw et al[51,51], 2009, 1335 cases (1.04%) were reviewed by a consensus panel: 606 (45.4%) were redirected for further assessment (US, biopsies), 71 cancers (7.3% of the total of 968 cancers) were identified. Similar to the study of Hofvind et al, 2009[48], the highest proportion of patients with calcifications were found in the group of discordant findings (32%). Outcomes after consensus review: sensitivity 90%, specificity 57%, and negative predictive value 99%. Comparing the recall rate and cancer detection rate between different recall policies (highest reader recall, unanimous recall only, discordant findings due to calcifications), the best results were obtained with the approach of only recalling the patients with discordant calcifications: increase from 98.98% to 99.66% for negative predictive value and only a small increase in recall rate (0.05%).

Inter-observer variability was assessed by Duijm et al[52], 2009. Different set-ups were compared: single-reading by radiologists, double-reading by radiologists, double-reading by radiographers and single reading by radiologist, double reading by a radiographer and secondary referral by two radiologists, double reading by a radiologist and a radiographer and referral of all positive findings. More details on each set-up can be found in the annexes, main findings of the study were a significant (7.3% relative) increase in sensitivity for the double reading groups. The different set-ups had its impact on cancer detection rate and recall rate. The highest sensitivity is obtained by a protocol that takes into account the interpretation of 4 readers combined with a referral of all positive findings, increasing the recall rate (and unfortunately the related cost and burden on patients). The benefits and possible harms of this scenario have to be further assessed before it can be recommended for implementation in a national screening program.

3.2.2.2. Single/Double reading versus CAD

Most studies about the performance of CAD in a screening population compare the clinical outcomes after double reading with the clinical outcomes after single-reading combined with CAD. If single reading combined with CAD results in a higher sensitivity compared to double reading, this kind of screening tool could be advantageous for implementation in clinical practice. The replacement of the double reader (or even the arbitration reader) by a software program could be beneficial for the screening units. If we look in detail at the primary studies, the results show that it is more complex than that. We included 5 primary studies. The summary of these studies can be found in appendix.

The study of Ciatto et al, 2003[53] shows an improvement in sensitivity but also a reduction in specificity. There is a slight, but not statistically significant higher number of detected cancer in CAD compared to double reading (90.0% versus 85.8%) but this is counterbalanced by an increased recall rate (CAD 11.4% versus double-reading 7.9%, p=0.003). A later study of Ciatto et al, 2006[54] , leads to the same conclusions: no statistically significant difference in sensitivity between double reading and CAD and reduced specificity, leading to an excess of false-positive marks.

Gilbert et al, 2008 [55]found in his equivalence trial (with matched-pair comparisons between cancer detection rates) that single-reading combined with CAD was equivalent to double reading (i.e. equivalence was defined as a 95% confidence interval that ruled out a difference of more than 10% in either direction in the rate of cancer detection): no statistically difference in cancer detection rate (sensitivity of 87.2% or 7.02 per 1000 women screened with CAD versus sensitivity of 87.7%, difference 0.50 % (c.i.−7.4% to +6.6%) or 7.06 per 1000 women screened with double reading) in contrast to a small but significant increase (p<0.001) in recall rate with CAD (3.9% compared to 3.4%).

A study within the national screening program in Australia (Cawson et al, 2009 [56]) also came to the same conclusion: differences in sensitivity between CAD and double reading was not statistically significant but results were very reader dependent.

These results are contradicted by the results of the study of Khoo et al, 2005[57]. This study, done within the framework of the national screening program in the UK, demonstrates an increased sensitivity by 1.3% with CAD compared to single reading (single-reading combined with CAD 91.5% versus single-reading 90.2%), but double reading increased the sensitivity by 8.2% (sensitivity of 98.4%), they did not formally assess if this difference was statistically significant. Recall rates (for arbitration and for assessment) were significant higher with CAD: 13.8% with CAD versus 10.5% with double reading for recall rate for arbitration and 6.1% with CAD versus 5.0% with double reading for recall rate for assessment.

3.2.3. Discussion

Both reviews ([39, 40]) come to the same conclusion: double reading increases the sensitivity and decreases the recall rate (if arbitration is used). Several countries, including Belgium, have implemented the double reading procedure in their national screening program. Nevertheless the advantages of independent double reading the same mammographic image, this extra reading procedure implies an increased workload. In the review of Dinnes et al, 2001[39] only indirect evidence was found that double reading may be more cost-effective compared to single reading. Due to the shortage of radiologists in some countries, researchers compared the reading performance of radiographers to the performance of radiologists, but this is not relevant for the Belgian situation. The approach how to handle with discordant interpretations (recall of all these findings, arbitration or consensus) has its impact on the recall rate and the cancer detection rate.

The CAD system is a supplementary tool to the interpretation of the radiologist (or image reader), aiming to increase the number of detected cancers. However, this small increase in sensitivity is counterbalanced by the significant increase in recall rate. The increased recall rate leads to patient distress, increased number of health care visits and a change in attitude towards screening mammography.

The way CAD is done implies a software-matic analysis of the images, putting marks on suspicious masses, followed by reading by a human reader and a decision on which marks are true-positives and which ones are false-positives. But the high number of positive marks per image makes that the human reader is overwhelmed, resulting in a decreased specificity, an increased time to read the images and an increased number of biopsies (of healthy women). An increased cancer detection rate is an advantage in screening for breast cancer, but the real added value of this increased rate is determined by the stage and type of the cancer that is missed by the first reader. Only 2 of the four studies in the review of Noble et al, 2008 [41]reported the type and stage of cancer, this limits the representativeness and generalisability of these finding. Also the clinical importance of these detected cancers is not assessed.

The interpretation of the results is limited by the poor internal validity of the primary studies in the review, caused by the retrospective design, the lack of blinding to clinical information (or lack of reporting about this blinding) and the lack of reporting about the case selection methods. Also the specificity and sensitivity of the CAD system could be overestimated due to the restricted follow-up period of one year. Slow-growing cancers can be missed due to this limited follow-up time and this will lead to false-negatives.

Apart from the methodological limitations, the set-up of a national screening program varies between countries, with differences in age, interval-screenings, etc.

Key points

- **Double reading:**
  - **increases the sensitivity compared to single reading**
  - **decreases the recall rate if arbitration is applied**
- **double reading is widespread used in national screening programs**
- **Single-reading combined with CAD:**
  - **Only a small increase in sensitivity**
  - **a significant increase of false-positive marks and increased recall rate**
- **CAD enhances the reading performance of the single reader but the clinical outcomes are comparable (or worse in case of recall rate) to double reading**

- **The high number of false-positive marks in CAD requires an additional arbitration, which decreases the specificity and enhances the implementation in a national screening program**

3.3. Full-field digital mammography as screening tool

3.3.1. Systematic reviews, meta-analyses, health technology assessments and evidence based guidelines

After critical appraisal, one systematic review on full-field digital mammography (FFDM)[42] was selected. However language restrictions made it difficult to fully understand the analysis of the authors. Therefore we decided to re-analyze the primary studies, but the search strategy and critical appraisal of the primary studies were maintained..

3.3.2. Primary Studies: randomized controlled trials, prospective cohort studies and cross-sectional studies included in the 1 selected systematic review

In the review of AETSA et al[42] 10 primary studies were included. The results of the primary studies are grouped per country in which the screening program was performed or by research group. First, attention was given to the findings about recall rate, cancer detection rate and biopsy rate, secondly to other variables and the presence of subgroups in the population of screened women. All studies compare full-field digital mammography (FFDM) to screen-film mammography (SFM) for cancer detection in a population of asymptomatic women, performed in a nationally organized screening program.

3.3.2.1. Lewin et al, 2001, 2002, Glueck, 2007

The results of the screening program in the United States are mentioned in the publications of Lewin et al, 2001, 2002 and Glueck et al, 2007[58-60].

In the study of Lewin et al, 2001[58], 3890 asymptomatic women of 40 years and older were examined (total of 4945 examinations) with both FFDM and SFM and re-examined after a minimum of 11 months. There were 1448 positive findings (findings recommended for evaluation by at least one of the two readers), of which 507 by FFDM, 746 by SFM and 195 on both. The next study of Lewin et al, 2002[59], is a similar analysis but with a larger number of examined women. Of the 6736 examinations of 4489 women of 40 years and older, 1467 were positive for additional examinations, of which 1345 were determined by SFM and 979 by FFDM and 293 by both modalities.

The recall rate of both studies is significantly lower for FFDM in comparison to SFM (p<.001): FFDM 11.5-11.8%, SFM: 13.8-14.9%.

The difference in number of biopsies between SFM and FFDM became more significant between both studies: SFM (83/152 and 87/181), FFDM (28/152 and 38/181) and both (31/152 and 56/181). No significant difference in cancer detection rate could be found between SFM and FFDM in both studies (total number of cancers of 35/152 in the first study and 42/181in the second study). In the study of Lewin et al, 2001[58], the calculated sensitivity for cancer detection confirms that there is no statistically significant difference between FFDM (60%, 31 of 35) and SFM (63% (22 of 35) (relative sensitivity of 95% (21 of 22) of FFDM to SFM).

The positive predictive value, defined as the fraction of recalled examinations that led to a diagnosis of breast cancer, was slightly lower for SFM (3.2-3.3%) than for FFDM (3.7-3.4%) but the difference is not statistically significant.

Results of these studies are biased because they do not take into account the interval cancers and by the high level of disagreement about the interpretation of the examinations (17% of all examinations had discordant interpretations), indicating reader variability.

In the study of Glueck et al, 2007[60], the data of the study of Lewin et al, 2002[59] were used for the comparison of the area under the curve (ROC) between SFM, FFDM and the combination of both. No difference in Roc could be found using the parametric tests. The non-parametric tests show a statistically significant difference between SFM versus combined (p=.008) and between FFDM versus combined test (p=.0008). No significant difference in ROC was found between SFM versus FFDM. These results indicate the increased cancer detection rate when using both modalities (83.7%) instead of SFM (65.3%) of FFDM (55.1%) as a standalone modality. The implementation of the strategy of using both modalities in mammographic screening has some financial and practical restraints, such as an increased cost and manpower. The authors could not determine whether the number of readers, the number of compressions

or the use of both modalities resulted in an increased cancer detection rate.

The main conclusion that can be drawn out the results of the three studies is the lack of difference in cancer detection rate between SFM and FFDM. Only the recall rate is significantly lower in FFDM.

3.3.2.2. The Oslo studies

Skaane et al, 2003, 2005, 2004, 2007[61-64] compared the performance of SFM and FFDM with soft-copy reading in the Norwegian population-based breast cancer screening program (Oslo I and Oslo II studies).

The results of the Oslo I (n= 3683 women underwent both SFM and FFDM) study[61, 63] found no difference in performance between SFM and FFDM: no statistically significant difference in cancer detection rate and in contrast to the studies of Lewin et al[58, 59], a slightly higher recall rate was found for FFDM. The higher recall rate could be explained by a learning curve effect of the readers.

The increased number of participants in the Oslo II study[62] (n= 14 436 women aged 50-69 years of which 10 391 women underwent SFM and 4 045 women FFDM) enlarges the minor differences found in the Oslo I study between SFM and FFDM: a higher (but not significant) cancer detection rate with FFDM and a significantly higher recall rate with FFDM in the group aged 50-69 years (p<.05). The lack of difference in PPV underlines the comparable performance of SFM and FFDM.

The Oslo II study follow-up results[64] (n= 13 912 women aged 50-69 years of which 9 903 underwent SFM and 4 009 underwent FFDM) show a significantly higher recall rate with FFDM and a significantly higher detection rate in FFDM, but PPVs are comparable.

In the Vestfold County study[65] (n= 18 239 women, aged 50-69years), as part of the Norwegian national screening program, no difference in recall rates were found. The authors state that recalls due to technically inadequate imaging was significantly lower with FFDM. The cancer detection rate was higher (but not statistically significant) with FFDM. Dependent on the type of tumour, the cancer detection rate varied: for invasive tumours no significant difference could be found between SFM and FFDM, but for ductal carcinoma in situ (DCIS) the detection rate was significantly higher with FFDM. In this study, Vigeland et al, 2007[65], found also a difference in PPV (18.5% in FFDM versus 15.1% in SFM, p=0.015). The authors conclude that the performance of FFDM is equal or even better than SFM, based on the higher cancer detection rate and the difference in PPV).

3.3.2.3. The DMIST study

The Digital Mammographic Imaging Screening Trial (DMIST) study [66] (n= 42 760 women underwent both SFM and FFDM) investigated the difference in diagnostic accuracy between SFM and FFDM. They did also a subgroup analysis and, adjusting for multiple comparisons using the Bonferoni method, set the significance level at p<0.003 for differences in area under the curve.. In the overall group of participants (without classification in age or risk groups) the diagnostic accuracy was similar between SFM and FFDM: no statistically significant difference in the area under the curve (AUC) (p=0.18). But in women under age of 50 years (diff in AUC 0.15; 95%CI 0.05-0.25;p=0.002), women with heterogeneously dense or extremely dense breasts (diff in AUC 0.11; 95%CI 0.04-0.18; p=0.003), pre- or perimenopausal women (diff in AUC 0.15; 95%CI 0.05-0.24; p=0.002), the diagnostic accuracy is significantly higher with FFDM. These results indicate FFDM may be of value for screening in specific target groups.

3.3.2.4. Del Turco et al, 2007, screening program in Florence, Italy

Analysis of the Italian screening program (n= 36 262 women of which 14 706 underwent FFDM and 21 556 women underwent SFM) by Del Turco et al, 2007[67], show a statistically significant higher recall rate with FFDM (4.56% versus 3.96%, p=0.01). This difference in recall rate was also found in a subgroup analysis of the age group of 50-59 years and in all breast density categories (only significant for the very dense breast (>75%) p=0.03). The recall rate due to poor technical quality was lower with FFDM.

Differences in cancer detection rate were found in subgroup analysis on type of abnormality, age group, breast density category and screening round: a higher detection rate with FFDM in women aged 50-59 years, significantly more cancer cases as well as micro-calcifications found in FFDM (p=0.007) and a higher detection rate with FFDM at incidence screening. But the overall analysis show no significant differences between

SFM and FFDM in cancer detection rate. The performance of SFM and FFDM is similar, but the additional cancers detected with FFDM compensated its higher recall rate, suggesting a higher sensitivity of FFDM, especially in specific groups, such as younger women and women with denser breasts.

In the overall conclusion of the review of AETSA the heterogeneity (and lack of significance in results) in performance for screening population is mentioned, but the higher accuracy of FFDM in specific groups (such as women with dense breasts) is underlined.

3.3.3. *Primary Studies: randomized controlled trials, prospective cohort studies and cross-sectional studies published 2008-2011*

The evidence found in the review of AETSA et al[42] is updated with a search for primary studies, published between 2008 and 2011. Fifteen primary studies were included in this report. The characteristics and main results are summarized in appendix.

Main finding in the primary studies is the heterogeneity in performance ranging from no difference between SFM and FFDM to a significant higher performance of FFDM. No studies were found in which SFM had a better performance than FFDM.

3.3.3.1. *Recall rate*

Similar to the above-mentioned studies of the review, a higher recall rate in FFDM is found in the majority of the primary studies[68-73]. Bluekens et al, 2010[68] (n= 312,414 screening mammograms of which 43,913 FFDM and 268, 501 SFM, women aged 50-75years), found after a peak of the referral rate and the false-positive results (due to pseudo-lesions and increased detection of benign microcalcifications) a decrease over time of the referral rate, but this rate was still higher in FFDM compared to SFM. This decrease over time could be explained by a learning curve. The authors recommend a training in digital screening for the image readers to obtain a stabilization of the increase in recall rate.

In contradiction to the above-mentioned studies, Sala et al, 2011[74] (n= 242 838 mammograms of which 171,191 SFM and 71, 647 FFDM, 103, 613 women aged 45-69years), found a higher recall rate (8.1% SFM vs 6.2% FFDM, p<.001) and false-positive rate (7.6% SFM vs 5.7% FFDM, p<.001)in SFM. The cancer detection rate did not differ between SFM and FFDM. Also Heddson et al, 2007[75] (n= 24,875 women) found a higher recall rate for SFM.

Vinnicombe, 2009[76] (n= 39,651 women, aged 50-70years) and Juel, 2010[77] (n= 14, 374 women, aged 49-70 years) found no increase in recall rate with FFDM.

Table 12 Overview of recall rates in primary studies

Author, year	Recall rate in SFM	Recall rate in FFDM
Bluekens, 2010	3.4%	4.3% (p<.001)
Sala, 2011	12.1%	11.7% (p=0.91)
Heddson, 2007	1.4%	1.0% (p<.001)
Vinnicombe, 2009	3.4%	3.2% (p=.44)
Juel, 2010	2.3%	2.4% (p>.05)

3.3.3.2. *Cancer detection rate*

Several authors [68, 69] mention the compensation of a higher recall rate by an increased cancer detection rate but the significance of the difference in cancer detection rate between SFM and FFDM is often lacking.

Domingo et al, 2011[78] (Spanish Screening Program, n= 242,838 mammograms of which 171,191 SFM and 71,647 FFDM from 103,613 women aged 45-69years) confirms the results of the review: a comparable performance between SFM and FFDM, without significant difference in tumor characteristics and cancer detection rate. Only the PPV for masses was significant higher in FFDM.

The conclusions of the study of Feeley et al, 2010[69] (n=107,818 women aged 50-64years) contradict the above-mentioned compensation of the recall rate by the increased cancer detection rate. The authors explain the increase in cancer detection rate, found in their study, by the improved detection of microcalcifications. The detection of such 'minimal sign lesions' may lead to an overtreatment.

In the study of Karssemeijer et al, 2009[70] (n= 367,600 mammograms of which 56,518 FFDM and 311,082 SFM, aged 50-75years) FFDM is combined with CAD (computer-aided detection), resulting in a similar detection performance as SFM alone. The detection of DCIS and microcalcifications was improved in a statistically significant way, but also the recall rate was increased.

The increased cancer detection rate with FFDM is also found in the study of Lipasti, 2010[79], Perry, 2010[80], Vernacchia, 2009[72], Hambly, 2009[73], Heddson, 2007[75].

Pisano et al, 2008[81] (n= 49,528 women) analyses more profound the impact of breast density, age, menopausal status on the comparison between SFM and FFDM, concluding similar findings as in the study of Pisano et al, 2005[66]: a better performance (area under the curve) of FFDM in pre- and perimenopausal women younger than 50 years with dense breasts.

In the study of Van Ongeval et al, 2010[71] and Juel, 2010[77], no difference in cancer detection rate could be found. The meta-analysis in the study of Vinnicombe et al, 2009[76] shows a slightly higher detection rate for FFDM, but no differences in recall rates or PPVs.

Table 13 Overview of cancer detection rate in primary studies

Author, year	cancer detection rate in SFM	Cancer detection rate in FFDM
Domingo, 2011	0.45%	0.43% (p=0.592)
Feeley, 2011	6.2 per 1000 women	7.2 per 1000 women (p=0.04)
Karssemeijer, 2009	0.62%	0.77% (p=.11)
Lipasti, 2010	0.406%	0.623%
Perry, 2010	2.8 per 1000 women	6.4 per 1000 women (p<.001)
Vernacchia, 2009	4.1 per 1000 women	7.9 per 1000 women (p=0.01)
Hambly, 2009	5.2 per 1000 women	6.3 per 1000 women (p=0.01)
Heddson, 2007	0.31%	0.49%, 0.38% (p=0.04)
Van Ongeval, 2010	0.64%	0.59% (p=0.56)
Juel, 2010	0.39%	0.48% (p>0.05)
Vinnicombe, 2009	0.72 per 100 women	0.68 per 100 women (p=.74)

3.3.3.3. Different systems of digital screening

The fourth edition of the "European protocol for the quality control of the physical and technical aspects of mammography screening" sets up minimum standards for quality control of mammography screening[82]. The quality control on the performance standards is built on three cornerstones of screening: the image quality, the minimum level of diagnostic information and the breast dose As Low As Reasonably Achievable (ALARA). The European commission developed a protocol for quality control and a protocol on dosimetry in Mammography (EUR16263).

In digital screening two systems can be distinguished[83]:

- the direct detection or digital radiography (DR)

The detector is integrated in the digital mammography unit and the images are directly shown on the screen. The DR systems incorporate also the photon-counting systems.

- the indirect detection of computed radiography (CR)

The imaging detector incorporates a phosphor to produce visible photons and a removable digital reader system is used, facilitating the implementation in SFM units.

Most above mentioned studies use a mix of both systems. Comparing CR systems and DR systems and the further investigation into new developing system (such as needle plates) falls out of the scope of this report..

3.3.4. Discussion

The heterogeneity in results hampers to draw one consistent conclusion about the clinical performance of FFDM. The difference in cancer detection rate ranges from no difference to a significant higher cancer detection rate with FFDM. Similar range in results can be seen in the recall rate. A majority of the studies found an increase in recall rate. Different explanations for this increase in recall rate are mentioned, such as the variability of the readers, the learning curve of the readers, the more precise detection of microcalcifications etc. Some authors suggest the hypothesis of compensation of the increased recall rate by the increased detection of cancers; others contradict this positive look on the increased variables and warn for the risk of overtreatment.

The quality of digital screening can be guaranteed by a specific training in reading of the digital mammographies. The importance of the provision of a specific training in reading of digital mammographies is emphasized by the Belgian experts.

Next to the clinical performance of a technical modality, other factors may influence the implementation in clinical practice. These factors, such as user friendliness, data storage, data exchange, etc are in advantage of FFDM. Nowadays the evolution towards the electronical medical file and the information exchange between health providers via internet, supports the integration of FFDM in a screening program. There are disadvantages however, like the high cost and difficulty of sharing of the digital images derived from another technology (Van Ongeval et al, 2007)[83].

As regards to the content of this report, we decided to restrict this study to the clinical performance of FFDM, in particular cancer detection rate and recall rate. Other performance indicators, such as the technical characteristics are not described in this report. In case of a overall view on the performance of FFDM, other aspects, such as technical characteristics, cost-effectiveness etc, should also be considered.

As regards to the methodological aspects of the above-mentioned studies, some aspects hamper the interpretation of the results. For example the difference in recall rate between the countries. Only the tendency towards an increase or decrease could be mentioned, the absolute numbers were to specific for each country.

In conclusion could be stated that all the authors agree on the better or at least similar performance of FFDM and support the integration of FFDM in the population-based screening programs as an equivalent to SFM.

Key points

- **Studies on the value of FFDM are conflicting and there is no convincing proof that it benefits the patient in a population based screening program.**
- **Some subgroup analyses indicate that performance of FFDM is better in premenopausal women and women with dens breasts**
- **FFDM and SFM can be seen as equivalent screening modalities**
- **Organisational aspects, such as data storage and image exchange, facilitate the implementation of FFDM in clinical practice**

3.4. Breast Ultrasound as a screening tool

3.4.1. Systematic reviews, meta-analyses, health technology assessments and evidence based guidelines

After selection and critical appraisal, four systematic reviews on the use of breast ultrasound in breast cancer screening, were selected. No good quality meta-analysis could be identified in the literature.

A summary of characteristics and results of the four reviews is presented in Table 35.

Although the 4 systematic reviews[43-46] applied critical appraisal, studies with important methodological flaws remained included. Reported sensitivity for ultrasound screening varies between 20% and 90.4% and reported range for specificity varies between 50% and 99.4%.

As it was not possible to conclude from the systematic reviews what is the most exact estimate of the accuracy of ultrasound, the original publications of the primary studies included were reviewed. Furthermore, available evidence was updated with primary studies published after the search date of the most recent systematic review. Findings of primary studies are discussed in the following paragraphs (3.4.2 and 3.4.3).

3.4.2. Primary Studies: randomized controlled trials, prospective cohort studies and cross-sectional studies included in the 4 selected systematic reviews

As the reported results in the selected reviews differ substantially, the 15 primary studies selected in at least one systematic review were separately reviewed and assessed using the QUADAS checklist for diagnostic accuracy studies.

A summary of the study characteristics of these reviews is presented in Table 36.

None of the systematic reviews could identify a RCT investigating the role of ultrasound in breast cancer screening. Included studies had a cross-sectional or cohort design.

After critical appraisal, we excluded two studies[84, 85]. The study by Trecate et al. [84] appeared a narrative of four case reports with hardly any information on ultrasound results and the study by Sim et al.[85] is a retrospective study without consecutive inclusion of patients.

3.4.3. Primary Studies: randomized controlled trials, prospective cohort studies and cross-sectional studies published 2007-2011

Literature search for primary studies published since 2007 revealed 12 studies reporting on ultrasound as a screening tool for breast cancer. Two more studies were identified through screening of the references of selected papers. Details of the 14 articles are summarized in Table 37.

After critical appraisal, two studies were excluded. The study by Youk et al[86] was excluded due to a dropout rate of more than 60%. The publication of Lenz et al.[87] was a retrospective study without clear consecutive inclusion of patients and without follow-up of patients with a negative test result.

3.4.4. Discussion

There are no randomized controlled trials or meta-analyses investigating the effect of ultrasound screening on accuracy of a screening program or on breast cancer related mortality.

In total, 25 primary cross-sectional and cohort studies were selected, with important differences regarding set-up, population included, reference standards and diagnostic threshold.

In six [88-93] of the 25 studies ultrasound was used incremental to mammography in patients with dense breast tissue and normal mammographic findings. In the other studies, ultrasound was used simultaneously, irrespective of the results of clinical breast examination and other imaging.

Importantly, in the majority of studies, single reading mammography was used as comparison. Double reading of mammography was used in seven studies[89, 90, 94-98].

None of the studies included women with average breast cancer risk and entirely fatty breasts. None of the selected studies investigated the use of ultrasound in an organized screening program with Western, unselected patients. Ten studies[88-93, 99-102] used dense breasts on mammography, defined is BIRADS-M ≥D2 or ≥ D3 as one of the inclusion criteria. Other studies included women with a moderate or high risk for breast cancer, based on family history and/ or genetic testing[94, 96, 98, 101, 103-111]. The study by Leconte et al. included women independent of their breast cancer risk, but excluded women with entirely fatty breasts and normal findings on mammography. The majority of studies also included symptomatic women[89-91, 99, 102, 112] and/or women with a personal family history[88-90, 92, 94, 96, 98, 100, 101, 105-112]. The study population in the studies by Hou et al. [104], Tohno et al.[97] and Honjo et al.[95] consisted of Asian women with a high proportion of dense breasts and may not be representative for the Western population.

All selected studies used fine needle aspiration (FNA) and/or biopsy to confirm true positives. The diagnostic threshold to consider an ultrasound result positive and to perform FNA or biopsy however differs between studies, mainly regarding lesions classified as BIRADS-US 3 (probably benign finding). Reference standard for true and false negatives was often not clearly defined but mainly consisted of comparison with other screening techniques (mammography and/or CBE and/or MRI) with or without follow-up for interval cancers. 15 of the 25 studies had a follow-up for interval cancers of varying duration and completeness. Calculations of sensitivity,

specificity and NPV without inclusion of interval cancers are not stated in the evidence tables.

Proportion of prevalent and incident rounds varies between studies, probably also affecting the heterogeneity between results. As prevalent rounds dominate study results, cancer detection rate and sensitivity will decrease when implemented in a screening program where the majority of examinations are incident screens.

Furthermore, the known operator dependence for ultrasound can explain partially heterogeneity of results and must be taken into account when implementation of ultrasound in a screening program is considered.

All these factors can explain why results from trials differ substantially. In the following tables, results are summarized to give a general overview. For details of the specific trials, we refer to the detailed evidence tables in appendix.

Incremental cancer detection rate

Although (incremental) detection rate can be an early indicator of the effectiveness of a screening program, it is subject to overdiagnosis bias[6].

Table 14 Reported incremental cancer detection rate of ultrasound screening

Author, year	Cancer detection rate of ultrasound screening used incremental to normal mammography
Kolb, 1998	0.3% for ultrasound incremental to single reading mammography
Buchberger, 2000	0.46% for ultrasound incremental to double reading mammography, 0.26% for patients without personal cancer history
Kaplan, 2001	0.3% for ultrasound incremental to single reading mammography
Crystal, 2003	0.42% (0.25% for average risk women) for ultrasound incremental to single reading mammography
Brancato, 2007	0.38 per 1000 women, 6.5% increase compared to single reading (?) mammography alone

Sensitivity

Twelve studies included interval cancers in the calculation of ultrasound sensitivity. Crystal et al. [92] achieved a sensitivity of 100%. However ultrasound was used incremental to negative mammography, no MRI was performed and follow-up was incomplete, as mentioned by the authors themselves. The sensitivity is thus probably highly overestimated. Reported sensitivity for the other studies which included follow-up for interval cancers is summarized below. Sensitivity for ultrasound varies between 17% and 67%, for mammography between 12.5% and 61.5% and between 48.1 and 86.7% for combination screening with ultrasound and mammography.

Table 15 Reported sensitivity for ultrasound, mammography and the combination of ultrasound + mammography per study. 95% confidence interval between []

Author, year	Sensitivity US	Sensitivity Mx	Sensitivity US + Mx
Warner, 2004	25% 1st round 57% incident round		
Kuhl, 2005	38.7%	32.3%	51.6%
Honjo, 2007	53.8%	61.5%	
Riedl, 2007	42%	50%	
Berg, 2008	50% [33.8-66.2]	50% [33.8-66.2]	77.5% [61.6-89.2]
Daguet, 2008	50% [15.7-84.3]	12.5% [0.3-52.7]	
Weinstein, 2009	17%	39%	
Kuhl, 2010	37% [20-57.5]	33.3% [17.2-53.9]	48.1%[29.1-67.6]
Kelly, 2010	67% [53-79]	40% [27.5-54]	81% [68-90]
Sardanelli, 2011	52% [37.4-66.3]	50% [35.5-64.5]	62.5% [47.4-76.0]
Corsetti, 2011			86.7%

Specificity

Crystal et al. [92] report a specificity of 94%, however the same comments as for sensitivity apply.

Table 16 Reported specificity for ultrasound, mammography and the combination of ultrasound + mammography per study. 95% confidence interval between []

Author, year	Specificity US	Specificity Mx	Specificity US + Mx
Warner, 2004	95% 1st round 96% incident rounds		
Kuhl, 2005	91%	97.1%	89.4%
Honjo, 2007	95.4%	92.1%	88.4%
Riedl, 2007	97%	97%	
Berg, 2008	91.8% [90.7-92.8]	95.5% [94.7-96.3]	89.4% [88.2-90.6]
Daguet, 2008	97.3% [94.1-98.9]	98.7% [?]	
Weinstein, 2009	88%	91%	
Kuhl, 2010	98% [98.2-99.3]	99.1% [98.5-99.5]	98.3% [97.5-98.8]
Kelly, 2010	89.9% [89.1-90.6]	95.2% [94.6-95.7]	98.7% [98.4-98.9]
Sardanelli, 2011	98.4% [97.5-99.1]	99.0%[98.2-99.5]	97.6% [96.4-98.5]

Positive predictive value, recall rate and biopsy rate

Although the European guidelines advise strongly against short term follow-up with repeat imaging after e.g. 6 months (desirable standard 0%, minimal standard < 1%)[113], many studies reported a significant number of such early recalls. These early recalls are included in the reported total recall rates below as it reflects the total morbidity generated by the screening investigations. When ultrasound was used incremental to a negative mammography or only the additional recalls or biopsies generated by ultrasound are reported, (I) is added behind the result.

Table. 17 Overview of reported recall rate for ultrasound, mammography and combined screening with ultrasound and mammography (Ms = months)

Author, year	recall US	recall Mx	recall US +Mx
Hou, 2002	12.9%		
Crystal, 2003	6.6%(I)		
Warner, 2004	5.1% US after 6ms		
Kuhl 2005	16.7% US after 6ms		
Brancato, 2007	2.1%(I)		
Honjo, 2007	4.8%		15.3%
Lehman, 2007	9%		
Berg, 2008	21.4%	12.7%	27.4%
Weinstein, 2009	13.9%		
Tohno, 2009	4%	4.3%	
Kuhl, 2010	19.8% US after 6ms		
Kelly, 2010	7.2%	4.8%	9.6%

In their study published in 1998, Kolb et al.[88] needed to perform 131 fine needle aspirations, 45 biopsies and 188 repeat ultrasounds to diagnose 11 cancers. Buchberger et al. [90] performed 242.4 ultrasounds, 3.4 fine needle aspirations, 6.4 core biopsies and 0.6 surgical biopsies for each detected cancer. European guidelines promote a recall rate of < 5 (acceptable < 7%) for the initial screening round and < 3% (acceptable < 5%) for the subsequent screening rounds[113].

Table 18 Overview of reported PPV for ultrasound, mammography and combined screening with ultrasound and mammography. 95% confidence interval between []

Author, year	PPV US	PPV Mx	PPV US +Mx
Buchberger, 1999	7.9%(I)		
Buchberger, 2000	13.7%(I)		
Kolb, 2002	20.5%(I)		
Warner, 2004	23% 1st round 44% 2nd round		
Kuhl, 2005	10.4%	23.3%	11.7%
Riedl, 2007	42.1%	61.5%	
Berg, 2008	6.5% [4.1-9.7]	7.6% [4.8-11.4]	7.3% [5.1-10.2]
Daguet, 2008	40% [12.2-73.8]	25% [0.6-80.6]	
Kuhl, 2010	35.7% [19.3-55.8]	39.1% [20.4-61.2]	32.5% [19.1-49.2]
Sardanelli, 2011	61.9% [45.6-76.4]	71.4% [53.7-85.4]	55.6% [41.4-69.1]

Table 19 Overview of reported biopsy rate for ultrasound and combined screening with ultrasound and mammography

Author, year	Biopsy rate US	Biopsy rate US +Mx
Kolb, 1998	1.9% (I)	
Kaplan, 2001	5.2% (I)	
O'Driscoll, 2001	6.1% (I)	6.7%
Hou, 2002	2.5%	
Crystal, 2003	2.5% (I)	
Corsetti, 2008	4.9%(I)	
Lehman, 2007	2.3%	
Weinstein, 2009	3.5%	
Corsetti, 2011	5.5% (I)	

Table 20 PPV of biopsies. 95% confidence interval between []

Author, year	PPV biopsies/FNA
Kolb, 1998	6.25%
Buchberger, 2000	9.9%
Kaplan, 2001	11.8%
O'Driscoll, 2001	10%
Kolb, 2002	10.3%
Hou, 2002	79.2%
Crystal, 2003	21.2%
Corsetti, 2008	11.1%
Lehman, 2007	25%
Berg, 2008	11.2% [7.8-15.6]
Kelly, 2010	38.4%

Guidelines advise a benign:malignant biopsy ratio of ≤ 1:1 (PPV biopsies ≥ 50%) for the initial screening round and even lower for the subsequent screening rounds (desirable ratio ≤ 0.2:1, acceptable ratio ≤ 0.5:1)[6]..

Negative predictive value (NPV)

Table 21 Reported NPV for US, mammography and combined screening by ultrasound and mammography. 95% confidence interval between []

Author, year	NPV US		NPV Mx		NPV US +Mx	
Warner, 2004	96% 1st round 98% incident rounds					
Riedl, 2007	96%		96.6%			
Daguet, 2008	98.2%	[95.3-99.5]	96.9%	[93.8-98.8]		
Kuhl, 2010	98.9%	[98.3-99.4]	98.9%	[98.2-99.2]	99.1%	[98.5-99.5]
Sardanelli, 2011	97.7%	[96.5-98.5]	97.6%	[96.5-98.5]	98.2%	[97.1-98.9]

Advanced cancer rate and mortality

None of the included studies reported on advanced cancer rate or mortality.

There are no data assessing the impact of implementing ultrasound in a screening program on breast cancer related mortality. As there are also no data on the ability of ultrasound to reduce the incidence of advanced breast cancer, it is not possible to estimate the proportion of overdiagnosis induced by ultrasound screening.

None of the identified studies investigated the role of ultrasound screening in unselected women aged 50 to 69 years, the population eligible for the general screening program in Belgium. Conclusions on the usefulness of ultrasound in this population can only be deduced from the findings in selected populations. As the prevalence of breast cancer in the general population is lower than in high risk populations, we can expect that the incremental cancer detection rate and the PPV of ultrasound will be lower. The higher sensitivity of mammography in this population and the use of double reading mammography in the screening program reinforce this effect.

In a high risk population and in women with dense breasts, it appears that ultrasound can detect additional, small cancers missed on mammography. In women with an elevated breast cancer risk, the sensitivity of screening rose from 32-50% to 52-81% when adding ultrasound to mammography. When used incremental to negative mammography, the reported cancer detection rate of ultrasound in women without personal cancer history is maximum 0.3%. The study by Brancato et al.[93] _ENREF_58 selected women by breast density (BIRADS D3-D4), without additional risk factors, and reported an additional detection rate of only 0.038%. The overall cancer detection rate by mammography in this study population was indeed comparable with the detection rate observed in population-based screening.

The improved detection rates come at the cost of a considerable number of false positive results and a high number of recalls and biopsies. In none of the studies performed in a western population, the total recall rate was lower than 7%, the rate considered acceptable by the European guidelines for breast cancer screening[6]. The reported percentages by Crystal et al.[92] and Brancato et al.[93] refer to an additional recall rate generated by ultrasound, to be added to the women recalled for further investigation after mammography. The recall rate reported by Warner et al.[106] and Kuhl et al.[94, 98] include only people recalled for repeat imaging after 6 months. Also the positive-negative biopsy ratio was far under the 1:1, the advisable ratio for first screening rounds[6], in all Western studies.

The final decision to use ultrasound as a screening tool for breast cancer will be a trade of between possible benefits and harms and depends on the risk to be diagnosed with breast cancer.

For women older than 50 years with an average breast cancer risk eligible for the population-based screening in Belgium, the use of ultrasound screening for breast cancer is not recommended. The possible detection of additional breast cancers by ultrasound does not weigh up to following the European guidelines for breast cancer screening[113]. The possible additional detection of breast cancer by ultrasound does not justify the significant risk for a false positive screening exam with additional investigations, anxiety and costs. This especially as the number of

additionally detected cancers and positive predictive value will be even lower than reported for selected populations. Furthermore, it is not clear if an increased cancer detection rate would result in a reduction of breast cancer related mortality as there are no data to estimate the contribution of overdiagnosis.

For women with a high breast cancer risk, possible gains and harms balance differently for several reasons. First, as the life-time risk to develop breast cancer is high, the number of women that potentially benefit from early detection and thus less invasive treatment and better prognosis, is higher. Second, the sensitivity of mammography screening appears to be lower in this patient group as the appearance of breast cancer on imaging maybe different and screening starts at an earlier age when breast tissue is generally more dense. Furthermore, the knowledge of being at high risk may lead to an increased acceptability of false positive results. However, also in a high risk population, ultrasound screening is hampered by inter-observer variability and the lack of data on the reduction of advanced stage cancers and mortality. Moreover, the availability of breast MRI surpasses the use of ultrasound, as will be discussed in the next section.

Special concern is raised on the use of ultrasound in women with dens breast tissue participating in the organized screening program aged 50-69 years. Dense breast tissue on mammography is a risk factor for breast cancer, as discussed in the previous chapter. As dense breast tissue can obscure the visibility of a cancerous lesion on mammography, the additional imaging by ultrasound is suggested. However, the available evidence suggests only limited benefit of ultrasound for women with dense breast tissue as the only identified risk factor. The studies that included specifically women with dense breast tissue all included women with personal or family history for breast cancer or symptomatic women, except the study by Brancato et al.[93] They detected only 0.38 cancers per thousand women by ultrasound in women with normal findings and dense breasts (BIRADS D3-D4) on mammography. It is not clear from their report if digital mammography and double reading were used; the detection rate of ultrasound may be even lower when applied in an organized screening program with high level quality assurance. This limited benefit has to be weighed against an additional recall rate of 2.1% in women with an only modestly increased breast cancer risk. Applying the Tice model[22] (see chapter on risk estimation) on white women, we calculated a risk of 9.7% to develop breast cancer between the age of 40y and 86y for women with breast density BIRADS D2 on mammography, compared with a risk of 14.2% and 16.6% for women with BIRADS D3 or D4 density respectively when no other risk factors are apparent. Hence, having dense breasts on mammography without other apparent risk factors increases indeed the risk for breast cancer but not to a level of high risk as defined by NICE[9].

Other factors complicate the implementation of ultrasound for breast cancer screening in women with dense breast tissue. The diagnosis of 'dense breast tissue' and assignment to the four BIRADS categories for breast density know a significant variability between different readers. In the screening program of the Belgian Communauté française et germanophone, only 53% of mammographies classified as BIRADS-M D3 by the first reader, were classified as D3 or D4 by the second reader too (Pr. Anne Vandenbroucke, centre communautaire de référence pour le dépistage des cancers, personal communication). The inter-observer variability was better in the Flemish program but still a considerable degree of disagreement exists. If two groups are considered, BIRADS 0,I or II versus BIRADS III or IV, first and second reader achieve an agreement of 81% (G. Vande Putte, personal communication). Furthermore, problems with inter-observer variability arise when implementing ultrasound screening in a decentralized, multicentre setting.[114-116] As the execution and the resulting stored images are operator dependent, the problem of inter-observer variability cannot easily be diminished by double reading procedures, as you would need to redo the entire exam. In conclusion, the use of ultrasound for women with dense breasts without other risk factors cannot be supported with currently available evidence.

Conclusion

Following the European guidelines for breast cancer screening, the balance of benefits and harms is insufficient to support the implementation of ultrasound in a screening population of average risk women, with or without dense breasts, as the expected number of additional cancers detected does not justify the additional harm generated by the high number of false positives and additional recall rate and the risk of overdiagnosis. Furthermore, there is no proof of a beneficial effect on breast cancer related mortality.

The use of ultrasound screening in women with high breast cancer risk can be considered as the prevalence of breast cancer is higher, the detection rate by mammography alone is lower and an increase of recall rate and false positives may be acceptable in this group of women. However, the value of ultrasound must be weighed against the use of MRI.

Key points

- **The use of ultrasound in breast cancer screening has been investigated in several cross-sectional and cohort studies; no randomized controlled trials or meta-analyses are available.**
- **There are no data to evaluate overdiagnosis, rate of advanced-stage breast cancers and breast cancer related mortality in population-based screening programs using breast ultrasound.**
- **There are no studies investigating the accuracy of ultrasound screening in average risk women aged 50-69 years participating in the population based screening program in Belgium.**
- **Extrapolated from data for women at high risk, ultrasound screening is not recommended in a population-based screening program as the recall rate and number of false positives is too high and the additional cancer detection rate is minimal.**
- **For women with dense breast tissue on mammography, the benefit-risk ratio does not support ultrasound screening if no other risk factors for breast cancer are identified in spite of the modestly increased risk compared to women with non-dense breast tissue.**
- **For women at high risk for breast cancer, the use of ultrasound screening can nevertheless considered as the prevalence of breast cancer is significantly higher, sensitivity of mammography is reduced and a low specificity may be accepted. However, the emergence of MRI may surpass the use of ultrasound.**
- **Problems with inter-observer agreement for the assessment of breast density on mammography and for the interpretation of ultrasound further hinder the implementation of breast ultrasound in a screening program.**

3.5. Breast MRI as screening tool

3.5.1. Systematic reviews, meta-analyses, health technology assessments and evidence based guidelines

After selection and critical appraisal, 4 systematic reviews on the use of breast MRI in breast cancer screening were selected.

A summary of the study characteristics of these reviews is presented in Table 38 in appendix.

The four selected reviews[43-45, 47] included studies with women with a high breast cancer risk based on mutation analysis, with or without personal cancer history. Included studies used different criteria to define high risk. Also diagnostic threshold, reference standard, follow-up and calculation methods differ significantly between studies.

In spite of the heterogeneity, the four systematic reviews all report a higher sensitivity for MRI (between 71.7-100%) versus mammography (0-59%) at the cost of a lower specificity (81-97.5% for MRI versus 93-99.8% for mammography). The same conclusions can be drawn for MRI versus ultrasound or versus the combination of ultrasound and mammography.

As reported by Irwig et al.[45] false positive rate, defined as % of patients requiring biopsy varied between 5 and 9% for MRI and between 1 and 7% for mammography.

The studies included in the systematic review by Davidson et al. reported a PPV for MRI between 32.3 and 50%, a NPV of 99-99.7% and an AUC between 0.83 and 0.89.

3.5.2. Primary Studies: randomized controlled trials, prospective cohort studies and cross-sectional studies

Literature search revealed 12 studies reporting on breast cancer screening using MRI. Two more studies were identified through screening of the references of the selected papers. Details of the 14 articles are summarized in Table 39.

After critical appraisal, four studies were no longer withheld. Yu et al.[117] performed a retrospective review of patients with high breast cancer risk screened with MRI. There was no consecutive inclusion of patients with the final decision to perform a MRI left to patient's and physician's discretion. Only 37% of eligible patients were included. Also in the studies by Lapierre-Combes et al.[118], Elmore et al.[119] and Shah et al.[120] significant flaws in patient inclusion are noted as only a non-specified selection of patients underwent screening tests.

3.5.3. Discussion

There are no randomized controlled trials or meta-analyses reporting on the influence of MRI screening for breast cancer on breast cancer related mortality published since 2007.

Eight out of ten selected articles performed MRI simultaneously with mammography and/or ultrasound and/or clinical breast examination. Price et al.[121] and Abramovici et al.[122] did not report on other screening techniques.

Eligibility criteria included high risk women in all studies, mainly based on genetic analysis for BRCA1, BRCA2 or p53 mutations and family tree analysis, using different models. Respectively four[110, 121-123] and three[110, 121, 122] studies also included patients with high risk lesions (e.g. LCIS, atypical hyperplasia) on previous biopsy or history of mantle field radiotherapy. A minority of patients in the study of Price et al.[121] were included with dense breasts, breast implants or patient preference as the only indication.

As for ultrasound, the diagnostic threshold for recall and biopsies, handling of intermediate BIRADS 3 results, proportion of prevalent and incident rounds and screening interval varies between studies.

Sensitivity

Table 22 Reported sensitivity for MRI, mammography and combined screening with MRI + mammography. 95% confidence interval between []

Author, year	Sensitivity MRI	Sensitivity Mx	Sensitivity MRI + Mx
Riedl, 2007	85%	50%	
Daguet, 2008	87.5% [47.4-99.7]	12.5% [0.3-52.7]	
Weinstein, 2009	71%	39%	
Kuhl, 2010	92.6% [84.2-98.7]	33.3% [17.2-53.9]	100% [85.8-100]
Sardanelli, 2011	91.3% [79.2-97.6]	50% [35.5-64.5]	93.2% [81.3-98.6]

Specificity

Table 23 Reported specificity for MRI, mammography and combined screening with MRI + mammography. 95% confidence interval between []

Author, year	Specificity MRI	Specificity Mx	Specificity MRI + Mx
Riedl, 2007	88%	97%	
Daguet, 2008	94.8% [91.4-97.5]	98.7% [?]	
Weinstein, 2009	79%	91%	
Kuhl, 2010	98.4% [95.9-98.9]	99.1% [98.5-99.5]	97.6% [96.7-98.2]
Sardanelli, 2011	96.7% [95.4-97.7]	99% [98.2-99.5]	96.3% [95-97.4]

Positive predictive value, recall rate and biopsy rate

Although the European guidelines advise strongly against short term follow-up with repeat imaging after e.g. 6 months (desirable standard 0%, minimal standard < 1%)[113, 113], many studies report a significant number of such a early recalls. These early recalls are included in the reported total recall rates below as it reflects the total morbidity generated by the screening investigations.

Table 24 Overview of reported PPV for MRI, mammography and combined screening with MRI and mammography. . 95% confidence interval between []

Author, year	PPV MRI	PPV Mx	PPV MRI + Mx
Riedl, 2007	48%	61.5%	
Daguet,2008	38.9%	25%	
Kuhl, 2010	48% [34.2-62.2]	39.1% [20.4-61.2]	40.2% [28.7-53.0]
Sardanelli, 2011	56% [44.1-67.5]	71.4 [53.7-85.4]	53.2% [41.5-64.7]

Table 25 Overview of reported recall rate for MRI

Author, year	Recall rate MRI
Lehman 2007	24%
Peters 2008	12.5% 1st round 7.5% 2nd round
Price, 2009	15% [10-20]
Weinstein 2009	22.6%
Kuhl, 2010	17% early recall
Abramovici 2011	11.4% 16% 1st round 7.3% incident rounds

Table 26 Overview of reported biopsy rate for MRI

Author, year	Biopsy rate MRI
Lehman, 2007	8.2%
Peters, 2008	7.9%
Daguet, 2008	12% 1st round 6-12% incident rounds
Price, 2009	13%
Weinstein, 2009	8.4%

European guidelines promote a recall rate of < 5 (acceptable < 7%) for the initial screening round and < 3% (acceptable < 5%) for the subsequent screening rounds.[113]

Table 27 PPV of biopsies generated by MRI

Author, year	PPV biopsies/FNA
Peters, 2008	9%
Daguet, 2008	FNA: 30% Biopsies: 58%
Price, 2009	30.4%

Guidelines advise a benign: malignant biopsy ratio of ≤ 1:1 (PPV biopsies ≥ 50%) for the initial screening round and even lower for the subsequent screening rounds (desirable ratio ≤ 0.2:1, acceptable ratio ≤ 0.5:1)[6].

The accuracy of MRI as a breast cancer screening tool has been investigated in cross-sectional and cohort studies only. No data on the impact on mortality or reduction of advanced-stage cancer are available.

Breast cancer screening using MRI has been tested only in high risk populations because of limited availability and too high costs to justify implementation in the general population. As mentioned before, the expected gain of adding MRI to mammography in this population is important. Women at high risk for breast cancer start screening early in life, when the sensitivity is reduced because of breast density and different phenotype of BRCA-related cancers[124, 125]. Furthermore, MRI can reduce the risk for radiation –induced cancer by limiting the number of mammographies or views per mammography needed.[126]

Systematic reviews and more recent published studies consistently show an increased sensitivity for MRI screening compared to mammography. This increase is more important compared to ultrasound. Reported sensitivity in this patients group varies between 68% and 100% for MRI compared to 52-81% for mammography and ultrasound combined and a maximal sensitivity of 50% for mammography.

The detection of additional cancers by the use of MRI is accompanied by a lower specificity and PPV. Recall rate for MRI is substantial, with recall rates higher than 20% reported by Lehman et al. and Weinstein et al. Positive predictive values lie between 39 and 56%.

Given the high incidence of breast cancer, the benefit-risk ratio appears to support the use of MRI for breast cancer screening in a high risk population. This in spite of the high recall rate as still significant PPV is achieved. It must be kept in mind however that there are no data to proof that the higher sensitivity of MRI will lead to a better prognosis and reduced mortality in the high risk population. Patients should be informed about the remaining risk for false positives and negatives and the uncertainty of long term benefits.

For women with average or raised breast cancer risk, the use of breast MRI in screening has not been investigated. In the Netherlands, a trial will be performed in women with dense breast tissue on mammography (van Gils et al. http://clinicaltrials.gov/ct2/show/NCT01315015?term=MRI+and+screening+AND+breast&rank=1). Until more data are available, the high recall rate as seen in women at high risk does not support the use of MRI for other women.

As data suggest results depend on technical quality of the procedure, experience of the centers, learning curves and double reading procedures[9], it is recommended that MRI screening programs are subjected to strict quality assurance procedures as is the case for mammography. The ideal time interval between screening rounds, the use of (cheap) ultrasound to shorten the interval and the significance of yearly mammography in addition to MRI are still matter of debate.

Key points

- **The use of MRI in breast cancer screening has been investigated only in women with a high risk for breast cancer.**
- **No RCT or meta-analysis has investigated the role of MRI in breast cancer screening.**
- **The effect of MRI screening on treatment morbidity and mortality has not been proven.**
- **Available evidence shows a significantly increased sensitivity compared to mammography or mammography and ultrasound combined. Reported sensitivity for MRI varies between 68% and 100% in a high risk population.**
- **Implementing MRI in breast cancer screening results in a high recall rate up to 24%. Positive predictive value is still as high as 39-56%**
- **Given the increased sensitivity and significant PPV, evidence supports the use of MRI in a high risk population. Patients should be informed about the risk for false positive and false negative results and the absence of data on long term benefits.**
- **The ideal time interval between screening rounds, the use of ultrasound to shorten the interval and the significance of yearly mammography in addition to MRI are still matter of debate.**
- **Breast cancer screening using MRI should be subjected to audit and strict quality assurance.**

3.6. Screening in women with average, raised and high breast cancer risk: summary

3.6.1. Breast cancer screening in women with average risk (life-time risk < 17%)

European guidelines[6] advise screening with mammography between the age of 50 and 69, as is implemented in the Belgian screening program.

To start screening earlier is not recommended, as is discussed in an earlier report of the KCE[1]. Breast cancer screening in older women is currently under investigation and conclusions will be reported shortly.

Mammography screening should be performed within a quality assured program following the guidelines of the European Union (not the scope of this report). The use of double reading by two independent readers with a consensus or arbitration based recall procedure. The use of film-screen or full-field digital mammography can be considered of similar accuracy for the general population.

There are no data to support the use of ultrasound or MRI for screening purposes in women with an average breast cancer risk. The high recall rate and high proportion of false positive examinations seen in a high risk population lead to a disadvantageous benefit-risk ratio.

3.6.2. Breast cancer screening in women with raised risk (life-time risk 17-30%)

Women with identified risk factors for breast cancer resulting in a life-time breast cancer risk between 17 and 30% are of special concern. The increased incidence of and often more aggressive nature of breast cancer in this group raise anxiety and distress and justify a more extensive screening program. However, to date, there is no high level evidence to support the recommendation of any additional screening techniques or other modifications from the screening program in the general population.

Several possible measures can be considered:

- To start screening at a younger age
- To increase the frequency of screening rounds (shorter interval)
- Use of ultrasound or MRI

However, data on additional detection rate, accuracy, false positives and long-term benefits (advanced-breast cancer rate, mortality) are very sparse for this specific group of patients. This should be discussed with all patients when screening outside the general screening program is offered.

Both the Dutch[127] and British guidelines[9] advice to start annual mammography screening at the age of 40 years in this group. As discussed above, the main argument is the similar prevalence as in the general population aged between 50 and 70 years, so a similar benefit-risk ratio can be expected. It is paramount that the annual mammographies should be performed within a quality assured program following the European guidelines as is the case for population-based screening. The younger women and women with dense breast tissue can especially benefit from the use of double reading procedures and full-field digital mammography (see above). As discussed in the previous paragraphs, the use of ultrasound is not recommended outside of a clinical trial setting in this group of patients, including women with very dense breast tissue, because of the high recall rate and high number of false positive results. The use of MRI has not been investigated in women with a raised breast cancer risk and is currently considered not feasible in this patient group because of high costs and limited availability.

3.6.3. Breast cancer screening in women with high risk (life-time risk > 30%)

Breast cancers occurring in women with a strong family history or other high risk factors are characterized by negative prognostic factors, a short sojourn time and appearance on a younger age than in breast cancers appearing in the general population.

Hence it is recommended to start screening early in life, generally at the age of 30 years based on incidence data per age-group. Families where breast cancer is diagnosed before the age 35 are advised to start screening even earlier, namely five years before the age of the youngest family member diagnosed with breast cancer.

Available data support the use of yearly MRI with a clearly raised sensitivity compared to mammography. As discussed above, a reasonable positive predictive value is achieved in spite of the high recall rate. An additional advantage is that the radiation dose of mammography can be

reduced. The use of ultrasound appears superfluous but it can be considered in between MRI screening rounds to shorten the screening interval.

As for patients with raised risk, it must be discussed with every participating patient that there is a risk for false positive results and a remaining chance for interval cancer. Furthermore, there is no proof of a beneficial effect on mortality. The use of screening also needs to be put in the context of other preventative measures such as prophylactic surgery and prevention by hormone therapy.

RECOMMENDATIONS[a]

Who should be considered at risk?

- **A risk assessment should first distinguish persons who have a risk that equals that of the general population and people who have a raised risk. This is essentially done with simple questions about the family history.**
- **A more in depth assessment is needed to classify women who are above population risk in order to give individual advise on screening strategy, genetic tests and prophylactic measures. Such an individual risk assessment and subsequent screening and treatment decisions should always be discussed with the women at risk, taking into account all possible advantages and limitations, uncertainties and alternatives.**

[a] These recommendations are under the sole responsibility of the KCE

HOW TO DEFINE THE INDIVIDUAL RISK [b]

A Family history is the strongest risk factor

1. Women can be categorised in 3 risk categories based on family history *(strong recommendation, moderate level of evidence)***.**

Average risk

- **Maximum one first-degree or second-degree relative diagnosed with breast cancer at older age than 40 years.**

Raised risk (that is, a 10-year risk of 3–8% for women aged 40–49 or a lifetime risk of 17% or greater but less than 30%):

- **one first-degree relative diagnosed with breast cancer at younger than age 40 years**

or

- **two first-degree or second-degree relatives diagnosed with breast cancer at an average age of older than 50 years**

or

- **three first-degree or second-degree relatives diagnosed with breast cancer at an average age of older than 60 years**

High risk (that is, a 10-year risk at age 40–49 years of greater than 8% or a lifetime risk of 30% or greater):

- **two first-degree or second-degree relatives diagnosed with breast cancer at younger than an average age of 50 years (at least one must be a first-degree relative)**

or

- **three first-degree or second-degree relatives diagnosed with breast cancer at younger than an average age of 60 years (at least one must be a first-degree relative)**

or

- **four relatives diagnosed with breast cancer at any age (at least one must be a first-degree relative)**

or

- **Jewish ancestry**

of

[b] Breast cancer of the women herself as a risk factor falls under follow up after treatment and is not part of the current report.

- **one of the following is present in the family history**
  - **bilateral breast cancer**
  - **male breast cancer**
  - **ovarian cancer**
  - **sarcoma in a relative younger than 45 years of age**
  - **glioma or childhood adrenal cortical carcinomas**
  - **complicated patterns of multiple cancers at a young age**
  - **very strong paternal history (four relatives diagnosed at younger than 60 years of age on the father's side of the family)**

2. Women with a high breast cancer risk based on the above mentioned criteria should be offered individual risk assessment in order to give individual advise on screening strategy, genetic tests and prophylactic measures.

- **Individual risk assessment consists of an in depth family history and can make use of computerized risk models such as the Gail model or the Tirer-Cuzick model only. Models integrating dense breast tissue, e.g. Tice-model, need further validation.**
- **Individual risk assessment should be done by professionals with sufficient skills and experience, with extensive counselling and sufficient attention to patient preferences and support.** *(weak recommendation, very low level of evidence)***.**

B Risk factors other than family history

3. Persons with a past history of mantle irradiation for Hodgkin lymphoma should be considered at high risk *(strong recommendation, moderate level of evidence)***.**

4. Women with very dense breast tissue (BIRADS 4) could be considered as raised risk (life-time risk +/-17 %) *(weak recommendation, very low level of evidence)***.**

5. Lobular and ductal atypical hyperplasia should be considered as high risk *(weak recommendation, low level of evidence)***.**

6. Other risk factors such as BIRADS 3, obesitas, alcohol intake, hormone replacement therapy, early menarche, nulliparity, oral contraceptives, or exogenous hormones (such asDiethylstilbestrol or DES) should be used only as an element integrated in comprehensive risk models as they are only moderately or modestly associated with breast cancer *(strong recommendation, low level of evidence)***.**

WHICH TECHNIQUES SHOULD BE USED?

7. Every screening mammography should be performed in a setting with adequate quality control following the European guidelines and evaluated with independent double reading. A consensus or arbitration procedure should be used in case of discordance. *(strong recommendation, high level of evidence)*.

8. The use of computer-aided detection is not recommended and cannot replace quality controlled mammography with double reading *(strong recommendation, very low level of evidence)*.

9. Film –screen and full-field digital mammography can both be used for screening purposes, with similar accuracy. The use of digital mammography can be beneficial for young women and women with dense breast tissue *(weak recommendation, low level of evidence)*.

10. Ultrasound screening is not recommended in a population-based screening program as the recall rate and number of false positives is too high and the additional cancer detection rate is minimal *(strong recommendation, low level of evidence)*.

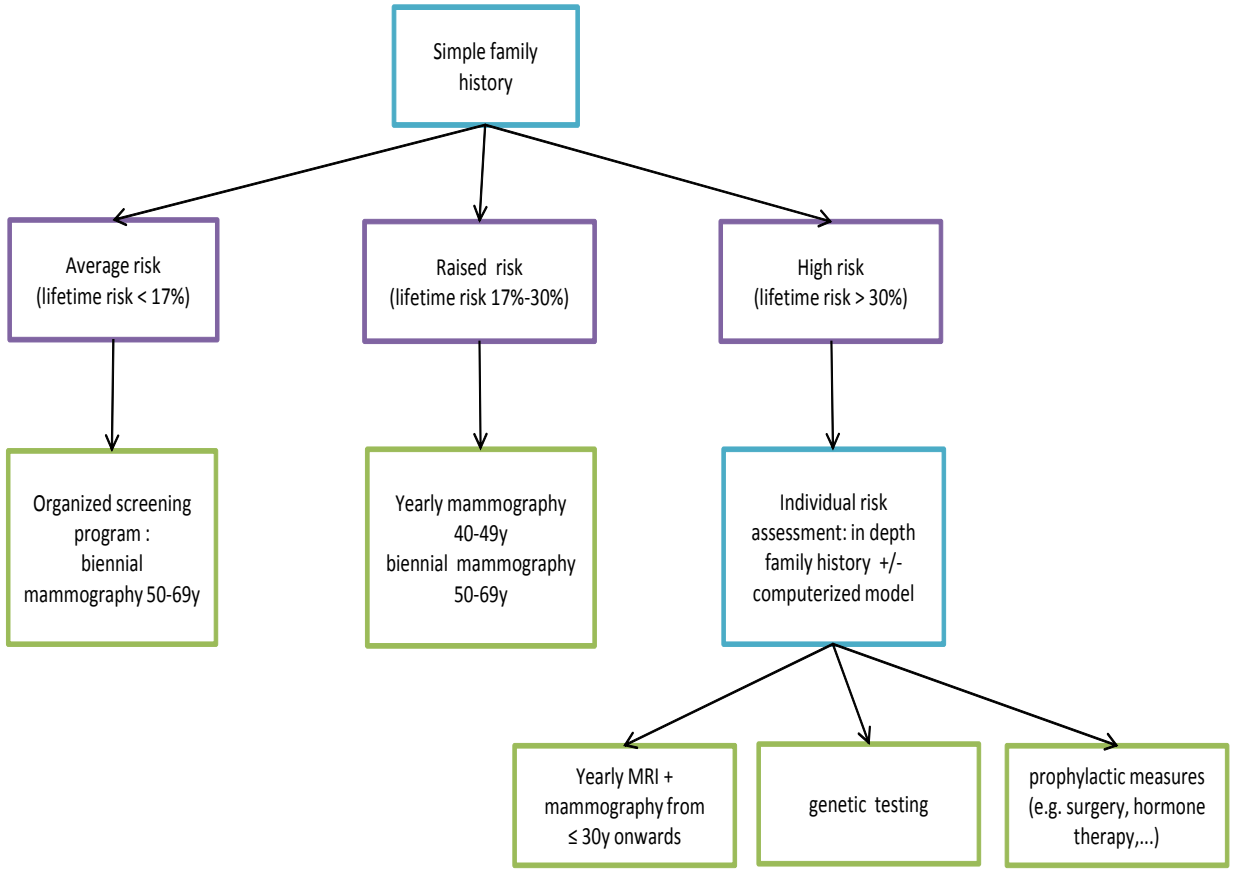
11. Currently available data do not support the use of ultrasound as a screening tool in women with dense breast tissue. In women with very dense breast (BIRADS 4) the screening by ultrasound is not recommended outside a clinical trial setting. *(strong recommendation, low level of evidence)*.

12. Women with raised risk or greater should be offered annual mammographic surveillance from age 40 – 49 years within a quality assured program following European guidelines. From the age of 50 to 69 years, women with a raised breast cancer risk can be included in the general screening program with biennial mammography *(weak recommendation, very low level of evidence)*.

13. For women at proven high risk for breast cancer, yearly MRI and mammography is recommended from the age of 30 years onwards or starting five years before the age of the youngest diagnosed family member with breast cancer (strong recommendation, very low level of evidence).The use of ultrasound can be considered to shorten the interval or as adjunct to a positive mammography or MRI *(weak recommendation, very low level of evidence).*

14. All women participating in screening should be informed about the risk for false positive results, the remaining risk for interval cancer and the absence of data on long term effects on mortality or morbitity for screening outside the population-based screening program, decisions should be taken in dialogue taking into account patients preferences *(strong recommendation, very low level of evidence)*.

Figure 4. Flow chart on the recommendations for screening per risk group



APPENDICES

APPENDIX 1. SEARCH STRATEGY

Appendix 1.1. Women at risk for breast cancer

Appendix 1.1.1. Risk assessment in general

Author	JOR
Project number	2010_03_02
Project name	Part: Technical screening methods in women with or without risk factors of breast cancer
Search questions (PICO,…)	Risk assessment for breast cancer
Keywords	"Breast Neoplasms"[Mesh]

Date 17 Jun 2011	Search for guidelines
Databases	National guidelines Clearinghouse, Guidelines international Network (GIN), SBU, NICE, DACEHTA, MSAC, MAS, HAS, AHRQ, BCBS, AETSA, AATRM, CCOHTA,ECRI, DIMDI, IQWIG
Search Strategy	breast
Note	Potentially relevant publication: 6

Date 17 Jun 2011	Search of HTA reports, systematic reviews and meta analysis
Cochrane database of systematic reviews (CDSR)	

Search Strategy	Search terms : breast neoplasms OR breast cancer
Note	Potentially relevant publication: 2 (80 results)

Date 17 Jun 2011	Search of HTA reports, systematic reviews and meta analysis
Center for review and dissemination databases CRD: DARE, NHS EED and HTA	
Search Strategy	Search terms : breast neoplasms OR breast cancer (in any field)
Note	Potentially relevant publication: 11 (1368 results)

Appendix 1.1.2. Family risk

Date 21/06/2011	Search of HTA reports, systematic reviews and meta analysis
MEDLINE via OVID	
Search Strategy	1 Breast/ or Breast Diseases.mp. [mp=protocol supplementary concept, rare disease supplementary concept, title, original title, abstract, name of substance word, subject heading word, unique identifier] (33657) 2 Neoplasms/ (232024) 3 61 and 62 (560) 4 exp Breast Neoplasms/ (184299) 5 (breast$ adj5 neoplas$).tw. (2601) 6 (breast$ adj5 cancer$).tw. (143736) 7 (breast$ adj5 carcin$).tw. (32599) 8 (breast$ adj5 tumo$).tw. (24854) 9 (breast$ adj5 metasta$).tw. (17368) 10 (breast$ adj5 malig$).tw. (7620) 11 exp Carcinoma, Ductal, Breast/ (9341) 12 1or 2 or 3 or 4 or 5 or 6 or 7 or 8 or 9 (218463) 13 brca1.tw. (6513) 14 brca2.tw. (3652) 15 familial.tw. (72491) 16 family histor$.ti. (3332) 17 hereditary.ti. (21822) 18 13 or 14 or 15 or 16 or 17 (101504) 19 12 and 18 (7707) 20 meta-analysis.mp,pt. or review.pt. or search:.tw. (1744857) 21 19 and 20 (1433) 22 limit 20 to yr="2003 -Current" (794)
Note	**Potentially relevant publications: 794**

Date 21/06/2011	Search of HTA reports, systematic reviews and meta analysis
Database Embase	
Search Strategy	#14 #12 AND #13 1031 #13 2003:py OR 2004:py OR 2005:py OR 2006:py OR 2007:py OR 2008:py OR 2009:py OR 2010:py OR 2011:py 7129881 #12 #6 AND #11 1362

	#11 #1 AND #9 7374 #10 #8 AND #9 237 #9 'risk assessment'/exp 255235 #8 #1 AND #6 AND (2003:py OR 2004:py OR 2005:py OR 2006:py OR 2007:py OR 2008:py OR 2009:py OR 2010:py OR 2011:py) AND 'review'/it 1296 #7 #1 AND #6 10321 #6 #2 OR #3 OR #4 OR #556114 #5 'family history'/exp 39471 #4 'familial cancer'/exp 8138 #3 brca2 7126 #2 brca1 10920 #1 'breast cancer'/exp 213153
Note	Potentially relevant publications: 1031

Appendix 1.1.3. Risk models

Date 22/06/2011	Search of HTA reports, systematic reviews and meta analysis
MEDLINE via OVID	
Search Strategy	1 Breast/ or Breast Diseases.mp. [mp=protocol supplementary concept, rare disease supplementary concept, title, original title, abstract, name of substance word, subject heading word, unique identifier] (33663) 2 Neoplasms/ (232056) 3 1 and 2 (560) 4 exp Breast Neoplasms/ (184326) 5 (breast$ adj5 neoplas$).tw. (2601) 6 (breast$ adj5 cancer$).tw. (143758) 7 (breast$ adj5 carcin$).tw. (32604) 8 (breast$ adj5 tumo$).tw. (24860) 9 (breast$ adj5 metasta$).tw. (17368) 10 (breast$ adj5 malig$).tw. (7623) 11 exp Carcinoma, Ductal, Breast/ (9344) 12 3 or 4 or 5 or 6 or 7 or 8 or 9 or 10 or 11 (218499) 13 exp genetics/ (142147) 14 exp genetic predisposition to disease/ (66277) 15 meta-analysis.mp,pt. or review.pt. or search:.tw. (1745158) 16 risk assessment/ (132224) 17 (assess$ adj2 risk$).ti. (9501) 18 ((classif$ or category$ or stratify$) adj2 risk$).tw. (3769) 19 model$.ti. (263955) 20 *models, theoretical/ (31374) 21 exp *models, statistical/ (23668) 22 exp *models, statistical/ (23668) 23 risk$.ti. (215612) 24 gail model$.tw. (127) 25 28 or 29 or 30 or 31 or 32 or 33 or 34 or 35 or 36 (609396) 26 12 and 37 (16035) 27 30 or 31 or 32 or 33 or 34 or 36 (301989) 28 12 and 39 (3658) 29 12 and 28 (4454) 30 39 and 41 (240)

Note	Potentially relevant publications: 240

Date 22/06/2011	Search of HTA reports, systematic reviews and meta analysis
Database Embase	
Search Strategy	#12 #1 AND #4 AND (2003:py OR 2004:py OR 2005:py OR 2006:py OR 2007:py OR 2008:py OR 2009:py OR 2010:py OR 2011:py) 519 #11 #1 AND #4 691 #4 'statistical model'/exp 72601 #1 'breast cancer'/exp 213211
Note	Potentially relevant publications: 519

Appendix 1.1.4. Risk factors

Date 22/06/2011	Search of HTA reports, systematic reviews and meta analysis
MEDLINE via OVID	
Search Strategy	1 exp genetics/ (142083) 2 Breast/ or Breast Diseases.mp. [mp=protocol supplementary concept, rare disease supplementary concept, title, original title, abstract, name of substance word, subject heading word, unique identifier] (33657) 3 Neoplasms/ (232024) 4 2 and 3 (560) 5 exp Breast Neoplasms/ (184299) 6 (breast$ adj5 neoplas$).tw. (2601) 7 (breast$ adj5 cancer$).tw. (143736) 8 (breast$ adj5 carcin$).tw. (32599) 9 (breast$ adj5 tumo$).tw. (24854) 10 (breast$ adj5 metasta$).tw. (17368) 11 (breast$ adj5 malig$).tw. (7620) 12 exp Carcinoma, Ductal, Breast/ (9341) 13 4 or 5 or 6 or 7 or 8 or 9 or 10 or 11 or 12 (218463) 14 meta-analysis.mp,pt. or review.pt. or search:.tw. (1744857) 15 alcohol$.mp. (250079) 16 exp diet/ or exp food/ (1014451) 17 hormone replacement therapy/ or estrogen replacement therapy/ (18671) 18 exp contraceptives, oral/ (39142) 19 parity/ (18804) 20 (nulliparous or nulliparity or childless$).mp. (6883) 21 menarche/ (3696) 22 menarche.tw. (5108) 23 exp obesity/ (109781) 24 (breast adj3 dens$).mp. (1216) 25 carcinoma in situ/ (12076) 26 ductal hyperplasia.mp. (812) 27 lobular hyperplasia.mp. (247) 28 sclerosing adenosis.mp. (282) 29 previous breast cancer.tw. (101) 30 neoplasms, second primary/ (8847) 31 15 or 16 or 17 or 18 or 19 or 20 or 21 or 22 or 23 or 24 or 25 or 26 or 27 or 28 or 29 or 30 (1422760)

	32 13 and 31 (21671) 33 14 and 32 (4324) 34 limit 33 to yr="2006 -Current" (1019)
Note	Potentially relevant publications: 1019

Date 22/06/2011	Search of HTA reports, systematic reviews and meta analysis
Database Embase	
Search Strategy	#26 #24 AND [2006-2011]/py AND 'review'/it736 #25 #24 AND [2006-2011]/py 3389 #24 #21 AND #23 6651 #23 #1 AND #22 25136 #22 'risk'/exp 1020546 #21 #2 OR #3 OR #4 OR #5 OR #6 OR #7 OR #8 OR #9 OR #11 OR #12 OR #13 OR #14 OR #15 OR #16 OR #17 OR #18 OR #19 OR #20 1112855 #20 ductal AND 'hyperplasia'/exp OR sclerosing AND adnenosis OR lobular AND 'hyperplasia'/exp 607 #19 'carcinoma'/exp AND in AND situ 40002 #18 'oral'/exp AND 'contraceptive'/exp 13669 #17 exp AND 'oral'/exp AND 'contraceptive'/exp AND agent 271 #16 'fat'/exp AND intake 5611 #15 'food'/exp 529257 #14 'diet'/exp 159733 #13 'hormone'/exp AND replacement OR 'estrogen'/exp AND replacement OR hrt 25181 #12 'hormone'/exp AND substitution 1979 #11 previous AND 'breast'/exp AND 'cancer'/exp 1341 #9 second AND 'cancer'/exp 119037 #8 'breast'/exp AND 'density'/exp 490 #7 'menarche'/exp 5832 #6 morbid AND 'obesity'/exp OR 'obesity'/exp 195181 #5 nullipar$ OR childless$ 892 #4 'parity'/exp 17529 #3 'alcohol'/exp AND 'drinking'/exp 2184 #2 alcohol.tw. OR 'alcoholism'/exp 85454 #1 'breast cancer'/exp 213211
Note	Potentially relevant publications: 519

Appendix 1.2. Technical methods for breast cancer screening

Appendix 1.2.1. Search strategy for systematic reviews, meta-analyses, HTA, evidence-based guidelines

Author	JEG
Project number	2010_03_02
Project name	Part: Technical screening methods in women with or without risk factors of breast cancer
Search questions (PICO,…)	Screening with mammography (single or double reading) compared with digital mammography (computer aid) and/or mammography + ultrasound and/or MRI (with or without mammography)
Keywords	"Breast Neoplasms"[Mesh]

Date 17 Jun 2011	Search for guidelines
Databases	National guidelines Clearinghouse, Guidelines international Network (GIN), CBO, Evidence-Based Medicine guidelines, Guidelines finder UK, New Zealand guidelines group, SIGN, NICE, HAS
Search Strategy	breast
Note	Potentially relevant publication: 3

Date 17 Jun 2011	Search of HTA reports, systematic reviews and meta analysis
Cochrane database of systematic review (CDSR)	
Search Strategy	Search terms : breast neoplasms OR breast cancer
Note	Potentially relevant publication: 1 (80 results)

Date 17 Jun 2011	Search of HTA reports, systematic reviews and meta analysis
Center for review and dissemination databases CRD: DARE, NHS EED and HTA	
Search Strategy	Search terms : breast neoplasms OR breast cancer (in any field)
Note	Potentially relevant publication: 40 (1364 results)

Date 18 Jul 2011	Search of HTA reports, systematic reviews and meta analysis
MEDLINE via OVID	
Search Strategy	1 breast/ or breast diseases/ (33409) 2 Neoplasms/ (233463) 3 1 and 2 (554) 4 exp Breast neoplasms/ (185557) 5 (breast$ adj5 neoplas$).tw. (2613) 6 (breast$ adj5 cancer$).tw. (145011) 7 (breast$ adj5 carcin$).tw. (32753) 8 (breast$ adj5 tumo$).tw. (25068) 9 (breast$ adj5 metasta$).tw. (17490) 10 (breast$ adj5 malig$).tw. (7682) 11 exp Carcinoma, Ductal, Breast/ (9412) 12 3 or 4 or 5 or 6 or 7 or 8 or 9 or 10 or 11 (220083) 13 screening.mp. or exp Mass Screening/ (297862) 14 exp "Early Detection of Cancer"/ (2471) 15 exp Diagnosis/ or exp Early Diagnosis/ (5419772) 16 13 or 14 or 15 (5555389) 17 12 and 16 (102358) 18 mammography.mp. or exp Mammography/ (24700) 19 exp Radiographic Image Enhancement/ (259928) 20 digital mammography.mp. (762) 21 exp Radiographic Image Interpretation,

	Computer-Assisted/ (7075) 22 exp Ultrasonography, Mammary/ or exp Ultrasonography/ (215504) 23 ultrasound.mp. (117206) 24 echography.mp. or Ultrasonography/ (62403) 25 exp Magnetic Resonance Imaging/ or mri.mp. (271087) 26 MRI.mp. (100631) 27 18 or 19 or 20 or 21 or 22 or 23 or 24 or 25 or 26 (746220) 28 17 and 27 (26282) 29 meta-analysis.pt,ti,ab,sh. (40049) 30 (meta anal$ or metaanal$).ti,ab,sh. (50797) 31 29 or 30 (50797) 32 (methodol$ or systematic$ or quantitativ$).ti,ab,sh. (617325) 33 ((methodol$ or systematic$ or quantitativ$) adj (review$ or overview$ or survey$)).ti,ab,sh. (31423) 34 (medline or embase or index medicus).ti,ab. (38933) 35 ((pool$ or combined or combining) adj (data or trials or studies or results)).ti,ab. (10324) 36 32 or 33 or 34 or 35 (649104) 37 review.pt,sh. (1625450) 38 36 and 37 (103368) 39 31 or 38 (140332) 40 28 and 39 (416)
Note	Potentially relevant publications: 416

Date 18 Jul 2011	Search of HTA reports, systematic reviews and meta analysis
Database **Embase**	
Search Strategy	#17. #13 AND #16 172 #16. #14 OR #15 78,582 #15. 'systematic review'/exp 42,641 #14. 'meta analysis'/exp 55,599 #13. #11 AND #12 6,914 #12. #1 AND #5 12,234 #11. #6 OR #7 OR #8 OR #9 OR #10 759,123 #10. 'nuclear magnetic resonance imaging'/exp 375,730 #9. 'echography'/exp 399,170 #8. 'echomammography'/exp 3,999 #7. 'digital mammography'/exp 542 #6. 'mammography'/exp 34,170 #5. #2 OR #3 OR #4 68,866 #4. 'mass radiography'/exp 240 #3. 'genetic screening'/exp 31,495 #2. 'cancer screening'/exp 37,736 #1. 'breast cancer'/exp 214,696 .....
Note	Potentially relevant publications: 172

Appendix 1.2.2. Search strategy for primary studies 2007-2011

Author	AND, LEV
Project number	2010_03_02
Project name	Part: Technical screening methods in women with or without risk factors of breast cancer
Search questions	Screening with mammography (single or double reading) compared with digital mammography (computer assisted) and/or ultrasound +/- mammography and/or MRI +/- mammography
Keywords	“Breast Neoplasms” [MesH]

Date 30 Aug 2011	Search for primary studies digital mammography / computer assisted reading published since 2007
Database	Medline Ovid
Search Strategy	1 Breast Diseases/ or Breast/ (33729) 2 Neoplasms/ (237507) 3 1 and 2 (557) 4 exp Breast Neoplasms/ (188326) 5 (breast$ adj5 neoplas$).tw. (2645) 6 (breast$ adj5 cancer$).tw. (147506) 7 (breast$ adj5 carcin$).tw. (33097) 8 (breast adj5 tumo$).tw. (25398) 9 (breast adj5 metasta$).tw. (17741) 10 (breast adj5 malig$).tw. (7756) 11 exp Carcinoma, Ductal, Breast/ (9550) 12 3 or 4 or 5 or 6 or 7 or 8 or 9 or 10 or 11 (223448) 13 mass screening.mp. or exp Mass Screening/ (89329) 14 exp "Early Detection of Cancer"/ (2689) 15 exp Diagnosis/ or exp Early Diagnosis/ (5498221) 16 13 or 14 or 15 (5502683) 17 screening.mp. or exp Mass Screening/ (302732) 18 14 or 15 or 17 (5636152) 19 12 and 18 (103969) 20 mammography.mp. or exp Mammography/ (25050) 21 exp Radiographic Image Enhancement/ (264935) 22 digital mammography.mp. (793) 23 exp Radiographic Image Interpretation, Computer-Assisted/ (7252) 24 20 or 21 or 22 or 23 (289611) 25 19 and 24 (21722) 26 25 (21722) 27 limit 26 to yr="2008 -Current" (3977) 29 exp *Mammography/ (12442) 30 exp diagnosis, computer-assisted/ or radiographic image enhancement/ (59307) 31 22 or 23 or 29 or 30 (69889) 32 19 and 31 (10872) 33 limit 32 to yr="2007 -Current" (2797) 34 exp *Mass Screening/ (47133) 35 exp *"Early Detection of Cancer"/ (1075) 36 exp *Early Diagnosis/ (1508) 37 34 or 35 or 36 (48449) 38 19 and 31 and 37 (2503)

	39 limit 38 to yr="2007 -Current" (629)
Note	Potentially relevant publications: 629

Date 30 Aug 2011	Search for primary studies Ultrasound published since 2008
Database	Medline Ovid
Search Strategy	1 Breast Diseases/ or Breast/ (33729) 2 Neoplasms/ (237507) 3 1 and 2 (557) 4 exp Breast Neoplasms/ (188326) 5 (breast$ adj5 neoplas$).tw. (2645) 6 (breast$ adj5 cancer$).tw. (147506) 7 (breast$ adj5 carcin$).tw. (33097) 8 (breast adj5 tumo$).tw. (25398) 9 (breast adj5 metasta$).tw. (17741) 10 (breast adj5 malig$).tw. (7756) 11 exp Carcinoma, Ductal, Breast/ (9550) 12 3 or 4 or 5 or 6 or 7 or 8 or 9 or 10 or 11 (223448) 13 mass screening.mp. or exp Mass Screening/ (89329) 14 exp "Early Detection of Cancer"/ (2689) 15 exp Diagnosis/ or exp Early Diagnosis/ (5498221) 16 13 or 14 or 15 (5502683) 17 screening.mp. or exp Mass Screening/ (302732) 18 14 or 15 or 17 (5636152) 19 12 and 18 (103969) 55 exp *Ultrasonography, Mammary/ or exp *Ultrasonography/ (107191) 56 19 and 55 (2245) 57 limit 56 to yr="2008 -Current" (524) 58 12 and 37 and 55 (72)
Note	Potentially relevant publications: 72

Date 30-8-2011	Search for primary studies Ultrasound published since 2008
Database	Medline Ovid
Search Strategy	1 Breast Diseases/ or Breast/ (33729) 2 Neoplasms/ (237507) 3 1 and 2 (557) 4 exp Breast Neoplasms/ (188326) 5 (breast$ adj5 neoplas$).tw. (2645) 6 (breast$ adj5 cancer$).tw. (147506) 7 (breast$ adj5 carcin$).tw. (33097) 8 (breast adj5 tumo$).tw. (25398) 9 (breast adj5 metasta$).tw. (17741) 10 (breast adj5 malig$).tw. (7756) 11 exp Carcinoma, Ductal, Breast/ (9550) 12 3 or 4 or 5 or 6 or 7 or 8 or 9 or 10 or 11 (223448) 34 exp *Mass Screening/ (47133) 35 exp *"Early Detection of Cancer"/ (1075) 36 exp *Early Diagnosis/ (1508) 37 34 or 35 or 36 (48449) 60 exp *Magnetic Resonance Imaging/ (104418)

	61 12 and 37 and 60 (84) 62 limit 61 to yr="2007 -Current" (48)
Note	Potentially relevant publications: 48
Date	30-08-2011
Database	Embase Primary studies
Search Strategy	#33. #32 AND ([article]/lim OR [article in press]/lim 134 OR [review]/lim) AND [embase]/lim AND [2007-2012]/py #32. #23 AND #27 134 #31. #30 AND ([article]/lim OR [article in press]/lim 50 OR [review]/lim) AND [embase]/lim AND [2008-2012]/py #30. #20 AND #27 62 #29. #28 AND ([article]/lim OR [article in press]/lim 522 OR [review]/lim) AND [embase]/lim AND [2007-2012]/py #28. #15 AND #27 522 #27. #1 AND #26 2,861 #26. #5 AND ([article]/lim OR [article in press]/lim 17,763 OR [review]/lim) AND [embase]/lim AND [2007-2012]/py #25. #24 AND ([article]/lim OR [article in press]/lim 134 OR [review]/lim) AND [embase]/lim AND [2007-2012]/py #24. #9 AND #23 293 #23. 'nuclear magnetic resonance imaging'/exp/mj 108,916 #22. #21 AND ([article]/lim OR [article in press]/lim 50 OR [review]/lim) AND [embase]/lim AND [2008-2012]/py #21. #9 AND #20 195 #20. #18 OR #19 128,744 #19. 'echography'/exp/mj 128,744 #18. 'echomammography'/exp/mj 1,627 #17. #16 AND ([article]/lim OR [article in press]/lim 522 OR [review]/lim) AND [embase]/lim AND [2007-2012]/py #16. #9 AND #15 2,758 #15. #13 OR #14 16,041 #14. 'digital mammography'/exp/mj 342 #13. 'mammography'/exp/mj 16,041 #12. #10 AND ([article]/lim OR [article in press]/lim 1,437 OR [review]/lim) AND [embase]/lim AND [2007-2012]/py #10. #8 AND #9 6,638 #9. #1 AND #5 12,359 #8. #6 OR #7 34,449 #7. 'digital mammography'/exp 569 #6. 'mammography'/exp 34,449 #5. #2 OR #3 OR #4 69,850

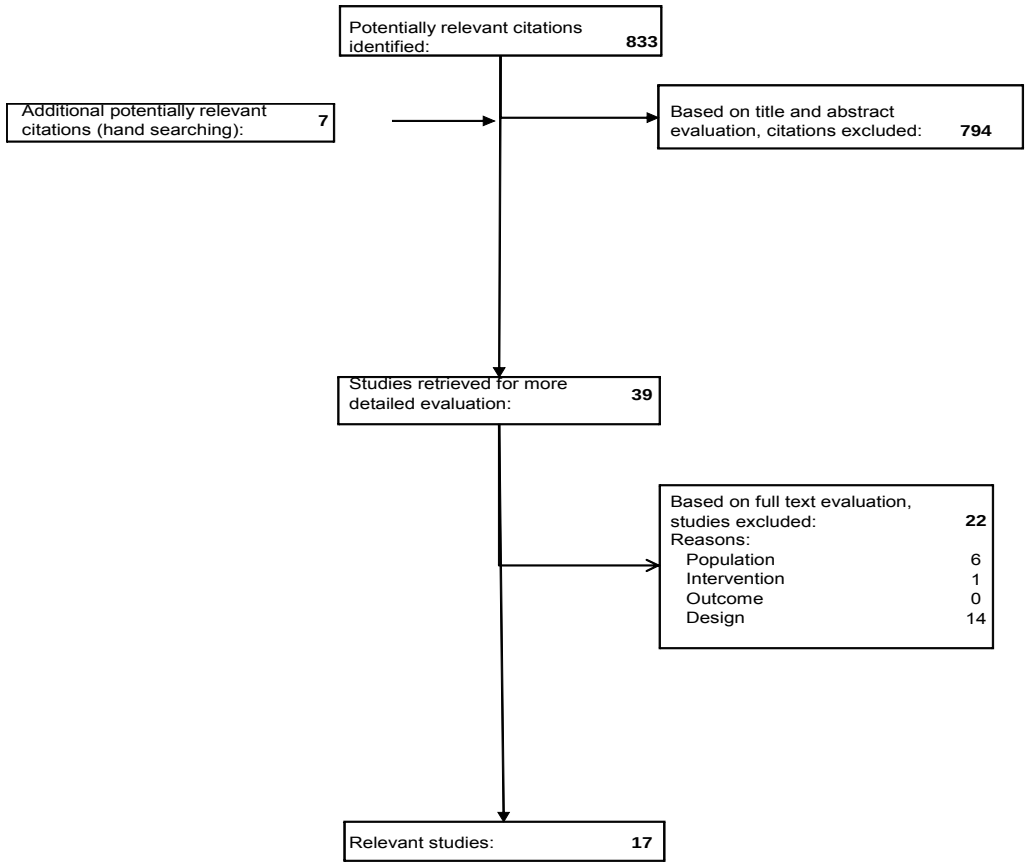
	#4. 'mass radiography'/exp 244 #3. 'genetic screening'/exp 32,147 #2. 'cancer screening'/exp 38,082 #1. 'breast cancer'/exp 217,239
Note	

Date	30-08-2011
Database	Cochrane Library (clinical trials)
Search Strategy	#15 MeSH descriptor Breast Neoplasms, this term only 7050 #16 MeSH descriptor Magnetic Resonance Imaging, this term only 3591 #17 MeSH descriptor Mass Screening, this term only 3325 #18 MeSH descriptor Early Detection of Cancer, this term only 129 #19 ((#17 OR #18) AND #16 AND #15), from 2007 to 2011 10 #20 MeSH descriptor Ultrasonography explode all trees 6304 #21 ((#17 OR #18) AND #15 AND #20), from 2008 to 2011 2 #22 MeSH descriptor Mammography, this term only 787 #23 MeSH descriptor Radiographic Image Enhancement, this term only 317 #24 MeSH descriptor Radiographic Image Interpretation, Computer-Assisted, this term only 167 #25 ((#17 OR #18) AND (#22 OR #23 OR #24) AND #15), from 2007 to 2011 0
Note	

APPENDIX 2. RESEARCH AND SELECTION RESULTS

Appendix 2.1. Women at risk for breast cancer

Appendix 2.1.1. Flow chart search risk models

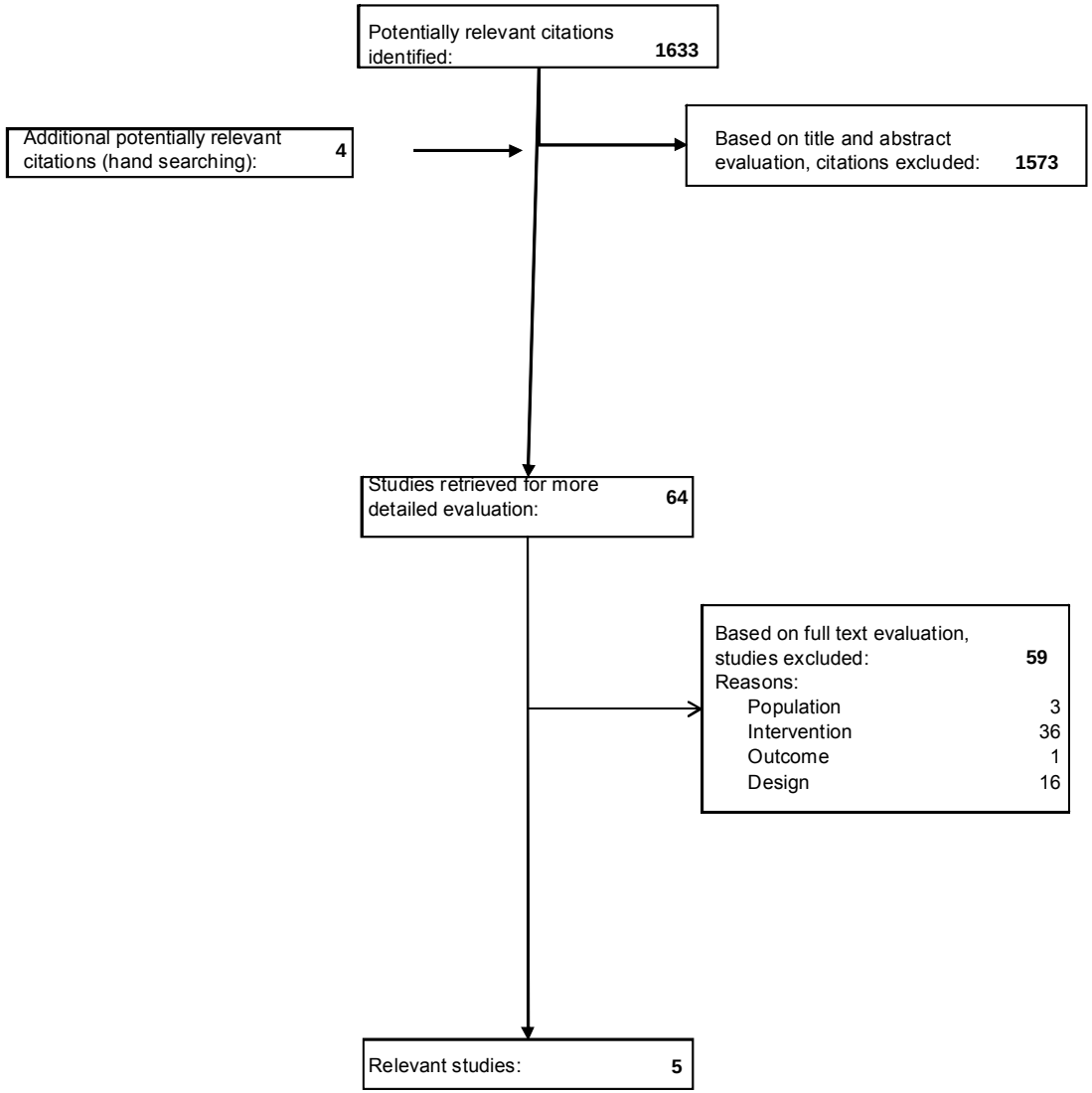


Reason for exclusion

Amir E, Freedman OC, Seruga B, Evans DG. Assessing women at high risk of breast cancer: a review of risk assessment models. J Natl Cancer Inst. 2010;102(10):680-91.	Design	Not a systematic review
Rao NY, Hu Z, Yu JM, Li WF, Zhang B, Su FX, et al. Evaluating the performance of models for predicting the BRCA germline mutations in Han Chinese familial breast cancer patients. Breast Cancer Res Treat. 2009;116(3):563-70.	Population	Chinese population
Pauw AD, Stoppa-Lyonnet D, Andrieu N, Asselain B. Estimation du risque individuel de cancer du sein: interet et limites des modeles de calcul de risque. Bull Cancer. 2009;96(10):979-88.	Design	Narrative review
Kurian AW, Gong GD, John EM, Miron A, Felberg A, Phipps AI, et al. Performance of prediction models for BRCA mutation carriage in three racial/ethnic groups: Findings from the Northern California breast cancer family registry. Cancer Epidemiol. Biomarkers Prev. 2009;18(4):1084-91.	Population	Performance tested on breast cancer patients
Jacobi CE, de Bock GH, Siegerink B, van Asperen CJ. Differences and similarities in breast cancer risk assessment models in clinical practice: which model to choose? Breast Cancer Res Treat. 2009;115(2):381-90.	Design	Narrative
Cook NR, Rosner BA, Hankinson SE, Colditz GA. Mammographic screening and risk factors for breast cancer. Am. J. Epidemiol. 2009;170(11):1422-32.	Intervention	Not on risk models
Amir E, Freedman O. Underestimation of risk by Gail model extends beyond women with atypical hyperplasia. J Clin Oncol. 2009;27(9):1526; author reply 7.	Design	Letter
Adams-Campbell LL, Makambi KH, Frederick WA, Gaskins M, Dewitty RL, McCaskill-Stevens W. Breast cancer risk assessments comparing Gail and CARE models in African-American women. Breast Journal. 2009;15(1):Sep-Oct.	Population	Afr Am population
Novotny J, Pecen L, Petruzelka L, Svobodnik A, Dusek L, Danes J, et al. Breast cancer risk assessment in the Czech female population--an adjustment of the original Gail model. Breast Cancer Res Treat. 2006;95(1):29-35.	Design	Not a validation but a calibratin on a chech case control study
Hoskins KF, Zwaagstra A, Ranz M. Validation of a tool for identifying women at high risk for hereditary breast cancer in population-based screening. Cancer. 2006;107(8):1769-76.	Design	Not a model validation study
Chen J, Pee D, Ayyagari R, Graubard B, Schairer C, Byrne C, et al. Projecting absolute invasive breast cancer risk in white women with a model that includes mammographic density. J. Natl. Cancer Inst. 2006;98(17):1215-26.	Design	Proposes new model but without validation

Antoniou AC, Durocher F, Smith P, Simard J, Easton DF, members IBp. BRCA1 and BRCA2 mutation predictions using the BOADICEA and BRCAPRO models and penetrance estimation in high-risk French-Canadian families. Breast Cancer Res. 2006;8(1):R3.		Too specific population
Lee EO, Ahn SH, You C, Lee DS, Han W, Choe KJ, et al. Determining the main risk factors and high-risk groups of breast cancer using a predictive model for breast cancer risk assessment in South Korea. Cancer Nurs. 2004;27(5):400-6.	Population	Asian population

Appendix 2.2.1. Flow chart risk factors



Reasons for exclusion

Yang XR, Chang-Claude J, Goode EL, Couch FJ, Nevanlinna H, Milne RL, et al. Associations of breast cancer risk factors with tumor subtypes: a pooled analysis from the Breast Cancer Association Consortium studies. J Natl Cancer Inst. 2011;103(3):250-63.	Outcome	compares ER and non ER tumors
Salagame U, Canfell K, Banks E. An epidemiological overview of the relationship between hormone replacement therapy and breast cancer. Expert Rev. Endocrinol. Metab. 2011;6(3):397-409.	Design	not a systematic review
Lynch BM, Neilson HK, Friedenreich CM. Physical activity and breast cancer prevention. Recent Results Cancer Res. 2011;186:13-42.	Intervention	Not useful for targeting of interventions
La Vecchia C. Infertility, ovulation, induced ovulation, and female cancers. Eur.J. Cancer Prev. 2011;20(3):147-9.	Design	not a systematic review
Kim J, Oktay K. Infertility as a risk factor of ovarian and breast cancer. Expert Rev. Obstet. Gynecol. 2011;6(2):153-61.	Design	not a systematic review
Howell A, Evans GD. Hormone replacement therapy and breast cancer. Recent Results in Cancer Research. 2011;188:115-24.	Design	not a systematic review
Gandini S, Boniol M, Haukka J, Byrnes G, Cox B, Sneyd MJ, et al. Meta-analysis of observational studies of serum 25-hydroxyvitamin D levels and colorectal, breast and prostate cancer and colorectal adenoma. Int J Cancer. 2011;128(6):1414-24.	Intervention	Not useful for targeting of interventions
Friedenreich CM. Physical activity and breast cancer: Review of the epidemiologic evidence and biologic mechanisms. Recent Results Cancer Res. 2011;188:125-39.	Intervention	Not useful for targeting of interventions
Dong JY, Zhang L, He K, Qin LQ. Dairy consumption and risk of breast cancer: A meta-analysis of prospective cohort studies. Breast Cancer Res. Treat. 2011;127(1):23-31.	Intervention	Not useful for targeting of interventions
Dong JY, Qin LQ. Dietary glycemic index, glycemic load, and risk of breast cancer: Meta-analysis of prospective cohort studies. Breast Cancer Res. Treat. 2011;126(2):287-94.	Intervention	Not useful for targeting of interventions
Basen-Engquist K, Chang M. Obesity and cancer risk: Recent review and evidence. Curr. Oncol. Rep. 2011;13(1):71-6.	Design	not a systematic review
Yin L, Grandi N, Raum E, Haug U, Arndt V, Brenner H. Meta-analysis: serum vitamin D and breast cancer risk. Eur J Cancer. 2010;46(12):2196-205.	Intervention	Not useful for targeting of interventions

Iodice S, Barile M, Rotmensz N, Feroce I, Bonanni B, Radice P, et al. Oral contraceptive use and breast or ovarian cancer risk in BRCA1/2 carriers: a meta-analysis. Eur J Cancer. 2010;46(12):2275-84.	Population	already high risk persons
Druesne-Pecollo N, Latino-Martel P, Norat T, Barrandon E, Bertrais S, Galan P, et al. Beta-carotene supplementation and cancer risk: a systematic review and metaanalysis of randomized controlled trials. Int J Cancer. 2010;127(1):172-84.	Intervention	Not useful for targeting of interventions
Cibula D, Gompel A, Mueck AO, La Vecchia C, Hannaford PC, Skouby SO, et al. Hormonal contraception and risk of cancer. Hum Reprod Update. 2010;16(6):631-50.	Design	no assesment of study quality
Buck K, Zaineddin AK, Vrieling A, Linseisen J, Chang-Claude J. Meta-analyses of lignans and enterolignans in relation to breast cancer risk. Am J Clin Nutr. 2010;92(1):141-53.	Intervention	Not useful for targeting of interventions
Brennan SF, Cantwell MM, Cardwell CR, Velentzis LS, Woodside JV. Dietary patterns and breast cancer risk: a systematic review and meta-analysis. Am J Clin Nutr. 2010;91(5):1294-302.	Intervention	Not useful for targeting of interventions
Boyd NF, Martin LJ, Bronskill M, Yaffe MJ, Duric N, Minkin S. Breast tissue composition and susceptibility to breast cancer. J Natl Cancer Inst. 2010;102(16):1224-37.	Design	not a systematic review
Alexander DD, Morimoto LM, Mink PJ, Lowe KA. Summary and meta-analysis of prospective studies of animal fat intake and breast cancer. Nutr. 2010;23(1):169-79.	Intervention	Not useful for targeting of interventions
Alexander DD, Morimoto LM, Mink PJ, Cushing CA. A review and meta-analysis of red and processed meat consumption and breast cancer. Nutr. 2010;23(2):349-65.	Intervention	Not useful for targeting of interventions
Velentzis LS, Cantwell MM, Cardwell C, Keshtgar MR, Leathem AJ, Woodside JV. Lignans and breast cancer risk in pre- and post-menopausal women: meta-analyses of observational studies. Br J Cancer. 2009;100(9):1492-8.	Intervention	Not useful for targeting of interventions
Taylor VH, Misra M, Mukherjee SD. Is red meat intake a risk factor for breast cancer among premenopausal women? Breast Cancer Res Treat. 2009;117(1):1-8.	Intervention	Not useful for targeting of interventions
Gaffield ME, Culwell KR, Ravi A. Oral contraceptives and family history of breast cancer. Contraception. 2009;80(4):372-80.	Population	already high risk persons

Farquhar C, Marjoribanks J, Lethaby A, Suckling JA, Lamberts Q. Long term hormone therapy for perimenopausal and postmenopausal women. Cochrane Database Syst Rev. 2009(2):CD004143.	Design	main focus on beneficial effect, only RCT's considered, which is not powerfull enough to estimate the risk of breast cancer
Enderlin CA, Coleman EA, Stewart CB, Hakkak R. Dietary soy intake and breast cancer risk. Oncol Nurs Forum. 2009;36(5):531-9.	Intervention	Not useful for targeting of interventions
Edefonti V, Randi G, La Vecchia C, Ferraroni M, Decarli A. Dietary patterns and breast cancer: A review with focus on methodological issue. Nutr. Rev. 2009;67(6):297-314.	Intervention	Not useful for targeting of interventions
Cohen JM, Hutcheon JA, Julien SG, Tremblay ML, Fuhrer R. Insufficient milk supply and breast cancer risk: a systematic review. PLoS ONE. 2009;4(12):e8237.	Intervention	Not useful for targeting of interventions
Boyle P, Boffetta P. Alcohol consumption and breast cancer risk. Breast Cancer Research. 2009;11(3).	Design	not a systematic review
Bertone-Johnson ER. Vitamin D and breast cancer. Ann Epidemiol. 2009;19(7):462-7.	Intervention	Not useful for targeting of interventions
Wu AH, Yu MC, Tseng CC, Pike MC. Epidemiology of soy exposures and breast cancer risk. Br J Cancer. 2008;98(1):9-14.	Intervention	Not useful for targeting of interventions
Velentzis LS, Woodside JV, Cantwell MM, Leathem AJ, Keshtgar MR. Do phytoestrogens reduce the risk of breast cancer and breast cancer recurrence? What clinicians need to know. Eur J Cancer. 2008;44(13):1799-806.	Intervention	Not useful for targeting of interventions
Thompson AK, Shaw DI, Minihane AM, Williams CM. Trans-fatty acids and cancer: the evidence reviewed. Nutr. 2008;21(2):174-88.	Intervention	Not useful for targeting of interventions

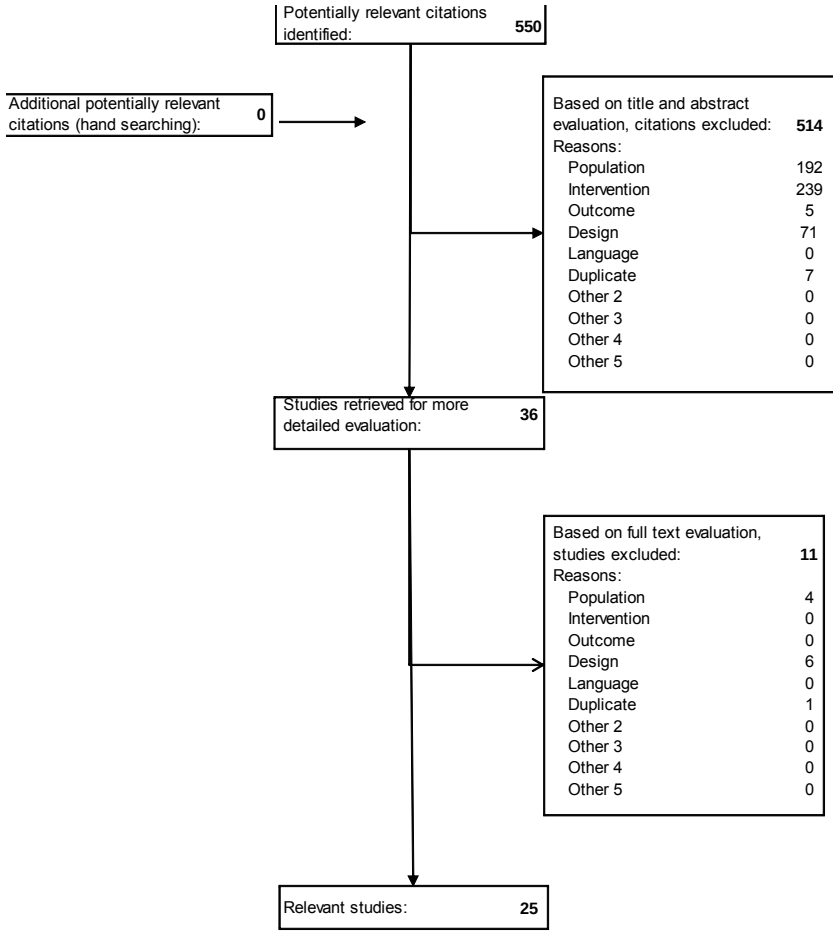
Pichard C, Plu-Bureau G, Neves-e Castro M, Gompel A. Insulin resistance, obesity and breast cancer risk. Maturitas. 2008;60(1):19-30.	Intervention	focus on fysiological mechanisms, not on RR estimation
Neves ECM. Association of ovarian and uterine cancers with postmenopausal hormonal treatments. Clin Obstet Gynecol. 2008;51(3):607-17.	Design	not a systematic review
Namer M, Luporsi E, Gligorov J, Lokiec F, Spielmann M. L'utilisation de deodorants/antitranspirants ne constitue pas un risque de cancer du sein. Bull Cancer. 2008;95(9):871-80.	Intervention	not usable for rintervention targetting
Messina MJ, Wood CE. Soy isoflavones, estrogen therapy, and breast cancer risk: analysis and commentary. Nutrition Journal. 2008;7(17).	Intervention	not usable for rintervention targetting
Gissel T, Rejnmark L, Mosekilde L, Vestergaard P. Intake of vitamin D and risk of breast cancer--a meta-analysis. J Steroid Biochem Mol Biol. 2008;111(3-5):195-9.	Intervention	not usable for rintervention targetting
Ginsburg OM, Martin LJ, Boyd NF. Mammographic density, lobular involution, and risk of breast cancer. Br J Cancer. 2008;99(9):1369-74.	Intervention	focus on fysiological mechanisms, not on RR estimation
Cuzick J. Hormone replacement therapy and the risk of breast cancer. Eur J Cancer. 2008;44(16):2344-9.	Design	not a systematic review
Colston KW. Vitamin D and breast cancer risk. Baillieres Best Pract Res Clin Endocrinol Metab. 2008;22(4):587-99.	Intervention	focus on fysiological mechanisms, not on RR estimation
Cetin I, Cozzi V, Antonazzo P. Infertility as a cancer risk factor - a review. Placenta. 2008;29:169-77.	Design	not a systematic review
Casey PM, Cerhan JR, Pruthi S. Oral contraceptive use and risk of breast cancer. Mayo Clin Proc. 2008;83(1):86-90; quiz -1.	Design	not a systematic review
Xue F, Michels KB. Diabetes, metabolic syndrome, and breast cancer: a review of the current evidence. Am J Clin Nutr. 2007;86(3):s823-35.	Design	not a systematic review
Qin LQ, Xu JY, Wang PY, Kazuhiko H. Effects of milk and its products on breast cancer risk: A review. Chin. J. Cancer Prev. Treat. 2007;14(17):1345-9.	Intervention	not usable for rintervention targetting

Nagata C, Mizoue T, Tanaka K, Tsuji I, Wakai K, Inoue M, et al. Alcohol drinking and breast cancer risk: an evaluation based on a systematic review of epidemiologic evidence among the Japanese population. Jpn J Clin Oncol. 2007;37(8):568-74.	Population	only japanese
Michels KB, Mohllajee AP, Roset-Bahmanyar E, Beehler GP, Moysich KB. Diet and breast cancer: a review of the prospective observational studies. Cancer. 2007;109(12 Suppl):2712-49.	Intervention	not usable for rintervention targetting
Britt K, Ashworth A, Smalley M. Pregnancy and the risk of breast cancer. Endocr Relat Cancer. 2007;14(4):907-33.		
Trock BJ, Hilakivi-Clarke L, Clarke R. Meta-analysis of soy intake and breast cancer risk. J Natl Cancer Inst. 2006;98(7):459-71.	Intervention	not usable for rintervention targetting
Qin LQ, Xu JY, Wang PY, Hoshi K. Soyfood intake in the prevention of breast cancer risk in women: a meta-analysis of observational epidemiological studies. J Nutr Sci Vitaminol (Tokyo). 2006;52(6):428-36.	Intervention	not usable for rintervention targetting
Mourits MJ, GH DEB. Exogenous steroids for menopausal symptoms and breast/endometrial cancer risk. International Journal of Gynecological Cancer. 2006;2:494-6.		
MacLean CH, Newberry SJ, Mojica WA, Khanna P, Issa AM, Suttorp MJ, et al. Effects of omega-3 fatty acids on cancer risk: a systematic review. JAMA. 2006;295(4):403-15.	Intervention	not usable for rintervention targetting
Kim YI. Does a high folate intake increase the risk of breast cancer? Nutr Rev. 2006;64(10 Pt 1):468-75.	Intervention	not usable for rintervention targetting
Boffetta P, Hashibe M. Alcohol and cancer. Lancet Oncol. 2006;7(2):149-56.	Design	not a systematic review
Boccardo F, Puntoni M, Guglielmini P, Rubagotti A. Enterolactone as a risk factor for breast cancer: a review of the published evidence. Clin Chim Acta. 2006;365(1-2):58-67.	Intervention	not usable for rintervention targetting
Velie EM, Nechuta S, Osuch JR. Lifetime reproductive and anthropometric risk factors for breast cancer in postmenopausal women. Breast Dis. 2005/2006;24(1):17-35.	Design	not a systematic review
Hankinson SE. Endogenous hormones and risk of breast cancer in postmenopausal women. Breast Dis. 2005/2006;24(1):3-15.	Intervention	not usable for rintervention targetting

Appendix 2.3. Technical methods for breast cancer screening

Appendix 2.3.1. Flow chart results search for SR, MA, HTA and guidelines

Figure 1: Flow chart of the literature selection process SR, MA, HTA, guidelines

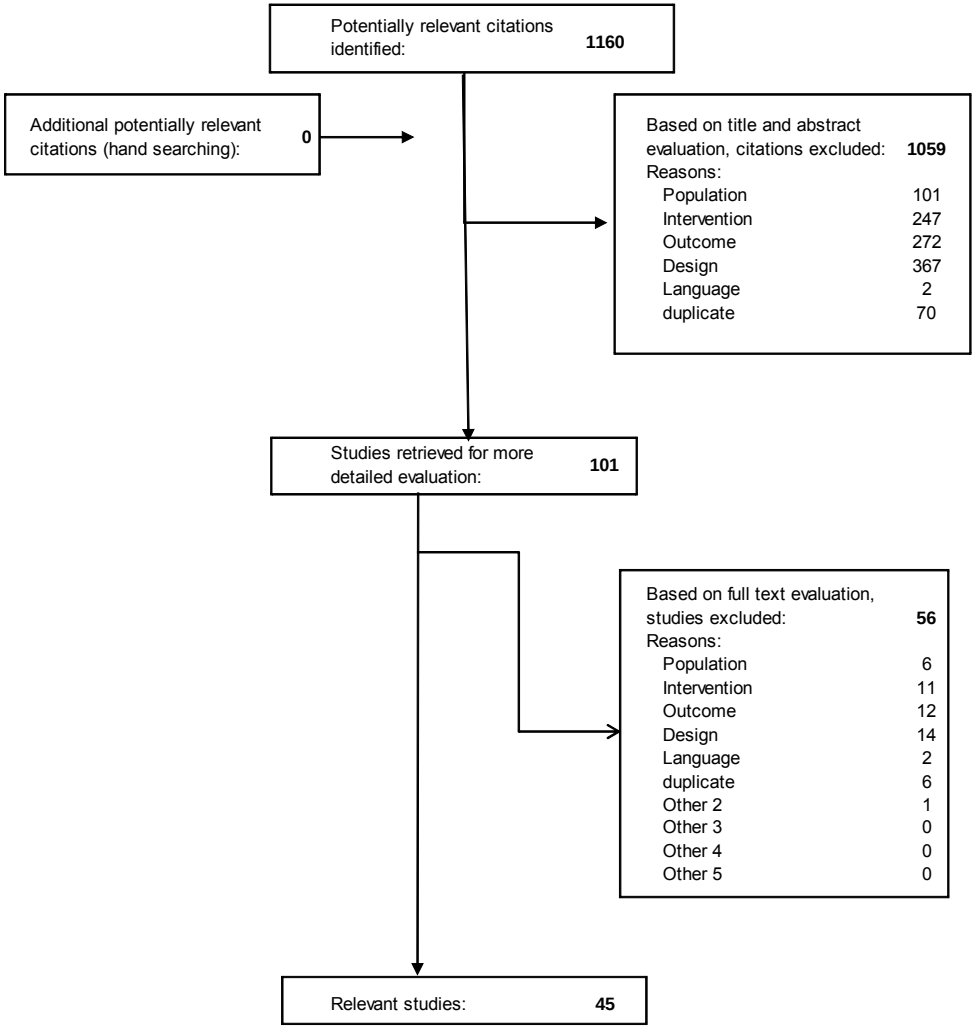


Appendix 2.3.2. Critical appraisal results for SR, MA, HTA and guidelines

Critical appraisal is also considered in evidence tables

Author + year	search question	search strategy	selection procedure	quality assessment	data-extraction	characteristics studies	meta-analysis	valid and applicable
Bermejo-Perez, 2008	±	Y	Y	Y	Y	Y	NA	Y
Bywood, 2004	Y	Y	N	±	N	Y	NA	±
Davidson, 2007	Y	Y	Y	Y	Y	Y	NA	Y
Dinnes, 2001	Y	Y	±	±	±	Y	NA	±
Dunfield, 2007	Y	Y	N	N	N	Y	NA	±
Granader, 2008	Y	Y	N	±	Y	Y	Y	±
Hailey, 2006	Y	Y	±	N	±	Y	Y	±
HAS, 2006	±	Y	±	N	±	Y	NA	±
Irwig, 2004	Y	Y	Y	Y	±	Y	NA	Y
Jansen-van der Weide, 2010	Y	Y	Y	Y	Y	Y	Y	Y
Lord, 2006	Y	Y	Y	Y	Y	Y	Y	Y
Medical AS, 2010	Y	Y	Y	N	±	Y	NA	±
Mundy, 2004	Y	Y	±	±	N	Y	NA	±
Nelson, 2005	Y	Y	Y	Y	Y	Y	NA	±
NICE, 2006	Y	Y	Y	N	Y	Y	NA	±
Noble, 2009	Y	Y	Y	Y	Y	Y	Y	Y
Nothacker, 2009	Y	Y	Y	Y	±	Y	NA	Y
Parella, 2005	Y	±	N	N	N	Y	NA	±
Ravert, 2009	±	Y	±	N	±	Y	NA	±
Taylor, 2008	Y	Y	±	N	Y	Y	±	±
Vinnicombe, 2009	±	±	±	N	Y	Y	Y	±
Warner, 2008	Y	Y	Y	N	Y	Y	Y	±
AETSA, 2007	Y	Y	Y	Y	Y	Y	NA	Y
Carreira, 2007: no detailed critical appraisel because of language restrictions								
CBO, 2008: no detailed information about design and methods received								

Appendix 2.3.3. Flowchart results search for primary studies 2007-2011



APPENDIX 3. EVIDENCE TABLES

Appendix 3.1. Women at risk for breast cancer

Table 28 breast cancer risk assessment

Reference	Search date	Recommendations/conclusions	Evidence base	Level of evidence
Nice 2004/2006 [9]	2006	Average risk Women can be cared for in primary care if the family history shows only one first-degree or second-degree relative diagnosed with breast cancer at older than age 40 years, provided that none of the following are present in the family history: • bilateral breast cancer • male breast cancer • ovarian cancer • Jewish ancestry • sarcoma in a relative younger than age 45 years • glioma or childhood adrenal cortical carcinomas • complicated patterns of multiple cancers at a young age • paternal history of breast cancer (two or more relatives on the father's side of the family).	Meta-analysis of observational studies/ case control study.	
		Raised risk (that is, a 10-year risk of 3–8% for women aged 40–49 or a lifetime risk of 17% or greater but less than 30%) Women who meet the following criteria should be offered secondary care and do not require referral to tertiary care: • one first-degree relative diagnosed with breast cancer at younger than age 40 years, or • two first-degree or second-degree relatives diagnosed with breast cancer at an average age of older than 50 years, or • three first-degree or second-degree relatives diagnosed with breast cancer at an average age of older than 60 years, or • a formal risk assessment (usually carried out in tertiary care) or a family history pattern is likely to give a 10-year risk of 3–8% for women aged 40–		

	49 years5, or a lifetime risk of 17% or greater but less than 30% provided that none of the following are present in the family history: • bilateral breast cancer • male breast cancer • ovarian cancer • Jewish ancestry • sarcoma in a relative younger than 45 years of age • glioma or childhood adrenal cortical carcinomas • complicated patterns of multiple cancers at a young age • very strong paternal history (four relatives diagnosed at younger than 60 years of age on the father's side of the family).
	High risk (that is, a 10-year risk at age 40–49 years of greater than 8% or a lifetime risk of 30% or greater, or a 20% or greater chance of a faulty BRCA1, BRCA2 or TP53 gene in the family) At least the following female breast cancers only in the family: − two first-degree or second-degree relatives diagnosed with breast cancer at younger than an average age of 50 years (at least one must be a first-degree relative), or − three first-degree or second-degree relatives diagnosed with breast cancer at younger than an average age of 60 years (at least one must be a first-degree relative), or − four relatives diagnosed with breast cancer at any age (at least one must be a first-degree relative). or • Families containing one relative with ovarian cancer at any age and, on the same side of the family: − one first-degree relative (including the relative with ovarian cancer) or second-degree relative diagnosed with breast cancer at younger than age 50 years, or − two first-degree or second-degree relatives diagnosed with breast cancer at younger than an average age of 60 years, or − another ovarian cancer at any age.

	or • Families containing bilateral cancer (each breast cancer has the same count value as one relative): − one first-degree relative with cancer diagnosed in both breasts at younger than an average age of 50 years, or − one first-degree or second-degree relative diagnosed with bilateral breast cancer and one first-degree or second-degree relative diagnosed with breast cancer at younger than an average age of 60 years. or • Families containing male breast cancer at any age and on the same side of the family, at least: − one first-degree or second-degree relative diagnosed with breast cancer at younger than age 50 years, or − two first-degree or second-degree relatives diagnosed with breast cancer at younger than an average age of 60 years. or • A formal risk assessment has given risk estimates of: − a 20% or greater chance of a BRCA1, BRCA2 or TP53 mutation being harboured in the family, or − a greater than 8% chance of developing breast cancer age 40–49 years, or − a 30% or greater lifetime risk of developing breast cancer.	
	All women satisfying referral criteria to secondary or specialist care (at raised risk or greater) should be offered mammographic surveillance from age 40 years.	
	Women who have been referred to a clinical genetics centre who are not known to have a genetic mutation should be offered an assessment of their 10-year breast cancer risk using a validated risk assessment tool (for example, Tyrer-Cuzick or BOADICEA6,7) to assess whether they are or will be eligible for MRI.	
	Women who are known to have a genetic mutation should be offered annual MRI surveillance if they are:	

	• BRCA1 and BRCA2 mutation carriers aged 30–49 years • TP53 mutation carriers aged 20 years or older. MRI surveillance should be offered annually when indicated. From 30–39 years: • to women at a 10-year risk of greater than 8%8 From 40–49 years: • to women at a 10-year risk of greater than 20%, or • to women at a 10-year risk of greater than 12% where mammography has shown a dense breast pattern9. 1.4.4.13 New Women who have not been tested but have a high chance of carrying a BRCA1 or TP53 genetic mutation should be offered annual MRI surveillance from 30–49 years if they are at: • a 50% risk of carrying one of these mutations in a tested family, or • a 50% risk of carrying a BRCA1 or TP53 mutation in an untested or inconclusively tested family with at least a 60% chance of carrying a BRCA1 or TP53 mutation (that is, a 30% risk of carrying one of these mutations themselves).	
	Computerised risk-assessment models can be helpful aids to risk assessment, but can be misleading and should not yet totally replace careful clinical assessment of family trees with a manual approach. (D)	
	1. Existing computer models (Gail, Claus, BRCAPRO) underestimate in a family history setting in terms of breast cancer risk prediction, although the manual Claus tables produce risks close to those seen in a screened familial risk population. (III) 2. One US study found that BRCAPRO predicted BRCA 1 & 2 mutation status better than genetic counsellors. (III) 3. The degree of correlation between different risk models is relatively poor. (III)	

I Study ID	II Method	III Patient characteristics	IV Intervention(s)	V Results primary outcome	VI Results secondary and other outcomes	VII Critical appraisal of study quality
Tice 2005 [20]	Design validation on cohort study Source of fundingThis work was supported in part by a NCI-funded Breast Cancer Surveillance Consortium co-operative agreement Sample size 81.777 Duration 5 years	Eligibility criteria: women age 35 years and older who had a reading of mammographic density associated with at least one of their mammograms taken prior to January 1, 2002. Prevalence of disease: 955 women were diagnosed with invasive breast cancer	Index test(s) Gail score BI-RADS breast density	Gail model: predictive accuracy (concordance index (c-index) 0.67; 95% CI 0.65–0.68) Gail model + breast density: 0.68 (95% CI .66–.70, p < 0.01 compared with the Gail model alone) Breast density alone: (c-index 0.67, 95% CI 0.65–0.68)		Only predictive accuracy reported, calibration not reported Case ascertainment with SEER Validation on a high quality cohort, US population may have different chararcteristics than the Belgian population.

I Study ID	II Method	III Patient characteristics	IV Intervention(s)	V Results primary outcome	VI Results secondary and other outcomes	VII Critical appraisal of study quality
Evans 2006	Validation on a Family History Evaluation and Screening Programme in Manchester, UK, . Sample size: 1,933 women a mean follow-up of 5.27	Eligibility criteria: women attending the above mentioned screening program Prevalence of disease: of which 52 developed cancer	Index test(s) Gail, Claus, BRUCAPRO IBIS(Cuzick-Tyrer)	Calibration The ratios of expected to observed numbers of breast cancers (95% confidence interval) were 0.48 (0.37–0.64) for the Gail model, 0.56 (0.43–0.75) for the Claus model, 0.49 (0.37–0.65) for the		Fairly small cohort with sufficient follow up, population may be closer to the Belgian population compared with the US data

				BRCAPRO model and 0.81 (0.62–1.08) for the Cuzick–Tyrer model (Accuracy AUC was 0.735 for the Gail model, 0.716 for the Claus model, 0.737 for the BRCAPRO model and 0.762 for the Cuzick–Tyrer model.		

I Study ID	II Method	III Patient characteristics	IV Intervention(s)	V Results primary outcome	VI Results secondary and other outcomes	VII Critical appraisal of study quality
Tice 2008 [22]	Design model development and validation on National Cancer Institute–funded Breast Cancer Surveillance Consortium (BCSC) Sample size 1 095 484 women Duration 5 years	Eligibility criteria: age 35 years or older who had had at least 1 mammogram with breast density measured by using the Breast Imaging Reporting and Data System (BI-RADS) classification system Prevalence of disease: 14 766 women diagnosed with invasive breast cancer	Index test(s) Tice score	The breast density model was well calibrated overall (expected–observed ratio, 1.03 [95% CI, 0.99 to 1.06]) It had modest discriminatory accuracy (concordance index, 0.66 [CI, 0.65 to 0.67]).		Case ascertainment with SEER Model development and validation on same database (but different samples in the same database, needs independent validation. population may have different chararcteristics than the Belgian population.

I Study ID	II Method	III Patient characteristics	IV Intervention(s)	V Results primary outcome	VI Results secondary and other outcomes	VII Critical appraisal of study quality
Decarli 2006 [24]	Design validation on cohort study Source of supported by contributions from the Associazione Italiana per la Ricerca sul Cancro and the Italian Ministry of Education Setting: Florence Italy Sample size N = 10 031, Duration 1993 – 2002	Eligibility criteria: women aged 35 – 64 years who resided in the Italian provinces of Florence and Prato, which are covered by the Florence Cancer RegistryPrevalence of disease: 194 women were diagnosed with invasive breast cancer	Index test(s) Score from Gail Model (GM) Score from Gail Model modified based on Italian Case control study (IT-GM) and (IT1-GM)	Calibration The overall E/O ratios were 0.96 (95% confi dence interval [CI] = 0.84 to 1.11) and 0.93 (95% CI = 0.81 to 1.08) for the IT-GM and the GM, respectively. The average age-specific concordance statistics: 58.6% (95% CI = 54.4% to 62.8%) for the IT-GM, 59.0% (95%CI = 54.8% to 63.2%) for the IT1-GM, and 58.8% (95% CI = 54.6% to 63.1%) for the GM		Validation on a high quality cohort, Italian population may have different chararcteristics than the Belgian population. Alternative model developed based on Italian case control study.

I Study ID	II Method	III Patient characteristics	IV Intervention(s)	V Results primary outcome	VI Results secondary and other outcomes	VII Critical appraisal of study quality
Barlow 2006 [23]	Design development and validation on cohort Source of funding: from the NCI through a BCSC cooperative Agreement Setting Seven mammography registries of the BCSC Sample size N = 1 000 000 Duration 5 year follow up	Eligibility criteria: women aged 35 – 84 years were included. Women with previous breast cancer were excluded. Women with breast augmentation were also excluded 11 638 women were diagnosed with breast cancer	'Barlow model' Developed using logistic regression	risk factors among premenopausal women : age, breast density, family history of breast cancer, and a prior breast procedure. For postmenopausal women: age, breast density, race, ethnicity, family history of breast cancer, a prior breast procedure, body mass index, natural menopause, hormone therapy, and a prior false-positive mammogram. . The c statistics were 0.631 (95% confi dence interval [CI] = 0.618 to 0.644) for premenopausal women and 0.624 (95% CI = 0.619 to 0.630) for postmenopausal women.		Validation and development on different samples of the same cohort, US population may have different characteristics than the Belgian population.

I Study ID	II Method	III Patient characteristics	IV Intervention(s)	V Results primary outcome	VI Results secondary and other outcomes	VII Critical appraisal of study quality
Chlebowski 2007 [25]	Design validation on cohort study funded by the National Heart, Lung, and Blood Institute, National Institutes of Health, Department of Health and Human Services. Setting 40 clinical centers in the United States Sample size N = 147 916, Duration 1993 – 2002	Eligibility criteria: Postmenopausal women, who were aged 50 – 79 years and unlikely to move or die within 3 years, were eligible Prevalence of disease: 3236 women were diagnosed with invasive breast cancer	Index test(s) Score from Gail Model (GM) Gail model also evaluated in for the prediction of both estrogen receptor [ER] – positive and ER-negative disease	Calibration The Gail model underestimated 5-year invasive breast cancer incidence by approximately 20% (P <.001), mostly among those with a low estimated risk. Accuracy AUC for the Gail model was 0.58 (95% confidence interval [CI] = 0.56 to 0.60). Discriminatory performance was better for the risk of ER-positive cancer (AUC = 0.60, 95% CI = 0.58 to 0.62) than for the risk of ER-negative cancer (AUC = 0.50, 95% CI = 0.45 to 0.54).		Validation on a high quality cohort, Case ascertainment with SEER US population may have different chararcteristics than the Belgian population. Clinical value of ER and non ER estimation not clear.

I Study ID	II Method	III Patient characteristics	IV Intervention(s)	V Results primary outcome	VI Results secondary and other outcomes	VII Critical appraisal of study quality
Chlebowski 2007 [25] **Crispo 2008** [28]	Design validation on cohortcohort study Not mentioned Setting 40 clinicalclinical Sample size Duration 1993 – 20022002	Eligibility criteria: Cases womenwho had invasive breast cancer	Index test(s) Score from Gail Model (GM) Gail model also evaluated in for the prediction of both estrogen receptor [ER] – positive and ER-negative disease	The concordanceAUC for the model was 0.5558 (95% CI 0.53–0.60). the model with SDR (0.5660, 95% CI 0.53–0.62) than forfor the riskrisk of 0.57 (95% CI 0.54).		Validation on a cas control, ItalianUS population may have different chararcteristics than the Belgian population.

I Study ID	II Method	III Patient characteristics	IV Intervention(s)	V Results primary outcome	VI Results secondary and other outcomes	VII Critical appraisal of study quality
Shonfeld 2010 [26]	Design validation on cohort study Supported by the Intramural Research Program of the National Institutes of Health and the National Cancer Institute. Setting United States Sample size	Eligibility criteria: Cohort 1 NIH-AARP, age 50 to 71 years Cohort 2 PLCO age 55 to 71 years Prevalence of disease: Cohort 1 NIH-AARP, 5,665women were diagnosed with invasive breast cancer	Index test(s) Score from Gail Model (GM) Score from Calibrated Gail, calibrated with 1995 to 2003 SEER invasive breast cancer incidence rates.	the Gail model significantly underpredicted the number of invasive breast cancers in NIH-AARP, with an expected-to-observed ratio of 0.87 (95% CI, 0.85 to 0.89), and in PLCO, expected-to-observed ratio of 0.86 (95% CI, 0.82 to 0.90). The updated model expected-to-observed ratio of 1.03 (95% CI, 1.00 to 1.05) in NIH-		Validation on 2 high quality cohort, Case ascertainment with SEER US population may have different characteristics than the Belgian population.

	Cohort 1 NIH-AARP, N = 200 000, Duration 1993 – 2002 Cohort 2 (PLCO) : 77 5000	Cohort 2 PLCO 2,223 women were diagnosed with invasive breast cancer		AARP and an expected-to-observed ratio of 1.01 (95% CI: 0.97 to 1.06) in PLCO.		

I Study ID	II Method	III Patient characteristics	IV Intervention(s)	V Results primary outcome	VI Results secondary and other outcomes	VII Critical appraisal of study quality
Crispo 2008 [28] Shonfeld 2010 [26]	Design validation on cohortcontrol study Not mentioned Setting Sample size Cases: 588 Controls1207	Eligibility criteria: CasesCases women, from the Breast Unit of the National Cancer Institute of Naples, who had invasive breast canceror Cohort 2 PLCO 2,223 women were diagnosed with invasivefor breast cancer	Index test(s) Score from Gail Model (GM) Score from Calibrated Gail, calibrated with 1995 to 2003 SEER invasive breast cancer incidence rates. SDR	the model with FDR was 0.5555 (95% CI 0.53–0.58),), the model with SDR (0.56, 95% CI 0.53–0.59), combination of FDR+SDR gave the concordance statistic of 0.57 (95% CI 0.54–0.60)		Validation on a cas control, ItalianItalian population may have different characteristicscharar cteristics than the Belgian population.

I Study ID	II Method	III Patient characteristics	IV Intervention(s)	V Results primary outcome	VI Results secondary and other outcomes	VII Critical appraisal of study quality
Vacec 2011 [27]	Design validation on cohort study funded by grant KG090134 from Susan G. Komen for the Cure United States (Vermont) Sample size Cohort 19,779 Average follow up time 7 years	Eligibility criteria: women aged 70 and older from Vermont (USA) Prevalence of disease: , 821 women were diagnosed with invasive breast cancer	Index test(s) Gail model, the Tice modification of the Gail model, the Barlow model, and the Vermont model	C-statistics were 0.54 (95% CI = 0.52–0.56) for the Gail model, 0.54 (95% CI = 0.51–0.56) for the Tice modification of the Gail model, 0.55 (95% CI = 0.53–0.58) for a model developed by Barlow and 0.55 (95% CI = 0.53–0.58) for a Vermont model. These results indicate that the models are not useful for assessing risk in women aged 70 and older.		Validation on a high quality cohort, US population may have different characteristics than the Belgian population.

Table 29 Attempts to improve models with genetic data

I Study ID	II Method	III Patient characteristics	IV Intervention(s)	V Results primary outcome	VI Results secondary and other outcomes	VII Critical appraisal of study quality
Wacholder 2010 [29]	Design validation on 5 studies done for other purposes (4 RCT and one Case control) Supported in part by the Intramural Research Program of the Division of Cancer Epidemiology and Genetics of the National Cancer Institute and by grants from the National Institutes of Health Setting: United States & poland) Sample size 5590 case subjects and 5998 control subjects	Eligibility criteria: Participants of Women's Health Initiative Observational Study,9 the American Cancer Society Cancer Prevention Study II Nutrition Cohort, the Prostate, Lung, Colorectal, and Ovarian Cancer Screening Trial and the Nurses' Health Study	Index test(s) Gail model, Gail model modified using 10 common genetic variants associated with breast cancer	AUC for a risk model with age, study and entry year, and four traditional risk factors was 58.0%; with the addition of 10 genetic variants, the AUC was 61.8%		Validation on 5 rather heterogeneous studies, however conclusion that adding common genetic variants only modestly improves the model remains robust.

I Study ID	II Method	III Patient characteristics	IV Intervention(s)	V Results primary outcome	VI Results secondary and other outcomes	VII Critical appraisal of study quality
Mealliffe 2010	Design validation on nested case–control study from the Women's Health Initiative (WHI) Clinical Trial Funding National Heart Lung and Blood Institute National Cancer Institute at the National Institutes of Health, Setting: United States Sample size 1664 case patients and 1636 control subjects	Eligibility criteria: White non Hispanic women, Participants of Women's Health Initiative Observational Study	Index test(s) Gail model, Gail risk single-nucleotide polymorphisms (SNP) risk and cobined.	Combined risk score was more discriminating, with area under the curve of 0.594 compared with area under the curve of 0.557 for Gail risk alone (P < .001). Classification also improved for 5.6% of case patients and 2.9% of control subjects, showing an NRI value of 0.085 (P = 1.0 × 1025). Focusing on women with intermediate Gail risk resulted in an improved NRI of 0.195 (P = 8.6 × 1025)		Validation on a case control studies, however conclusion that adding common genetic variants only modestly improves the model remains robust.

I Study ID	II Method	III Patient characteristics	IV Intervention(s)	V Results primary outcome	VI Results secondary and other outcomes	VII Critical appraisal of study quality
Pankratz 2008	Design validation on Mayo Benign Breast Disease cohort Funding Supported by DOD Center of Excellence Grant Susan G. Komen Breast Cancer Foundation Setting: United States Sample size 9,376 subjects in cohort, of whom 331 with atypias median follow-up of 14.6 years	Eligibility criteria: Women presenting with benign breast disease	Index test(s) Gail Model	58 of 331 (17.5%) patients had developed invasive breast cancer, 1.66 times more than the 34.9 predicted by the Gail model (95% CI, 1.29 to 2.15; P _ .001). For individual women, the concordance between predicted and observed outcomes was low, with a concordance statistic of 0.50 (95% CI, 0.44 to 0.55).		Specific subgroup

I Study ID	II Method	III Patient characteristics	IV Intervention(s)	V Results primary outcome	VI Results secondary and other outcomes	VII Critical appraisal of study quality
Boughey 2010 [32]	Design validation on Mayo Benign Breast Disease cohort Funding Supported by Mayo Clinic Breast Cancer Specialized Program of Research Excellence Setting: United States Sample size 9,376 subjects in cohort, of whom 331 with atypias median follow-up of 14.6 years	Eligibility criteria: Women presenting with benign breast disease	Index test(s) Tyrer-Cuzick (International Breast Cancer Intervention Study) Model	The observed-to-predicted ratio was 0.53 (95% CI, 0.37 to 0.75). Concordance statistic was 0.540, revealing that the Tyrer-Cuzick model did not accurately distinguish, on an individual level, between women who developed invasive breast cancer and those who did not		Specific subgroup

gene mutation prediction models

I Study ID	II Method	III Patient characteristics	IV Intervention(s)	V Results primary outcome	VI Results secondary and other outcomes	VII Critical appraisal of study quality
Kang 2006 [33]	Design cross sectional validation of risk prediction in genetic centre funded by the Kathleen Cuningham Foundation, National Breast Cancer Foundation, National Health and Medical Research Council (NHMRC) Setting Family cancer clinics at St Vincent's and Westmead Hospitals, Sydney Sample size 380 families	Eligibility criteria: high risk participants in genetic clinics in sydney	Index test(s) BRCAPRO, Manchester, Penn and the Myriad-Frank	All 7 models showed similar AUC : Manchester 0.759 0.688 0.831 BRCAPRO 0.743 0.672 0.814 Myriad 0.753 0.680 0.827 Penn 0.757 0.686 0.827 all models have high false-negative and false-positive rates using 10 % probability thresholds used to refer for mutation testing		Results only valid amongst participants in genetic clinic.

I Study ID	II Method	III Patient characteristics	IV Intervention(s)	V Results primary outcome	VI Results secondary and other outcomes	VII Critical appraisal of study quality
Ruodgari 2007 [34]	Design validation of risk prediction in genetic clinics Funding NIH Setting USA (mayo genetic clinic Sample size 200 families	Eligibility criteria: 275 Scottish families tested for BRCA1/2 mutations in genetic clinics in	Index test(s) Four probability estimation models including COS, Manchester scoring system (MSS), BOADICEA and Tyrer–Cuzick (T–C)	COS and MSS models demonstrated the greatest sensitivities and area under ROC curves for the majority offamily structures. They also showed the highest sensitivities (91–92%) and AUCs (76–78%) for the entire dataset overall. However, BOADICEA and T–C had the highest specificities for the majority of the family structures. BOADICEA and T–C generated the best estimates for the prevalence of mutations in the population		Results only valid amongst participants in genetic clinic.

I Study ID	II Method	III Patient characteristics	IV Intervention(s)	V Results primary outcome	VI Results secondary and other outcomes	VII Critical appraisal of study quality
Antinou 2008 [35, 36]	Design validation of risk prediction in genetic clinics centers. Funding NCI Cancer Genetics Network + divers Setting USA (Sample size 2140 families	Eligibility criteria: 1934 families tested for BRCA1/2 mutations.	Index test(s) BRCAPRO, IBIS, the Manchester scoring system and Myriad tables,	All models showed similar AUC : BRCAPRO=0.76, IBIS=0.74)Myriad=0.72) .		Results only valid amongst participants in genetic clinic.

I Study ID	II Method	III Patient characteristics	IV Intervention(s)	V Results primary outcome	VI Results secondary and other outcomes	VII Critical appraisal of study quality
Antinou 2008 [36]	Design validation of risk prediction in genetic clinics supported by a grant from the UK Department of Health. Setting 6 UK genetic clinics	Eligibility criteria: 1934 families seen in cancer genetics clinics in the UK in whom an index patient had been screened for BRCA1 and/or BRCA2 mutations.	Index test(s) carrier prediction algorithms BOADICEA, BRCAPRO, IBIS, the Manchester scoring system and Myriad tables,	calibration Only BOADICEA well calibrated (only for BOADICEA no statistically significant difference E/O. All models underestimate probability in low risk		Results only valid amongst participants in genetic clinic.

	Sample size 2140 families			population Accuracy: receiver operating characteristic curve statistics: BOADICEA=0.77, BRCAPRO=0.76, IBIS=0.74, Manchester=0.75, Myriad=0.72)		

I Study ID	II Method	III Patient characteristics	IV Intervention(s)	V Results primary outcome	VI Results secondary and other outcomes	VII Critical appraisal of study quality
Panchal 2008 37	Design cross sectional validation of risk prediction in genetic centre No source of funding was used for this study Setting Family cancer Toronto, Canada 100 carriers and 200 non-carriers	Eligibility criteria: high risk participants in genetic clinics in Canada	Index test(s) BRCAPRO, Manchester, Penn II, Myriad II, FHAT, IBIS and BOADICEA models	BRCAPRO, Penn II, Myriad II, FHAT and BOADICEA models all have similar AUCs of approximately 0.75 for BRCA status. The Manchester and IBIS models have lower AUCs (0. and 0.47 respectively). At a 10 % testing threshold, the sensitivities and specificities for a BRCA mutation were, respectively, as		Results only valid amongst participants in genetic clinic.

				follows: BRCAPRO (0.75, 0.62), Manchester (0.58,0.71), Penn II (0.93,0.31), Myriad II (0.71,0.63), FHAT (0.70,0.63), IBIS (0.20,0.74), BOADICEA (0.70, 0.65)		

I Study ID	II Method	III Patient characteristics	IV Intervention(s)	V Results primary outcome	VI Results secondary and other outcomes	VII Critical appraisal of study quality
Lindor 2010 [8]	Design validation of risk prediction in genetic clinics Funding not mentioned Setting USA (mayo genetic clinic Sample size 200 families	Eligibility criteria: 200 families seen in Mayocancer genetics clinics in whom an index patient had been screened for BRCA1 and/or BRCA2 mutations.	Index test(s) LAMBDA, ; BRCAPRO, a ; modified Couch tables Myriad II tables	All models gave similar areas under the ROC curve of 0.71 to 0.76. All models except LAMBDA substantially under-predicted the numbers of carriers. All models were too dispersed		Results only valid amongst participants in genetic clinic.

I Study ID	II Method	III Patient characteristics	IV Intervention(s)	V Results primary outcome	VII Critical appraisal of review quality
NZHTA 2007[11]NZHTA 2007[11]	1. design systematic review 2. Government of New Zealand 3. Search date nov 2005 4. Searched databases Medline and Embase databases, the Cochrane Database of Systematic Reviews, the DARE and HTA databases Included study designs 1. Systematic reviews. 2. Clinical studies with a control group for the time period beyond key systematic reviews 5. Number of included studies 139	Women who were assessed for breast cancer	1. risk factors for breast cancer	1. Effect size primary outcome(s) • past history of breast cancer (RR between 2.8 and 7.4) • ductal hyperplasia RR 1.5 - 2 • Atypical ductal carcinoma RR 4 • lobular carcinoma RR 6-10 • ductal carcinoma *in situ* RR 8 - 10 • increased breast density RR 4 • alcohol intake (10% for 10g alcohol/day, 25% for 25g alcohol/day and 55% for 50g alcohol/day) • nulliparity relative risk estimates decrease by approximately 0.09 for each additional birth • post menopausal obesity RR 1.12 • for the association with the overweight category and 1.25 for the obese category • hormone replacement therapy RR 1.2-1.4 • current or recent use of oral contraceptives 1. current users: RR 1.24 (95% CI 1.15-1.33) 2. 1-4 years after stopping: RR 1.16 (95% CI 1.08-1.23) 3. 5-9 years after stopping: RR 1.07 (95% CI 1.02-1.13) 4. >10 years after stopping: RR 1.01 (95% CI 0.96-1.05). • high total energy intake 1.4-2.1 For some risk factors the level of increased risk was	High quality review The majority of studies included in the review used the case-control design. Case control studies are characterized by susceptibility to selection bias and recall bias.

				difficult to determine: • early menarche (likely to be relatively modest) • xenoestrogens • phytoestrogens • stilboestrol.		

I Study ID	II Method	III Patient characteristics	IV Intervention(s)	V Results primary outcome	VI Results secondary and other outcomes	VII Critical appraisal of review quality
Vrieling 2010 [14]	Design systematic review of observational studies funded by the Deutsche Krebshilfe Search date March 2010	Adult women	Index test(s) BMI or other measure of wheight Comparing the highest versus the lowest categories of adult weight gain ER: estrogen receptor status PR progesterone receptor status	risk for ER+PR+ and ER+ tumors combined (11 studies; RE = 2.03; 95% CI 1.62, 2.45). Statistically significant heterogeneity (p heterogeneity = 0.002) was shown between REs for a mixed population of pre- and postmenopausal women combined (4 studies; RE = 1.54; 95% CI 0.86, 2.22) and for postmenopausal women only (7 studies; RE = 2.33; 95% CI 2.05, 2.60). .	Risk for ER-PR- tumors among postmenopausal women (7 studies; RE = 1.34; 95% CI 1.06, 1.63), but statistically significantly different from risk for ER+PR+ tumors(p for heterogeneity\0.0001). No associations were observed for ER+PR- tumors whereas risk for ER-PR+ tumors could not be assessed.	Clinical implications of receptor status of the tumors unclear

I Study ID	II Method	III Patient characteristics	IV Intervention(s)	V Results primary outcome	VII Critical appraisal of review quality
Cummings 2009[7] **Cummings 2009**[7] **Update of McCormack 2006**[15]	Search date 2008 Design systematic review of observational studies funded by the National Cancer Institute and the Daniel and Phyllis Da Costa Fund Meta-analysis 47 prospective studies.	Adult women	Index test(s) 3 different measures of breast density: Wolfe grade BI-RADS % Of breast area that is dense	Wolfe grade N1 (fatty) 1 (reference) P1 1.76 (1.41 to 2.19) P2 3.05 (2.54 to 3.66) Dy (most dense) 3.98 (2.53 to 6.27) BI-RADS 1 (fatty) 1 (reference) 2 (scattered densities) 2.03 (1.61 to 2.56) 3 (heterogeneously dense) 2.95 (2.32 to 3.73) 4 (extremely dense) 4.03 (3.10 to 5.26) % Of breast area that is dense <5 1 (reference) 5 – 24 1.74 (1.50 to 2.03) 25 – 49 2.15 (1.87 to 2.48) 50 – 74 2.92 (2.55 to 3.34) ≥ 75 4.20 (3.61 to 4.89)	The meta-analysis by McCormack et al. was included in this meta-analysis. All studies were adjusted for age; studies that further adjust for body mass index or weight observed somewhat stronger associations

I Study ID	II Method	III Patient characteristics	IV Intervention(s)	V Results primary outcome	VI Results secondary and other outcomes	VII Critical appraisal of review quality
Key et al, 2006 [128]	Search date 2005 Design systematic review & meta-analysis of observational studies Study funded by the Department of Health in England 98 studies were included, involving 75,728 and 60,653 cases in drinker versus nondrinker and dose–response analyses, respectively.	Adult women	Index test(s) Alcohol use	excess risk associated with alcohol drinking was 22% (95% CI: 9–37%); each additional 10 g ethanol/day was associated with risk higher by 10% (95% CI: 5–15%). There was no evidence of publication bias. Risk did not differ significantly by beverage type or menopausal status.	Estimated population attributable risks were 1.6 and 6.0% in USA and UK, respectively	Considerable heterogeneity in effects measures, meta- regression was used but did not help to explain the heterogeneity.

I Study ID	II Method	III Patient characteristics	IV Intervention(s)	V Results primary outcome	VI Results secondary and other outcomes	VII Critical appraisal of review quality
Kahlenborn et al, 2006 [16]	Search date 2006 Systematic review and meta-analysis Supported by National Institutes of Health (US) 34 eligible studies	Adult women, nulliparous, parous or multiparous.	Index test(s) Oral contraceptive use (OC)	Use of OCs was associated with an increased risk of premenopausal breast cancer in general (OR, 1.19; 95% CI, 1.09-1.29) and across various patterns of OC use. Among studies that provided data on nulliparous and parous women separately, OC use was associated with breast cancer risk in both parous (OR,1.29; 95% CI, 1.20-1.40) and nulliparous (OR, 1.24; 95% CI, 0.92-1.67) women.	Longer duration of use did not substantially alter risk in nulliparous women (OR, 1.29; 95% CI, 0.85-1.96). Among parous women, the association was stronger when OCs were used before first full-term pregnancy (FFTP) (OR, 1.44; 95% CI, 1.28-1.62) than after FFTP (OR, 1.15; 95% CI, 1.06-1.26). The association between OC use and breast cancer risk was greatest for parous women who used OCs 4 or more years before FFTP (OR,1.52; 95% CI, 1.26-1.82)	Only case control studies in meta-analysis DerSimonian-Laird random effects model used but no other measure or exploration of heterogeneity.

Appendix 3.2. Technical methods for breast cancer screening

Appendix 3.2.1. Double reading and computer-aided detection Mammography

Systematic reviews

Table 30 Double reading and computer-aided detection mammography: systematic reviews

Reference	Methodology	Patient characteristics	Intervention(s)	Results primary outcome	Results secondary and other outcomes	Critical appraisal of review quality
Dinnes et al, 2001[39]	• SR • Funding: UK Department of Health R&D Division • Search date: between April 1991 and July 1999 • Databases: Medline, CINAHL, DHSS, BIOSIS, Embase, BIDS, CancerLit, NHS EED, CCTR, Dissertation abstracts, PASCAL, Conference Papers Index, SIGLE, Health Star, EconLit • Study design: prospective and retrospective cohort studies • N included studie: 10 cohort studies	Eligibility criteria Asymptomatic women undergoing mammography for routine breast cancer screening Patient characteristics: - Age range: 50-70y	Single (SR) versus double reading (DR)	Recall rate - DR with unilateral recall: increase (between 38 and 149 per 10 000 women screened) - DR with consensus or arbitration: decrease (between 61 and 269 per 10 000 women screened) - DR either unilateral of consensus: overall increase Cancer detection rate - DR: increase (range +2.9 to +11.2 per 10 000 women screened)	Interval cancers: increased with longer follow-up Higher proportions of small and early stage cancers for DR Number of mammographic views: - Single-view: increase in detection (4.4 to 6.9 per 10 000) - Two-view: increase in detection (3.0 to 4.4 per 10 000)	• Impact of experience of reader unknown • Data insufficient to quantify difference between SR and DR • DR can cause a delay in delivery of screening results • Full report (mentioned in article) not found • Quality assessment of studies not defined • Selection criteria not explained

				Sensitivity: increase in DR Specificity: - decrease with unilateral recall - increase with consensus of mixed recall "DR with consensus reduces recall rates and increases specificity, whereas unilateral recall increases recall rate"		
Taylor et al, 2008[40]	• SR • Funding: partly by NHS Breast Screening Programme • Search date: until 2007 • Databases: Google Scholar, Biotech, CINAHL, Embase, HMIC, Psychinfo, Web of Science, Science Direct, British Library, recent proceedings of relevant conferences, previous systematic	Eligibility criteria: Asymptomatic women undergoing mammography for routine breast cancer screening Patient characteristics	Computer aids vs human second reading (single reading (SR), double reading (DR))	Cancer detection rate - CAD: no sign increase and no pooled effect (odds ratio of 1.04, 95% CI: 0.96-1.13) - DR: - individually effects not sign, but pooled estimate sign (95% CI 1.06-1.14; χ^2 (1)=23.5, p<0.001)	Number needed to treat for DR with arbitration: 2222 women scrrened for each additional cancer detected	• Possible drop of specificity in unmatched studies on CAD

review and its references • Study designs: prospective and retrospective studies intervention incorporated in routine screening work • N included studies: 10 studies CAD vs SR, 17 studies SR vs DR			- Arbitration/ consensus studies: odds ratio 1.08 (95% CI:1.02-1.15; χ^2 (1)=6.2, p=0.012) ➔ Extra 0.44 cancers detected per 1000 women Recall Rate - CAD: increase but strong evidence of heterogeneity, pooled estimate sign (odds ratio 1.13 (95% CI: 1.05-1.23) - DR: - heterogeneity between and within each Group - mixed and unilateral studies: increase - arbitration studies: decrease (odds ratio 0.94, 95% CI 0.92-0.96; χ^2		

				(1)=30.1, p<0.001), reduction of 2.67 per 1000 (95% CI: -1.72,-3.62; z=5.49, p<0.001) “ clear difference on recall rate, which is significantly better for double reading with arbitration than for CAD+ importance of arbitration/consensus in DR”		
Noble, 2008 (CAD)[41]	• SR • Funding: ECRI Institute (independent not-for-profit health research organization) • Search date: until 25 September 2008 • Databases: Medline, Embase, Cochrane Library, bibliographies and reference lists, gray literature • Study design: prospective and retrospective cohort studies • N included studies= 7 (392 015 women)	Eligibility criteria: Asymptomatic women undergoing mammography for routine breast cancer screening Patient characteristics: • mean and median age ranged from 49-60 years	Computer-assisted detection (CAD) Vs mammography	- pooled sensitivity 86.0% (95% CI 84.2-87.6%) (sensitivity of primary studies: 72.2%, 84.0%, 90.4%) - pooled specificity 88.2% (95%CI 88.1-88.3%) (specificity of primary studies: 87.2%, 89.7%, 92.3%)	Total recall rate: 96% (95%CI 93.9-97.3%) Incremental cancer detection rate: 50 women per 100 000 screened (95% CI 30-80) Proportion of women recalled and diagnosed with cancer: 4.1% (95% CI 2.7-6.3%) Additional recalls of healthy women: 1190 (95%CI 1090-1290)	• heterogeneity for sensitivity and specificity but quantitatively robust to sensitivity analyses • increase of recall rate and biopsy rate of healthy women • retrospective design and lack of blinding (to clinical information) limits internal validity • slow-growing cancers (false negatives) may not be detected by reference

				Additional biopsies of healthy women 80 (95%CI 60-110), 65.1% (95%CI 52.3-76.0%)	standard, clinical outcome at one year follow-up, which will overestimate CAD's sensitivity and specificity

Table 31 Double reading and computer-aided detection mammography: primary studies, update 2007-2011

Reference	Methodology	Patient characteristics	Intervention(s)	Results primary outcome	Results secondary and other outcomes	Critical appraisal of review quality
Hofvind et al, 2009[48]	• retrospective cohort study • Part of Norwegian breast Cancer Screening Program	- All Norwegian women, aged 50-69years - Two-view mammography (24-month interval) - 1 033 870 screenings, 5978 cancers (5.4 cancers per 1000) - 1791 interval cancers (1.7 per 1000) - Five point scale for probability of cancer - Discordant: reader 1: score 1+ reader 2: score 2 or higher	discordant findings vs concordant findings in double reading (DR) and use of consensus or arbitration	Score 1: 92.6% by both readers Discordant in 5.3% (54 447/1 033 870) Concordant positive: 2.1% (21 928/ 1 033 870) At consensus: 66.8% of discordant and 17.9% of concordant dismissed Rate of agreement of detected cancers: 41.3% Microcalcifications: higher in disc (24.9% vs 17.7%, p<.001) Mass or density with	- Recall rate: 3.5% ➔ No diff between disc (1.75%) and conc (1.74%) (p=.71) - Use of SFM 97%, FFDM 3% Discordant cancers: - 23.6% (1326/5611) cancers detected - 24.6% for age 50-54y - 21.7% for age 65-69y ➔ Only sign higher proportion of disc cancers in incident screening vs prevalent screenings	• Interobserver variability in mammography screening • Interpretation of microcalcifications may require additional skill building

		- Concordant: both readers score 2 or higher		microcalcifications: less common in disc (11.1% vs 15.4%, p<.001) DCIS: higher in disc (23.9% vs 15.7%, p<.001) Lobular cancers: lower in conc (7.3% vs 9.1%, p=.035) “ independent DR with consensus has the potential to increase the cancer detection rate. Microcalcifications are more common in disc findings.”	(p=0.011) Breast density: higher association of dis for extremely dense breasts than for fatty or scattered dense breast patterns (Odds ratio 1.58, 95% CI:1.24-2.00)	
Caumo et al, 2010[49]	• Retrospective cohort study •	- 7660 double readings - FFDM with delayed double reading	Role of third reader in discordant double readings	Recall rate: 6.8% - 43.5% conc - 56.5% disc After arbitration of disc: 72.4% neg, 27.6% pos→ 6 cancers Cancer detection rate: 49 cancers - conc: 43 - disc: 6 “arbitration could decrease recall rate”	Neg arbitration: - 2.8% absolute and 40.9% relative reduction of recall rate - 0.13% absolute and 2.0% relative reduction of cancer detection rate PPV: - Disc: 2.0% - Conc: 18.6%	• Analysis of recalls

Shaw et al, 2009[51]	• Retrospective cohort study • Part of Irish National Breast Screening Program	- Independent double reading of mammograms - Consensus panel when readers disagreed - 128 569 screenings performed, 1 % (1335 cases) discussed by panel - Analysis of consensus review	Consensus review of discordant findings in double reading	Recall rate of mammograms reviewed in consensus: 45.39% Overall recall rate of 4.41% Cancer detection rate: 71 cancers Sensitivity for 6-year study period: 90% Specificity for 6-year study period: 57% “ recall after discordant findings could potentially increase cancer detection rate by 0.6 per 1000 but would increase recall rate by 12.69% and number of false-positives by 15.37%”	PPV for consensus recall: 11.7% Calcifications: 32% Asymmetry: 10.5% Architectural distorsion: 9.86% Mass: 8.33% DCIS: 34% vs 18.9% in overall study group NPV: 99% Highest reader recall method (recall after 1 pos finding): - increase in referral rate of 12.69% (from 4.41% to 4.97%) - increase in false-pos (15.37%) - increase in cancer detection rate from 7.47 per 1000 to 7.53 per 1000 Unanimous recall only: - decrease in recall rate with 10.66% (to 3.94%) - decrease of false-pos by 11.39%	• Non-uniform review panel (change in membership) • Different levels of experience of readers • Consensus review: forum for discussion and educational tool

					Calcifications: - 10.04% reason for referral+disagreement - Highest PPV (32%) - Recall of all patn with disc calc: NPV increase from 98.98% to 99.66% but minimal effect on recall rate (0.05% increase)	
Duijm et al, 2009[52]	• Prospective cohort study	- 21 screening radiographers - 8 radiologists - 106 093 screenings mammograms, double read by 2 radiographers and 2 radiologists - 2-year follow-up	Inter-observer variability and effect of type and number of readers on outcome	Single radiologist reading: - Mean cancer detection rate: 4.64 per 1000 screens (95% CI: 4.23-5.05) - Sensitivity: 63.9% (95% CI: 60.5-67.3) Two radiologists reading - Sensitivity: 68.6% (95% CI: 65.3-71.9) - Increase in referral rate: 1.24% to 1.36%	Variation in performance single-reading	• no blinding of readers • during study conversion from SFM to FFDM • influence of inter-observer variability • delicate balance between referral rate and cancer detection rate

				- Increase in cancer detection rate: 4.64 to 4.98 Radiologist double reading+ radiographer+ pos finding radiographer read by radiologist: - Sensitivity: 73.2% (95% CI: 70.1-76.4) - Increase in referral rate: 1.24 to 1.96% - Increase in detection rate: 4.64 to 5.46 - Decrease in PPV: 37.4% to 27.9% Triple reading by 1 radiologist + 2 radiographers: - Sensitivity: 75.2% (95% CI:: 72.1-78.2) Quadruple reading by 2 radiographers and 2 radiologists: - sensitivity: 76.9% (95% CI: 73.9-79.9) - highest referral rate: 2.04% - highest detection		

				rate: 5.58 “ triple reading by 1 radiologist and 2 radiographers may replace radiologist double reading”		
Duijm et al, 2008[129]	• prospective cohort study	- Period A: 66 225 mammograms double-read by 2 radiologists - Period B: 78 325 mammograms double read by 2 radiographers in addition to 2 radiologists	Additional double reading by radiographers	Period A: - Referral of 678 women, 1.02% - 322 cancers, 4.86 per 1000 Period B: - Referral of 1122 women - 411 cancers - Decrease PPV biopsy (57.8% vs 69.7%) - Increase detection rate (4.86 to 5.25) - Decrease of PPV referral (36.6% vs 47.5%) - Larger proportion DCIS (27.6% vs 16.1%)	Cancer detection rate: - Radiographers: 4.5 per 1000 - Radiologists: 5.26 per 1000	• Increased referral rate resulted in higher detection rate but also in drop of PPV of referral • No blinding of radiologists • Potential additional diagnostic cost
Bennett et al, 2006[130]	• Review • Databases: Pubmed • Number of studies: 8 studies	- Asymptomatic women in screening program	Single reading (SR) with computer-aided detection (CAD) vs double reading (DR)	Heterogeneity in results: - Four studies found no stat sign diff between		• Only Pubmed as database • Heterogeneity in study designs and in results

				sensitivity and specificity - Other studies: DR more sensitive but SR with CAD more specific "limited evidence that SR with CAD did not perform as well as DR"		
Ciatto et al, 2005[131]	• Retrospective cohort study • Part of Florence Screening program	- 177 631 mammograms, double read - 11 trained radiologists - Asymptomatic women, age 50-69years - Biennal mammogrpahy	Double reading (DR) versus single-reading (SR)	Referral rate: - Reader 1: 2.89% - Reader 2: 3.15% - Both: 3.59% ➔ Increase of 0.70% Cancer detection rate: - Reader 1: 670 - Reader 2: 695 - 61 detected by one reader ➔ Increase in detection rate 0.024% "Detecting 43 additional cancers required 177 631 additional readings and 1250 additional referrals"	Cancers detected by second reader expected to be smaller	• No blinding of readers • Fatigue and loss of attention of first reader • Doubling of number and workload of radiologists
Ciatto et al, 2005[50]	• Retrospective cohort study	- 195 872 screening	Arbitration of discordant findings in	Arbitration neg: 60. 8% (741 cases)		• Arbitration substantially

	• Part of Florence Screening program	mammograms, 7529 positives, 3976 discordant, 1217 arbitrations by third reader - Five-grade scale	double reading (DR)	Arbitration pos: 39.2% (476 cases) After pos arbitration - cancer detection rate: 30 cancers - PPV: 6.3% After neg arbitration - 311 directed to follow-up - 2 cancers detected (0.64%) Sensitivity: 86.3% NPV: 99.3% Referral rate: decrease from 3.82% to 2.59% (relative decrease 32.1%, absolute decrease 60.8%) “for each missed cancer due to false-negative arbitration, 151 unnecessary recalls would have been saved”		reduces recall rates in discordant readings • No complete follow-up available
Ciatto et al, 2006[54]	• Retrospective study	- 108 mammograms - 33 cancers, missed by reader 1 but detected by	Computer-aided detection (CAD) in cancers detected by one reader in double reading (DR)	CAD: - Sensitivity: 51.5% - Specificity control cases: 18.6% - PPV: 21.7% - Benign mass in		• Retrospective simulation • CAD poorly specific and generates excess

		reader 2 - 75 case controls - Total of 108 cases read by CAD and 1 reader		105 controls and in 45 cancer cases - Malignant mass in 16 cancer cases - PPV for masses 9.6% - PPV for microcalcifications: 10.3% Radiologist: - Sensitivity: 74.7% - Recall rate 14.2% ➔ No sign diff in sensitivity but CAD poorly specific and excess false-positives " some limitations in the use of CAD as substitution for conventional DR"		false-positives
Ciatto et al, 2003[53]	• Retrospective cohort study	- 120 mammograms - 31 interval cancers - 19 radiologists	Computer-aided detection (CAD) versus single-reading (SR) versus double reading (DR)	CAD: - detection of 340 sites (average 2.8 per case or 1.06 per film, 132 microcalcifications, 208 opacities) - sensitivity 51.6% - compared to DR: not sign less sensitive (42.1 vs		• aim of CAD is not diagnosis but alerting reader to specific areas for second review • intraobserver inconsistency

				46.1%, p=0.07)+ more specific (recall rate 23.9 vs 26.1%, p=0.04) Radiologists: - increase of sensitivity (from 11 to 88% "CAD increased sensitivity but increased recall rate"		
Duijm et al, 2004[132]	• prospective cohort study • part of Dutch Nationwide breast Cancer Screening Program	- asymptomatic women in biennal screening program, aged 50-75years - arbitration panel of 3 radiologists - 65 779 women	Effect of arbitration on discordant findings in double reading (DR)	DR agreement: - referral: 498 cases (0.8%) - no referral: 64 949 cases (98.7%) DR disagreement: - 332 cases (0.5%) After consensus DR: disagreement on 183 cases (0.3%) Arbitration: - 89 of 183 cases - 20 cancers (22%) "if all 183 cases were referred, referral rate would have increased from 1.5% to 1.7% and number	Overall biopsy detection rate: 4.9cases per 1000 Biopsy rate: 6.5 biopsies per 1000 Risk: - 588 cases per 1000 with agreement for referral - 1.5 cases per 1000 with agreement for non-referral - 93 cases per 1000 with discrepant reading IF: no referral of 183 cases (with	• Arbitration seems nt useful • Number of mammographic views varies between screening rounds • Availability of previous screenings results may influence reader (and referral and detection rates)

				of cancers detected would have increased from 4.4 to 4.5 per 1000 women" "women should be referred for further diagnostic assessment whenever two independent readers do not reach a consensus"	disagreement): cancer detection rate 5-7% lower IF referral of all 183 cases or all 332 discrepant cases: cancer detection rate: 1.4-1.8% higher But: 74.2% increase false-positives	
Liston et al, 2003[133]	• Retrospective cohort study • Part of National Health Service Breast Screening Program (NHSBSP)	- 177 167 women aged 50-64 years - Mammograms double read - Third reader arbitration	Double reading (DR) versus single reading (SR)	Cancer detection rate: 1072 cancers Cancer detection rate after third arbitration: 8.1% (87/1072), of which 73 invasive and 14 in situ 80 cancers missed by 1st reader, 7 by 2nd reader " policy SR has to be reviewed versus DR in the NHSBSP"	Wide variation in recall rate (3.7-6.0%) Dr with arbitration detected 32% more small invasive cancers with two mammographic views and 73% more with single oblique views	• Shortage of radiologists in UK • Fine dividing line between overcalling women for assessment and missing small cancers
Gilbert et al, 2006[134]	• Retrospective cohort study	- 10 096 mammograms - Women aged 50-65y	Computer-aided detection (CAD) and single reading (SR) versus double reading (DR)	Cancer detection rate: 230 cancers and 85 interval cancers - SR+ CAD: 49.1%	Only cancer cases: 85% agreement between SR+CAD and DR	• Large sample size • Success of CAD higly dependent on specificity of

				- DR: 42.6% ➔ Mean diff 6.5% (95%CI: 1.1-11.9%, p=.02)+ relative increase of 15% Recall rate: - SR+CAD: 8.6% - DR: 6.5% ➔ Relative increase of 32% "performance SR+CAD higher than DR (higher detection rate) but higher recall rate"	For normal cases: 91% agreement between SR+CAD and DR, recall rate sign higher for SR+CAD (7.7% versus 5.7%)(p=.001) For all cancer cases (with interval cancers): 84% agreement between SR+CAD and DR	prompts • Large number of false prompts may lead to reader fatigue and reduced performance • 70% of cases single view mammograms • Difference in experience level of readers
Mucci et al, 1999[135]	• Retrospective cohort study • Part of NHSBSP (UK)	- Two view mammograms - Asymptomatic women in national screening program	Third reader as arbitration in double reading	398 disagreements between 1st and 2nd reader: 196 (49%) recalled, 202 (51%) to screening 196 women recalled: - 49 (25%) cytology - 9 (4.6%) biopsy: 5 benign, 4 malignant " third reader arbitration reduces recall rate without reduction in cancer detection"	1 interval cancer in 3year follow-up	• No blinding of readers • Double reading reduces observer errors

Georgian-Smith et al, 2007[136]	• retrospective cohort study	- 6381 screening mammograms - Asymptomatic women, aged - 1ste reader read all mammograms and reinterpreted with the use of CAD - 2nd reader: double reader of mammograms - Screen-film unit	Single reading (SR) with computer-aided detection (CAD) versus double reading (DR)	1st reader: - Recall rate: 475 (7.4%) - Biopsies in 70/475 (14.7% - 13 malignancies (18.6%) - Cancer detection rate: 2.04 per 1000 women screened SR + CAD: - Recall rate: additional 30 cases (0.47%) - Biopsies in 3/30 (10%) - No malignancy 2nd reader: - Recall rate: additional (to 1st reader) 34 cases (0.53%) - Biopsies in 5/34 (14.7%) - 2 malignancies (40%) ➔ Relative increase in cancer detection rate of 15.4% (2.35 per 1000) between 1st and 2nd reader ➔ No sign diff between 2nd	CAD and 2nd reader: detection of additional cancers but markings missed by 1st reader False-negatives: 3 within 12 months False-positive marking rate CAD: 11 968 false-pos marks (rate of 99.7%) PPV: 0% for CAD and 40% for 2nd reader Overall cancer detection rate for three readers: 2.35 per 1000	• No blinding of second reader to findings of 1st reader • Small sample size • CAD used on analogue films

				reader and CAD in recall rate (p=0.70) and cancer detection rate (p=0.50)		
Gromet et al, 2008[137]	• Retrospective cohort study	- 231 221 screening mammograms - Double reading: third reader for arbitration - Experienced mammographers - If 1st reader pos finding, 2nd reader neg finding, recall of the woman - Mean age DR: 53.8y (11.5 SD) - Mean age SR+ CAD: 53.5y (11.1SD) - Sign diff in average age but small (0.3years)(p<0.0001)	Single reading (SR) with computer-aided detection (CAD) versus double reading (DR)	Recall rate: - 1st reader: 10.2% - 2nd reader: 11.9% - SR+CAD: 10.6% Sensitivity: - 1st reader: 81.4% - 2nd reader: 88.0% - SR+CAD: 90.4% Cancer detection rate: - 1st reader: 4.12 per 1000 - 2nd reader: 4.46 per 1000 - SR+CAD: 4.2 per 1000 ➔ DR: - 38 add cancers (reduction false neg to 68) - Increase sensitivity from 81.4% to 88% - PPV decrease from 4.1% to 3.7% - Cancer detection rate increase by 0.34 per 1000	SR+CAD vs SR: - Sens: sign increase for CAD (90.4% vs 81.4%)(p<0.0001) - Recall rate: sign increase for CAD (10.6% vs 10.2%)(p<0.0001) - PPV or detection rate: no sign diff PPV3 (% biopsies resulting in diagnosis of cancer): - 1st reader: 30.6% - 2nd reader: 22.1% SR+CAD: 27.8%	• Diff in age groups • Longer time between examinations is associated with increased cancer detection rate, increased recall rate and increased sensitivity • DR: increase in sens and detection rate but increase in recall rate and more negative biopsies+ costly for manpower

				(from 4.12 to 4.46 per 1000) - Higher recall rate (11.9% vs 10.2%) ➔ SR+CAD vs DR: no sign diff in sens, detection rate and PPV but sign lower recall rate with CAD (10.6% vs 11.9%) (p<0.0001)		
Ciatto et al, 2003[53]	• Retrospective cohort study	- 150 mammograms - 10 radiologists	Single reading (SR) versus computer-aided detection (CAD)	Overall cancer detection rate: 17 cancers (170 cancer cases: 10x17) SR: - Detection rate: 146/170 (85.8%) - Recalls: 106/1330 (7.9%) CAD - Sensitivity 94.1% (16/17) - Detection rate 153/170 (90.0%) - Recalls: 152:1330 (11.4%) ➔ Increase in sensitivity but also in specificity (higher recall rate)	Marking of 767 sites for second review: sens for calcifications 100% (6/6), for opacities (90.9% (10/11)	• Blinding of readers to test results • Initial evaluation of performance of CAD • Sample not representative for screening (higher prevalence of cancer cases)

Destounis et al, 2004[138]	• Retrospective cohort study • Community-based practice	- 64 442 women - All mammograms double read by 2 independent radiologists - 519 histologically proved cancers (175 diagnostic, 344 screening) - 52 false-negative findings analyzed with CAD	Computer-aided detection (CAD) compared to double reading (DR)	52 false-negative cancers (30 minimal cancers) CAD: - 218 marks (average of 4.5 marks per case) - 75% of marks indicating cancer - 37/52 cancers detected (71%) on prior mammo False-neg rate: 61 of 318 (19%)	CAD has potential to decrease false-negative rate	• No blinding of readers • Cancer visibility depending of capacity of imaging system • False marks lead to increase in recall rate, radiologists' workload, operating expenses • Preliminary results
Khoo et al, 2005[57]	• Prospective cohort study • UK National Breast Screening Programme	- 6111 women in screening program - Mean age 58.4years - Routine screening every 3 years - Independently double read by 12 readers - Recall after arbitration	Double reading (DR) versus single-reading (SR) with computer-aided detection (CAD)	Cancer detection rate: 62 cancers in 61 women Cad detected 51/61 cancers (84%) Sensitivity: - SR: 90.2% (95%CI: 83.0-95.0%) - SR+CAD: 91.5% (95%CI: 85.0%-96.0%) - DR: 98.4% (95%CI: 91-100%) ➔ No sign diff in cancer detection rate ➔ Higher recall rate	12 cancers missed on SR: 9 correctly prompted by CAD but 7 overruled by reader False prompt rate: 1.59 per case	• Increase of recall both to arbitration and to assessment • Readers reject some true prompts • Low specificity, readers more likely to ignore correct prompts • No follow-up period in study design • Training in CAD necessary?

				for CAD (increase of 5.8%) "CAD increases sensitivity of SR by 1.3%, whereas DR increases sensitivity by 8.2%"		
Gilbert et al, 2008[55]	• Equivalence trial with matched-pair comparisons • Prospective study	- 31 057 women - Film mammography - DR, SR +CAD, SR+CAD+DR -	Single-reading (SR) with computer-aided detection (CAD) compared to double reading (DR)	Cancer detection rate: - SR + CAD: 8 cancers (6.8 per 1000) - SR+CAD+DR: 227 cancers (8.0 per 1000) - DR: 12 cancers (10.4 per 1000) ➔ Detection rates similar Recall rate: - SR+CAD: 3.9% - DR: 3.4% ➔ Small sign diff (p<0.001) SR+CAD: - Sens: 87.2% - Spec: 96.9% - PPV: 18.0% DR: - Sens: 87.7% - Spec: 97.4% - PPV: 21.1%	No sign diff in pathological findings between SR+CAD and DR	• Large trial • No bias by difference in experience of reader • Additional cost of CAD equipment, costs associated with increased recall rate • Potential saving in reader time • Use of screen film in study, performance of CAD in digital mammography not examined

				"SR+CAD could be an alternative to DR and could improve cancer detection rate"		
Cawson et al, 2009[56]	• Retrospective study • Case mix study • BreastScreen Australia	- independent double reading with arbitration (reader A, reader B) - 157 invasive cancers mixed with normal cases (total 1569) - 1569 film-screen mammograms - Women aged 50-69y - Screening every two year	Single reading(SR-with computer-aided detection (CAD) versus double reading (DR)	Sensitivity - DR: 90.4% - CAD-RA: 86.6% - CAD-RB: 94.3% - CAD: 93% ➔ No sign diff between CAD and DR (p=0.20) After CAD: reader's sens increased 1.9% (95% CI: 0.4-5.5%) but specificity dropped 0.2% and 0.8% (not sign)	Arbitration after DR decreased spec 4.7% Mean prompts per case with CAD: 2.1 AUC: - CAD-RB: 0.96 - CAD-RA: 0.94 - DR: 0.95 Size of cancers not sign diff between CAD and DR	• Shortage of radiologists • Readers rejected most positive prompts • Role of experience of reader in accepting or rejecting prompts
Taylor et al, 2004[139]	• Retrospective cohort study	- 35 readers read 120 films (including 44 cancers)	Computer-aided detection (CAD) versus double reading (DR)	Sensitivity - SR: 0.77% - SR+CAD: 0.80% - DR: 0.81% ➔ CAD: sens increase but not sign ➔ DR: increase compared to SR Specificity: - SR: 0.85% - SR+CAD: 0.86% - DR: 0.88% ➔ CAD: spec increase but not		• Mix of cancer cases which were missed by one of the readers in study design • Readers will ignore a sign % of correctly placed prompts

				sign		
Yang et al, 2007[140]	• Retrospective cohort study	- Digital mammograms of 103 women with breast cancers (mean age: 51y, range 35-69) - Normal mammograms of 100 women (mean age 54y, range 35-75y)	Computer-aided detection (CAD) versus double reading (DR) in FFDM	CAD in cancer cases - 442 marks - 182 masses (of which 84 true-pos, 98 false-pos) - 260 microcalcifications (of which 208 true-pos, 52 false-pos) ➔ Overall false-pos mark rate per patient: 1.45 ➔ 99/103 correctly marked (96.1%, 95%CI: 90.1-98.8%) CAD in normal cases - Mean false-pos marks per patient:1.80 "CAD correctly marked 96.1% asymptomatic breast cancers with acceptable false-positive marks (1.8 per patient)"	Sensitivity CAD in fatty breast group: 95% (59/62) Sensitivity CAD in dense breast group: 98% (40/41) ➔ No sign diff (p=.766)	• Large number of false-pos marks can hinder usefulness of CAD by distracting the interpreting radiologist • Small sample size
Skaane et al, 2007[141]	• Retrospective study	- 3683 women underwent both SFM and FFDM with independent DR	Computer-aided detection (CAD) and double reading (DR) in SFM and FFDM	DR with FFDM: - CAD cancer detection27/29 at baseline, 10/10 subsequent		• Goal of CAD: reducing number of false-negatives • Potential benefit of 36% in FFDM

		- 55 biopsy-proven cancers: 29 at baseline, 10 interval , 16 at second screening round - Mean age 58.2years		➔ Sens: 94% vs 64% (DR with FFDM) ➔ Sign diff (p=0.006) DR with SFM: - CAD cancer detection: 27/29 at baseline, 6/10 subsequent ➔ Sens: 85% vs 77% DR with SFM ➔ No sign diff (p=0.57) “CAD has the potential for increasing the cancer detection rate”		with soft-copy reading • Learning curve effect in FFDM • Suboptimal reading environment in FFDM soft-copy review

Appendix 3.2.2. Full-field digital mammography

Systematic reviews

Table 32 full-field digital mammography: systematic reviews

Reference	Methodology	Patient characteristics	Intervention(s)	Results primary outcome	Results secondary and other outcomes	Critical appraisal of review quality
AETSA, 2007[42]	• SR • Funding: Ministerio de sanidad y politica social (Andalucia, Spain) • search date: 1995-2007 • databases: Medline, Pre-Medline, Embase, Cochrane Library Plus, Centre for reviews and Dissemination, INAHTA, National Guidelines Clearinghouse, ECRI Institute, Sumsearch, Tripdatabase • study design: RCTs, cross-sectional studies, prospective and retrospective cohorts • N included studies: 6 studies (11 reports)	*Eligibility criteria*: asymptomatic adult women, breast cancer screening, Digital mammography (DM) versus traditional mammography (TM) and outcomes: diagnostic performance; intermediate results (like recall rate) or final outcomes as mortality *Patient characteristics*: asymptomatic women Age:40 - 70 years - N: 3683 - 324763	digital mammography *versus* traditional mammography (or combination DM and TM) *Reference standard:* biopsies (in all studies) and interval cancer during follow up (not in all studies) during 1 or 2 years	*Sensitivity, specificity, PPV, NPV, LR+, LR-* Divergence in results according to studies *Sensitivity:* DM: 35-70 TM: 45-83 *VPP* DM:3-21 TM:3-22 No statistical difference for sensitivity between DM and TM in the most valid studies *ROC Curve* For Bi-rads scale categories of malignity: statistically significant bigger area (Higher sensitivity) of MD in	Cancer detection rate: No difference *Interval cancer rate*: no difference *Carcinoma in situ*: divergent results (equivalent in one study and higher percentage with DM in another study) *Recall rate*: No difference except higher recall for DM in subgroup 50-69 (in one study) *Biopsies*: Divergent results (No difference in one good quality study and higher percentage in one lower quality study)	• Primary studies in detail

				women aged less than 50 or with high density or who were perimenopausal (one prospective cohort study/opportunistic screening) *PPV:* No difference *Specificity*: divergent results	*Mortality:* No studies found *Safety:* not include in selected outcomes	

Primary studies derived from systematic review

Table 33 full-field digital mammography: primary studies derived from systematic reviews

Reference	Methodology	Patient characteristics	Intervention(s)	Results primary outcome	Results secondary and other outcomes	Critical appraisal of review quality
Lewin, 2001[58]	• Prospective, cohort study • Grant from the U.S. Army Breast Cancer Research and Materiel Command • "Colorado-Massachusetts study"	• Eligibility criteria: asymptomatic women presenting for screening mammography, at least 40years old • Patient characteristics: - Mean age: 55.5y ±9.8 - Number of BRCA women: - 4945 examinations in 3890 women (1055 women enrolled twice)	Full-field digital mammography (FFDM) vs screen-film mammography (SFM)	• Recall rate: FFDM: 11.5% (568 of 4945) SFM: 13.8% (685 of 4945) - Positive biopsy rate FFDM: 30% (21 of 69) SFM: 19% (22 of 114) - Positive predictive value (= fraction of recalled examinations that led to a diagnosis of	• Sensitivity (comparator: additional imaging, prior images, biopsy) FFDM: 60% (21 of 35), SFM: 63% (22 of 35) Relative sensitivity of FFDM to SFM: 95% (21 of 22)	• Screening population results in low cancer rate, which decreases power to detect differences between modalities • Large reader variability, disagreements on 821 of the 4945 examinations (17% of total, 79% of positive examinations)

				breast cancer) FFDM: 3.7% (21 of 568) SFM: 3.2% (22 of 685) "no difference in cancer detection rate has yet been observed between FFDM and SFM. FFDM has so far led to fewer recalls than SFM"		
Lewin, 2002[59]	• Prospective cohort study • Grant from the U.S. Army Breast Cancer Research and Materiel Command	• All (asymptomatic) women 40 years and older who presented for screening mammography • 6736 paired examinations on 4489 subjects (1665 subjects enrolled twice, 291 three times) • Average age: 55.6 years	Full-field digital mammography (FFDM) vs screen-film mammography (SFM)	- Higher recall rate for SFM: 14.9%, FFDM 11.8% - Positive predictive value lower for SFM (33/1001, 3.3%) than for FFDM (27/793, 3.4%) "no significant difference in cancer detection, FFDM resulted in fewer recalls"	- Statistically difference in number of biopsies (p<0.001): 87 on SFM, 38 on FFDM, 56 on both - Number of detected cancers: SFM 9, FFDM 15 (18 on both) but difference not statistically significant (p>0.1) - No significant difference (p=0.18) in area under the curve (= threshold for positivity)	• Digital mammography technology, used in study is dated • Data on recall rate could be biased by lower threshold for recall used with SFM

					between SFM (0.80) and FFDM (0.74)	
Glueck, 2007[60]	Re-analysis of study of Lewin, 2002 Grant from National Cancer Institute	• 6736 paired mammograms performed in 4489 women	Women received both full-field digital (FFDM) and screen-film mammography (SFM)	Total cancer detection rate: 49 - SFM 32 (65.3%) - FFDM 27 (55.1%) - Both 18 (83.7%) - 8 interval cancers Significant increase in proportion of cancers detected by combined modality " using two mammograms, one film and one digital, significantly increases the detection of breast cancer"	No significant difference in ROC curves between FFDM, SFM or combined with parametric tests But: significant difference in ROC curves between SFM versus combined, and FFDM versus combined with non-parametric tests	• Definite conclusion about benefit of one modality or combined can not be drawn, due to differences between parametric and non-parametric tests • Based on clinical trial data, increased cancer detection rate cannot be explained by number of readers, number of compressions or use of two different modalities
Skaane, 2003[61]	Prospective cohort study (Oslo I study) Participants from the Norwegian Breast Cancer Screening Program	• 3683 women • Aged 50-69 years, mean age 58.2years • Women underwent both FFDM and SFM • Group of women that	Full-field digital mammography Vs Screen-film mammography with soft-copy reading in	Recall rate: - FFDM 4.6% (168 of 3683 cases) - SFM 3.5% (128 of 3683		• Reluctancy to implementation of full-field digital mammography with soft-copy reading is inferior spatial resolution,

		only underwent SFM used as control population	a population-based mammography screening program Independent double reading of images (five-point rating scale for probability of cancer) Reference test: biopsy	cases) Positive predictive values: - PPV1 (cancers among recalls): 20% for SFM and 12% for FFDM - PPV2 (cancers detected after cytology): 46% for SFM and 39% for FFDM Cancer detection rate: 31 - SFM 28 - FFDM 23 ➔ No significant difference between both modalities - 0.84% (31/3683)vs 0.40% (25/6249) in control population		user-unfriendliness of soft-copy display for routine use in screening setting • Divergence with study results from Lewin et al: no confirmation of lower recall rate of FFDM • Reader variability

				" There was no statistically significant difference in cancer detection rate between screen-film and full-field digital mammography. Full-field digital mammography is comparable to screen-film mammography in population-based screening"		
Skaane, 2005[63]	Prospective cohort study (Oslo I study): follow-up and final results	• 3683 women in screening program • Mean age 58.2years • All women underwent both SFM and FFDM	Screen-film (SFM) Versus Full-field digital mammography (FFDM) with soft-copy reading Reference test: needle biopsy	Total cancer detection rate: 31 - SFM: 28 (detection rate 0.76%) - FFDM: 23 (detection rate 0.62%) - Both 20 (65%) ➔ No significant difference in cancer detection No significant difference in cancer detection after recall	Positive interpretation: SFM 442 cases and FFDM 612 cases Total of 31 cancers detected in initial screening round (detection rate 0.84%) 10 interval cancers detected 16 cancers detected in subsequent screening round	• Learning curve effect for FFDM • Inter-observer variation • (in)experience of readers in soft-copy reading • Double reading by consensus or arbitration increases cancer detection with a reduction of recalls but cancers may be dismissed

				Recall rate: - SFM 3.5% (128 of 3683 cases) - FFDM 4.6% (168 of 3683 cases) "There is no statistically significant difference in cancer detection rate between SFM and FFDM with soft-copy reading in a mammography screening program."	(2 years later) False-negative interpretations: 31% (22 of 72) at SFM and 47% (34 of 72) at FFDM True positive scores: 69%(50/72) on SFM, 53% (38/72) on FFDM	
Skaane, 2004[62]	Prospective cohort study (Oslo II study)	• 25 263 women • 45-69 years • Screening program • Women underwent SFM or FFDM • Independent double reading with use of five-point rating scale for probability of cancer	Screen-film mammography (SFM) Versus Full-field digital mammography (FFDM) with soft-copy reading Comparison between two age groups (45-49y and 50-69y)	Cancer detection rate: - Total 120 (detection rate 0.48%) - SFM: 73 in 17911 women (detection rate 0.41%) - FFDM: 41 in 6997 women (detection rate 0.59%)	Cancer detection rate in subgroups: Group 50-69 years: 56 in 10304 women SFM (detection rate 0.54%), 33 in 3985 women FFDM (detection rate 0.54%) Group 45-49 years: 17 in 7607 women SFM (detection rate	• learning curve effect • influence of reading environments

				➔ Difference in cancer detection rate approached significance (p=.06) Recall rate Group 50-69years - SFM 253 (2.5%) of 1.304 - FFDM 153 (3.8%) of 3985 Group 45-49 years: - SFM 231 (3.0%) of 7607 - FFDM112 (3.7%) of 3012 ➔ significantly higher at FFDM than at SFM in group 50-69years, not in group 45-49 years No significantly difference in	0.22%), 8 in 3012 women FFDM (detection rate 0.27%) Positive predictive value (PPV) Group 50-69years: 56(22.1%) of 253 for SFM and 33 (21.6%) of 153 for FFDM Group 45-49 years: 17 (7.4%) of 231 for SFM, 8 (7.1%) of 112 for FFDM ➔ differences non significant	

				positive predictive value " FFDM allowed a higher cancer detection rate than did SFM in the group aged 50-69years, although difference did not reach statistical significance. SFM and FFDM are comparable techniques for population-based screening mammography programs."		
Skaane, 2007[64]	Prospective, cohort study (follow-up and final results Oslo II study)	• 23 929 women 45-69 years (13912 in 50-69 years and 10017 in 45-49 years) • underwent SFM (n= 16 985) or FFDM (n= 6944) • follow-up for 1.5years (group 45-49years) and 2.0years (group 50-69years)	SFM vs FFDM	Recall rate: - FFDM 4.2% - SFM 2.5% Cancer detection rate - FFDM 41 (0.59%) of 6944 cases - SFM 64 (0.38%) of 16985 cases No significant difference in PPVs	Overall true positive score 73 (0.43%) of 16 985 cases at SFM and 44 (0.63%) of 6944 cases at FFDM) ➔ higher true-positive score at FFDM statistically significant (p=.03) Sensitivity 77.4%	• significantly higher recall rate at FFDM than at SFM in Oslo II study important difference with Lewin et al (significantly lower recall rate) • double reading can help increase cancer detection rate by 10-15%

				Interval cancer rate - FFDM 17.4 - SFM 23.6 Group 50-69years:30 (38%) of 80 cases at SFM and 10 (24%) of 42 cases at FFDM Group 45-49years: 40 in SFM and 12 in FFDM "FFDM resulted in a significantly higher cancer detection rate than did SFM. PPVs were comparable for both."	at FFDM and 61.5% at SFM Specificity 96.5% FFDM and 97.9% SFM	
Vigeland, 2007[65]	Prospective cohort study Regional comparison within Norwegian Breast Cancer Screening Programme	• 18239 women aged 50-69 years (mean age 58.9years) underwent FFDM • 324763 women underwent SFM • Population-based screening	Full-field digital mammography Versus Screen-film mammography with soft-copy reading	Cancer detection rate: - FFDM 0.77% (140 of 18239) - SFM 0.65% (2105 of 324763 cases) ➔ FFDM significantly higher detection rate	Positive predictive value (PPV): 16.6% (140 of 843) for FFDM and 13.5% (2105 of 15537) for SFM ➔ FFDM significantly higher PPV	• Lower technical recall for FFDM

				for DCIS than SFM (no difference for invasive cancers) Recall rates: - FFDM 4.09% (746 of 18239) - SFM 4.16% (13520 of 324764) ➔ No significant difference " FFDM performed better than or equal to SFM"		
Pisano, 2005[66]	Prospective cohort study (DMIST) Grants from the National Cancer Institute	• 42 760 women underwent both FFDM and SFM in random order • Mean age 54.9years	Digital mammography Versus film mammography	Diagnostic accuracy (mean area under the curve): 0.78 ±0.02 for FFDM and 0.74±0.02 for SFM ➔ No significant difference After 455 days of follow-up: FFDM Sensitivity: 0.41±0.03 Specificity:	Under age of 50 years performance significantly better for FFDM than for SFM compared to women older than 50years, Women under age of 50: AUC FFDM 0.84±0.03, AUC SFM 0.69±0.05.	Use of a seven-point scale of malignancy

				0.98±0.001 PPV: 0.12±0.01 SFM Sensitivity: 0.41±0.03 Specificity: 0.98±0.001 PPV: 0.13±0.01 →no sign difference After 365days of follow-up FFDM Sensitivity:0.70±0.03 Specificity:0.92±0.001 PPV: 0.05±0.004 SFM Sensitivity: 0.66±0.03 Specificity: 0.92±0.001 PPV: 0.05±0.003 “overall diagnostic accuracy of FFDM and SFM is similar, but FFDM is more accurate in women under the age of 50	Difference 0.15 (95CI 0.05-0.25), p=0.002 Women with dense or extremely dense breasts: AUC FFDM 0.78±0.03, AUC SFM 0.68±0.03. Difference 0.11 (95CI 0.04-0.18), p=0.003 Premenopausal or perimenopausal women: AUC FFDM 0.82±0.03, AUC SFM 0.67±0.05. Difference 0.15 (95CI 0.05-0.24), p=0.002 No significant difference between SFM and FFDM for women 50years or older, women with fatty breasts or scattered fibroglandular

				years, women with dense breasts and pre- or perimenopausal women	densities and postmenopausal women	
Del Turco, 2007[67]	Retrospective cohort study	• 2 cohorts of women 50-69 years old • 14385 women per cohort • Screening program in mobile unit	Digital mammography Versus Film-screen mammography	Recall rate: SFM 3.96% FFDM 4.56% ➔ Sign difference (p=0.01) Detection rate: 188 cancers detected FFDM 84 (0.58%) SFM 104 (0.72) ➔ Detection rate higher for FFDM but not sign difference (p=0.14) PPV: FFDM 15.9% SFM 14.7% ➔ No sign difference in PPV (p=0.65) "FFDM may be	FFDM sign more recalls because of radiologic abnormalities, sign less recall because of poor technical quality Sign higher recall rate because of microcalcifications No diff for masses or distorsions Sign higher recall for women 50-59 years and for women with very dens breasts	Recall rate in FFDM lower due to better imaging quality and opportunity for postprocessing Higher detection rate with FFDM in younger women and women with denser breasts related with lower sensitivity of SFM

				more effective than SFM in contemporary screening practice mobile units. The data indicate that FFDM depicts more tumors than does SFM."		

Table 34: Study characteristics primary studies digital screening in breast cancer screening

Reference	Methodology	Patient characteristics	Intervention(s)	Results primary outcome	Results secondary and other outcomes	Critical appraisal of review quality
Bluekens et al, 2010[68]	• Retrospective cohort study • Dutch screening programme	- Total of 312 414 screenings mammograms (43 913 FFDM and 268 501 SFM) - Mean age of referred women: 58.5years for SFM and 57.4 years for FFDM	Referral pattern after FFDM in population-based breast cancer screening programme	Higher recall rate in FFDM: Initial screening round: from 3.4% to 4.3% Subsequent rounds: from 1.0% to 1.7% Significant increase in cancer detection (p=.010): Initial screening round: 7.6% FFDM vs 6.0% SFM Subsequent rounds: 5.5% FFDM vs 4.9% SFM	Referral rate decreases and stabilises on long-term effect	• Learning curve effect • Training in digital screening recommended
Domingo et al, 2011[78]	• Retrospective cohort study • Grants from Instituto	- 103 613 asymptomatic women	SFM vs FFDM	PPV1 (at least one further assessment): - SFM: 5.5%	No sign diff in tumour characteristics	• A short period of use of FFDM • No information

	de Salud Carlos III Feder	- 45-69 years - 242 838 screenings mammograms (171 191 SFM, 71 647 FFDM)		- FFDM: 7.0% PPV2 (invasive procedures): - SFM: 19.3% - FFDM: 36.9% No sign diff in cancer detection rate "DM has a similar diagnostic precision to SFM and fewer adverse effects. The differences in tumour characteristics and higher rates for DCIS suggest an advance in early detection"	% DCIS in 1ste screening round - SFM: 15.8% - FFDM: 18.5% % DCIS in successive rounds - SFM: 15.7% - FFDM: 23.2%	about breast density
Feeley et al, 2011[69]	• Restrospective cohort study	- 107 818 women - 53 803 SFM, 54 015 FFDM - Age women: 50-64 years	SFM vs FFDM Reference test: biopsy	Recall rate - SFM 3.52% - FFDM 4.21% ➔ Sign higher recall rate for FFDM (p<0.0001) Overall cancer detection rate - SFM 6.2 per 1000 women screened	PPVs of B3 and B3/B4 diagnosis: non-sign higher for SFM PPV of B4: non-sign higher for FFDM Recall rate sign higher for microcalcifications, architectural distorsion and	• Increased detection of nminimal sign lesions may increase the number of atypical diagnoses • Risk of overtreatment without reducing breast-cancer mortality • Possibility of bias during overlap

				- FFDM 7.2 per 1000 women screened ➔ Sign higher in FFDM (p=0.04) PPV1 and PPV2: similar with SFM and FFDM "FFDM resulted in a higher cancer detection rate, especially for microcalcifications, but higher recall and open biopsy rates"	asymmetry (p<0.0001 each) with FFDM Cancer detection rate sign higher with FFDM for microcalcifications (p<0.001), invasive cancers (p=0.03), pure DCIS)p=0.003)	period when both methods were used
Karssemeijer et al, 2009[70]	• Cohort study • Grant from the European Community in the 5th Framework Information Society Technologies program	- 367 600 screening examinations: 56 518 FFDM, 311 082 SFM - Asymptomatic women in population-based screening program - 50-75years - Mean age first screening round: 51.3y FFDM, 51.9y SFM - Mean age subsequent	FFDM vs CAD with SFM	First screening round Cancer detection rate - SFM: .62% - FFDM: .77% Recall rate - SFM 2.32% - FFDM 4.41% Subsequent screening rounds: Cancer detection rate - SFM: .49% - FFDM: .55%	First screening round DCIS detection - SFM .12% - FFDM .22% PPV recall - SFM: 26.8% - FFDM 17.4% Recall on microcalcifications: - SFM:	• Concurrent comparison within the cneters • All readers involved in FFDM and SFM reading, risk of bias due to reading skill differences minimized • Slightly difference in mean age of women • Screening interval FFDM shorter than

		screening rounds: 61.6y FFDM, 62.7y SFM		Recall rate - SFM: 1.17% - FFDM: 1.70% ➔ Sign higher recall with FFDM in both screening rounds (both p<.001)	19.0% - FFDM: 39.3% Subsequent screening rounds: DCIS detection - SFM .08% - FFDM .12% PPV recall - SFM 43.1% - FFDM 30.4% Recall on microcalcifications: - SFM: 21.6% - FFDM 41.2% ➔ Sign increase in recall based on microcalcifications with FFDM ➔ PPV decreased with FFDM for all lesion types	SFM
Lipasti et al, 2010[79]	• Retrospective cohort study • Finnish population-	- 27 593 women SFM, 23 440 women FFDM	SFM vs FFDM	Cancer detection - SFM: 0.406%,	PPV: - SFM: 26% - FFDM:	• 8years difference in study periods of SFM and FFDM

	based screening program			tumor-like masses - FFDM:0.623 %, parenchymal distorsions, asymmetric densities, calcifications, masses with calcifications Recall rate: similar in both groups	36%	
Perry et al, 2011[80]	• Cohort study • London Breast Institute	- 14 946 screening mammograms: 5010 FFDM, 9936 SFM	SFM vs FFDM	Cancer detection rate - SFM: 2.8 per 100 women screened (28/9 936) - FFDM: 6.4 per 1000 women screened (32/5 010) Recall rate - SFM: 5.0% - FFDM: 7.3% ➔ Sign higher for FFDM (p<0.001) Microcalcifications	Women <50years Cancer detection rate - SFM: 1.4 per 1000 - FFDM: 4.3 per 1000 ➔ Sign higher for FFDM (p=0.02) Recall rate: - SFM: 5.3% - FFDM: 7.3% ➔ Sign higher for FFDM (p=0.009) PPV: - SFM: 2.6%	• Almost half of the screenings performed in women younger than 50years • Better performance of FFDM in women with denser breast tissue and under 50 years • No distinction between first and subsequent screening rounds

				- SFM: 14%, 0.4 per 1000) - FFDM: 31% (2.0 per 1000) " Cancer detection rates were significantly higher for FFDM than for SFM, especially for women <50 and cancers detected as clustering microcalcifications"	- FFDM: 5.9% ➔ Not sign (p=0.1) Women > 50years Cancer detection rate - SFM: 4.0 per 1000 - FFDM: 8.6 per 1000 ➔ Sign higher for FFDM (p=0.002) Recall rate - SFM: 4.7% - FFDM: 7.2% ➔ Sign higher for FFDM (p=0.001) PPV - SFM: 8.5% - FFDM: 11.9% ➔ Not sign	
Pisano et al, 2008[81]	• Retrospective cohort stud • Analysis of	- Women underwent both SFM and FFDM - 42 760 women	FFDM vs SFM	Women <50 years - AUC: sign diff (p=.0015) between	Women <50years with dense breasts: - all lesion types more	• DMIST seven point scale • Exploratory

	population subgroups in DMIST	- Division in 10 subgroups based on age, breast density and menopausal status		FFDM (0.791) and SFM (0.544) - Sensitivity: sign diff (p=.0013) between FFDM (0.591) and SFM (0.273) - PPV: sign diff (p=0.0005) between FFDM (0.033) and SFM (0.015) ➔ Improved accuracy with FFDM for pre-and perimenopausal women younger than 50 years with dense breasts Women aged 65 years and older with non-dense breasts - AUC sign diff (p=.0025) between FFDM (0.705) and SFM (0.877) - PPV sign diff (p=0.0055)	detected with FFDM - FFDM depicted more cancers Women >65years with fatty breasts: - all lesion types more detected with SFM - SFM depicted more cancers	analyses

				between FFDM (0.092) and SFM (0.127) ➔ Improved diagnostic accuracy with SFM for women aged 65years and older with fatty breasts		
Sala et al, 2009[142]	• Retrospective cohort study • Grants from the Health Ministry of Spain (Fondo de Investigacion de Sanitaria)	- 12 958 women with SFM, mean age 59.6y - 6074 women with FFDM, mean age 59.5y - Age: 50-69years	SFM vs FFDM	Overall recall rate - SFM: 5.5% - FFDM 4.2% ➔ Sign lower in FFDM (p<.001) Recall rate subsequent screening rounds - SFM 3.6% - FFDM 2.4% ➔ Sign lower in FFDM (p<.001) Overall cancer detection rate: similar in both (0.4%) Cancer detection rate	Proportion of invasive cancers higher in FFDm, but not sign Overall false positive rate - SFM 5.1% - FFDM 3.8% ➔ Sign lower in FFDM (p<.001) False-pos rate in first screening round: no diff False pos rate in subsequent	• Results on recall rate in contrast to previous studies • No information on breast density • Small size of study population • Reduction of false positives could reduce adverse effects of screening programs

				in first screening round - SFM 0.4% (12 cancers) - FFDM 1.1% (14 cancers) → sign higher in FFDM (p=.009) Cancer detection rate in subsequent screening rounds - SFM 0.4% - FFDM 0.2% ➔ Not sign diff PPV - SFM 7.5% (95% CI 5.81%-9.68%) - FFDM 9.7% (95% CI 6.68%-13.97%) - In first screening round: sign higher in FFDM (p=.002)	screening rounds: - SFM 3.2% - FFDM 2.1% ➔ Sign lower in FFDM (p<.001) % women US and fine needle aspiration lower in FFDM but no diff for women core biopsy	
Sala et al, 2011[74]	• Cohort study • Grants from the	- 103 613 women, aged 45-69y	SFM vs FFDM	Overall recall rate - SFM 8.1%	False positives - SFM 7.6%	• Contradictory results with

	Health Ministry of Spain (Fondo de Investigacion Sanitaria)	- 242 838 mammograms (171 191 SFM, 71 647 FFDM)		- FFDM 6.2% ➔ higher in SFM (p<.001) Recall rate at first screening round - SFM 12.1% - FFDM 11.7% ➔ Higher in SFM (p=.091) Recall rate at successive screening rounds - SFM 5.0% - FFDM 4.6% ➔ Higher in SFM (p<.001) Overall cancer detection rate: 1080 cancers - SFM 0.45% (770) - FFDM 0.43% (310) ➔ No sign diff in both screening rounds	- FFDM 5.7% False positives resulting in invasive procedures - SFM 3.0% (1st screening round), 1% (subs screening rounds) - FFDM 1.7% (1ste screening round), 0.45% (subs screening rounds) ➔ Sign higher in SFM (p<.001) in both screening rounds Sign increased risk of a false-pos recall in SFM DCIS - SFM 13.2% (1st screening round),	previous studies • No information about breast density • No information about false-negatives

				PPV - SFM 5.6% - FFDM 7% "cancer detection did not differ, recall rate and false-positive risk were lower with FFDM"	13.5% (subs screening round) - FFDM 17.4% (1st screening round), 18.8% (subs screening round) ➔ Higher in FFDM for both screening rounds Cancer detection risk increased with age	
Van Ongeval et al, 2010[71]	• Cohort study on decentralized screening program in Belgium (first reading in local unit, second reading in centralized organization)	- 34 680women (11 355 FFDM, 23 325 SFM) - Second control group: 14 7690 women in 1st round and 16 4476 women in subs round	FFDM vs SFM	Recall rate - SFM 2.10% (2.40% 1st screening, 1.58% subs screening) - FFDM1.58% (2.64% 1st screening, 1.20% subs screening) ➔ Sign lower in FFDM in sub	DCIS - SFM 0.16% - FFDM 0.07% ➔ FFDM sign lower (p=0.02)	• No registration of interval cancers by Flemish government • Results of small number of centers • Training in reading and regular update of individual parameters important key to a successful screening program

				screening (p=0.03) but not in 1st screening (p=0.43) Cancer detection rate - SFM 0.64% (150 cancers) (0.60% in 1st screening, 0.72% in subs round) - FFDM 0.59% (67 cancers) (0.63% in 1st screening, 0.57% in subs round) ➔ No sign diff PPV - SFM 30.67% (24.86% in 1st round, 45.93% in subs round) - FFDM 34.90% (24.05% in 1st round, 48.00% in subs round)	

				➔ No sign diff “ FFDM: high cancer detection rate and without increase of the recall rate”		
Vernacchia et al, 2009[72]	• Cohort study • Conversion from SFM to FFDM in small community-based practice	- 26 386 mammograms	SFM vs FFDM	Recall rate - FFDM: Increase during audit 1(5.9%) and 2 (10.2%), but decrease in audit 3 (7.5%) and 4 (9.0%) ➔ Sign increase ➔ Decrease over time but remained higher than SFM Cancer detection rate - Audit 1 (4.1 cancers per 1000) - Audit 2 (7.9 cancers per 1000) ➔ Sign increase (p=0.012) - Audit 3 (5.1		• Outlier in radiologist’ interpretations • Small center • Conversion to FFDM, only comparison before/after conversion • No information on breast density

				cancers per 1000) ➔ No sign diff between 1 and 3 ➔ After high increase return to level that is higher than SFM but not sign		
Vinnicombe et al, 2009[76]	• Cohort study • Results from UK Breast Screening Program (CELBSS study) and systematic review	- 39 651 women underwent 40 198 screening examinations - Median age 58 years - FFDM group younger, Caucasian and self-referrals	FFDM vs SFM	Cancer detection rate (263 cancers, 0.65 per 100 mammograms) - SFM 0.65 per 100 mammo - FFDM 0.68 per 100 mammo ➔ No sign diff Recall rate: 4.5% ➔ No diff between SFM and FFDM PPV: 14.5% ➔ No diff between SFM and FFDM	No diff between SFM and FFDM in proportion of detected tumors, histologic grades and tumor size	• No randomization

				" FFDM is performing at least as well as SFM"		
Hambly et al, 2009[73]	• Retrospective cohort study	- Total of 188 823 mammograms of 146 114 women (35 204 FFDM, 153 619 SFM) - Women underwent or SFM or FFDM - Age 50-64years	SFM versus FFDM	Recall rate - Total 3.2% (6 135) - SFM 3.1% (4729/153 619) - FFDM 4.0% (1406/35 204) ➔ Sign diff (p<0.001) Cancer detection rate - Total: 5.4 per 1000 (1013 cancers) - SFM: 5.2 per 1000 (792/153 619) - FFDM: 6.3 per 1000 (221/35 204) ➔ Sign diff (p=0.01)	PPV (number recalled for assessment): - SFM: 16.7% - FFDM: 15.7% ➔ No sign diff (p=0.383) Biopsy rate: - SFM: 35.9% - FFDM: 33.4% ➔ No sign diff (p=0.09) PPV2 (number recalled for biopsy) - SFM 46.6% - FFDM 47% ➔ No sign diff (p=0.93) Microcalcifications - SFM: 1.3 per 1000 - FFDM: 1.9 per 1000	• Five point rating scale for probability of cancer • Higher recall rate due to improved conspicuity of abnormalities, degree of unfamiliarity • Possible bias during randomization • Short period of study (2006-2007)

					➔ Sign higher for FFDM (p=0.01) DCIS - SFM: 0.7 per 1000 - FFDM: 1.2 per 1000 ➔ Sign higher for FFDM (p=0.009) Architectural distorsion: - SFM 0.7 per 1000 - FFDM 1.0 per 1000 ➔ Sign higher for FFDm (p=0.03) No sign diff in tumor size between SFM and FFDM	
Heddson et al, 2007[75]	• Retrospective cohort study	- 52 172 two-view mammography examinations: 50% SFM, 19% PC-DR, 31% CR - 24 875 women - Mean age: 58.9y: 58.0y	SFM vs photon-counting direct radiography (PC-DR) vs computed radiography	Cancer detection rate - SFM0.31% (81/25 901 - PC-DR: 0.49% (48/9841) - CR: 0.38% (63/16 430) ➔ Sign higher for PC-DR vs	PPV - SFM: 22% - PC-DR: 47% - CR: 39% Average glandular dose: - SFM:	• Digital mammography attractive: image acquisition, display and storage, saving time and effort • Patients not assigned on a randomization

		for SFM, 60.4y for PC-DR, 59.4y for CR ➔ Sign age diff (p<0.001)		SFM (p=0.04) Recall rate: - SFM: 1.4% - PC-DR: 1.0% - CR: 1.0% ➔ SFM sign higher than 2 others (p= 0.003, p<0.001) " PC-DR and CR had a high rate of cancer detection, a low recall rate and a high PPV in addition to a lower average glandular dose than SFM... valid alternative to SFM"	1.1mGy - PC-DR: 0.28mGy - CR: 0.92mGy ➔ PC-DR: 75% dose reduction, CR: 16% dose reduction	scheme • Sign age diff between groups • Bias due to cases for which double reading occurred
Juel et al, 2010[77]	• Retrospective cohort study • Part of Norwegian Breast Cancer Screening Program	- Age 49-70y - Mean age SFM: 57.84y - Mean age FFDM: 57.83y	SFM vs FFDM using photon-couting detector	Recall rate - SFM: 2.3% (174/7442) - FFDM: 2.4% (168/6932) ➔ No sign diff (0.779)	PPV (abnormal mammography) - SFM: 16.7% (29/174) - FFDM: 19.6% (33/168) ➔ No sign diff	• Five-point rating scale for probability of cancer • Learning curve effect • Digital systems: lower object contrast threshold

				Cancer detection rate - SFM: 0.39% (29/7442), 4.1/1000 - FFDM: 0.48% (33/6932), 4.8/1000 ➔ No sign diff (p=0.508) " a trend of higher cancer detection rate and PPV for FFDM but differences not significant but sign lower recall rate dus to technically inadequate images and sign lower average glandular dose"	(p=0.566) PPV (biopsy) - SFM: 39.2% (29/74) - FFDM: 44.0% (33/75) ➔ No sign diff (p=0.668) Recall rate due to technically inadequate mammograms: - SFM: 0.3% (19/7442) - FFDM: 0.01% (1/6932) Interval cancer rate: - SFM: 12.1/10 000 negative screened, 22.0% (9/41) - FFDM: 11.6/10 000, 19.6%	+ inadequate settings of windowing and levelling (contrast and brightness) • Small number of women screened

					(8/41) Average glandular dose for one breast: - SFM: 2.17mGy (95% CI: 2.00-2.34) - FFDM: 1.25mGy (95% CI: 1.16-1.34) Tumor characteristics: no sign diff Breast density: higher in SFM - Entirely fat (cat 1): 12.2% SFM, 15.2% FFDM - Heterogeneously or extremely dense (cat 3 or 4): 38.0% SFM, 21.9%	

					FFDM	

Appendix 3.2.3. Ultrasound

Table 35 Study characteristics systematic reviews ultrasound in breast cancer screening

I Study ID	II Method	III Patient characteristics	IV Intervention(s)	V Results primary outcome	VI Results secondary and other outcomes	VII Critical appraisal of review quality
Bermejo-Perez 2008[43]	• Design: SR • Funding: Andalusian HTA agency • Search date: 1996-2005 • Searched databases: MEDLINE, EMBASE, Cochrane Library, Clinicaltrials.gov, National Research Register of the National Health Service, Centre for Reviews and Dissemination databases, websites related to study topics and references of included studies • Included study designs: prospective cohort studies • 3 studies included	• Eligibility criteria: Asymptomatic BRCA1- & BRCA2- carriers with or without personal cancer history • Patient characteristics: total number of women: 43-236. Mean age 38.9-46.6 years	• Index test: ultrasound (within program with other modalities) • Diagnostic threshold: ultrasound BIRADS-US ≥4 or use of specific scale • Reference standard: pathology (biopsies) + follow-up for interval cancers (not in all studies)	• Sensitivity: 20-33% • Specificity: 91.2-96%	• Total number of cancers detected: 5-22	• Level of evidence: low • Results critical appraisal: methodological problems in all studies mainly related to gold standard and work-up selection bias. No blinding. Management of doubtful results not reported. Total number of cancers diagnosed in trials low.
Davidson 2007[44]	• Design: SR • Funding: New Zealand Ministry of Health • Search date: 1996-June 2006 • Searched databases:	• Eligibility criteria: asymptomatic women with high breast cancer risk, with or without known genetic	• Index test: ultrasound (together with other screening modalities)	• Sensitivity: 33.3-86.4% • Specificity: 90.5-99.4%	• Total number of cancers detected: 3-43 • Cancer detection rate: 20-32 per 1000 women under	• Level of evidence: low • Results critical appraisal: verification bias. US prone to inter-

	MEDLINE, EMBASE, Current Contents, NZ National Bibliographic database, NZ Ministry of health website, NZ university and medical library catalogues, NZHTA in-house collection, references of obtained material • Included study designs: prospective cohort studies • 4 studies included	mutation, with or without personal cancer history. Different risk stratification strategies used. • Patient characteristics: total number of women: 23-935. mean age 41.7-48.6.	• Diagnostic threshold: BIRADS-US ≥4/ not documented • Reference standard: pathology (biopsies) +/- follow-up for interval cancers (not in all studies)	• PPV:11.2-29.2% • NPV: 96.7%-98%	surveillance • Tumour characteristics: not specified for ultrasound only	observer variability. US often used in combination with other tests. Blinding not in all studies. Total number of cancers diagnosed in trials low. Short FU or high number lost of FU. Results Asian population may not be applicable to Western populations.
Irwig 2004[45]	• Design: SR • Funding: NHMRC • Search date: 1966-2002 • Searched databases: Medline, references of obtained material, experts contacted • Included study designs: cohort studies • 5 studies included	• Eligibility criteria: asymptomatic women with high breast cancer risk or young age or dense breast tissue. One study used ultrasound only if mammography normal • Patient characteristics: total number of women: 150-8970. Mean age: 42-54.7y	• Indextest: Ultrasound (together with other screening modalities) • Diagnostic threshold: not reported • Reference standard: pathology +/- follow-up for interval cancers (not in all studies)	• Sensitivity: 50-90.4% • Specificity: not reported	• Total number of cancers detected: 2-182	• Level of evidence: low • Results critical appraisal: small populations, no data on interval cancers, ultrasound highly operator dependent •
Nothacker 2009[46]	• Design: SR • Funding: German cancer Aid, German Society of Senology	• Eligibility criteria: asymptomatic women with negative	• Indextest: Ultrasound incremental to mammograph	• Sensitivity: 75.3%% • Specificity: 96.8%%	• Cancer detection rate: diagnosis of invasive cancer in 0.32% of women	• Level of evidence: moderate • Results critical appraisal: no

	• Search date: 2000-August 2008 • Searched databases: Pubmed, DARE, Cochrane-database 'Cochrane Reviews' and 'clinical Trials'. • Included study designs: cohort studies • 6 studies included	mammographic screening with dense breast tissue (BIRADS-US 2-4) • Patient characteristics: total number of women: 1517-13547. Median age: 47.6-60.7y	y • Diagnostic threshold: different for all studies • Reference standard: pathology +/- follow-up (only in 2 studies)	• PPV:2-28 % • NPV: 99.7%	screened. Highest proportion cancers diagnosed in BIRADS-US 3-4 women. • Tumour characteristics: median tumour size: 9-11mm. invasive cancers:81-100%. Node negatieve cancers: 86-100% • Biopsy rate: 2.3-4.7%. PPV of biopsies: 8.4-13.7%	information on FU in most studies, only 2 studies adequate FU. One study no consecutive inclusion.

Table 36 Study characteristics primary studies ultrasound in breast cancer screening included in systematic reviews

I Study ID	II Method	III Patient characteristics	IV Intervention(s)	V Results primary outcome	VI Results secondary and other outcomes	VII Critical appraisal of study quality
Kolb 1998[88]	• Design: prospective cohort • Source of funding: not stated • Setting: single centre, USA • Sample size: 11220, of whom 3626 asymptomatic with dense breast tissue. • Duration: Jan 1995-April 1997	• Eligibility criteria: asymptomatic women with dense breast (BIRADS-M D2-D4) and normal findings on CBE and mammography, with or without personal (16.8%) or family (21.7%) history of breast cancer. • Patient characteristics: mean age not stated. For 90% of patients prior mammograms were available for comparison.	• Index test(s): ultrasound – incremental to single reading mammography (double reading in retrospect) • Reference standard: diagnostic threshold for biopsy not clear. Positive defined by biopsy. No info on false negatives (interval cancers).	• Sensitivity, specificity, PPV, NPV: not calculated-	• 11/3626 (0.3%) women in screening group diagnosed with cancer. • To diagnose 11 cancers, 131 FNA, 45 biopsies and 188 repeat US after 4-6 months were performed • Mean size of US detected cancers: 11.9mm, 89% stage 0 or 1. • 5/11 (45.5%) US detected cancers in women with personal cancer history • No info on interval cancers, QoL, mortality	• Level of evidence: low • Dropouts: 19/273(7%) of pts in close FU lost of FU, no info for other patients • Results critical appraisal: patients prospectively included. Single reading mammography. Repeat CBE after US with retrospective exclusion of palpable cancers. No blinding. No info on interval cancers.
Buchberger 1999 [89]	• Design: prospective cohort • Funding: not stated • Setting: single centre, Austria • Sample size: 6800 (6113 screening + 687 symptomatic)	• Eligibility criteria: asymptomatic women with negative double reading mammography + women with palpable mass or mammographically identified mass. Dense breast tissue BIRADS-M D2-D4. Women with or without personal	• Index test(s): Ultrasound – incremental to double reading mammography • Reference standard: FNA/biopsy for positives, mammography (or FU US) for	• For total group: 94/103 (91%) cancers seen on US, of which 28 not on mammography (no FU for interval	• Negative-positive biopsy ratio 11.6/1. If benign looking lesions not biopsied: 7.7/1 • For each detected lesion, 243 initial US, 11 FU US, seven core needle biopsies, five fine-needle aspiration biopsies and 1 surgical biopsy had to	• Level of evidence: low • Dropouts: not stated • Results critical appraisal: symptomatic women and women with personal cancer history included. 48% of US detected cancers in

	• Duration: 1996-1998	cancer history. • Patient characteristics: mean age screening group: 48y	negatives (no FU interval cancers)	cancers) • PPV: 7.9% if all lesions biopsied	be performed. • No data on interval cancers (no FU), no data on QoL, mortality	screening group were in women with personal cancer history! No blinding, 2e CBE after US. No info on prevalent/incident rounds. No FU for interval cancers.
Buchberger 2000[90]	• Design: prospective cohort • Source of funding: not stated • Setting: single centre, Austria • Sample size: 8970 (8103 screening + 867 symptomatic) • Duration: 1996-2000	• Eligibility criteria: asymptomatic women with negative double reading mammography + women with palpable mass or mammographically identified mass. Dense breast tissue BIRADS-M D2-D4. Women with or without personal cancer history. • Patient characteristics: mean age. 49y • Prevalence of disease: 9.9%	• Index test(s): ultrasound – incremental to double reading mammography • Reference standard: Ultrasound scored positive if classified as indeterminate or malignant. FNA/biopsy for positives, mammography (or FU US) for negatives (no FU interval cancers)	• PPV: 13.7%	• Cancer detection rate: 0.46%. 15/32 (46.9%) detected in women with personal cancer history • 269 biopsies and 136 FNA performed to detect 40 cancers. (113 benign looking lesions also biopsied) • Negative-to-positive biopsy ratio 10.1:1 • For each cancer detected, 242.4 US, 3.4 ,FNA, 6.4 core biopsies and 0.6 surgical biopsies had to be performed. • 75% of lesions detected by US ≤ 10mm. Mean size: 9.1mm • No data on interval cancers (no FU), no data on QoL, mortality	• Level of evidence: low • Dropouts: not stated • Results critical appraisal: symptomatic women and women with personal cancer history included. 47% of US detected cancers in screening group were in women with personal cancer history! No blinding, 2e CBE after US with retrospective exclusion. No info on prevalent/incident rounds. No FU for interval cancers. •
Kaplan	• Design:	• Eligibility criteria:	• Index test(s):	.	• Cancer detection rate:	• Level of evidence:

2001[91]	prospective cohort • Source of funding: not stated • Setting: Single centre, USA • Sample size: 1862 women • Duration: 1998-2000	asymptomatic women with dense breast tissue BIRADS-M D3-D4 and negative mammography. Women with abnormal mammography or CBE are included for other quadrants. • Patient characteristics: age range 35-87y	ultrasound – incremental to single reading mammography • Reference standard: FNA/biopsy for positives, mammography (or FU US) for negatives. Incomplete FU on interval cancers.		0.3% • 97/1862 (5,2%) pts underwent at least 1 biopsy or FNA, ≥ 72/1862 (3,9%) FU US after 6 months. 2 pts had diagnostic MRI. • PPV biopsies: 6/51 (11,8%) cancers • Imprecise data on FU and interval cancers after 1 year: no interval cancers detected • Tumour characteristics: mean diameter 9mm, 100% node negative. • No data on QoL, mortality	moderate • Dropouts: 5/57 biopsy results not known. Imprecise data on FU • Results critical appraisal: Single reading mammography. No blinding. No data on prevalent/incident rounds. Symptomatic women included. Incomplete FU
O'Driscoll 2001[103]	• Design: prospective cross-sectional • Source of funding: Kodax bursary from the Royal College of Radiologists • Setting: single centre, UK • Sample size: 149 • Duration: April 1999-June 2000	• Eligibility criteria: asymptomatic women with moderate breast cancer risk, based on family history • Patient characteristics: mean age 45.15y (range 30-69y), 61% had previous mammography films available.	• Index test(s): ultrasound + single reading mammography • Reference standard: biopsy for positives, FU for false negatives.	• Sensitivity, specificity, PPV, NPV: not calculated-	• 1 (0.7%) cancer detected • 10/149 (6.7%) underwent biopsy, 9/10 based on US only • PPV of biopsies: 10% • 1 interval cancer with mean FU 13.7 months.	• Level of evidence: low • Dropouts: no info • Results critical appraisal: small sample size, low number of cancers detected. Single reading mammography. Blinding for US and mammography reading. No clear diagnostic threshold for biopsy

Kolb 2002[100]	• Design: prospective cross-sectional • Source of funding: not stated • Setting: single centre, USA • Sample size: 14278 examinations in 5418 women for US • Duration: January 1995 – September 2000	• Eligibility criteria: asymptomatic women with dense breast tissue (BIRADS-M D2-D4), with (16.7%) or without personal cancer history. Women with family history included. Women with abnormal mammography included. Total high risk women: 26.5%. • Patient characteristics: mean age 54.7y. 84% previous mammography films available •	• Index test(s): ultrasound + single reading mammography • Reference standard: biopsy for positives, negative findings on biopsy and other investigations for negatives (no FU for interval cancers)	<u>Ultrasound</u> • PPV: 20.5%	• 12193 US in 4897 women with normal mammography lead to 320 biopsies (1.9% of US) and diagnosis of 33 cancers (10.3%) • Cancer detection rate for US only: 0.23%, increasing with breast density categories and for high risk women. 37% of cancers in women with dense breast tissue detected by US only. • Mean size US detected cancers: 14.7mm; 61% stage 0 or 1. 89% of cancers found by US only node negative. • No data on recall rate, interval cancers, QoL, mortality	• Level of evidence: low • Dropouts: no info • Results critical appraisal: no blinding. Repeat CBE if abnormal findings. Indication for biopsy not fully reproducible. Not clear if palpable cancers were included in calculations. Interval cancers not included in calculation sensitivity, specificity and accuracy. Single reading mammography
Hou 2002[104]	• Design: cross-sectional • Source of funding: none • Setting: single centre, Taiwan • Sample size: 935 • Duration: May 1994-Aug 2001	• Eligibility criteria: asymptomatic women, high risk as relatives of breast cancer patients, ≥ 35y • Patient characteristics: mean age 48.6y	• Index test(s): ultrasound + single reading mammography • Reference standard: Biopsy of all lesions BIRADS-M or BIRADS-US ≥		• 121/935 (12.9%) abnormal US, 24 (2.5%) biopsies of which 19 (79.2%) malignancies. • 1 interval cancer with median FU time 41.8 months. • Mean size of detected cancers: 12mm	• Level of evidence: low • Dropouts: none • Results critical appraisal: single reading mammography. No blinding; interval cancers not included in

			4 and all solid lesions not obviously looking benign on ultrasound. Other imaging to define false negatives (FU cancer not included in calculations.		• No data on recall rate, QoL, mortality	calculation of sens & spec. Taiwanese high risk populations may not be representative for Western screening population.
Podo 2002[105]	• Design: cross-sectional • Source of funding: not stated • Setting: multi centre, Italy • Sample size: 105 • Duration: June 2000-March 2002	• Eligibility criteria: asymptomatic women and men with a high breast cancer risk, with or without BRCA mutation, with (38%) or without personal cancer history • Patient characteristics: mean age. 46.0y. 100% prevalent screen, 14/105 (13%) incident round	• Index test(s): ultrasound + mammography/ MRI • Reference standard: biopsy/FNA for positives, other imaging + 2y FU for negatives (but no info on FU)	• - Sensitivity, specificity, PPV, NPV: not calculated- • Only 1/8 (12.5%) cancers detected by ultrasound	• 8/105 (7.6%)women diagnosed with cancer. 5/8 (62.5%) women have personal cancer history • No data on biopsy rate, recall rate, interval cancers, QoL, mortality	• Level of evidence: low • Dropouts: only 13% incident round • Results critical appraisal: small sample size, 38% with personal cancer history. No blinding. No info on single/double reading mammography. No info on interval cancers, incomplete FU and no explanation on drop-outs.
Leconte 2003[112]	• Design: prospective cross-sectional • Source of funding: not stated	• Eligibility criteria: asymptomatic women with (24%) or without previous surgery for breast cancer. 3% of women had a palpable	• Index test(s): ultrasound (tissue harmonic imaging) + single reading		• 50 non-palpable cancers in 47 pts detected, in total cancer diagnosed in161 patients. 16/50 non-palpable cancers	• Level of evidence: low • Dropouts: no info • Results critical appraisal: no blinding. Single

	• Setting: single centre, Belgium • Sample size: 4236 • Duration: April 2000-March 2001	lesion and were included in the analysis for possible other lesions only. Women with entirely fatty breasts and normal mammography excluded. • Patient characteristics: median age 60y (range 41-87)	mammography • Reference standard: FNA if new or enlarged atypical cyst, core biopsy if FNA insufficient and for lesions only visible on mammography. True negatives defined by other imaging, no FU interval cancers.		detected by ultrasound only. • 25/50 cancers detected in patients with personal cancer history or symptoms. • Mean size US detected cancers: 10mm (range 2-30mm) • No data on biopsy rate, recall rate, interval cancers, QoL, mortality	reading mammography. FU and symptomatic patients included. No info on prevalent / incident rounds. Only 25 of 50 non-palpable cancers detected in screening patients. Interval cancers not included in calculations.
Crystal 2003[92]	• Design: prospective cohort • Source of funding: not stated • Setting: single centre, Israel • Sample size:1517 • Duration: Jan 2000-Jan 2002	• Eligibility criteria: asymptomatic women with dense breast tissue (BIRADS-M D2-D4) and negative mammography, with or without personal/family history of breast cancer • Patient characteristics: mean age 52.1y. 318/1517 considered high risk based on personal or family cancer history.	• Index test(s): ultrasound – incremental to single reading mammography • Reference standard: biopsy/FNA for positives. Biopsy for all solid lesions, FNA of complex cysts in selected cases. Not clear if interval cancers in calculations (none detected, incomplete FU	• Sensitivity: 100% • Specificity: 94.4%	• Cancer detection rate: 0.42%. For average risk women: 0.25% • 38/1517 (2.5%) biopsy or FNA. For average risk women: 2.3% • 62/1517 (4.1%) FU US after 6 months • In 8/38 (21.1%) biopsies/FNA cancer diagnosed. 4/8 cancers in high risk pts. No cancers detected in BIRADS 2 women. • Mean size tumours diagnosed: 9.6 mm (range 4-12mm). 1 LN positive.	• Level of evidence: low • Dropouts: no info • Results critical appraisal: no blinding. Probably single reading mammography. Retrospective review of CBE/mammography after US: patients excluded if positive in retrospect. Incomplete FU for interval cancers (range 8-30 months).

			8-30 months)		• No interval cancers detected with FU range 8-30 months • No data on QoL, mortality	
Trecate 2003[84]	• *Design: cross-sectional* • *Source of funding: not stated* • *Setting: single centre, Italy* • *Sample size: 23* • *Duration: 7 month period*	• *Eligibility criteria: asymptomatic and symptomatic (3/23) women with high breast cancer risk (>50%), with (5/23) or without BRCA mutation, with or without personal cancer history* • *Patient characteristics: age range 30-61y* • *Prevalence of disease*	• *Index test(s): ultrasound + mammography/ MRI* • *Reference standard: biopsy/FNA for positives, other imaging for negatives*	• *Not stated for US. 0/4 cancers detected by US??*	• *No separate results for US*	• *Level of evidence: very low* • *Dropouts: none* • *Results critical appraisal: small sample size, symptomatic patients included. No blinding. No FU for interval cancers. No data specific for US.*
Sim 2004[85]	• *Design: retrospective cross-sectional* • *Source of funding: not stated* • *Setting: single centre, the Netherlands* • *Sample size: 84* • *Duration: 1994-2001*	• *Eligibility criteria: asymptomatic women with high breast cancer risk (>15%) for whom sufficient FU data were available, with or without personal cancer history.* • *Patient characteristics: mean age 42.4y for biopsied women.*	• *Index test(s): ultrasound + mammography/ MRI* • *Reference standard: Imaging positive if BIRADS score ≥ 4. Confirmation by histopathology or 2y FU*	*For Ultrasound:* • *Sens: 83.3%* • *Spec: 65.5%* • *PPV: 50%* • *NPV: 90.5%* • *Accuracy: 70.7%* *For US + mammography:* • *Sens: 92.9%* • *Spec: 62.5%* • *PPV: 52%* • *NPV: 95.2%* • *Accuracy: 71.7%*	• *Malignancy in 31.3% of biopsies (based on all imaging performed), benign-malignant ratio thus 2 to 1*	• *Level of evidence: very low* • *Dropouts: 66/245 women excluded for insufficient FU* • *Results critical appraisal: Small sample size. No consecutive inclusion, only selected patients had ultrasound, retrospective. No blinding. No info on handling of intermediate results. Definition of true/false negatives*

						and cases included in calculations unclear.
Warner 2004[106]	• Design: prospective cross-sectional • Source of funding: Canadian Breast Cancer Research Alliance, Terry Fox Foundation, International Breast MRI consortium • Setting: single centre, Canada • Sample size: 236 • Duration: Nov 1997-March 2003	• Eligibility criteria: asymptomatic BRCA 1 or BRCA 2 mutation carriers with (30%) or without personal breast cancer history ≤ 91 kg. • Patient characteristics: mean age 46.6y (range 26.4-64.8y)	• Index test(s): ultrasound + CBE/single reading mammography/ MRI • Reference standard: Biopsy if one of the modalities suspicious. Additional diagnostic studies other than biopsies not included in definition of false positives. False negatives defined by cancers detected by other modalities + interval cancers during 3y FU!	For 1st round, US: • Sens: 25% • Spec: 95% • PPV:23% • NPV: 96% For 2nd round US: • Sens: 57% • Spec:96% • PPV:44% • NPV:98%	• 22 cancers in 21 women detected by 4 modalities • 33% of detected cancers in women with personal cancer history. • 1 interval cancer detected after 3rd screening round, 1 DCIS diagnosed in prophylactic mastectomy specimen. • 16/22 (73%) invasive cancers • After 1st round: 5.1% of US resulted in FU US after 6 months • No data on biopsy rate for US only • No data on QoL, mortality	• Level of evidence: moderate • Dropouts: 31/236 (13.2%) women left study before completing 3rd round. FU continued as much as possible. • Results critical appraisal: single reading mammography. Blinding for other modalities. 10 MRI-guided biopsies excluded. Additional diagnostic studies other than biopsies not included in definition of false positives. 33% of detected cancers in women with personal cancer history.
Kuhl 2005[94]	• Design: prospective cross-sectional • Source of funding: Förderverein für	• Eligibility criteria: asymptomatic women with life time breast cancer risk ≥ 20%, with or without personal cancer history, with or	• Index test(s): ultrasound + CBE/double reading mammography/ MRI	For women without personal cancer history: <u>US</u> • Sens: 38.7% • Spec: 91%	• 43 cancers diagnosed in 41 women, of which 12 in 11 women with history of breast cancer, of which 3 classified as	• Level of evidence: moderate • Dropouts: 49 women lost of FU after first round not included in sample

	Radiologie an der Universität Bonn, German Cancer Aid. • Setting: single centre, Germany • Sample size: 529 • Duration: Feb 1996 - Feb 2002	without proven mutation. • Patient characteristics: mean age. 41.7y. 139/529 (26.3%) with personal cancer history • Prevalence of disease: 26.5/1000 women	• Reference standard:validation of positive findings by histology, of negative findings by FU (mean FU 5.3y) . For BIRADS-US 4-5 biopsy indicated except if benign correlate on mammo or MRI. US BIRADS 3: short term FU +/- biopsy if still BIRADS-US ≥ 3	• PPV:10.4% US + mammography • Sens 51.6% • Spec: 89.4% • PPV:11.7% mammography: • Sens: 32.3%% • Spec: 97.1% • PPV: 23.3%%	local recurrence • 16.7% of ultrasound BIRADS 3 with 6 month FU US recommended • False positive diagnosis (BIRADS 4-5) on US in 134 women, 78 not biopsied because of benign correlate on mammo or MRI • No data on biopsy rate, interval cancers, QoL, moratlity	of 529 • Results critical appraisal: no blinding for CBE, blinding for other imaging. Biopsy rate and recall rate influenced by other modalities. BIRADS-3 with 6 month FU not considered positive result. Lobular carcinoma in situ considered benign. Mean FU 5.3y.
Corsetti 2008[99]	• Design: prospective cross-sectional • Source of funding: not stated • Setting: single centre, Italy • Sample size: 9157 • Duration: Jan 2000-Feb2007	• Eligibility criteria: asymptomatic and symptomatic women with negative mammography and dense breast tissue (BIRADS-M D3-D4) • Patient characteristics: mean age. 52y	• Index test(s): ultrasound + single reading mammography. • Reference standard: Biopsy if BIRADS-US ≥ 3 for positives. No FU for interval cancers	• Sensitivity, specificity, PPV, NPV: not calculated	• 37 cancers detected by US only in asymptomatic subjects. Incremental detection rate 0.40% for asymptomatic women. • 33/37 mammograms retrospectively reviewed (blinded): 8/33 (24%) positive • Additional investigations in 449/9157 (4.9%) subjects. 490 FNA, 24 core biopsies and 133	• Level of evidence: low • Dropouts: no info • Results critical appraisal: single reading mammography. No blinding. No FU for interval cancers. No info on prevalent/incident rounds or on number of rounds per woman. Symptomatic women included.

					surgical biopsies performed. • benign findings in399/449 (88.9%) women (symptomatic included). • No data on interval cancers, QoL, mortality	False positives calculated on surgical biopsies only.

Table 37 Study characteristics primary studies ultrasound in breast cancer screening published after 2008

Brancato 2007[93]	• Design: prospective cohort • Source of funding: not stated • Setting: single centre, Italy • Sample size:5227 • Duration: January 2003-December 2006	• Eligibility criteria: asymptomatic women with normal screening mammography en dense breasts BIRADS-M D3-D4 • Patient characteristics: no details given	• Index test: ultrasound incremental to normal mammography • Reference standard: cytology or biopsy for BIRADS-US 3-5 lesions. No FU for interval cancers.		• 2/5227 women diagnosed with cancer, cancer detection rate 0.38 per 1000 women, a 6.5% increase compared to mammography alone • Recall rate 2,1% • Total cost per detected cancer: 145 496.53 EUR • No info on biopsy rate, interval cancers, mortality	• Level of evidence • Dropouts: only 20% of eligible pts had US • Results critical appraisal: not all consecutive patients included due to organisational problems. Symptomatic patients appear also included. Probably single reading mammography. No info on interval cancers.
Honjo 2007[95]	• Design: cross-sectional • Source of funding: Ministry of Health, labour and Welfare Japan • Setting: multi-	• Eligibility criteria: asymptomatic women ≥ 40y from general population • Patient characteristics: not stated	• Index tests: CBE + ultrasound + double reading mammography • Reference standard: diagnostic	<u>Ultrasound</u> • Sens: 53.8% • Spec: 95.4% <u>Mammography</u> • Sens: 61.5% • Spec: 92.1% <u>US + Mx</u>	• Recall rate 15.3% for combined examinations, 4.8% for ultrasound • Detection rate overall: 0.29%, • No data on biopsy	• Level of evidence • Dropouts: not stated • Results critical appraisal: Asian population, young women included. Blinding for other

	centre, Japan • Sample size: 3455 • Duration: October 1999-March 2000	• Prevalence of disease: 36.0/100 000 women age-standardized incidence rate	threshold unclear. Biopsy for positives, FU for interval cancers	• Sens: 84.6% • Spec: 88.4%	rate, mortality. Info on interval cancers unclear	imaging. Diagnostic threshold and definitions of true/false positives and negatives not clear.
Lehman 2007[107]	• Design: cross-sectional • Source of funding: NCI + Office of Women's health, gadolinium-based contrast delivered by companies • Setting: multicentre, USA • Sample size:195 • Duration: November 2002 – April 2003	• Eligibility criteria: women > 25y with high breast cancer risk based on genetic analysis or family history • Patient characteristics: mean age 45.4y, 24.7% personal history of breast cancer	• Index tests: CBE + US+ single reading mammography + MRI • Reference standard: Positive exam = BIRADS-US ≥ 3. Biopsy and other imaging to define true/false positives and negatives. No FU for interval cancers		• Recall rate US: 9% • Biopsy rate US: 2.3% • PPV biopsies: 25% • Diagnostic yield US: 0.3% • Additional cancer yield (not detected by MRI or mammography): 0% • Tumour characteristics: one T2N1M0 detected by ultrasound • No info on interval cancers, mortality	• Level of evidence • Dropouts: 24/195 (12.3%) • Results critical appraisal: maximum delay between imaging 90 days. Blinding for other imaging. (Probably) single reading mammography. No FU for interval cancers.
Riedl 2007[108]	• Design: cross-sectional • Source of funding: Medizinisch-Wissenschaftlicher Fonds des Bürgemeister der Bundesshauptsadt and Jubiläums Fonds der Österreichischen Nationalbank • Setting: single centre, Austria	• Eligibility criteria: asymptomatic women with high breast cancer risk based on genetic analysis or family history, with or without personal cancer history • Patient characteristics: median age 41y (range 22-80y)	• Index test(s): single reading mammography + ultrasound + MRI • Reference standard: histopathology + FU. Biopsy for all BIRADS 4-5 lesions on at least 1 imaging. BIRADS 3: FU 6 months, considered	<u>US</u> • Sens: 42% • Spec:97% • PPV: 42.1% • NPV: 96% <u>Mammography</u> • Sens: 50% • Spec: 97% • PPV:61.5% • NPV: 96.6%	• 1 interval cancer detected. • All US detected cancers also detected on mammography • False positive rate US: 68% • No data on biopsy rate, total recall rate, mortality	• Level of evidence • Dropouts: 8% • Results critical appraisal: single reading mammography. Blinding for other imaging. PPV and NPV calculated per breast not per woman.

	• Sample size:327 • Duration: 1999-2006		negative exam.			
Berg 2008[109]	• Design: RCT (order of screening investigations randomized) • Source of funding: Avon foundation, National Cancer Institute • Setting: multicentre; USA, Canada, Argentinia • Sample size: 2725 • Duration: April 2004-February 2006	• Eligibility criteria: asymptomatic women with elevated breast cancer risk, based on personal cancer history, Gail/Claus model, chest RT or gene mutation and heterogeneously or extremely dense breast tissue in at least 1 quadrant • Patient characteristics: median age: 55y (range 25-91y), 53.09% with personal cancer history, 21% current chemoprevention. 73% had previous mammography • Prevalence of disease: unknown	• Index test(s): ultrasound + single reading mammography in randomized order • Reference standard: biopsy proven cancer (in situ or invasive) within 1y for disease positive, no cancer diagnosis within 1y FU for disease negative. BIRADS 3 lesions considered negative.	<u>US + mammography</u> • Sens: 77.5% • PPV: 7.3% • AUC: 0.91 <u>US</u> • PPV: 6.5% • AUC: 0.80 <u>mammography</u> • Sens:50% • PPV: 7.6% • AUC: 0.78	• 31 cancers detected, diagnostic yield 11.8 per 1000 women. 12/31 cancers seen on US only, increased yield of US 4.2 per 1000 women. • Tumour characteristics US detected cancers: median size 10mm (range 5-40mm). 8/9 (89%) cancers node negative • 8 interval cancers diagnosed + 1 excluded (?) • % positive biopsies 22.6% for mammography, 8.9% for US, 11.2% for US + mammography • Total recall rate: 12.7% for mammography, 21.4% for ultrasound, 27.4% for US + mammography • No info on mortality	• Level of evidence • Dropouts: 172 (75 no complete reference standard) • Results critical appraisal: single reading mammography. Blinding for other imaging. Majority of cancers detected in women with personal cancer history.
Daguet	• Design: cross-	• Eligibility criteria:	• Index tests: MRI +	<u>US</u>	• 7 cancers detected	• Level of evidence

2008 [96]	sectional • Source of funding: not stated • Setting: single centre, France • Sample size: 85 • Duration: December 2000-February 2006	women with BRCA 1 or 2 or p53 (1/85) mutation, with (43.5%) or without personal cancer history • Patient characteristics: mean age 43y (range 27-65y), 24.7% pBSO	US + double reading mammography • Reference standard: FNA or biopsy for BIRADS ≥ 3. If BIRADS 3 on MRI only, short term FU. Biopsy and repeat imaging + FU to define true/false positives and negatives. BIRADS 3 defined as negative exam.	• Sens: 50% • Spec: 97.3% • PPV: 40% • NPV: 98.2% <u>Mammography</u> • Sens: 12.5% • Spec: 98.7% • PPV: 25% • NPV: 96.9%	on screening imaging, 1 interval cancer. 4/8 cancers detected in women with personal cancer history. 2/7 screening detected cancers palpable on CBE. • No info on total recall rate, biopsy rate, mortality	• Dropouts: 1 woman with interval cancer excluded as she did not have mammography. 6/85 (7%) quit study • Results critical appraisal: no clear consecutive inclusion of patients. No blinding. Maximum interval between mammography and MRI 6 months (median 12 days). 50% of cancers detected in women with personal cancer history
Weinstein 2009 [110]	• Design: cross-sectional • Source of funding: National Institutes of Health • Setting: single centre USA • Sample size:612 • Duration: May 2002-July 2007	• Eligibility criteria: BRCA mutation carriers, women with ≥ 25% life time risk of breast cancer, previous LCIS or atypical hyperplasia or chest wall radiotherapy. Women with recent breast cancer included for contralateral breast. All women normal FSM 180d before	• Index tests: FFDM + US + MRI • Reference standard: biopsy for all BIRADS 4-5 lesions (consensus of all imaging). Biopsy + 2y FU to define true/false positives and negatives. BIRADS 0 and 3 for each modality considered 'positive'	<u>US</u> • Sens: 17% • Spec: 88% <u>FFDM</u> • Sens: 39% • Spec:91% <u>MRI</u> • Sens: 71% • Spec: 79%	• Overall cancer yield 3%, cancer yield US: 0.5% • Recall rate US: 79/567 (13.9%) • Biopsy rate US: 20/567 (3.5%) • No info on mortality	• Level of evidence • Dropouts: 3/612 (0.5%) • Results critical appraisal: women with personal cancer history included. Initial reading with blinding followed by consensus evaluation of all imaging modalities. Only prevalent round. Probably no

		intervention. • Patient characteristics: median age 41y (range 27-81y), 41.2% cancer in contralateral breast					interval cancers detected, unclear in text. No info on single or double reading mammography.
Tohno 2009[97]	• Design: cross-sectional • Source of funding: not stated • Setting: general screening program, Japan • Sample size: 48294 • Duration: April 2004-March 2006	• Eligibility criteria: asymptomatic women from the general population, aged 30-69y • Patient characteristics (e.g. age, tumour characteristics, stage, etc.)	• Index tests: ultrasound + double reading mammography • Reference standard: diagnostic threshold and reference standard not clearly defined in text.	• Sensitivity, specificity, PPV, NPV: not calculated	• Recall rate US: 4%, recall rate mammography: 4.3% • Cancer detection rate US: 0.15%, detection rate mammography 0.21% • 1/3 cancers detected by US or mammography only • No data on biopsy rate, interval cancers, mortality	• Level of evidence • Dropouts: not stated • Results critical appraisal: Asian population, young women included. No info on blinding. Diagnostic threshold and methodology calculation not described in text. No data on interval cancers.
Kuhl 2010[98]	• Design: cross-sectional • Source of funding: German Cancer Aid Society • Setting: multi-centre, Germany • Sample size: 687 • Duration: October 2002-December 2005	• Eligibility criteria: women with high breast cancer risk based on mutation analysis or family history, with (27%) or without personal cancer history • Patient characteristics: median age 44y (range 25-71y)	• Index tests: MRI + US + double reading mammography • Reference standard: BIRADS 1-2-3 taken as negative, BIRADS 4-5 taken as positive. Biopsy + 1y FU to define true/false positives and negatives.	<u>US</u> • Sens: 37% • Spec:98 % • PPV: 35.7% • NPV: 98.9% <u>US+ Mx</u> • Sens: 48.1% • Spec: 98.3% • PPV: 42.5% • NPV: 99.1% <u>Mx</u> • Sens: 33.3% • Spec:99.1% • PPV: 39.1%	• 27 cancers detected, 9/27 (33.3%) in women with personal cancer history. 25/27 detected by MRI • No interval cancers detected • 21/27 (77%) cancers ≤ 10mm • 136/687 (19.8%) BIRADS 3 diagnosis on US, requiring short-term FU	• Level of evidence • Dropouts: 38/725 (5.2%) • Results critical appraisal: No info on blinding. 1/3 cancers detected in women with personal cancer history, CBE positive in 110 screening rounds. Double reading mammography

				• NPV: 98.9%	• No cancer detected at half-yearly screening round with CBE + US • No data on mortality	
Kelly 2010[101]	• Design: cross-sectional • Source of funding: Sonocine, Inc. Two authors are (majority) shareholders of the company • Setting: multicentre, USA • Sample size:4419 • Duration: January 2003 – July 2007	• Eligibility criteria: asymptomatic women with BIRADS-D D3-4 dense breasts, family or personal cancer history and/or implants. Very obese women excluded. • Patient characteristics: median age 53y (range24-89y), 10% personal cancer history, 11% implants	• Index test: automated whole-breast ultrasound + single reading mammography • Reference standard: additional imaging for BIRADS 3 (based on US + mammography). Biopsy for BIRADS 4-5. Biopsy + FU to define true/false positives and negatives.	<u>US</u> • Sens: 67% • Spec: 89.9% <u>US + Mx</u> • Sens: 81% • Spec: 98.7% <u>mammography</u> • Sens: 40% • Spec: 95.2%	• 57 cancers diagnosed, 18/57 (31.6%) in women with personal cancer history. 40% of cancers diagnosed by US only. 11/57 (19%) interval cancers • Recall rate 7.2% for US, 4.8% for mammography and 9.6% for US + mammography • PPV for biopsies generated by US: 38.4%. For biopsies generated by mammography: 39% • No info on mortality	• Level of evidence • Dropouts: 6 paired examinations incomplete, 50pts excluded, 11 biopsies excluded. 1y FU available for 80% of pts. • Results critical appraisal: 1434/6425 (22.3%) US performed as FU of previous abnormal findings. Single reading mammography. Blinding. Obese patients excluded. Part of women alternated US and mammography every 6 months.
Youk 2011[86]	• *Design: prospective cohort* • *Source of funding: Yonsei University College of Medicine* • *Setting: single centre, South Korea*	• *Eligibility criteria: women with dense breasts BIRADS D3-4 and negative single reading mammography with (61.8%) or without personal cancer history. 9.2%*	• *Index test: ultrasound incremental to negative single reading mammography* • *Reference standard: biopsy for BIRADS-US 4-*	• *Sensitivity, specificity, PPV, NPV: not calculated*	• *43/1507 (2.9%) pts diagnosed with cancer. 22 cancers detected by diagnostic ultrasound, 10 cancers detected in patients with personal cancer*	• *Level of evidence* • *Dropouts: 2313/3820 (60%) excluded due to lack of FU or confirmation histology* • *Results critical appraisal: US as*

	• *Sample size: 1507* • *Duration: July 2001-june 2005*	*diagnostic examinations included.* • *Patient characteristics: median age 47y (range 21-74y)*	*5 lesions, 6 month FU for BIRADS-US 3 lesions. Biopsies and 2y FU to define true/false positive and negative cases.*		*history.* • *No interval cancers during at least 2y FU.* • *Total recall rate for BIRADS-US 3-4-5: 19.5%* • *For screening pts without personal cancer history: 11.4% BIRADS 4-5, PPV of biopsies 20.4%, 22.4/1000 cancer detection rate, mean size cancer 13mm, 12.5% node positive* • *No info on mortality*	*adjunct to single reading mammography. > 60% dropouts and > 60% women with personal cancer history, 9.2% diagnostic examinations.*
Sardanelli 2011 [111]	• Design: cross-sectional • Source of funding: Italian Ministry of health • Setting: multicentre, Italy • Sample size: 501 • Duration: June 2000-January 2007	• Eligibility criteria: asymptomatic women with high breast cancer risk based on mutation analysis or family history, with (43.5%) or without personal cancer history. • Patient characteristics: median age 45y (range 22-79y)	• Index tests: CBE + US + mammography + MRI • Reference standard: biopsy for BIRADS 4-5 on any imaging or positive CBE. Short term FU for BIRADS 3. BIRADS 3 considered negative exam. Biopsy/FNA and 1 y FU to define true/false positives and negatives	US • Sens: 52% • Spec: 98.4% • PPV: 61.9% • NPV: 97.7% US+mammography • Sens: 62.5% • Spec: 97.6% • PPV: 55.6% • NPV: 98.2%	• 49 cancers detected through screening, 3 interval cancers. 29 (56%) cancers diagnosed in women with personal cancer history • No info on recall rate, mortality	• Level of evidence • Dropouts 85%,67%,46% underwent 2nd, 3rd and 4th round respectively • Results critical appraisal: probably single reading mammography. 56% of cancers diagnosed in women with personal cancer history. No info on blinding. Handling of missing data not clear from text, may

						influence results.
Corsetti 2011[102] (update Corsetti 2008)	• Design: prospective cross-sectional • Source of funding: not stated • Setting: single centre, Italy • Sample size: 8865 • Duration: 2001-2006	• Eligibility criteria: asymptomatic and symptomatic women with negative mammography and dense breast tissue (BIRADS-M D3-D4) • Patient characteristics: mean age. 20% prevalent screens	• Index test(s): ultrasound + single reading mammography. • Reference standard: Biopsy if BIRADS-US ≥ 3 for positives, 1 year FU for interval cancers	<u>US + mammography</u>: • Sens: 86.7%	• 21 interval cancers diagnosed in 1y FU, meaning 1.07/1000 negative screening examinations • Additional testing (mostly fine needle biopsy) due to false positive ultrasound in women with dense breasts: 5.5% • No data on total recall rate, mortality	• Level of evidence: Dropouts: no info • Results critical appraisal: update Corsetti 2008, see critical appraisal 2008. single reading mammography. No blinding. Symptomatic women included. Additional imaging and short term FU not reported.
Lenz 2011[87]	• *Design: prospective cohort* • *Source of funding: Fonden for Faglig Udvikling i Speciallaegepraksis* • *Setting: single centre, Denmark* • *Sample size: 1428* • *Duration: 1997-2007*	• *Eligibility criteria: women > 40y, women with high breast cancer risk or on patient's request. Symptomatic patients included.* • *Patient characteristics: no info*	• *Index tests: CBE + US +/- mammography* • *Reference standard: mammography and biopsy/FNA for all solid tumours and not simple cysts. Biopsy/FNA and 1y FU to define true/false positives and negatives.*	*<u>Ultrasound</u>* • *Sensitivity: 89%*	• *25/28 (89%) seen by ultrasound. 13/25 (52%) non-palpable.* • *Mean size of detected tumours 11mm (range 4-30mm)*	• *Level of evidence* • *Dropouts: no info* • *Results critical appraisal: Not clear if all consecutive patients were included. No info on blinding, limited info on mammography. 48% of cancers detected in symptomatic patients.*

Appendix 3.2.4. MRI

Table 38 Study characteristics systematic reviews MRI in breast cancer screening

I Study ID	II Method	III Patient characteristics	IV Intervention(s)	V Results primary outcome	VI Results secondary and other outcomes	VII Critical appraisal of review quality
Bermejo-Perez 2008[43]	• Design: SR • Funding: Andalusian HTA agency • Search date: 1996-2005 • Searched databases: MEDLINE, EMBASE, Cochrane Library, Clinicaltrials.gov, National Research Register of the National Health Service, Centre for Reviews and Dissemination databases, websites related to study topics and references of included studies • Included study designs: prospective and retrospective cohort-design • Number of included studies: 8	• Eligibility criteria: Asymptomatic BRCA1- & BRCA2- carriers with or without personal cancer history • Patient characteristics: Total number of women included: 24-236. Mean age 38.9-46.6 years	• Index test: MRI • Diagnostic threshold: BIRADS 3-4 or use of specific scale • Reference standard: pathology (biopsies) +/- follow-up for interval cancers	<u>MRI:</u> Sensitivity: 77-100% Specificity: 81-97.5% <u>Mammography:</u> Sensitivity: 0-50% Specificity: 96.9-99.8% <u>US</u> Sensitivity: 20-33% Specificity: 91.2-96%	Total number of cancers detected: 1-22	• Level of evidence: low • Results critical appraisal: methodological problems in all studies mainly related to gold standard and work-up selection bias. No blinding. Management of doubtful results not reported. Total number of cancers diagnosed in trials low.
Davidson 2007[44]	• Design: SR • Funding: New Zealand Ministry of Health • Search date: 1996-June 2006 • Searched databases: MEDLINE, EMBASE,	• Eligibility criteria: asymptomatic women with high breast cancer risk, with or without known genetic mutation,	• Index test: MRI • Diagnostic threshold: BIRADS ≥4/not reported	<u>MRI</u> • Sensitivity: 71.1-90.7% • Specificity: 81-97.2% • PPV: 32.3-	• Total number of cancers detected: 1-51 • Tumour characteristics: mean size: 11-20mm. 0-33.3%	• Level of evidence: low • Results critical appraisal: verification bias. MRI used in program with other

	Current Contents, NZ National Bibliographic database, NZ Ministry of health website, NZ university and medical library catalogues, NZHTA in-house collection, references of obtained material • Included study designs: retrospective and prospective cohort studies • 10 studies included	with or without personal cancer history. Different risk stratification strategies used. • Patient characteristics: total number of women: 23-1909. mean age: 40-46.6y	• Reference standard: pathology (biopsies) +/- follow-up for interval cancers	50% • NPV: 99-99.7% • AUC 0.83-0.89 <u>Mammography</u> • Sensitivity: 32.6-40% • Specificity: 93-99.8%	node positive.	imaging techniques. No data on FU or short FU. Blinding not in all studies. No data on mortality, no comparison with no survaillance
Irwig 2004[45]	• Design: SR • Funding: NHMRC • Search date: 1966-2002 • Searched databases: Medline, references of obtained material, experts contacted • Included study designs: cohort studies • 4 studies included	• Eligibility criteria: asymptomatic women with high breast cancer risk or dense breast tissue. Different risk stratification strategies used. One study used MRI only if mammography normal • Patient characteristics: total number of women: 105-196. Mean age: 39-43y	• Indextest: MRI • Diagnostic threshold: not reported • Reference standard: pathology +/- follow-up.	<u>MRI</u> • Sensitivity: 100% • Specificity: not reported • False positive rate (% requiring biopsy): 5-9% <u>Mammography</u> • Sensitivity: 0-46% • Specificity: not reported • False positive rate (% requiring	• Total number of cancers detected: 6-12	• Level of evidence: low • Results critical appraisal: small populations, low number of cancers detected. No full assessment of accuracy, no reports on interval cancers.

				biopsy): 1-7%		
Lord 2007[47]	• Design: SR • Funding: Department of Health, Commonwealth of Australia • Search date: 1966-March 2007 • Searched databases: medline, Pre-Medline, EMBASE, the Cochrane Library, websites of HTA agencies • Included study designs: cohort studies • 5 studies included	• Eligibility criteria: asymptomatic women with high breast cancer risk, with or without personal cancer history. Different risk stratification strategies used. MRI used as incremental test to mammography +/- US, CBE. • Patient characteristics: total number of women:236-649. Mean age: 40-47y	• Indextest: MRI • Diagnostic threshold: BIRADS ≥3 or ≥4 • Reference standard: pathology +/- follow-up.	<u>Screening strategy with MRI</u> • Sensitivity: 86-100% • Specificity: 91-97% <u>Mammography + US</u> • Sensitivity: 49-67% • Specificity: 89% <u>Mammography</u> • Sensitivity: 25-59% • Specificity: 93-99.8%	• Total cancers detected 1st year: 1-6% • 74-78 additional recalls per 1000 screening rounds • Tumour characteristics: 15-32% of cancers ≥ 20mm. 8-23% node positive. •	• Level of evidence: low • Results of critical appraisal: no report on consecutive inclusion. Only three studies reported on FU and interval cancers (false negatives)

Table 39 Study characteristics primary studies MRI in breast cancer screening published 2007-2011

Hagen 2007[143]	• Design: cross-sectional • Source of funding: not stated • Setting: multicentre, Norway • Sample size: 554 • Duration:2002-2005	• Eligibility criteria: asymptomatic BRCA1 or 2 mutation carriers, with or without personal cancer history • Patient characteristics: mean age 41y, 50.3% pBSO	• Index test(s): mammography +/- US + MRI • Reference standard: BIRADS 3 for short term FU, BIRADS 4-5 for biopsy. Biopsy + FU for interval cancers to define true/false positives and negatives. Median FU 0.5y	<u>MRI</u> • Sens: 68% <u>Mammography +/- US</u> • Sens: 33.3%	• 20 cancers detected at screening, 5 interval cancers with a median FU 0.5y • Cancer detection rate at prevalent round: 2.7% • No info on recall rate, biopsy rate and mortality	• Level of evidence • Dropouts: 445/554 underwent screening (80%) • Results critical appraisal: no info on single or double reading mammography. No info on blinding. Median FU of 0.5y only.
Lehman 2007[107]	• Design: cross-sectional • Source of funding: NCI + Office of Women's health, gadolinium-based contrast delivered by companies • Setting: multicentre, USA • Sample size:195 • Duration: November 2002 – April 2003	• Eligibility criteria: women > 25y with high breast cancer risk based on genetic analysis or family history, with (24.7%) or without personal cancer history. • Patient characteristics: mean age 45.4y,	• Index tests: CBE + US+ single reading mammography + MRI • Reference standard: BIRADS ≥ 3 considered positive exam. Biopsy and other imaging to define true/false positives and negatives. No FU for interval cancers	•	• Recall rate MRI: 24% • Biopsy rate MRI: 8.2% • PPV biopsies: 43% • Diagnostic yield MRI: 3.5% • Additional cancer yield (not detected by US or mammography): 2.3% • Tumour characteristics: 4/5 T0-1, 1/5 T2. 1/5 node positive • No info on interval cancers, mortality	• Level of evidence • Dropouts: 24/195 (12.3%) • Results critical appraisal: maximum delay between imaging 90 days. Blinding for other imaging. Probably single reading mammography. No FU for interval cancers. Contrast delivered by companies.
Riedl 2007[108]	• Design: cross-sectional • Source of funding:	• Eligibility criteria: asymptomatic women with high breast cancer risk	• Index test(s): single reading mammography + ultrasound + MRI	<u>MRI</u> • Sens: 85% • Spec:88%	• 1 interval cancer detected. • 43% of cancers	• Level of evidence • Dropouts: 8% • Results critical

	Medizinisch-Wissenschaftlicher Fonds des Bürgemeister der Bundesshauptsadt and Jubiläums Fonds der Österreichischen Nationalbank • Setting: single centre, Austria • Sample size:327 • Duration: 1999-2006	based on genetic analysis or family history, with or without personal cancer history • Patient characteristics: median age 41y (range 22-80y)	• Reference standard: Biopsy for all BIRADS 4-5 lesions on at least 1 imaging. BIRADS 3: FU 6 months. Biopsy + FU to define true/false positives and negatives. BIRADS 3 considered negative exam.	• PPV: 48%% • NPV: 99.4% <u>Mammography</u> • Sens: 50% • Spec: 97% • PPV:61.5% • NPV: 96.6%	detected by MRI only • No info on total recall rate, biopsy rate, mortality	appraisal: single reading mammography. Blinding for other imaging. PPV and NPV calculated per breast not per woman.
Peters 2008[123]	• Design: cross-sectional • Source of funding: The Cancer Council WA, Friends of Breast Cancer Research • Setting: single centre, Australia • Sample size:72 • Duration: June 2002-October 2005	• Eligibility criteria: women with high breast cancer risk based on genetic analysis (7%) or family history or previous high risk lesion on biopsy (7%) • Patient characteristics: mean age 39y (range 25-50y)	• Index tests: MRI + US + CBE + mammography • Reference standard: biopsy for BIRADS ≥ 3. Biopsy to define true positives, no FU for interval cancers	• Sensitivity, specificity, PPV, NPV: not calculated	• Recall rate for MRI 12.5% in first round, 7.5% in second round • Biopsy rate: 11/139 (7.9%) 1/11 cancerous (9%) • 4/139 lesions detected by MRI only, for short term FU • No info on interval cancers, mortality	• Level of evidence • Dropouts: 72/102 (71%) consented, 5/72 (7%) no second round • Results critical appraisal: no info in blinding. No data on results other imaging than MRI. No FU for interval cancers.
Daguet 2008 [96]	• Design: cross-sectional • Source of funding: not stated • Setting: single centre, France • Sample size: 85 • Duration: December 2000-	• Eligibility criteria: women with BRCA 1 or 2 or p53 (1/85) mutation, with (43.5%) or without personal cancer history • Patient characteristics:	• Index tests: MRI + US + double reading mammography • Reference standard: FNA or biopsy for BIRADS ≥ 3. If BIRADS 3 on MRI only, short	<u>MRI</u> • Sens: 87.5% • Spec: 94.8% • PPV: 38.9% • NPV: 99.5% <u>Mammography</u> • Sens: 12.5% • Spec: 98.7% • PPV: 25%	• 7 cancers detected on screening imaging, 1 interval cancer. 4/8 cancers detected in women with personal cancer history. 2/7 screening detected cancer palpable on	• Level of evidence • Dropouts: 1 woman with interval cancer excluded as she did not have mammography. 6/85 (7%) quit study • Results critical appraisal: no clear

	February 2006	mean age 43y (range 27-65y), 24.7% pBSO	term FU. Biopsy and repeat imaging + FU to define true/false positives and negatives. BIRADS 3 defined as negative exam.	• NPV: 96.9%	CBE. • Biopsy/FNA rate MRI: 12% prevalent round, 6%,12% and 10% incident rounds. • PPV FNA: 30%, PPV biopsies 58% • Short term FU imaging after MRI: 27% (all benigne findings during FU) • No data on mortality	consecutive inclusion of patients. No blinding. Maximum interval between mammography and MRI 6 months (median 12 days). 50% of cancers detected in women with personal cancer history
Yu 2008 [117]	• *Design: retrospective cross-sectional* • *Source of funding: not stated* • *Setting: single centre* • *Sample size: 1019 eligible patients* • *Duration: April 1999-July 2006*					• *Level of evidence* • *Dropouts* • *Results critical appraisal: study excluded because retrospective analysis without consecutive inclusion of patients. Only 37% of eligible patients underwent MRI.*
Shah 2009 [120]	• *Design: cross-sectional* • *Source of funding: Cancer Genetics Network, Marjorie Cohen foundation, QVC Network-Fashion Footwear Association* • *Setting: single centre, USA*	• *Eligibility criteria: asymptomatic women ≥ 25y with BRCA1 or 2 mutation or > 75% risk of mutation, with (43%) or without personal cancer history* • *Patient characteristics:*	• *Index tests: MRI + mammography* • *Reference standard: biopsy for BIRADS 4-5 lesions, 6 month FU for BIRADS 3 lesions. Biopsy + FU to define positives and negatives. Median*	• *Sensitivity, specificity, PPV, NPV: not calculated*	• *11 cancers with 283 MRI's and 282 mammographies in 93 women* • *5/11 cancers detected in women with prior breast cancer.* • *2/11 interval cancer in women without personal cancer*	• *Level of evidence* • *Dropouts: 1 excluded for ovarian cancer diagnosis at the start of the study* • *Results critical appraisal: interval mammography-MRI up to 3 months accepted. No info on blinding,*

	• *Sample size: 94* • *Duration: February 2003 – September 2005*	*median age 47y (range 28-72y), 86% pBSO.* • *Prevalence of disease: unknown*	*FU 3.2y*		• *history* • *9/11 invasive cancers, 2/11 DCIS* • *7/9 invasive cancers node negative.* • *No data recall rate, biopsy rate, mortality*	*probably absent. 43% personal cancer history, 5//11 cancers detected in this group. Not all patients systematically underwent all index tests.*
Price 2009[121]	• Design: cohort • Source of funding: assistance of Suros Surgical Systems for MRI-guided biopsies • Setting: single centre, Australia • Sample size: 171 • Duration: January 2005-June 2008	• Eligibility criteria: women with moderate or high breast cancer risk based on gene mutation, family history, histology of previous biopsy, previous radiotherapy (1) or dense breasts, implants, 'other' • Patient characteristics: 41/171 (24%) personal cancer history. 21% dense breasts, of whom 19% without other risk factor.	• Index test: MRI • Reference standard: histology for all BIRADSS 4-5 lesions. For BIRADS 3 lesions, histology if possible, otherwise short term FU. Biopsy + FU to define true/false positives.	• Sensitivity, specificity, PPV, NPV: not calculated.	• 7 malignancies detected in 171 patients, cancer yield 4.0% • Recall rate: 15% • Biopsy rate: 13% • 7/23 (30.4%) biopsies positive, benign to malignant ratio 2:1 • 6/6 node negative • 2 interval cancers diagnosed • No info on mortality	• Level of evidence • Dropouts: only 35/171 completed second round • Results critical appraisal: 24% personal cancer history. No info on other imaging, no info on blinding.
Lapierre-Combes 2009[118]	• *Design: retrospective cohort* • *Source of funding: not stated* • *Setting: single centre, France*	• *Eligibility criteria: women with normal screening mammography and ultrasound, high breast cancer risk, dense breast tissue*	• *Index test: MRI*	• *-*		• *Level of evidence* • *Dropouts* • *Results critical appraisal: no clear consecutive inclusion of patients. 41% of recruited*

	• *Sample size: 51* • *Duration: October 2003 – June 2007*	*or symptomatic patients with discordance clinical-radiological findings (41%)* • *Patient characteristics: mean age 51y (range 33-71y). 9/51 (17.6%) with personal cancer history* • *Prevalence of disease: unknown*				*patients symptomatic.*
Weinstein 2009 [110]	• Design: cross-sectional • Source of funding: National Institutes of Health • Setting: single centre USA • Sample size:612 • Duration: May 2002-July 2007	• Eligibility criteria: BRCA mutation carriers, women with ≥ 25% life time risk of breast cancer, previous LCIS or atypical hyperplasia or chest wall radiotherapy. Women with recent breast cancer included for contralateral breast. All women normal FSM 180d before intervention. • Patient characteristics: median age 41y (range 27-81y), 41.2% cancer in contralateral breast	• Index tests:: single reading FFDM + US + MRI • Reference standard: biopsy for all BIRADS 4-5 lesions (consensus of all imaging). Biopsy + 2y FU to define true/false positives and negatives. BIRADS 0 and 3 for each modality considered 'positive'	<u>MRI</u> • Sens: 71% • Spec: 79% <u>FFDM</u> • Sens: 39% • Spec:91% <u>US</u> • Sens: 17% • Spec: 88%	• Overall cancer yield 3%, cancer yield MRI 2.1% • Recall rate MRI: 129/571 (22.6%) • Biopsy rate MRI: 48/571 (8.4%) • No info on mortality	• Level of evidence • Dropouts: 3/612 (0.5%) • Results critical appraisal: women with personal cancer history included. Single reading mammography. Initial reading with blinding followed by consensus evaluation of all imaging modalities. Only prevalent round. Probably no interval cancers detected, unclear in text.

Elmore 2010[119]	• *Design: cross-sectional* • *Source of funding: not stated* • *Setting: single centre, USA* • *Sample size: 200* • *Duration:January 2005-December 2008*	• *Eligibility criteria: asymptomatic women with high breast cancer risk > 20% life time risk based on mutation analysis, Gail score or previous chest radiotherapy without personal cancer history* • *Patient characteristics: median age 45y (range 18-76y). 32/104 (30.8%) pts gail score <20%, indication MRI unclear* • *Prevalence of disease: unknown*	• *Index test: MRI +/- mammography* • *Reference standard: biopsy + additional imaging. No FU for interval cancers.*	• *Sensitivity, specificity, PPV, NPV: not calculated*	• *25% recall for futher investigations of suspicious or indeterminate lesions* • *21/200 (10.5%) pts underwent biopsy* • *4/21 (19%) of biopsies positive cancer diagnosis* • *Cancer detection rate 1.5% for MRI; 0.8% for mammography*	• *Level of evidence* • *Dropouts: no info* • *Results critical appraisal: no info on consecutive inclusion of patients. No info on blinding, other investigations. No info on interval cancers. Inclusion criteria not respected in 30% of patients.*
Kuhl 2010[98]	• Design: cross-sectional • Source of funding: German Cancer Aid Society • Setting: multi-centre, Germany • Sample size: 687 • Duration: October 2002-December 2005	• Eligibility criteria: women with high breast cancer risk based on mutation analysis or family history, with (27%) or without personal cancer history • Patient characteristics: median age 44y (range 25-71y)	• Index tests: MRI + US + double reading mammography • Reference standard: BIRADSS 1-2-3 taken as negative, BIRADS 4-5 taken as positive. Biopsy + 1y FU to define true/false positives and negatives.	<u>MRI</u> • Sens: 92.6% • Spec:98.4% • PPV: 48% • NPV: 99.9% <u>MRI + Mx</u> • Sens: 100% • Spec: 97.6% • PPV: 40.2% • NPV: 100% <u>MRI+US</u> • Sens: 92.6% • Spec:98.5% • PPV: 50% • NPV: 99.9%	• 27 cancers detected, 9/27 (33.3%) in women with personal cancer history. 25/27 detected by MRI • No interval cancers detected • 21/27 (77%) cancers ≤ 10mm • 118/687 (17%) BIRADS 3 diagnosis on MRI, requiring short-term FU • No cancer detected	• Level of evidence • Dropouts: 38/725 (5.2%) • Results critical appraisal: No info on blinding. 1/3 cancers detected in women with personal cancer history, CBE positive in 110 screening rounds.

						at half-yearly screening round with CBE + US • No data on mortality	
Sardanelli 2011 [111]	• Design: cross-sectional • Source of funding: Italian Ministry of health • Setting: multicentre, Italy • Sample size: 501 • Duration:June 2000-January 2007	• Eligibility criteria: asymptomatic women with high breast cancer risk based on mutation analysis or family history, with (43.5%) or without personal cancer history. • Patient characteristics: median age 45y (range 22-79y)	• Index tests: CBE + US + mammography + MRI • Reference standard: biopsy for BIRADS 4-5 on any imaging or positive CBE. Short term FU for BIRADS 3. BIRADS 3 considered negative exam. Biopsy/FNA and 1 y FU to define true/false positives and negatives	<u>MRI</u> • Sens: 91.3% • Spec: 96.7% • PPV: 56% • NPV: 99.6% <u>MRI+Mx</u> • Sens: 93.2% • Spec: 96.3% • PPV: 53.2% • NPV: 99.7% <u>Mx</u> • Sens:50% • Spec 99% • PPV 71.4% • NPV 97.6%	• 49 cancers detected through screening, 3 interval cancers. 29 (56%) cancers diagnosed in women with personal cancer history • 3.3% incidence per woman-year • No info on recall rate, biopsy rate, mortality	• Level of evidence • Dropouts 85%,67%,46% underwent 2nd, 3rd and 4th round respectively • Results critical appraisal: probably single reading mammography. 56% of cancers diagnosed in women with personal cancer history. No info on blinding. Handling of missing data not clear from text, may influence results
Abramovici 2011 [122]	• Design: retrospective cohort • Source of funding: not stated • Setting: single centre, USA • Sample size: 650 • Duration: September 2007-December 2008	• Eligibility criteria: asymptomatic women with high breast cancer risk based on family history, previous radiotherapy or previous biopsy, with (41.5%) or without personal cancer history • Patient characteristics:	• Index test: MRI • Reference standard: BIRADS 3 defined as positive exam, shot term FU advised. Biopsy for all BIRADS 4-5 lesions.	• Sensitivity, specificity, NPV: not calculated.	• Total recall rate short term FU included: 11.4%, for prevalent round 16%, for incident rounds 7.3% • PPV of BIRADS 4-5 lesions 11.1% in prevalent round and 18.8% in incident rounds. • No info on cancer detection rate,	• Level of evidence • Dropouts: not stated • Results critical appraisal: retrospective cohort, not clearly consecutive inclusion. 41.5% women with personal cancer history. No info on other imaging; No info on interval

	mean age 51y (range 25-81y)	biopsy rate, interval cancers, mortality	cancers.

APPENDIX 4. SUPPLEMENTARY TABLES

Table 40 Eligible population per year per region and province, IMA data - Period 2006-2007

REGIONS	PROVINCES	Number of eligible women in 2006 *	Number of eligible women in 2007 * *	***Eligible population***
Undetermined region		*30.710*	*31.086*	*31.664*
Flemish region	Antwerp	462.830	467.257	**473.991**
	Fl. Brabant	291.621	294.582	**298.793**
	West Flanders	323.604	325.464	**330.636**
	East Flanders	388.013	391.827	**397.510**
	Limburg	219.253	221.462	**224.312**
	Total	**1.685.321**	**1.700.592**	**1.725.242**
Region Brussels Capital		**238.246**	**241.966**	**245.770**
Walloon region	Wal.Brabant	99.760	101.047	**102.341**
	Hainaut	352.041	355.208	**361.076**
	Liège	281.855	284.257	**288.739**
	Luxemburg	59.692	60.055	**61.091**
	Namur	123.580	124.908	**126.795**
	Total	**916.928**	**925.475**	**940.042**
Belgium		**2.871.205**	**2.899.119**	**2.942.718**

* Women born between 1927 and 1971
** Women born between 1928 and 1972

Table 41 Eligible and excluded populations with the reason for their exclusion

			Eligible Population (a)	Women excluded because dead	Women excluded because of missing data	Excluded Population (b)	Intermediate Population (a-b)	Women taken from the lower age category**	- women put into the higher age category**	- Women excluded because they fell outside the agegroup***	Study population
Flemish region	Antwerp	34 ans	11.161	2	0	2	11.159	0	5.580	5.580	0
		35-39 ans	59.474	77	0	77	59.397	5.580	6.024	0	58.953
		40-44 ans	65.061	142	0	142	64.919	6.024	6.531	0	64.412
		45-49 ans	63.785	208	0	208	63.577	6.531	6.034	0	64.074
		50-54 ans	57.762	311	0	311	57.451	6.034	5.478	0	58.007
		55-59 ans	51.839	412	0	412	51.427	5.478	5.215	0	51.690
		60-64 ans	45.722	545	0	545	45.177	5.215	3.873	0	46.519
		65-69 ans	40.584	711	0	711	39.873	3.873	4.078	0	39.668
		70-74 ans	41.036	1.268	0	1.268	39.768	4.078	4.014	0	39.832
		75-79 ans	37.567	1.820	0	1.820	35.747	4.014	3.236	3.236	36.525
		Total	473.991	5.496	0	5.496	468.495			8.816	459.680
	Flemish Brabant	34 ans	7.172	5	0	5	7.167	0	3.584	3.584	0
		35-39 ans	37.644	40	0	40	37.604	3.584	3.818	0	37.370
		40-44 ans	41.679	92	0	92	41.587	3.818	4.069	0	41.336
		45-49 ans	40.554	125	0	125	40.429	4.069	3.933	0	40.565
		50-54 ans	36.524	201	0	201	36.323	3.933	3.411	0	36.845
		55-59 ans	33.014	276	0	276	32.738	3.411	3.155	0	32.994
		60-64 ans	28.377	325	0	325	28.052	3.155	2.392	0	28.815
		65-69 ans	24.767	385	0	385	24.382	2.392	2.570	0	24.204
		70-74 ans	25.325	627	0	627	24.698	2.570	2.538	0	24.730
		75-79 ans	23.737	1.014	0	1.014	22.723	2.538	2.045	2.045	23.216
		Total	298.793	3.090	0	3.090	295.703			5.629	290.075
	West Flanders	34 ans	7.032	5	0	5	7.027	0	3.514	3.514	0
		35-39 ans	38.205	53	0	53	38.152	3.514	3.907	0	37.759
		40-44 ans	42.629	91	0	91	42.538	3.907	4.236	0	42.209
		45-49 ans	41.109	149	0	149	40.960	4.236	3.983	0	41.213
		50-54 ans	38.770	208	0	208	38.562	3.983	3.726	0	38.819
		55-59 ans	36.456	259	0	259	36.197	3.726	3.688	0	36.235
		60-64 ans	34.449	353	0	353	34.096	3.688	2.880	0	34.904
		65-69 ans	31.586	494	0	494	31.092	2.880	3.208	0	30.764
		70-74 ans	31.339	810	0	810	30.529	3.208	3.119	0	30.618
		75-79 ans	29.061	1.218	0	1.218	27.843	3.119	2.507	2.507	28.455
		Total	330.636	3.640	0	3.640	326.996			6.021	320.976
	East Flanders	34 ans	9.497	3	0	3	9.494	0	4.747	4.747	0
		35-39 ans	50.830	66	0	66	50.764	4.747	5.123	0	50.388
		40-44 ans	53.937	103	0	103	53.834	5.123	5.346	0	53.611
		45-49 ans	51.475	170	0	170	51.305	5.346	4.982	0	51.669
		50-54 ans	46.803	268	0	268	46.535	4.982	4.505	0	47.012
		55-59 ans	43.741	359	0	359	43.382	4.505	4.280	0	43.607
		60-64 ans	38.368	429	0	429	37.939	4.280	3.082	0	39.137
		65-69 ans	34.388	623	0	623	33.765	3.082	3.485	0	33.362
		70-74 ans	35.949	1.054	0	1.054	34.895	3.485	3.478	0	34.902
		75-79 ans	32.522	1.549	0	1.549	30.973	3.478	2.741	2.741	31.710
		Total	397.510	4.624	0	4.624	392.886			7.488	385.398
	Limburg	34 ans	5.059	0	0	0	5.059	0	2.530	2.530	0
		35-39 ans	27.398	24	0	24	27.374	2.530	2.813	0	27.091
		40-44 ans	31.799	51	0	51	31.748	2.813	3.262	0	31.299
		45-49 ans	31.694	100	0	100	31.594	3.262	3.058	0	31.798
		50-54 ans	28.531	155	0	155	28.376	3.058	2.722	0	28.712
		55-59 ans	25.304	201	0	201	25.103	2.722	2.364	0	25.461
		60-64 ans	20.824	221	0	221	20.603	2.364	1.865	0	21.102
		65-69 ans	19.436	305	0	305	19.131	1.865	1.980	0	19.016
		70-74 ans	18.475	517	0	517	17.958	1.980	1.774	0	18.164
		75-79 ans	15.792	725	0	725	15.067	1.774	1.374	1.374	15.467
		Total	224.312	2.299	0	2.299	222.013			3.904	218.110

Region of Brussels capital	Brussels-Capital	34 ans	7.524	4	0	4	7.520	0	3.760	3.760	0
		35-39 ans	36.488	66	0	66	36.422	3.760	3.351	0	36.831
		40-44 ans	33.025	86	0	86	32.939	3.351	3.151	0	33.139
		45-49 ans	31.081	143	0	143	30.938	3.151	2.959	0	31.130
		50-54 ans	28.921	214	0	214	28.707	2.959	2.850	0	28.816
		55-59 ans	26.529	276	0	276	26.253	2.850	2.631	0	26.472
		60-64 ans	22.657	336	0	336	22.321	2.631	1.929	0	23.023
		65-69 ans	19.519	457	0	457	19.062	1.929	1.886	0	19.105
		70-74 ans	19.872	721	0	721	19.151	1.886	1.960	0	19.077
		75-79 ans	20.154	1.026	0	1.026	19.128	1.960	1.829	1.829	19.259
		Total	245.770	3.329	0	3.329	242.441			5.589	236.852
Walloon region	Walloon Brabant	34 ans	2.581	2	0	2	2.579	0	1.290	1.290	0
		35-39 ans	13.159	6	0	6	13.153	1.290	1.342	0	13.101
		40-44 ans	14.168	27	0	27	14.141	1.342	1.417	0	14.066
		45-49 ans	13.952	57	0	57	13.895	1.417	1.328	0	13.984
		50-54 ans	12.985	63	0	63	12.922	1.328	1.248	0	13.002
		55-59 ans	12.601	94	0	94	12.507	1.248	1.298	0	12.457
		60-64 ans	10.103	119	0	119	9.984	1.298	766	0	10.516
		65-69 ans	7.949	126	0	126	7.823	766	747	0	7.842
		70-74 ans	7.618	192	0	192	7.426	747	772	0	7.401
		75-79 ans	7.225	343	0	343	6.882	772	625	625	7.029
		Total	102.341	1.029	0	1.029	101.312			1.915	99.398
	Hainaut	34 ans	9.035	11	0	11	9.024	0	4.512	4.512	0
		35-39 ans	44.833	86	0	86	44.747	4.512	4.339	0	44.920
		40-44 ans	46.757	146	0	146	46.611	4.339	4.629	0	46.321
		45-49 ans	47.534	250	0	250	47.284	4.629	4.581	0	47.332
		50-54 ans	45.805	370	0	370	45.435	4.581	4.369	0	45.647
		55-59 ans	44.258	458	0	458	43.800	4.369	4.375	0	43.794
		60-64 ans	33.694	509	0	509	33.185	4.375	2.523	0	35.037
		65-69 ans	28.342	642	0	642	27.700	2.523	2.829	0	27.394
		70-74 ans	29.862	1.115	0	1.115	28.747	2.829	3.088	0	28.488
		75-79 ans	30.956	1.767	0	1.767	29.189	3.088	2.813	2.813	29.464
		Total	361.076	5.354	0	5.354	355.722			7.325	348.397
	Liège	34 ans	6.884	4	0	4	6.880	0	3.440	3.440	0
		35-39 ans	35.165	72	0	72	35.093	3.440	3.529	0	35.004
		40-44 ans	37.867	100	0	100	37.767	3.529	3.714	0	37.582
		45-49 ans	38.019	190	0	190	37.829	3.714	3.728	0	37.815
		50-54 ans	35.522	240	0	240	35.282	3.728	3.435	0	35.575
		55-59 ans	34.422	345	0	345	34.077	3.435	3.355	0	34.157
		60-64 ans	26.781	414	0	414	26.367	3.355	2.277	0	27.445
		65-69 ans	24.509	531	0	531	23.978	2.277	2.335	0	23.920
		70-74 ans	24.890	916	0	916	23.974	2.335	2.495	0	23.814
		75-79 ans	24.680	1.375	0	1.375	23.305	2.495	2.156	2.156	23.644
		Total	288.739	4.187	0	4.187	284.552			5.596	278.956
	Luxemburg	34 ans	1.399	3	0	3	1.396	0	698	698	0
		35-39 ans	7.055	11	0	11	7.044	698	740	0	7.002
		40-44 ans	7.905	19	0	19	7.886	740	787	0	7.839
		45-49 ans	8.109	44	0	44	8.065	787	795	0	8.057
		50-54 ans	7.568	51	0	51	7.517	795	739	0	7.573
		55-59 ans	7.053	81	0	81	6.972	739	716	0	6.995
		60-64 ans	5.505	81	0	81	5.424	716	498	0	5.642
		65-69 ans	5.337	108	0	108	5.229	498	549	0	5.178
		70-74 ans	5.566	174	0	174	5.392	549	528	0	5.413
		75-79 ans	5.594	275	0	275	5.319	528	503	503	5.344
		Total	61.091	847	0	847	60.244			1.201	59.043
	Namur	34 ans	3.215	0	0	0	3.215	0	1.608	1.608	0
		35-39 ans	15.872	18	0	18	15.854	1.608	1.631	0	15.831
		40-44 ans	17.079	49	0	49	17.030	1.631	1.703	0	16.958
		45-49 ans	16.927	70	0	70	16.857	1.703	1.660	0	16.900
		50-54 ans	15.913	120	0	120	15.793	1.660	1.569	0	15.884
		55-59 ans	15.433	166	0	166	15.267	1.569	1.530	0	15.306
		60-64 ans	11.229	183	0	183	11.046	1.530	887	0	11.689
		65-69 ans	10.412	219	0	219	10.193	887	1.061	0	10.019
		70-74 ans	10.444	349	0	349	10.095	1.061	1.055	0	10.101
		75-79 ans	10.271	557	0	557	9.714	1.055	912	912	9.857
		Total	126.795	1.731	0	1.731	125.064			2.520	122.545
Undetermined region	*Undetermined*	*34 ans*	*954*	*0*	*954*	*954*	*0*	*0*	*0*	*0*	*0*
		35-39 ans	*4.567*	*2*	*4.567*	*4.567*	*0*	*0*	*0*	*0*	*0*
		40-44 ans	*3.925*	*5*	*3.925*	*3.925*	*0*	*0*	*0*	*0*	*0*
		45-49 ans	*3.276*	*11*	*3.276*	*3.276*	*0*	*0*	*0*	*0*	*0*
		50-54 ans	*2.797*	*11*	*2.797*	*2.797*	*0*	*0*	*0*	*0*	*0*
		55-59 ans	*2.840*	*22*	*2.840*	*2.840*	*0*	*0*	*0*	*0*	*0*
		60-64 ans	*3.045*	*26*	*3.045*	*3.045*	*0*	*0*	*0*	*0*	*0*
		65-69 ans	*3.401*	*38*	*3.401*	*3.401*	*0*	*0*	*0*	*0*	*0*
		70-74 ans	*3.597*	*70*	*3.597*	*3.597*	*0*	*0*	*0*	*0*	*0*
		75-79 ans	*3.262*	*102*	*3.262*	*3.262*	*0*	*0*	*0*	*0*	*0*
		Total	*31.664*	*287*	*31.664*	*31.664*	*0*			*0*	*0*

Table 42 Study population per region and per 5 year age-band, IMA data - period 2006-2007

REGIONS	AGE	Study population
Flemish region	35-39 years	211.561
	40-44 years	232.867
	45-49 years	229.319
	50-54 years	209.395
	55-59 years	189.987
	60-64 years	170.477
	65-69 years	147.014
	70-74 years	148.246
	75-79 years	135.373
	Total	**1.674.239**
Region Brussels Capital	35-39 years	36.831
	40-44 years	33.139
	45-49 years	31.130
	50-54 years	28.816
	55-59 years	26.472
	60-64 years	23.023
	65-69 years	19.105
	70-74 years	19.077
	75-79 years	19.259
	Total	**236.852**
Walloon region	35-39 years	115.858
	40-44 years	122.766
	45-49 years	124.088
	50-54 years	117.681
	55-59 years	112.709
	60-64 years	90.329
	65-69 years	74.353
	70-74 years	75.217
	75-79 years	75.338
	Total	**908.339**
Belgium	35-39 years	364.250
	40-44 years	388.772
	45-49 years	384.537
	50-54 years	355.892
	55-59 years	329.168
	60-64 years	283.829
	65-69 years	240.472
	70-74 years	242.540
	75-79 years	229.970
	Total	**2.819.430**

Table 43 Study population per province, IMA data - period 2006-2007

REGIONS	PROVINCES	Population studied
Flemish region	Antwerp	459.680
	Fl. Brabant	290.075
	West Flanders	320.976
	East Flanders	385.398
	Limburg	218.110
	Total	**1.674.239**
Region Brussels Capital		**236.852**
Walloon region	Wal.Brabant	99.398
	Hainaut	348.397
	Liège	278.956
	Luxemburg	59.043
	Namur	122.545
	Total	**908.339**
Belgium	**Total**	**2.819.430**

Table 44 Study population and coverage with screening mammography (mammotest) and diagnostic mammography per region and per 5 year age-band, IMA data - period 2006-2007

REGIONS	AGE	*study population*	coverage by screening mammography	coverage by diagnostic mammography	total coverage
Flemish region	35-39 years	*211.561*	0%	12%	**12%**
	40-44 years	*232.867*	0%	28%	**28%**
	45-49 years	*229.319*	0%	34%	**34%**
	50-54 years	*209.395*	51%	23%	**73%**
	55-59 years	*189.987*	44%	22%	**66%**
	60-64 years	*170.477*	43%	19%	**63%**
	65-69 years	*147.014*	38%	17%	**56%**
	70-74 years	*148.246*	0%	18%	**18%**
	75-79 years	*135.373*	0%	8,2%	**8,2%**
	Total	***1.674.239***	**19%**	**21%**	**40%**
Region Brussels Capital	35-39 years	*36.831*	0%	15%	**15%**
	40-44 years	*33.139*	0%	40%	**40%**
	45-49 years	*31.130*	0%	48%	**48%**
	50-54 years	*28.816*	9,5%	47%	**56%**
	55-59 years	*26.472*	9,5%	45%	**55%**
	60-64 years	*23.023*	9,9%	41%	**51%**
	65-69 years	*19.105*	9,2%	38%	**47%**
	70-74 years	*19.077*	0%	33%	**33%**
	75-79 years	*19.259*	0%	18%	**18%**
	Total	***236.852***	**3,9%**	**36%**	**40%**
Walloon region	35-39 years	*115.858*	0%	19%	**19%**
	40-44 years	*122.766*	0%	42%	**42%**
	45-49 years	*124.088*	0%	50%	**50%**
	50-54 years	*117.681*	8,7%	50%	**58%**
	55-59 years	*112.709*	9,1%	48%	**58%**
	60-64 years	*90.329*	9,5%	43%	**53%**
	65-69 years	*74.353*	9,5%	40%	**49%**
	70-74 years	*75.217*	0%	30%	**30%**
	75-79 years	*75.338*	0%	15%	**15%**
	Total	***908.339***	**4%**	**39%**	**43%**
Belgium	35-39 years	*364.250*	0%	15%	**15%**
	40-44 years	*388.772*	0%	33%	**33%**
	45-49 years	*384.537*	0%	40%	**40%**
	50-54 years	*355.892*	33%	34%	**67%**
	55-59 years	*329.168*	29%	33%	**62%**
	60-64 years	*283.829*	30%	29%	**59%**
	65-69 years	*240.472*	27%	26%	**53%**

Table 45 Study population and coverage with screening mammography (mammotest) and diagnostic mammography per region, province and per age-band, IMA data - period 2006-2007

REGION=Flemish region

AGE	PROVINCE	*study population*	coverage by screening mammography	coverage by diagnostic mammography	total coverage
35-40 years	Antwerp	*58.953*	0%	12%	**12%**
	Fl. Brabant	*37.370*	0%	15%	**15%**
	West Flanders	*37.759*	0%	11%	**11%**
	East Flanders	*50.388*	0%	11%	**11%**
	Limburg	*27.091*	0%	12%	**12%**
	Total	***211.561***	**0%**	**12%**	**12%**
40-49 years	Antwerp	*128.486*	0%	28%	**28%**
	Fl. Brabant	*81.901*	0%	39%	**39%**
	West Flanders	*83.422*	0%	26%	**26%**
	East Flanders	*105.280*	0%	33%	**33%**
	Limburg	*63.097*	0%	29%	**29%**
	Total	***462.186***	**0%**	**31%**	**31%**
50-69 years	Antwerp	*195.884*	42%	22%	**64%**
	Fl. Brabant	*122.858*	37%	27%	**64%**
	West Flanders	*140.722*	47%	14%	**62%**
	East Flanders	*163.118*	44%	22%	**66%**
	Limburg	*94.291*	56%	16%	**72%**
	Total	***716.873***	**45%**	**21%**	**65%**
70-74 years	Antwerp	*39.832*	0%	17%	**17%**
	Fl. Brabant	*24.730*	0%	22%	**22%**
	West Flanders	*30.618*	0%	15%	**15%**
	East Flanders	*34.902*	0%	18%	**18%**
	Limburg	*18.164*	0%	16%	**16%**
	Total	***148.246***	**0%**	**18%**	**18%**
75-79 years	Antwerp	*36.525*	0%	8,2%	**8,2%**
	Fl. Brabant	*23.216*	0%	10%	**10%**
	West Flanders	*28.455*	0%	6,9%	**6,9%**
	East Flanders	*31.710*	0%	8,6%	**8,6%**
	Limburg	*15.467*	0%	7,4%	**7,4%**
	Total	***135.373***	**0%**	**8,2%**	**8,2%**
Total	Antwerp	*459.680*	18%	21%	**39%**
	Fl. Brabant	*290.075*	16%	27%	**43%**
	West Flanders	*320.976*	21%	16%	**37%**
	East Flanders	*385.398*	19%	22%	**41%**
	Limburg	*218.110*	24%	18%	**43%**
	Total	***1.674.239***	**19%**	**21%**	**40%**

REGION=Region Brussels-Capital

AGE	PROVINCE	study population	coverage by screening mammography	coverage by diagnostic mammography	total coverage
35-40 years		*36.831*	0%	15%	**15%**
40-49 years		*64.269*	0%	44%	**44%**
50-69 years		*97.416*	9,5%	43%	**53%**
70-74 years		*19.077*	0%	33%	**33%**
75-79 years		*19.259*	0%	18%	**18%**
Total		***236.852***	**3,9%**	**36%**	**40%**

REGION=Walloon region

AGE	PROVINCE	study population	coverage by screening mammography	coverage by diagnostic mammography	total coverage
35-40 years	Wal.Brabant	*13.101*	0%	20%	**20%**
	Hainaut	*44.920*	0%	20%	**20%**
	Liège	*35.004*	0%	17%	**17%**
	Luxemburg	*7.002*	0%	15%	**15%**
	Namur	*15.831*	0%	18%	**18%**
	Total	***115.858***	**0%**	**19%**	**19%**
40-49 years	Wal.Brabant	*28.050*	0%	53%	**53%**
	Hainaut	*93.653*	0%	47%	**47%**
	Liège	*75.397*	0%	43%	**43%**
	Luxemburg	*15.896*	0%	41%	**41%**
	Namur	*33.858*	0%	46%	**46%**
	Total	***246.854***	**0%**	**46%**	**46%**
50-69 years	Wal.Brabant	*43.817*	13%	48%	**61%**
	Hainaut	*151.872*	9,1%	46%	**55%**
	Liège	*121.097*	7,5%	47%	**54%**
	Luxemburg	*25.388*	9,9%	43%	**53%**
	Namur	*52.898*	9,6%	46%	**56%**
	Total	***395.072***	**9,1%**	**46%**	**55%**
70-74 years	Wal.Brabant	*7.401*	0%	35%	**35%**
	Hainaut	*28.488*	0%	29%	**29%**
	Liège	*23.814*	0%	30%	**30%**
	Luxemburg	*5.413*	0%	25%	**25%**
	Namur	*10.101*	0%	29%	**29%**
	Total	***75.217***	**0%**	**30%**	**30%**
75-79 years	Wal.Brabant	*7.029*	0%	17%	**17%**
	Hainaut	*29.464*	0%	14%	**14%**
	Liège	*23.644*	0%	15%	**15%**
	Luxemburg	*5.344*	0%	14%	**14%**
	Namur	*9.857*	0%	14%	**14%**
	Total	***75.338***	**0%**	**15%**	**15%**
Total	Wal.Brabant	*99.398*	5,6%	43%	**48%**
	Hainaut	*348.397*	3,9%	39%	**43%**
	Liège	*278.956*	3,2%	38%	**41%**
	Luxemburg	*59.043*	4,3%	35%	**39%**
	Namur	*122.545*	4,1%	38%	**43%**
	Total	***908.339***	**4,0%**	**39%**	**43%**

Table 46 absolute numbers of women with a diagnostic (MD) and screening mammography (MT) per region and per province

REGIONS	PROVINCES	Nbr daily episodes DM	*Nbr women with DM>=1*	Nbr daily episodes MT	*Nbr women avec MT>=1*
Flemish region	Antwerp	130.451	*102.466*	83.228	*83.208*
	Fl. Brabant	102.943	*81.968*	46.033	*46.021*
	West Flanders	69.630	*57.865*	66.669	*66.654*
	East Flanders	113.801	*93.016*	71.584	*71.563*
	Limburg	52.446	*43.269*	53.095	*53.090*
	Total	**469.271**	***378.584***	**320.609**	***320.536***
Region Brussels Capital		**113.632**	***87.243***	**9.291**	***9.283***
Walloon region	Wal.Brabant	57.837	*43.584*	5.616	*5.612*
	Hainaut	184.237	*138.068*	13.777	*13.760*
	Liège	138.871	*107.868*	9.075	*9.061*
	Luxemburg	26.399	*21.067*	2.517	*2.517*
	Namur	63.637	*48.139*	5.057	*5.053*
	Total	**470.981**	***358.726***	**36.042**	***36.003***
Belgium	Total	**1.053.884**	**824.553**	**365.942**	***365.822***

Table 47 Absolute numbers of women with a diagnostic (MD) and screening mammography (MT) per region and 5 year age-band

REGIONS	AGE	Nbr daily episodes DM	*Nbr women avec DM>=1*	Nbr daily episodes MT	*Nbr women avec MT>=1*
Flemish region	35-39 years	29.161	*25.988*	0	*0*
	40-44 years	76.178	*65.552*	0	*0*
	45-49 years	93.403	*77.440*	0	*0*
	50-54 years	73.346	*57.354*	105.927	*105.905*
	55-59 years	61.314	*47.059*	84.210	*84.188*
	60-64 years	49.574	*37.749*	73.873	*73.852*
	65-69 years	37.948	*28.884*	56.599	*56.591*
	70-74 years	33.675	*26.952*	0	*0*
	75-79 years	14.672	*11.606*	0	*0*
	Total	**469.271**	***378.584***	**320.609**	***320.536***
Region Brussels Capital	35-39 years	6.433	*5.659*	0	*0*
	40-44 years	17.046	*13.289*	0	*0*
	45-49 years	19.991	*14.921*	0	*0*
	50-54 years	18.474	*13.982*	2.735	*2.734*
	55-59 years	16.467	*12.416*	2.518	*2.515*
	60-64 years	12.969	*9.786*	2.283	*2.280*
	65-69 years	9.796	*7.502*	1.754	*1.754*
	70-74 years	8.057	*6.256*	0	*0*
	75-79 years	4.399	*3.432*	1	*0*
	Total	**113.632**	***87.243***	**9.291**	***9.283***
Walloon region	35-39 years	25.328	*21.613*	0	*0*
	40-44 years	66.286	*51.829*	0	*0*
	45-49 years	82.045	*61.943*	0	*0*
	50-54 years	81.671	*60.952*	10.195	*10.190*
	55-59 years	76.040	*56.641*	10.220	*10.207*
	60-64 years	54.530	*40.803*	8.565	*8.556*
	65-69 years	40.829	*30.886*	7.062	*7.050*
	70-74 years	29.581	*22.717*	0	*0*
	75-79 years	14.671	*11.342*	0	*0*
	Total	**470.981**	***358.726***	**36.042**	***36.003***
Belgium	35-39 years	60.922	*53.260*	0	*0*
	40-44 years	159.510	*130.670*	0	*0*
	45-49 years	195.439	*154.304*	0	*0*
	50-54 years	173.491	*132.288*	118.857	*118.829*
	55-59 years	153.821	*116.116*	96.948	*96.910*
	60-64 years	117.073	*88.338*	84.721	*84.688*
	65-69 years	88.573	*67.272*	65.415	*65.395*
	70-74 years	71.313	*55.925*	0	*0*
	75-79 years	33.742	*26.380*	1	*0*
	Total	**1.053.884**	***824.553***	**365.942**	***365.822***

Table 48 Number and % of women with one mammography (mammographic examination, M.E.) in the period 2006-2007, number and % of women with one mammography in 2006 and one in 2007 and number and % of women with more than one mammography either in 2006, 2007 or both

		denominator Number of examined women	**Number of women with one M.E. in 2006-2007**	% of women with one M.E. in 2006-2007	**Number of women with one M.E. in 2006 and one M.E. in 2007**	% of women with one M.E. in 2006 and one M.E. in 2007	**Number of women with several M.E. in 2006 and/or several M.E. in 2007**	% of women with several M.E. in 2006 and/or several M.E. in 2007
Flemish region	35-40 ans	*25.860*	22.976	*89%*	2.350	*9,1%*	560	*2,2%*
	40-49 ans	*157.731*	119.958	*76%*	20.461	*13%*	3.878	*2,5%*
	50-69 ans	*452.023*	122.519	*27%*	38.436	*8,5%*	7.302	*1,6%*
	70-74 ans	*26.484*	20.537	*78%*	5.063	*19%*	980	*3,7%*
	75-79 ans	*11.154*	8.497	*76%*	2.277	*20%*	458	*4,1%*
	Total	***673.252***	**294.487**	***44%***	**68.587**	***10%***	**13.178**	***2,0%***
Region Brussels Capital	35-40 ans	*5.643*	4.924	*87%*	577	*10%*	144	*2,6%*
	40-49 ans	*28.328*	20.051	*71%*	7.153	*25%*	1.009	*3,6%*
	50-69 ans	*51.239*	30.837	*60%*	10.954	*21%*	1.687	*3,3%*
	70-74 ans	*6.214*	4.642	*75%*	1.307	*21%*	275	*4,4%*
	75-79 ans	*3.372*	2.543	*75%*	687	*20%*	152	*4,5%*
	Total	***94.796***	**62.997**	***66%***	**20.678**	***22%***	**3.267**	***3,4%***
Walloon region	35-40 ans	*21.566*	18.258	*85%*	2.617	*12%*	697	*3,2%*
	40-49 ans	*114.218*	82.275	*72%*	26.942	*24%*	4.494	*3,9%*
	50-69 ans	*217.275*	132.015	*61%*	47.751	*22%*	8.787	*4,0%*
	70-74 ans	*22.536*	16.641	*74%*	4.814	*21%*	1.111	*4,9%*
	75-79 ans	*11.146*	8.328	*75%*	2.281	*20%*	562	*5,0%*
	Total	***386.741***	**257.517**	***67%***	**84.405**	***22%***	**15.651**	***4,0%***
Belgium	35-40 ans	*53.069*	46.158	*87%*	5.544	*10%*	1.401	*2,6%*
	40-49 ans	*300.277*	222.284	*74%*	54.556	*18%*	9.381	*3,1%*
	50-69 ans	*720.537*	285.371	*40%*	97.141	*13%*	17.776	*2,5%*
	70-74 ans	*55.234*	41.820	*76%*	11.184	*20%*	2.366	*4,3%*
	75-79 ans	*25.672*	19.368	*75%*	5.245	*20%*	1.172	*4,6%*
	Total	***1.154.789***	**615.001**	***53%***	**173.670**	***15%***	**32.096**	***2,8%***

Table 49 Medical imaging following diagnostic mammography per age-band and per region, IMA data - Period 2006-2007

AGE	REGION	*N**	% followed with a senological bilan (DM+ECHO)	% followed by a diagnostic mammography.	% followed by an echography	% followed by only an echography	% followed by MRI	% folowed by a punction or biopsie
35-40 years	Flemish region	*12.297*	**0,0%**	0,0%	**88%**	0,0%	1,8%	**3,7%**
	Region Brussels capital	*2.602*	**0,0%**	0,0%	**91%**	0,0%	0,8%	**5,1%**
	Walloon region	*10.689*	**0,0%**	0,0%	**94%**	0,0%	1,2%	**7,0%**
	Belgium	***25.588***	**0,0%**	**0,0%**	**91%**	**0,0%**	**1,5%**	**5,2%**
40-49 years	Flemish region	*78.851*	**0,0%**	0,0%	**85%**	0,0%	1,6%	**3,5%**
	Region Brussels capital	*16.952*	**0,0%**	0,0%	**88%**	0,0%	0,7%	**3,9%**
	Walloon region	*69.344*	**0,0%**	0,0%	**92%**	0,0%	0,9%	**5,5%**
	Belgium	***165.147***	**0,0%**	**0,0%**	**88%**	**0,0%**	**1,2%**	**4,4%**
50-69 years	Flemish region	*94.630*	**0,0%**	0,0%	**79%**	0,0%	1,7%	**3,4%**
	Region Brussels capital	*27.388*	**0,0%**	0,0%	**81%**	0,0%	1,0%	**3,5%**
	Walloon region	*119.617*	**0,0%**	0,0%	**88%**	0,0%	1,0%	**4,3%**
	Belgium	***241.635***	**0,0%**	**0,0%**	**84%**	**0,0%**	**1,3%**	**3,8%**
70-74 years	Flemish region	*15.749*	**0,0%**	0,0%	**65%**	0,0%	1,3%	**4,1%**
	Region Brussels capital	*3.906*	**0,0%**	0,0%	**72%**	0,0%	1,3%	**3,9%**
	Walloon region	*14.661*	**0,0%**	0,0%	**83%**	0,0%	1,1%	**4,3%**
	Belgium	***34.316***	**0,0%**	**0,0%**	**74%**	**0,0%**	**1,2%**	**4,1%**
75-79 years	Flemish region	*6.919*	**0,0%**	0,0%	**67%**	0,0%	1,7%	**5,2%**
	Region Brussels capital	*2.154*	**0,0%**	0,0%	**72%**	0,0%	0,8%	**4,1%**
	Walloon region	*7.164*	**0,0%**	0,0%	**83%**	0,0%	1,0%	**5,5%**
	Belgium	***16.237***	**0,0%**	**0,0%**	**75%**	**0,0%**	**1,3%**	**5,2%**
Total	Flemish region	*208.446*	**0,0%**	0,0%	**80%**	0,0%	1,6%	**3,6%**
	Region Brussels capital	*53.002*	**0,0%**	0,0%	**82%**	0,0%	0,9%	**3,8%**
	Walloon region	*221.475*	**0,0%**	0,0%	**89%**	0,0%	1,0%	**4,8%**
	Belgium	***482.923***	**0,0%**	**0,0%**	**85%**	**0,0%**	**1,3%**	**4,2%**

Table 50 Medical imaging following screening mammography (mammotest) per age-band and per region, IMA data - Period 2006-2007.

AGE	REGION	*N**	% followed with a senological bilan (DM+ECHO)	% followed by a diagnostic mammography.	% followed by an echography	% followed by only an echography	% followed by MRI	% followed by a punction or biopsie
40-49 years	Flemish region	*6*	**0,0%**	0,0%	**0,0%**	0,0%	0,0%	**0,0%**
	Region Brussels capital	*4*	**0,0%**	0,0%	**25%**	25,0%	0,0%	**0,0%**
	Walloon region	*1*	**0,0%**	0,0%	**0,0%**	0,0%	0,0%	**0,0%**
	Belgium	***11***	**0,0%**	**0,0%**	**9,1%**	**9,1%**	**0,0%**	**0,0%**
50-69 years	Flemish region	*158.756*	**2,1%**	2,3%	**4,4%**	2,3%	0,3%	**0,9%**
	Region Brussels capital	*4.830*	**1,3%**	1,9%	**5,4%**	4,1%	0,2%	**1,4%**
	Walloon region	*21.667*	**5,6%**	6,1%	**9,6%**	4,0%	0,4%	**2,1%**
	Belgium	***185.253***	**2,5%**	**2,8%**	**5,0%**	**2,5%**	**0,3%**	**1,1%**
70-74 years	Flemish region	*29*	**0,0%**	17%	**0,0%**	0,0%	0,0%	**0,0%**
	Region Brussels capital	*4*	**0,0%**	0,0%	**0,0%**	0,0%	0,0%	**0,0%**
	Walloon region	*13*	**23%**	31%	**23%**	0,0%	0,0%	**0,0%**
	Belgium	***46***	**6,5%**	**20%**	**6,5%**	**0,0%**	**0,0%**	**0,0%**
75-79 years	Region Brussels capital	*1*	**0,0%**	0,0%	**0,0%**	0,0%	0,0%	**0,0%**
	Belgium	***1***	**0,0%**	**0,0%**	**0,0%**	**0,0%**	**0,0%**	**0,0%**
Total	Flemish region	*158.791*	**2,1%**	2,3%	**4,3%**	2,3%	0,3%	**0,9%**
	Region Brussels capital	*4.839*	**1,3%**	1,9%	**5,4%**	4,1%	0,2%	**1,4%**
	Walloon region	*21.681*	**5,6%**	6,1%	**9,6%**	4,0%	0,4%	**2,1%**
	Belgium	***185.311***	**2,5%**	**2,8%**	**5,0%**	**2,5%**	**0,3%**	**1,1%**

Table 51 Punctures, biopsies and surgery following diagnostic mammography, Belgium, 2006.

			Punctures/ biopsies after mammograpy		Surgery after punctures/ biopsies		
AGE	**REGION**	*Nb ref*	Nbr [a]	**%**	Nbr [b]	**[b/a]**	**%**
35-40 years	Flemish region	*12.297*	458	3,7%	138	30%	1,1%
	Region Brussels capital	*2.602*	132	5,1%	13	9,8%	0,5%
	Walloon region	*10.689*	745	7,0%	122	16%	1,1%
	Belgium	***25.588***	**1.335**	**5,2%**	**273**	**20%**	**1,1%**
40-49 years	Flemish region	*78.851*	2.789	3,5%	887	32%	1,1%
	Region Brussels capital	*16.952*	665	3,9%	109	16%	0,6%
	Walloon region	*69.344*	3.828	5,5%	571	15%	0,8%
	Belgium	***165.147***	**7.282**	**4,4%**	**1.567**	**22%**	**0,9%**
50-69 years	Flemish region	*94.630*	3.176	3,4%	1.513	48%	1,6%
	Region Brussels capital	*27.388*	962	3,5%	273	28%	1,0%
	Walloon region	*119.617*	5.098	4,3%	1.291	25%	1,1%
	Belgium	***241.635***	**9.236**	**3,8%**	**3.077**	**33%**	**1,3%**
70-74 years	Flemish region	*15.749*	638	4,1%	400	63%	2,5%
	Region Brussels capital	*3.906*	154	3,9%	69	45%	1,8%
	Walloon region	*14.661*	624	4,3%	223	36%	1,5%
	Belgium	***34.316***	**1.416**	**4,1%**	**692**	**49%**	**2,0%**
75-79 years	Flemish region	*6.919*	359	5,2%	246	69%	3,6%
	Region Brussels capital	*2.154*	88	4,1%	45	51%	2,1%
	Walloon region	*7.164*	396	5,5%	178	45%	2,5%
	Belgium	***16.237***	**843**	**5,2%**	**469**	**56%**	**2,9%**
Total	Flemish region	*208.446*	7.420	3,6%	3.184	43%	1,5%
	Region Brussels capital	*53.002*	2.001	3,8%	509	25%	1,0%
	Walloon region	*221.475*	10.691	4,8%	2.385	22%	1,1%
	Belgium	***482.923***	**20.112**	**4,2%**	**6.078**	**30%**	**1,3%**

Table 52 Punctures, biopsies and surgery following screening mammography (mammotest), Belgium, 2006

			Punctions/ biopsies after exam ref.		Surgery after punctions/ biopsies		
AGE	**REGION**	*Nb ref*	Nbr [a]	**%**	Nbr [b]	**% [b/a]**	**%**
40-49 years	Flemish region	*6*	0	/	0	/	0
	Region Brussels capital	*4*	0	/	0	/	0
	Walloon region	*1*	0	/	0	/	0
	Belgium	***11***	**0**	**/**	**0**	**/**	**0**
50-69 years	Flemish region	*158.756*	1.463	0,9%	799	55%	1%
	Region Brussels capital	*4.830*	68	1,4%	19	28%	0%
	Walloon region	*21.667*	464	2,1%	122	26%	1%
	Belgium	***185.253***	**1.995**	**1,1%**	**940**	**47%**	**1%**
70-74 ans	Flemish region	*29*	0	/	0	/	0
	Region Brussels capital	*4*	0	/	0	/	0
	Walloon region	*13*	0	/	0	/	0
	Belgium	***46***	**0**	**/**	**0**	**/**	**0**
75-79 ans	Region Brussels capital	*1*	0	/	0	/	0
	Belgium	***1***	**0**	**/**	**0**	**/**	**0**
Total	Flemish region	*158.791*	1.463	0,9%	799	55%	1%
	Region Brussels capital	*4.839*	68	1,4%	19	28%	0%
	Walloon region	*21.681*	464	2,1%	122	26%	1%
	Belgium	***185.311***	**1.995**	**1,1%**	**940**	**47%**	**1%**

Table 53 Evolution of diagnostic mammographies and screening mammographies (mammotest) per 100 000 from the period 2002 to 2007 by region and age group, Belgium

		Diagnostic mammography						Mammotest					
		2002	2003	2004	2005	2006	2007	2002	2003	2004	2005	2006	2007
Flemish region	35-40 years	6.091	6.022	5.818	6.279	5.961	5.944	0	0	0	0	0	0
	40-49 years	15.674	16.174	16.346	17.245	17.953	18.685	0	0	0	0	0	0
	50-69 years	15.209	15.260	14.880	15.357	15.549	15.595	15.325	17.360	18.008	18.680	22.225	22.274
	70-74 years	7.906	8.709	9.430	10.340	11.185	12.508	0	0	0	0	0	0
	75-79 years	4.448	4.742	5.057	5.650	6.369	7.042	0	0	0	0	0	0
	Total	**12.518**	**12.800**	**12.793**	**13.456**	**13.847**	**14.249**						
Region Brussels Capital	35-40 years	8.093	7.745	7.855	7.863	7.425	7.481	0	0	0	0	0	0
	40-49 years	25.751	25.268	26.531	27.203	27.712	29.104	0	0	0	0	0	0
	50-69 years	29.897	28.931	29.751	29.524	30.250	29.972	1.051	2.196	3.122	3.352	4.998	4.503
	70-74 years	18.397	19.373	19.933	20.510	21.140	21.458	0	0	0	0	0	0
	75-79 years	11.275	11.280	12.436	12.656	13.592	14.404	0	0	0	0	0	0
	Total	**22.769**	**22.335**	**23.213**	**23.386**	**23.916**	**24.267**						
Walloon region	35-40 years	9.574	9.499	9.661	9.599	9.610	9.249	0	0	0	0	0	0
	40-49 years	26.854	27.353	28.207	28.592	29.187	29.490	0	0	0	0	0	0
	50-69 years	31.381	30.924	31.380	31.917	32.867	32.081	1.164	5.406	4.444	4.113	5.510	3.577
	70-74 years	16.315	17.248	18.409	19.059	20.163	20.144	0	0	0	0	0	0
	75-79 years	8.851	9.382	10.328	10.714	11.431	12.144	0	0	0	0	0	0
	Total	**23.796**	**23.961**	**24.702**	**25.208**	**26.001**	**25.810**						
Belgium	35-40 years	7.364	7.274	7.222	7.480	7.265	7.157	0	0	0	0	0	0
	40-49 years	20.140	20.540	21.005	21.705	22.351	22.996	0	0	0	0	0	0
	50-69 years	21.636	21.459	21.463	21.902	22.391	22.149	9.581	12.239	12.379	12.690	15.376	14.724
	70-74 years	11.568	12.372	13.168	13.912	14.784	15.572	0	0	0	0	0	0
	75-79 years	6.604	6.929	7.500	7.954	8.659	9.320	0	0	0	0	0	0
	Total	**17.045**	**17.224**	**17.527**	**18.085**	**18.614**	**18.818**						

Table 54 Evolution of biopsies and punctures per 100 000 women from the period 2002 to 2007 by region and age group, Belgium

		Biopsy						Punctures					
		2002	2003	2004	2005	2006	2007	2002	2003	2004	2005	2006	2007
Flemish region	35-40 years	16	11	15	18	20	18	328	323	334	349	352	396
	40-49 years	30	34	43	36	26	37	805	852	861	905	953	986
	50-69 years	42	35	42	38	31	42	877	939	948	972	990	1.024
	70-74 years	31	26	21	32	20	35	373	463	509	558	660	673
	75-79 years	20	19	26	18	19	35	311	362	423	493	523	551
	Total	**32**	**29**	**35**	**33**	**26**	**37**	**688**	**741**	**760**	**796**	**831**	**865**
Region Brussels Capital	35-40 years	66	68	89	94	47	83	540	624	561	599	526	543
	40-49 years	154	198	215	221	105	198	1.322	1.425	1.515	1.339	1.456	1.511
	50-69 years	170	179	233	172	135	214	1.474	1.532	1.630	1.302	1.560	1.483
	70-74 years	134	178	169	127	151	197	966	1.140	1.029	881	1.206	1.155
	75-79 years	114	102	134	131	64	146	689	739	823	797	859	893
	Total	**142**	**160**	**192**	**166**	**109**	**183**	**1.175**	**1.261**	**1.315**	**1.125**	**1.285**	**1.269**
Walloon region	35-40 years	37	31	37	20	21	34	813	830	849	781	895	866
	40-49 years	74	53	74	69	30	63	2.100	2.092	2.169	2.084	2.091	2.060
	50-69 years	93	77	69	62	35	71	2.021	2.154	2.214	2.028	2.046	1.932
	70-74 years	60	44	40	58	31	53	1.127	1.113	1.210	1.262	1.222	1.219
	75-79 years	30	45	31	49	21	41	659	791	850	921	922	983
	Total	**72**	**59**	**60**	**57**	**30**	**60**	**1.675**	**1.746**	**1.815**	**1.721**	**1.745**	**1.692**
Belgium	35-40 years	27	23	29	26	23	30	499	510	517	510	542	561
	40-49 years	54	54	67	62	34	59	1.269	1.300	1.336	1.319	1.358	1.373
	50-69 years	69	61	66	57	41	66	1.295	1.382	1.415	1.343	1.381	1.358
	70-74 years	50	44	39	48	34	53	672	731	776	807	880	879
	75-79 years	32	36	37	38	24	47	466	543	603	662	683	720
	Total	**55**	**50**	**57**	**52**	**34**	**56**	**1.049**	**1.110**	**1.148**	**1.123**	**1.164**	**1.166**

Table 55 Evolution of number of Halsted and mastectomies per 100 000 women from the period 2002 to 2007 by region and age group, Belgium

		Halsted						Mastectomies					
		2002	2003	2004	2005	2006	2007	2002	2003	2004	2005	2006	2007
Flemish region	35-40 years	9	7	7	6	8	6	18	19	18	18	20	19
	40-49 years	23	24	22	21	20	22	25	30	30	34	33	32
	50-69 years	55	50	41	38	32	36	31	35	34	40	45	40
	70-74 years	55	55	53	47	43	45	25	16	26	22	20	29
	75-79 years	64	74	51	64	54	55	22	16	22	28	23	25
	Total	**41**	**39**	**33**	**32**	**29**	**31**	**26**	**28**	**29**	**33**	**35**	**33**
Region Brussels Capital	35-40 years	8	5	11	11	8	2	5	2	5	2	8	5
	40-49 years	17	12	18	23	14	15	19	11	29	25	20	17
	50-69 years	53	39	36	26	24	28	37	47	62	38	38	44
	70-74 years	44	36	47	34	20	46	44	41	18	14	40	26
	75-79 years	77	37	48	29	39	40	54	14	38	43	24	30
	Total	**38**	**27**	**29**	**24**	**20**	**23**	**30**	**27**	**39**	**27**	**27**	**28**
Walloon region	35-40 years	9	4	5	4	3	0	3	12	12	11	6	4
	40-49 years	17	19	15	8	10	11	27	18	27	26	18	20
	50-69 years	41	34	30	22	22	20	33	28	31	34	27	30
	70-74 years	45	39	26	25	24	22	21	17	24	28	30	28
	75-79 years	46	49	31	29	22	21	16	21	21	17	24	19
	Total	**31**	**28**	**22**	**17**	**17**	**15**	**25**	**22**	**26**	**27**	**22**	**23**
Belgium	35-40 years	9	6	7	6	7	4	12	15	15	14	14	13
	40-49 years	20	22	20	17	16	18	25	25	29	31	28	27
	50-69 years	50	44	37	32	28	30	32	34	36	38	39	37
	70-74 years	51	48	44	39	35	38	25	18	25	23	25	28
	75-79 years	59	62	44	49	42	42	23	18	23	26	23	23
	Total	**37**	**35**	**30**	**27**	**24**	**25**	**26**	**26**	**29**	**31**	**30**	**29**

Table 56 Evolution of partial mastectomies and tumorectomies from the period 2002 to 2007 by region and age group, Belgium

		Partial mastectomies						Tumorectomies					
		2002	2003	2004	2005	2006	2007	2002	2003	2004	2005	2006	2007
Flemish region	35-40 years	49	44	38	36	33	50	224	184	156	160	134	158
	40-49 years	93	116	103	116	111	107	310	268	225	224	223	219
	50-69 years	215	235	210	199	205	207	355	295	249	230	220	224
	70-74 years	134	140	165	132	176	171	134	122	127	119	137	121
	75-79 years	119	112	130	133	135	136	107	88	87	94	95	100
	Total	**143**	**158**	**147**	**144**	**149**	**151**	**284**	**240**	**206**	**198**	**192**	**195**
Region Brussels Capital	35-40 years	46	48	42	41	33	32	98	139	81	103	63	78
	40-49 years	110	126	155	151	139	162	225	193	196	168	131	131
	50-69 years	251	283	288	242	293	283	202	178	185	126	131	157
	70-74 years	224	320	235	251	343	265	130	128	136	88	111	124
	75-79 years	219	159	250	199	203	242	118	88	96	102	49	35
	Total	**178**	**198**	**208**	**184**	**208**	**207**	**178**	**163**	**160**	**128**	**112**	**125**
Walloon region	35-40 years	35	40	54	43	52	47	153	141	116	116	112	102
	40-49 years	136	130	135	153	141	152	241	227	203	184	182	169
	50-69 years	246	273	260	240	267	246	215	201	189	163	158	135
	70-74 years	196	198	201	217	204	187	121	102	99	112	75	81
	75-79 years	121	149	161	167	139	186	67	60	71	63	53	61
	Total	**172**	**186**	**186**	**183**	**189**	**186**	**191**	**178**	**165**	**150**	**142**	**129**
Belgium	35-40 years	44	43	43	38	39	47	190	166	136	141	120	132
	40-49 years	108	122	118	131	123	126	281	249	216	207	202	196
	50-69 years	228	251	233	216	232	226	297	255	224	200	192	189
	70-74 years	162	174	182	169	198	183	130	116	119	115	115	109
	75-79 years	129	129	152	150	142	161	94	78	82	85	77	82
	Total	**156**	**170**	**165**	**160**	**167**	**167**	**245**	**213**	**189**	**177**	**169**	**168**

Table 57 Delay (days) between diagnostic and screening mammographies, percentile for the region of Flanders

		Diagnostic mammographies *followed by complementary tests*						***Mammotests*** *followed by complementary tests*					
		N	**P 10**	**P 25**	**P 50**	**P 75**	**P 90**	*N*	**P 10**	**P 25**	**P 50**	**P 75**	**P 90**
Outpatient	35-40 years	*13.563*	0	0	0	0	0	/	/	/	/	/	/
Diagnostic	40-49 years	*63.331*	0	0	0	0	0	*470*	21	26	34	43	57
Mammography	50-69 years	*52.438*	0	0	0	0	0	*2.706*	18	23	32	43	56
	70-74 years	*10.703*	0	0	0	0	0	*1*	0	0	0	0	0
	75-79 years	*4.235*	0	0	0	0	0	/	/	/	/	/	/
	Total	***144.270***	**0**	**0**	**0**	**0**	**0**	***3.177***	**18**	**24**	**32**	**43**	**56**
Inpatient	35-40 years	*15*	7	13	20	30	36	/	/	/	/	/	/
Diagnostic	40-49 years	*93*	8	13	26	39	62	*13*	35	38	49	63	74
Mammography	50-69 years	*144*	13	17	26	39	60	*221*	30	36	47	62	76
	70-74 years	*49*	14	20	25	31	48	/	/	/	/	/	/
	75-79 years	*24*	13	15	22	29	45	/	/	/	/	/	/
	Total	***325***	**11**	**16**	**24**	**37**	**58**	***234***	**30**	**36**	**47**	**62**	**76**
Ultrasound	35-40 years	*12.001*	0	0	0	0	0	/	/	/	/	/	/
	40-49 years	*54.085*	0	0	0	0	0	*942*	16	24	31	42	60
	50-69 years	*42.196*	0	0	0	0	0	*4.813*	14	22	30	42	56
	70-74 years	*7.106*	0	0	0	0	0	/	/	/	/	/	/
	75-79 years	*2.951*	0	0	0	0	0	/	/	/	/	/	/
	Total	***118.339***	**0**	**0**	**0**	**0**	**0**	***5.755***	**14**	**22**	**30**	**42**	**57**
MRI	35-40 years	*238*	4	9	17	33	50	/	/	/	/	/	/
	40-49 years	*904*	5	10	20	35	55	*66*	21	35	44	59	76
	50-69 years	*717*	5	9	17	32	50	*406*	24	32	43	59	76
	70-74 years	*138*	6	10	16	27	41	/	/	/	/	/	/
	75-79 years	*61*	7	11	18	25	44	/	/	/	/	/	/
	Total	***2.058***	**5**	**9**	**18**	**33**	**51**	***472***	**24**	**33**	**43**	**59**	**76**
Poncture or	35-40 years	*494*	0	0	6	17	38	/	/	/	/	/	/
biopsy	40-49 years	*2.086*	0	0	6	18	36	*172*	21	26	38	55	70
	50-69 years	*1.660*	0	0	6	15	32	*1.117*	16	24	35	49	65
	70-74 years	*480*	0	1	6	12	22	/	/	/	/	/	/
	75-79 years	*323*	0	0	6	13	25	/	/	/	/	/	/
	Total	***5.043***	**0**	**0**	**6**	**15**	**33**	***1.289***	**17**	**24**	**35**	**50**	**66**

Table 58 Delay (days) between diagnostic and screening mammographies, percentile for region of Brussels-capital

		Diagnostic mammographies *followed by complementary tests*						***Mammotests*** *followed by complementary tests*					
		N	**P 10**	**P 25**	**P 50**	**P 75**	**P 90**	*N*	**P 10**	**P 25**	**P 50**	**P 75**	**P 90**
Outpatient Diagnostic Mammography	35-40 years	*3.041*	0	0	0	0	0	/	/	/	/	/	/
	40-49 years	*11.180*	0	0	0	0	0	*4*	43	44	48	54	56
	50-69 years	*14.768*	0	0	0	0	0	*86*	18	28	36	56	71
	70-74 years	*2.304*	0	0	0	0	0	/	/	/	/	/	/
	75-79 years	*1.217*	0	0	0	0	0	/	/	/	/	/	/
	Total	***32.510***	**0**	**0**	**0**	**0**	**0**	***90***	**18**	**28**	**39**	**56**	**71**
Inpatient Diagnostic Mammography	40-49 years	*16*	25	30	43	64	84	/	/	/	/	/	/
	50-69 years	*35*	23	26	36	54	65	*5*	45	57	64	72	73
	70-74 years	*8*	13	24	35	43	60	/	/	/	/	/	/
	75-79 years	*3*	18	18	27	47	47	/	/	/	/	/	/
	Total	***62***	**22**	**27**	**36**	**52**	**66**	***5***	**45**	**57**	**64**	**72**	**73**
Ultrasound	35-40 years	*2.775*	0	0	0	0	0	/	/	/	/	/	/
	40-49 years	*9.672*	0	0	0	0	0	*13*	24	43	51	60	83
	50-69 years	*11.721*	0	0	0	0	0	*233*	11	24	37	57	73
	70-74 years	*1.598*	0	0	0	0	0	/	/	/	/	/	/
	75-79 years	*849*	0	0	0	0	0	/	/	/	/	/	/
	Total	***26.615***	**0**	**0**	**0**	**0**	**0**	***246***	**12**	**24**	**38**	**57**	**75**
MRI	35-40 years	*24*	5	9	15	30	46	/	/	/	/	/	/
	40-49 years	*82*	6	8	20	41	54	*1*	43	43	43	43	43
	50-69 years	*125*	6	11	20	39	57	*7*	42	42	55	71	72
	70-74 years	*21*	2	11	28	49	65	/	/	/	/	/	/
	75-79 years	*13*	7	12	14	23	28	/	/	/	/	/	/
	Total	***265***	**6**	**9**	**19**	**38**	**57**	***8***	**42**	**43**	**51**	**64**	**72**
Poncture or biopsy	35-40 years	*150*	0	0	0	7	21	/	/	/	/	/	/
	40-49 years	*442*	0	0	0	5	22	*2*	20	20	32	43	43
	50-69 years	*460*	0	0	0	12	28	*39*	15	25	35	54	63
	70-74 years	*93*	0	0	5	24	37	/	/	/	/	/	/
	75-79 years	*46*	0	0	6	17	33	/	/	/	/	/	/
	Total	***1.191***	**0**	**0**	**0**	**10**	**28**	***41***	**17**	**25**	**35**	**51**	**63**

Table 59 Delay between diagnostic and screening mammographies, percentile for region of Walloon region

		Diagnostic mammographies *followed by complementary tests*						***Mammotests*** *followed by complementary tests*					
		N	**P 10**	**P 25**	**P 50**	**P 75**	**P 90**	*N*	**P 10**	**P 25**	**P 50**	**P 75**	**P 90**
Outpatient Diagnostic Mammography	35-40 years	*10.877*	0	0	0	0	0	/	/	/	/	/	/
	40-49 years	*44.125*	0	0	0	0	0	*47*	24	34	43	59	76
	50-69 years	*62.417*	0	0	0	0	0	*773*	20	27	36	53	70
	70-74 years	*7.862*	0	0	0	0	0	/	/	/	/	/	/
	75-79 years	*3.982*	0	0	0	0	0	/	/	/	/	/	/
	Total	***129.263***	**0**	**0**	**0**	**0**	**0**	***820***	**20**	**28**	**37**	**53**	**71**
Inpatient Diagnostic Mammography	35-40 years	*6*	28	34	36	42	56	/	/	/	/	/	/
	40-49 years	*26*	11	16	28	35	66	*1*	35	35	35	35	35
	50-69 years	*70*	16	23	32	51	69	*12*	21	29	50	55	58
	70-74 years	*16*	14	22	32	43	71	/	/	/	/	/	/
	75-79 years	*8*	12	17	30	45	61	/	/	/	/	/	/
	Total	***126***	**14**	**21**	**31**	**45**	**66**	***13***	**21**	**35**	**48**	**54**	**58**
Ultrasound	35-40 years	*10.253*	0	0	0	0	0	/	/	/	/	/	/
	40-49 years	*40.923*	0	0	0	0	0	*75*	24	29	41	56	76
	50-69 years	*54.648*	0	0	0	0	0	*1.217*	20	28	37	53	70
	70-74 years	*6.450*	0	0	0	0	0	/	/	/	/	/	/
	75-79 years	*3.295*	0	0	0	0	0	/	/	/	/	/	/
	Total	***115.569***	**0**	**0**	**0**	**0**	**0**	***1.292***	**20**	**28**	**37**	**54**	**70**
MRI	35-40 years	*115*	4	9	16	27	45	/	/	/	/	/	/
	40-49 years	*374*	3	8	17	28	49	*2*	54	54	57	59	59
	50-69 years	*590*	6	10	18	31	50	*47*	25	36	50	70	76
	70-74 years	*77*	7	10	17	31	47	/	/	/	/	/	/
	75-79 years	*45*	4	12	19	33	56	/	/	/	/	/	/
	Total	***1.201***	**5**	**9**	**17**	**29**	**49**	***49***	**25**	**37**	**50**	**69**	**76**
Poncture or biopsy	35-40 years	*726*	0	0	0	0	7	/	/	/	/	/	/
	40-49 years	*2.282*	0	0	0	0	12	*19*	14	26	43	57	81
	50-69 years	*2.336*	0	0	0	0	17	*241*	17	24	36	54	72
	70-74 years	*350*	0	0	0	3	18	/	/	/	/	/	/
	75-79 years	*257*	0	0	0	1	11	/	/	/	/	/	/
	Total	***5.951***	**0**	**0**	**0**	**0**	**14**	***260***	**17**	**24**	**36**	**54**	**72**

Table 60 Delay between biopsy and surgery after diagnostic mammography per region and age-group

		Within the month		Between 1 and 3 month		Between 3 and 6 month		Between 6 and 9 month		Between 9 and 12 month		After more than 12 month	
		Nbr	*Pct*	Nbr	*Pct*	Nbr	*Pct*	Nbr	*Pct*	Nbr	Pct	Nbr	*Pct*
Flemish region	35-40 years	153	*72%*	27	*13%*	20	*9,4%*	5	*2,3%*	4	*1,9%*	4	*1,9%*
	40-49 years	1.005	*74%*	182	*13%*	101	*7,5%*	29	*2,1%*	14	*1,0%*	23	*1,7%*
	50-69 years	1.771	*80%*	253	*11%*	100	*4,5%*	50	*2,3%*	20	*0,9%*	13	*0,6%*
	70-74 years	412	*83%*	52	*10%*	16	*3,2%*	6	*1,2%*	5	*1,0%*	5	*1,0%*
	75-79 years	231	*81%*	39	*14%*	9	*3,2%*	3	*1,1%*	1	*0,4%*	1	*0,4%*
	Total	**3.572**	***78%***	**553**	***12%***	**246**	***5,4%***	**93**	***2,0%***	**44**	***1,0%***	**46**	***1,0%***
Region Brussels Capital	35-40 years	9	*36%*	5	*20%*	9	*36%*	2	*8,0%*	/	/	/	/
	40-49 years	99	*44%*	69	*31%*	33	*15%*	15	*6,7%*	4	*1,8%*	4	*1,8%*
	50-69 years	254	*51%*	158	*32%*	56	*11%*	19	*3,8%*	2	*0,4%*	9	*1,8%*
	70-74 years	54	*48%*	46	*41%*	8	*7,1%*	4	*3,6%*	/	/	/	/
	75-79 years	36	*55%*	25	*38%*	3	*4,5%*	/	/	/	/	2	*3,0%*
	Total	**452**	***49%***	**303**	***33%***	**109**	***12%***	**40**	***4,3%***	**6**	***0,6%***	**15**	***1,6%***
Walloon region	35-40 years	81	*40%*	59	*29%*	35	*17%*	15	*7,4%*	7	*3,5%*	5	*2,5%*
	40-49 years	512	*51%*	287	*28%*	97	*9,6%*	63	*6,3%*	19	*1,9%*	30	*3,0%*
	50-69 years	1.135	*54%*	720	*34%*	127	*6,0%*	75	*3,6%*	17	*0,8%*	31	*1,5%*
	70-74 years	163	*49%*	126	*38%*	23	*6,9%*	10	*3,0%*	3	*0,9%*	6	*1,8%*
	75-79 years	126	*54%*	87	*37%*	8	*3,4%*	6	*2,6%*	2	*0,9%*	4	*1,7%*
	Total	**2.017**	***52%***	**1.279**	***33%***	**290**	***7,5%***	**169**	***4,4%***	**48**	***1,2%***	**76**	***2,0%***
Belgium	35-40 years	243	*55%*	91	*21%*	64	*15%*	22	*5,0%*	11	*2,5%*	9	*2,0%*
	40-49 years	1.616	*62%*	538	*21%*	231	*8,9%*	107	*4,1%*	37	*1,4%*	57	*2,2%*
	50-69 years	3.160	*66%*	1.131	*24%*	283	*5,9%*	144	*3,0%*	39	*0,8%*	53	*1,1%*
	70-74 years	629	*67%*	224	*24%*	47	*5,0%*	20	*2,1%*	8	*0,9%*	11	*1,2%*
	75-79 years	393	*67%*	151	*26%*	20	*3,4%*	9	*1,5%*	3	*0,5%*	7	*1,2%*
	Total	**6.041**	***65%***	**2.135**	***23%***	**645**	***6,9%***	**302**	***3,2%***	**98**	***1,0%***	**137**	***1,5%***

Table 61 Delay between biopsy and surgery after screening mammography per region

	Within the month		Between 1 and 3 month		Between 3 and 6 month		Between 6 and 9 month		Between 9 and 12 month		After more than 12 month	
age 50-69 years	Nbr	*Pct*	Nbr	*Pct*	Nbr	*Pct*	Nbr	*Pct*	Nbr	Pct	Nbr	*Pct*
Flemish region	1.010	*87%*	115	*9,9%*	22	*1,9%*	9	*0,8%*	7	*0,6%*	2	*0,2%*
Region Brussels-Capital	21	*51%*	16	*39%*	2	*4,9%*	1	*2,4%*	1	*2,4%*	/	/
Walloon region	112	*52%*	89	*41%*	6	*2,8%*	8	*3,7%*	/	/	1	*0,5%*
Belgium	**1.143**	***80%***	**220**	***15%***	**30**	***2,1%***	**18**	***1,3%***	**8**	***0,6%***	**3**	***0,2%***

Table 62 Delay between biopsy and surgery after screening or diagnostic mammography per region and age-group

		Within the month		Between 1 and 3 month		Between 3 and 6 month		Between 6 and 9 month		Between 9 and 12 month		After more than 12 month	
		Nbr	*Pct*	Nbr	*Pct*	Nbr	*Pct*	Nbr	*Pct*	Nbr	Pct	Nbr	*Pct*
Flemish region	35-40 years	153	*72%*	27	*13%*	20	*9,4%*	5	*2,3%*	4	*1,9%*	4	*1,9%*
	40-49 years	1.005	*74%*	182	*13%*	101	*7,5%*	29	*2,1%*	14	*1,0%*	23	*1,7%*
	50-69 years	2.781	*82%*	368	*11%*	122	*3,6%*	59	*1,7%*	27	*0,8%*	15	*0,4%*
	70-74 years	412	*83%*	52	*10%*	16	*3,2%*	6	*1,2%*	5	*1,0%*	5	*1,0%*
	75-79 years	231	*81%*	39	*14%*	9	*3,2%*	3	*1,1%*	1	*0,4%*	1	*0,4%*
	Total	**4.582**	***80%***	**668**	***12%***	**268**	***4,7%***	**102**	***1,8%***	**51**	***0,9%***	**48**	***0,8%***
Region Brussels Capital	35-40 years	9	*36%*	5	*20%*	9	*36%*	2	*8,0%*	/	/	/	/
	40-49 years	99	*44%*	69	*31%*	33	*15%*	15	*6,7%*	4	*1,8%*	4	*1,8%*
	50-69 years	275	*51%*	174	*32%*	58	*11%*	20	*3,7%*	3	*0,6%*	9	*1,7%*
	70-74 years	54	*48%*	46	*41%*	8	*7,1%*	4	*3,6%*	/	/	/	/
	75-79 years	36	*55%*	25	*38%*	3	*4,5%*	/	/	/	/	2	*3,0%*
	Total	**473**	***49%***	**319**	***33%***	**111**	***11%***	**41**	***4,2%***	**7**	***0,7%***	**15**	***1,6%***
Walloon region	35-40 years	81	*40%*	59	*29%*	35	*17%*	15	*7,4%*	7	*3,5%*	5	*2,5%*
	40-49 years	512	*51%*	287	*28%*	97	*9,6%*	63	*6,3%*	19	*1,9%*	30	*3,0%*
	50-69 years	1.247	*54%*	809	*35%*	133	*5,7%*	83	*3,6%*	17	*0,7%*	32	*1,4%*
	70-74 years	163	*49%*	126	*38%*	23	*6,9%*	10	*3,0%*	3	*0,9%*	6	*1,8%*
	75-79 years	126	*54%*	87	*37%*	8	*3,4%*	6	*2,6%*	2	*0,9%*	4	*1,7%*
	Total	**2.129**	***52%***	**1.368**	***33%***	**296**	***7,2%***	**177**	***4,3%***	**48**	***1,2%***	**77**	***1,9%***
Belgium	35-40 years	243	*55%*	91	*21%*	64	*15%*	22	*5,0%*	11	*2,5%*	9	*2,0%*
	40-49 years	1.616	*62%*	538	*21%*	231	*8,9%*	107	*4,1%*	37	*1,4%*	57	*2,2%*
	50-69 years	4.303	*69%*	1.351	*22%*	313	*5,0%*	162	*2,6%*	47	*0,8%*	56	*0,9%*
	70-74 years	629	*67%*	224	*24%*	47	*5,0%*	20	*2,1%*	8	*0,9%*	11	*1,2%*
	75-79 years	393	*67%*	151	*26%*	20	*3,4%*	9	*1,5%*	3	*0,5%*	7	*1,2%*
	Total	**7.184**	***67%***	**2.355**	***22%***	**675**	***6,3%***	**320**	***3%***	**106**	***1%***	**140**	***1,3%***

Table 63 Delay (days) between diagnostic tests: Diagnostic mammography (DM), Screening mammography (MT), mean and percentile, for Belgium and region of Flanders

Belgium

	DM-DM							DM-MT							MT-DM						
	N	Mean	P 10	P 25	P 50	P 75	P 90	*N*	Mean	P 10	P 25	P 50	P 75	P 90	*N*	Mean	P 10	P 25	P 50	P 75	P 90
35-40 year	*6.921*	364	179	304	370	435	539	/	/	/	/	/	/	/	/	/	/	/	/	/	/
40-49 year	*63.619*	381	218	342	377	440	533	*1.588*	398	235	322	389	476	582	*598*	57	21	28	38	56	139
50-69 year	*113.614*	373	214	341	371	420	517	*9.485*	376	204	288	367	458	573	*19.042*	228	24	36	172	398	518
70-74 year	*13.445*	362	199	334	368	405	493	/	/	/	/	/	/	/	*9*	201	8	45	130	363	498
75-79 year	*6.353*	358	197	329	366	403	486	/	/	/	/	/	/	/	/	/	/	/	/	/	/
Total	***203.952***	**374**	**210**	**340**	**371**	**425**	**521**	***11.073***	**379**	**210**	**294**	**370**	**462**	**573**	***19.649***	**223**	**24**	**36**	**147**	**393**	**513**

Region of Flanders

	DM-DM							DM-MT							MT-DM						
	N	Mean	P 10	P 25	P 50	P 75	P 90	*N*	Mean	P 10	P 25	P 50	P 75	P 90	*N*	Mean	P 10	P 25	P 50	P 75	P 90
35-40 year	*2.887*	364	183	311	369	429	532	/	/	/	/	/	/	/	/	/	/	/	/	/	/
40-49 year	*24.087*	375	205	336	372	433	533	*1.410*	394	232	316	385	471	574	*535*	55	21	28	36	54	125
50-69 year	*44.757*	364	213	341	369	404	490	*7.088*	369	196	280	362	451	569	*13.673*	223	23	35	160	392	513
70-74 year	*5.944*	361	220	341	367	398	476	/	/	/	/	/	/	/	*4*	221	8	12	190	431	498
75-79 year	*2.675*	361	210	340	367	399	485	/	/	/	/	/	/	/	/	/	/	/	/	/	/
Total	***80.350***	**367**	**210**	**339**	**370**	**412**	**506**	***8.498***	**374**	**203**	**285**	**365**	**455**	**570**	***14.212***	**217**	**23**	**34**	**125**	**388**	**510**

Table 64 Delay (days) between diagnostic tests: Diagnostic mammography (DM), Screening mammography (MT), mean and percentile, for region of Brussels-capital and Walloon region

Region of Brussels-capital

	DM-DM							DM-MT							MT-DM						
	N	Mean	P 10	P 25	P 50	P 75	P 90	*N*	Mean	P 10	P 25	P 50	P 75	P 90	*N*	Mean	P 10	P 25	P 50	P 75	P 90
35-40 year	*719*	365	177	309	371	446	532	/	/	/	/	/	/	/	/	/	/	/	/	/	/
40-49 year	*8.137*	390	255	354	380	445	532	*46*	459	348	363	448	550	606	*5*	113	43	45	51	56	371
50-69 year	*12.565*	382	231	349	375	430	524	*552*	409	249	330	399	494	581	*715*	310	36	105	360	444	530
70-74 year	*1.577*	365	194	333	369	415	508	/	/	/	/	/	/	/	/	/	/	/	/	/	/
75-79 year	*832*	361	207	327	368	406	498	/	/	/	/	/	/	/	/	/	/	/	/	/	/
Total	***23.830***	**382**	**231**	**349**	**376**	**434**	**525**	***598***	**413**	**254**	**335**	**406**	**498**	**584**	***720***	**309**	**36**	**98**	**359**	**444**	**530**

Walloon Region

	DM-DM							DM-MT							MT-DM						
	N	Mean	P 10	P 25	P 50	P 75	P 90	*N*	Mean	P 10	P 25	P 50	P 75	P 90	*N*	Mean	P 10	P 25	P 50	P 75	P 90
35-40 year	*3.315*	363	175	297	371	439	548	/	/	/	/	/	/	/	/	/	/	/	/	/	/
40-49 year	*31.395*	384	222	343	378	441	536	*132*	420	286	336	406	494	597	*58*	70	24	35	49	76	175
50-69 year	*56.292*	377	211	338	371	431	530	*1.845*	389	227	309	380	467	575	*4.654*	229	27	42	146	401	525
70-74 year	*5.924*	361	190	325	367	408	504	/	/	/	/	/	/	/	*5*	186	45	50	130	208	495
75-79 year	*2.846*	354	186	313	365	406	483	/	/	/	/	/	/	/	/	/	/	/	/	/	/
Total	***99.772***	**377**	**209**	**337**	**372**	**434**	**531**	***1.977***	**391**	**229**	**311**	**381**	**468**	**576**	***4.717***	**227**	**27**	**41**	**138**	**400**	**523**

Method to estimate proportion opportunistic screening amongst women undergoing diagnostic mammography.

Let

a = proportion surgery in the group organised screening mammography

b = observed proportion surgery in the group ‘diagnostic’ mammography

X = proportion opportunistic screening

c = assumed proportion surgery in the symptomatic group.

we assume that the proportion surgery in the group organised screening mammography is representative for the proportion surgery in the group opportunistic screening

then the observed proportion surgery in the group ‘diagnostic’ mammography consist of a part mammography for opportunistic screening and a part ‘true’ diagnostic screenings and following equation holds:

b = ax + (1-x)c

We let c vary from 5 % to 30 % and we end up with one unknown x

solving x in function of a, b and c gives:

x = (b – c)/(a-c)

APPENDIX 5. GRADE: THE STRENGTH OF RECOMMENDATION AND LEVEL OF EVIDENCE

Short explanation on the GRADE approach:

GRADE is an approach developed by the Grading of Recommendations Assessment, Development and Evaluation (GRADE) Working Group. GRADE offers four levels of evidence quality: high, moderate, low, and very low. Randomised trials begin as high quality evidence and observational studies as low quality evidence. Quality may be downgraded as a result of limitations in study design or implementation, imprecision of estimates (wide confidence intervals), variability in results, indirectness of evidence, or publication bias. Quality may be upgraded because of a large or very large magnitude of effect, a dose-response gradient, and if all plausible biases would reduce an apparent treatment effect. A special approach is used for evidence concerning diagnostic studies, where a evidence form good quality diagnostic studies are graded high level, provided the linked with clinical outcomes is sufficiently direct. The Grade approach also give an indication on the strength of the recommendation, based on whether *(a)* the evidence is high quality and the desirable effects clearly outweigh the undesirable effects, or *(b)* there is a close or uncertain balance. There are limitations to formal grading of recommendations. Like the quality of evidence, the balance between desirable and undesirable effects reflects a continuum. Some arbitrariness will therefore be associated with placing particular recommendations in categories such as “strong” and “weak.” GRADE is the result of a judgement, and the developpers warn against a too mechanical application of the approach.

Explanation on how GRADE was applied to the recommendations in this report:

In the following section we explain the arguments underlying the way we accorded a strenght of recommendation and a level of evidence to the different recommendations.

1. Family history is the strongest risk factor

Women can be categorised in 3 risk categories based on family history *(strong recommendation, moderate level of evidence)***.**

Evidence based on a meta-analysis of observational studies showing a strong effect of family history, so the low level usually attributed to observational studies was upgraded.

2. Women with a high breast cancer risk based on the above mentioned criteria are eligible for individual risk assessment in order to give individual advise on screening strategy, genetic tests and prophylactic measures. Individual risk assessment consists of an in depth family history and can make use of computerized risk models such as the Gail model or the Tirer-Cuzick model only. Models integrating dense breast tissue, e.g. Tice-model, need further validation. Individual risk assessment should be done by professionals with appriopriate training and skills with extensive counselling and sufficient attention to patient preferences and support. *(weak recommendation, very low level of evidence)***.**

There is now direct evidence of the benefit of the proposed strategy, validation studies of the different risk prediction models are inconclusive or non existent. Therefore we downgraded the evidence to very low.

B. Risk factors other than family history

3. Persons with a past history of mantle irradiation for Hodgkin lymphoma should be considered at high risk *(strong recommendation, moderate level of evidence)***.**

Observational evidence upgraded because the observed effect was large.

4. Women with very dense breast tissue (BIRADS 4) could be considered as raised risk (life-time risk +/-17 %) *(weak recommendation, very low level of evidence)***.**

Observational evidence downgraded because of imprecision: the point estimate of effect was just enough to place this women in the raised risk category but confience intervals are compatibel with average risk, moreover, there are considerable problems with the reproducibility of the radiological assessment of dense breast.

5. Lobular and ductal atypical hyperplasia should be considered as high risk *(weak recommendation, low level of evidence)***.**

This was based on observational data that were neither upgraded nor downgraded.

6. Other risk factors such as BIRADS 3, obesitas, alcohol intake, hormone replacement therapy, early menarche, nulliparity, oral contraceptives, or exogenous hormones should be used only as an element integrated in comprehensive risk models as they are only moderately or modestly associated with breast cancer (strong recommendation, low level of evidence)**.**

This was based on observational data that were neither upgraded nor downgraded.

Which techniques should be used?

7. Every screening mammography should be performed in a setting with adequate quality control following the European guidelines and evaluated with independent double reading.. A consensus or arbitration procedure should be used in case of discordance. *(strong recommendation, high level of evidence)***.**

High quality evidence from validation studies implications for patient outcomes are sufficiently direct and consistent to justify a the fact the the default high level was not downgraded.

8. The use of computer-aided detection is not recommended and cannot replace quality controlled mammography with double reading *(strong recommendation, very low level of evidence)***.**

Validation studies downgraded for heterogeneity of the estimates, indirectness as comparisons are made with single reading and imprecision of the estimates.

9. Film –screen and full-field digital mammography can both be used for screening purposes, with similar accuracy. The use of digital mammography can be beneficial for young women and women with dense breast tissue *(weak recommendation, low level of evidence)***.**

Validation studies downgraded for an indirect link with a clinical benefit and heterogeneity of the digital mammography techniques used.

10. Ultrasound screening is not recommended in a population-based screening program as the recall rate and number of false positives is too high and the additional cancer detection rate is minimal *(strong recommendation, low level of evidence)***.**

Validation studies downgraded for indirectness as available studies are not conducted in preselected patients, with as consquence that implications and because of heterogeneity of the estimates.

11. Currently available data do not support the use of ultrasound in women with dense breast tissue on mammography outside a clinical trial setting. *(strong recommendation, low level of evidence)***.**

Validation studies downgraded for indirectness as there are considerable problems with the reproducibility of the assessment of dense breast and low quality of the primary studies.

12. Women with raised risk or greater should be offered annual mammographic surveillance from age 40 – 49 years within a quality assured program following European guidelines. From the age of 50 to 69 years, women with a raised breast cancer risk can be included in the general screening program with biennial mammography (weak recommendation, very low level of evidence).

There is no direct evidence that this recommendation actually improves clinical outcomes.

13. For women at high risk for breast cancer, yearly MRI and mammography is recommended from the age of 30 years onwards or starting five years before the age of the youngest diagnosed family member with breast cancer (strong recommendation, very low level of evidence).The use of ultrasound can be considered to shorten the interval or as adjunct to a positive mammography or MRI *(weak recommendation, very low level of evidence)***.**

There is no direct evidence that this recommendation actually improves clinical outcomes.

14. All women participating in screening should be informed about the risk for false positive results, the remaining risk for interval cancer and the absence of data on long term benefits for screening outside the population-based screening program, decisions should be taken in dialogue taking into account patients preferences *(strong recommendation, very low level of evidence)***.**

There is no direct evidence that this recommendation actually improves clinical outcomes.

REFERENCES

1. Mambourg F, Robays J, Camberlin C, Vlayen J, Gailly J. Borstkankerscreening tussen 40 en 49 jaar. Good Clinical Practice (GCP). Brussel: Federaal Kenniscentrum voor de Gezondheidszorg (KCE); 2010 07/07/2010. KCE reports 129A (D/2010/10.273/28) Available from: http://kce.fgov.be/index_nl.aspx?SGREF=14842&CREF=16579
2. Guyatt G, Oxman AD, Akl EA, Kunz R, Vist G, Brozek J, et al. GRADE guidelines: 1. Introduction-GRADE evidence profiles and summary of findings tables. Journal of clinical epidemiology. 2011;64(4):383-94.
3. Schunemann HJ, Oxman AD, Brozek J, Glasziou P, Jaeschke R, Vist GE, et al. Grading quality of evidence and strength of recommendations for diagnostic tests and strategies. BMJ. 2008;336(7653):1106-10.
4. Fabri V RA, Boutsen M. Programma Borstkankerscreening. Vergelijking van de eerste drie rondes 2002-2003, 2004-2005, 2006-2007. Verslag nr. 7 van het Intermutualistisch Agentschap. Brussel: Intermutualistisch Agentschap; 2010. Available from: http://www.nic-ima.be/library/documents/quality%20projects/MA%207%20NL%20rapport.pdf
5. Barlow WE, Lehman CD, Zheng Y, Ballard-Barbash R, Yankaskas BC, Cutter GR, et al. Performance of diagnostic mammography for women with signs or symptoms of breast cancer. J Natl Cancer Inst. 2002;94(15):1151-9.
6. Cancer IAfRo. Breast Cancer Screening. Vainio H, Bianchini F, editor. Lyon: IARC Press; 2002.
7. Cummings SR, Tice JA, Bauer S, Browner WS, Cuzick J, Ziv E, et al. Prevention of breast cancer in postmenopausal women: approaches to estimating and reducing risk. J Natl Cancer Inst. 2009;101(6):384-98.
8. Lindor NM, Johnson KJ, Harvey H, Shane Pankratz V, Domchek SM, Hunt K, et al. Predicting BRCA1 and BRCA2 gene mutation carriers: comparison of PENN II model to previous study. Fam Cancer. 2010;9(4):495-502.
9. NICE. The classification and care of women at risk of familial breast cancer in primary, secondary and tertiary care. 2006.
10. NICE. The classification and care of women at risk of familial breast cancer in primary, secondary and tertiary care. 2004.
11. Weir R, Day P, Ali W. Risk factors for breast cancer in women: a systematic review of the literature (Structured abstract). 2007:361.
12. Claus EB, Risch N, Thompson WD. Autosomal dominant inheritance of early-onset breast cancer. Implications for risk prediction. Cancer. 1994;73(3):643-51.
13. Familial breast cancer: collaborative reanalysis of individual data from 52 epidemiological studies including 58,209 women with breast cancer and 101,986 women without the disease. Lancet. 2001;358(9291):1389-99.
14. Vrieling A, Buck K, Kaaks R, Chang-Claude J. Adult weight gain in relation to breast cancer risk by estrogen and progesterone receptor status: a meta-analysis. Breast Cancer Res Treat. 2010;123(3):641-9.
15. McCormack VA, dos Santos Silva I. Breast density and parenchymal patterns as markers of breast cancer risk: a meta-analysis. Cancer Epidemiol Biomarkers Prev. 2006;15(6):1159-69.
16. Kahlenborn C, Modugno F, Potter DM, Severs WB. Oral contraceptive use as a risk factor for premenopausal breast cancer: a meta-analysis. Mayo Clin Proc. 2006;81(10):1290-302.
17. Henderson TO, Amsterdam A, Bhatia S, Hudson MM, Meadows AT, Neglia JP, et al. Systematic review: surveillance for breast cancer in women treated with chest radiation for childhood, adolescent, or young adult cancer. Ann Intern Med. 2010;152(7):444-55; W144-54.
18. Zhou W-B, Xue D-Q, Liu X-A, Ding Q, Wang S. The influence of family history and histological stratification on breast cancer risk in

women with benign breast disease: a meta-analysis. J Cancer Res Clin Oncol. 2011;137(7):1053-60.

19. Amir E, Freedman OC, Seruga B, Evans DG. Assessing women at high risk of breast cancer: a review of risk assessment models. J Natl Cancer Inst. 2010;102(10):680-91.
20. Tice JA, Cummings SR, Ziv E, Kerlikowske K. Mammographic breast density and the Gail model for breast cancer risk prediction in a screening population. Breast Cancer Res. Treat. 2005;94(2):115-22.
21. Chen J, Pee D, Ayyagari R, Graubard B, Schairer C, Byrne C, et al. Projecting absolute invasive breast cancer risk in white women with a model that includes mammographic density. J. Natl. Cancer Inst. 2006;98(17):1215-26.
22. Tice JA, Cummings SR, Smith-Bindman R, Ichikawa L, Barlow WE, Kerlikowske K. Using clinical factors and mammographic breast density to estimate breast cancer risk: development and validation of a new predictive model.[Summary for patients in Ann Intern Med. 2008 Mar 4;148(5):I34; PMID: 18316748]. Ann Intern Med. 2008;148(5):337-47.
23. Barlow WE, White E, Ballard-Barbash R, Vacek PM, Titus-Ernstoff L, Carney PA, et al. Prospective breast cancer risk prediction model for women undergoing screening mammography. J Natl Cancer Inst. 2006;98(17):1204-14.
24. Decarli A, Calza S, Masala G, Specchia C, Palli D, Gail MH. Gail model for prediction of absolute risk of invasive breast cancer: Independent evaluation in the Florence-European Prospective Investigation Into Cancer and Nutrition cohort. J. Natl. Cancer Inst. 2006;98(23):1686-93.
25. Chlebowski RT, Anderson GL, Lane DS, Aragaki AK, Rohan T, Yasmeen S, et al. Predicting risk of breast cancer in postmenopausal women by hormone receptor status. J Natl Cancer Inst. 2007;99(22):1695-705.
26. Schonfeld SJ, Pee D, Greenlee RT, Hartge P, Lacey JV, Jr., Park Y, et al. Effect of changing breast cancer incidence rates on the calibration of the Gail model. J Clin Oncol. 2010;28(14):2411-7.
27. Vacek PM, Skelly JM, Geller BM. Breast cancer risk assessment in women aged 70 and older. Breast Cancer Res. Treat. 2011:1-9.
28. Crispo A, D'Aiuto G, De Marco M, Rinaldo M, Grimaldi M, Capasso I, et al. Gail model risk factors: impact of adding an extended family history for breast cancer. Breast J. 2008;14(3):221-7.
29. Wacholder S, Hartge P, Prentice R, Garcia-Closas M, Feigelson HS, Diver WR, et al. Performance of common genetic variants in breast-cancer risk models.[Erratum appears in N Engl J Med. 2010 Dec 2;363(23):2272]. N Engl J Med. 2010;362(11):986-93.
30. Mealiffe ME, Stokowski RP, Rhees BK, Prentice RL, Pettinger M, Hinds DA. Assessment of clinical validity of a breast cancer risk model combining genetic and clinical information. J Natl Cancer Inst. 2010;102(21):1618-27.
31. Pankratz VS, Hartmann LC, Degnim AC, Vierkant RA, Ghosh K, Vachon CM, et al. Assessment of the accuracy of the Gail model in women with atypical hyperplasia. J Clin Oncol. 2008;26(33):5374-9.
32. Boughey JC, Hartmann LC, Anderson SS, Degnim AC, Vierkant RA, Reynolds CA, et al. Evaluation of the Tyrer-Cuzick (International Breast Cancer Intervention Study) model for breast cancer risk prediction in women with atypical hyperplasia. J Clin Oncol. 2010;28(22):3591-6.
33. Kang HH, Williams R, Leary J, Ringland C, Kirk J, Ward R. Evaluation of models to predict BRCA germline mutations. Br. J. Cancer. 2006;95(7):914-20.
34. Roudgari H, Miedzybrodzka ZH, Haites NE. Probability estimation models for prediction of BRCA1 and BRCA2 mutation carriers: COS compares favourably with other models. Fam Cancer. 2008;7(3):199-212.
35. Parmigiani G, Chen S, Iversen ES, Jr., Friebel TM, Finkelstein DM, Anton-Culver H, et al. Validity of models for predicting BRCA1 and BRCA2 mutations.[Summary for patients in Ann Intern Med. 2007 Oct 2;147(7):I38; PMID: 17909202]. Ann Intern Med. 2007;147(7):441-50.
36. Antoniou AC, Hardy R, Walker L, Evans DG, Shenton A, Eeles R, et al. Predicting the likelihood of carrying a BRCA1 or BRCA2

mutation: validation of BOADICEA, BRCAPRO, IBIS, Myriad and the Manchester scoring system using data from UK genetics clinics. J Med Genet. 2008;45(7):425-31.

37. Panchal SM, Ennis M, Canon S, Bordeleau LJ. Selecting a BRCA risk assessment model for use in a familial cancer clinic. BMC Med Genet. 2008;9:116.
38. Carney PA, Miglioretti DL, Yankaskas BC, Kerlikowske K, Rosenberg R, Rutter CM, et al. Individual and combined effects of age, breast density, and hormone replacement therapy use on the accuracy of screening mammography. Ann Intern Med. 2003;138(3):168-75.
39. Dinnes J, Moss S, Melia J, Blanks R, Song F, Kleijnen J. Effectiveness and cost-effectiveness of double reading of mammograms in breast cancer screening: findings of a systematic review. Breast. 2001;10(6):455-63.
40. Taylor P, Potts HW. Computer aids and human second reading as interventions in screening mammography: two systematic reviews to compare effects on cancer detection and recall rate. Eur J Cancer. 2008;44(6):798-807.
41. Noble M, Bruening W, Uhl S, Schoelles K. Computer-aided detection mammography for breast cancer screening: systematic review and meta-analysis. Arch Gynecol Obstet. 2009;279(6):881-90.
42. AETSA. Diagnostic performance of digital mammography in breast cancer screening.

. Sevilla: Agencia de Evaluación de Tecnologías Sanitarias de Andalucía; 2007. Available from: http://www.juntadeandalucia.es/salud/aetsa

43. Bermejo-Perez MJ, Marquez-Calderon S, Llanos-Mendez A. Cancer surveillance based on imaging techniques in carriers of BRCA1/2 gene mutations: a systematic review. Br J Radiol. 2008;81(963):172-9.
44. Davidson E, Hancock S. Surveillance of women at high risk of breast cancer: a tech brief. . Christchurch: New Zealand Health technology assessment; 2007. Available from: http://nzhta.chmeds.ac.nz/#tech
45. Irwig L, Houssami N, van Vliet C. New technologies in screening for breast cancer: a systematic review of their accuracy. Br J Cancer. 2004;90(11):2118-22.
46. Nothacker M, Duda V, Hahn M, Warm M, Degenhardt F, Madjar H, et al. Early detection of breast cancer: benefits and risks of supplemental breast ultrasound in asymptomatic women with mammographically dense breast tissue. A systematic review. BMC Cancer. 2009;9:335.
47. Lord SJ, Lei W, Craft P, Cawson JN, Morris I, Walleser S, et al. A systematic review of the effectiveness of magnetic resonance imaging (MRI) as an addition to mammography and ultrasound in screening young women at high risk of breast cancer. Eur J Cancer. 2007;43(13):1905-17.
48. Hofvind S, Geller BM, Rosenberg RD, Skaane P. Screening-detected breast cancers: discordant independent double reading in a population-based screening program. Radiology. 2009;253(3):652-60.
49. Caumo F, Brunelli S, Tosi E, Teggi S, Bovo C, Bonavina G, et al. On the role of arbitration of discordant double readings of screening mammography: experience from two Italian programmes. Radiol Med. 2011;116(1):84-91.
50. Ciatto S, Ambrogetti D, Risso G, Catarzi S, Morrone D, Mantellini P, et al. The role of arbitration of discordant reports at double reading of screening mammograms. J Med Screen. 2005;12(3):125-7.
51. Shaw CM, Flanagan FL, Fenlon HM, McNicholas MM. Consensus review of discordant findings maximizes cancer detection rate in double-reader screening mammography: Irish National Breast Screening Program experience. Radiology. 2009;250(2):354-62.
52. Duijm LE, Louwman MW, Groenewoud JH, van de Poll-Franse LV, Fracheboud J, Coebergh JW. Inter-observer variability in mammography screening and effect of type and number of readers on screening outcome. Br J Cancer. 2009;100(6):901-7.
53. Ciatto S, Del Turco MR, Risso G, Catarzi S, Bonardi R, Viterbo V, et al. Comparison of standard reading and computer aided detection

(CAD) on a national proficiency test of screening mammography. European journal of radiology. 2003;45(2):135-8.

54. Ciatto S, Ambrogetti D, Collini G, Cruciani A, Ercolini E, Risso G, et al. Computer-aided detection (CAD) of cancers detected on double reading by one reader only. BREAST. 2006;15(4):528-32.
55. Gilbert FJ, Astley SM, Gillan MG, Agbaje OF, Wallis MG, James J, et al. Single reading with computer-aided detection for screening mammography. The New England journal of medicine. 2008;359(16):1675-84.
56. Cawson JN, Nickson C, Amos A, Hill G, Whan AB, Kavanagh AM. Invasive breast cancers detected by screening mammography: a detailed comparison of computer-aided detection-assisted single reading and double reading. Journal of Medical Imaging & Radiation Oncology. 2009;53(5):442-9.
57. Khoo LA, Taylor P, Given-Wilson RM. Computer-aided detection in the United Kingdom National Breast Screening Programme: prospective study. Radiology. 2005;237(2):444-9.
58. Lewin JM, Hendrick RE, D'Orsi CJ, Isaacs PK, Moss LJ, Karellas A, et al. Comparison of full-field digital mammography with screen-film mammography for cancer detection: results of 4,945 paired examinations. Radiology. 2001;218(3):873-80.
59. Lewin JM, D'Orsi CJ, Hendrick RE, Moss LJ, Isaacs PK, Karellas A, et al. Clinical comparison of full-field digital mammography and screen-film mammography for detection of breast cancer. AJR. 2002;American Journal of Roentgenology. 179(3):671-7.
60. Glueck DH, Lamb MM, Lewin JM, Pisano ED. Two-modality mammography may confer an advantage over either full-field digital mammography or screen-film mammography. Academic radiology. 2007;14(6):670-6.
61. Skaane P, Young K, Skjennald A. Population-based mammography screening: comparison of screen-film and full-field digital mammography with soft-copy reading--Oslo I study. Radiology. 2003;229(3):877-84.
62. Skaane P, Skjennald A. Screen-film mammography versus full-field digital mammography with soft-copy reading: randomized trial in a population-based screening program--the Oslo II Study. Radiology. 2004;232(1):197-204.
63. Skaane P, Skjennald A, Young K, Egge E, Jebsen I, Sager EM, et al. Follow-up and final results of the Oslo I Study comparing screen-film mammography and full-field digital mammography with soft-copy reading. Acta Radiol. 2005;46(7):679-89.
64. Skaane P, Hofvind S, Skjennald A. Randomized trial of screen-film versus full-field digital mammography with soft-copy reading in population-based screening program: follow-up and final results of Oslo II study. Radiology. 2007;244(3):708-17.
65. Vigeland E, Klaasen H, Klingen TA, Hofvind S, Skaane P. Full-field digital mammography compared to screen film mammography in the prevalent round of a population-based screening programme: the Vestfold County Study. European radiology. 2008;18(1):183-91.
66. Pisano ED, Gatsonis C, Hendrick E, Yaffe M, Baum JK, Acharyya S, et al. Diagnostic performance of digital versus film mammography for breast-cancer screening. The New England journal of medicine. 2005;353(17):1773-83.
67. Del Turco MR, Mantellini P, Ciatto S, Bonardi R, Martinelli F, Lazzari B, et al. Full-field digital versus screen-film mammography: comparative accuracy in concurrent screening cohorts. AJR. 2007;American Journal of Roentgenology. 189(4):860-6.
68. Bluekens AMJ, Karssemeijer N, Beijerinck D, Deurenberg JJM, van Engen RE, Broeders MJM, et al. Consequences of digital mammography in population-based breast cancer screening: initial changes and long-term impact on referral rates. European Radiology. 2010;20(9):2067-73.
69. Feeley L, Kiernan D, Mooney T, Flanagan F, Hargaden G, Kell M, et al. Digital mammography in a screening programme and its implications for pathology: a comparative study. Journal of clinical pathology. 2011;64(3):215-9.
70. Karssemeijer N, Bluekens AM, Beijerinck D, Deurenberg JJ, Beekman M, Visser R, et al. Breast cancer screening results 5 years after introduction of digital mammography in a population-based screening program. Radiology. 2009;253(2):353-8.

71. Van Ongeval C, Van Steen A, Vande Putte G, Zanca F, Bosmans H, Marchal G, et al. Does digital mammography in a decentralized breast cancer screening program lead to screening performance parameters comparable with film-screen mammography? European radiology. 2010;20(10):2307-14.
72. Vernacchia FS, Pena ZG. Digital mammography: its impact on recall rates and cancer detection rates in a small community-based radiology practice. AJR. American journal of roentgenology. 2009;193(2):582-5.
73. Hambly NM, McNicholas MM, Phelan N, Hargaden GC, O'Doherty A, Flanagan FL. Comparison of digital mammography and screen-film mammography in breast cancer screening: a review in the Irish breast screening program. AJR. 2009;American Journal of Roentgenology. 193(4):1010-8.
74. Sala M, Salas D, Belvis F, Sanchez M, Ferrer J, Ibanez J, et al. Reduction in false-positive results after introduction of digital mammography: analysis from four population-based breast cancer screening programs in Spain. Radiology. 2011;258(2):388-95.
75. Heddson B, Ronnow K, Olsson M, Miller D. Digital versus screen-film mammography: a retrospective comparison in a population-based screening program. European journal of radiology. 2007;64(3):419-25.
76. Vinnicombe S, Pinto Pereira SM, McCormack VA, Shiel S, Perry N, Dos Santos Silva IM. Full-field digital versus screen-film mammography: comparison within the UK breast screening program and systematic review of published data. Radiology. 2009;251(2):347-58.
77. Juel IM, Skaane P, Hoff SR, Johannessen G, Hofvind S. Screen-film mammography versus full-field digital mammography in a population-based screening program: The Sogn and Fjordane study. Acta radiologica. 2010;51(9):962-8.
78. Domingo L, Romero A, Belvis F, Sanchez M, Ferrer J, Salas D, et al. Differences in radiological patterns, tumour characteristics and diagnostic precision between digital mammography and screen-film mammogrpahy in four breast cancer screening programmes in Spain. Eur Radiol. 2011;21.
79. Lipasti S, Anttila A, Pamilo M. Mammographic findings of women recalled for diagnostic work-up in digital versus screen-film mammography in a population-based screening program. Acta radiologica. 2010;51(5):491-7.
80. Perry NM, Patani N, Milner SE, Pinker K, Mokbel K, Allgood PC, et al. The impact of digital mammography on screening a young cohort of women for breast cancer in an urban specialist breast unit. European radiology. 2011;21(4):676-82.
81. Pisano ED, Hendrick RE, Yaffe MJ, Baum JK, Acharyya S, Cormack JB, et al. Diagnostic accuracy of digital versus film mammography: exploratory analysis of selected population subgroups in DMIST. Radiology. 2008;246(2):376-83.
82. European Commission. European guidelines for quality assurance in breast cancer screening and diagnosis, Fourth edition. 2010.
83. Van Ongeval C, Jacobs J, Bosmans H. Artifacts in digital mammography. Jbr-Btr: Organe de la Societe Royale Belge de Radiologie. 2008;91(6):262-3.
84. Trecate G, Vergnaghi D, Bergonzi S, De Simone T, Fengoni E, Costa C, et al. Breast MRI screening in patients with increased familial and/or genetic risk for breast cancer: a preliminary experience. Tumori. 2003;89(2):125-31.
85. Sim LS, Hendriks JH, Fook-Chong SM. Breast ultrasound in women with familial risk of breast cancer. Ann Acad Med Singapore. 2004;33(5):600-6.
86. Youk JH, Kim EK, Kim MJ, Kwak JY, Son EJ. Performance of hand-held whole-breast ultrasound based on BI-RADS in women with mammographically negative dense breast. Eur Radiol. 2011;21(4):667-75.
87. Lenz S. Breast ultrasound in office gynecology--ten years of experience. Ultraschall Med. 2011;32 Suppl 1:S3-7.

88. Kolb TM, Lichy J, Newhouse JH. Occult cancer in women with dense breasts: detection with screening US--diagnostic yield and tumor characteristics. Radiology. 1998;207(1):191-9.
89. Buchberger W, DeKoekkoek-Doll P, Springer P, Obrist P, Dunser M. Incidental findings on sonography of the breast: clinical significance and diagnostic workup. AJR Am J Roentgenol. 1999;173(4):921-7.
90. Buchberger W, Niehoff A, Obrist P, DeKoekkoek-Doll P, Dunser M. Clinically and mammographically occult breast lesions: detection and classification with high-resolution sonography. Semin Ultrasound CT MR. 2000;21(4):325-36.
91. Kaplan SS. Clinical utility of bilateral whole-breast US in the evaluation of women with dense breast tissue. Radiology. 2001;221(3):641-9.
92. Crystal P, Strano SD, Shcharynski S, Koretz MJ. Using sonography to screen women with mammographically dense breasts. AJR Am J Roentgenol. 2003;181(1):177-82.
93. Brancato B, Bonardi R, Catarzi S, Iacconi C, Risso G, Taschini R, et al. Negligible advantages and excess costs of routine addition of breast ultrasonography to mammography in dense breasts. Tumori. 2007;93(6):562-6.
94. Kuhl CK, Schrading S, Leutner CC, Morakkabati-Spitz N, Wardelmann E, Fimmers R, et al. Mammography, breast ultrasound, and magnetic resonance imaging for surveillance of women at high familial risk for breast cancer. J Clin Oncol. 2005;23(33):8469-76.
95. Honjo S, Ando J, Tsukioka T, Morikubo H, Ichimura M, Sunagawa M, et al. Relative and combined performance of mammography and ultrasonography for breast cancer screening in the general population: a pilot study in Tochigi Prefecture, Japan. Jpn J Clin Oncol. 2007;37(9):715-20.
96. Daguet E, Malhaire C, Hardit C, Athanasiou A, El Khoury C, Thibault F, et al. [MR breast screening in patients with genetic mutation]. J Radiol. 2008;89(6):783-90.
97. Tohno E, Ueno E, Watanabe H. Ultrasound screening of breast cancer. Breast Cancer. 2009;16(1):18-22.
98. Kuhl C, Weigel S, Schrading S, Arand B, Bieling H, Konig R, et al. Prospective multicenter cohort study to refine management recommendations for women at elevated familial risk of breast cancer: the EVA trial. J Clin Oncol. 2010;28(9):1450-7.
99. Corsetti V, Houssami N, Ferrari A, Ghirardi M, Bellarosa S, Angelini O, et al. Breast screening with ultrasound in women with mammography-negative dense breasts: evidence on incremental cancer detection and false positives, and associated cost. Eur J Cancer. 2008;44(4):539-44.
100. Kolb TM, Lichy J, Newhouse JH. Comparison of the performance of screening mammography, physical examination, and breast US and evaluation of factors that influence them: an analysis of 27,825 patient evaluations. Radiology. 2002;225(1):165-75.
101. Kelly KM, Dean J, Comulada WS, Lee SJ. Breast cancer detection using automated whole breast ultrasound and mammography in radiographically dense breasts. Eur Radiol. 2010;20(3):734-42.
102. Corsetti V, Houssami N, Ghirardi M, Ferrari A, Speziani M, Bellarosa S, et al. Evidence of the effect of adjunct ultrasound screening in women with mammography-negative dense breasts: interval breast cancers at 1 year follow-up. Eur J Cancer. 2011;47(7):1021-6.
103. O'Driscoll D, Warren R, MacKay J, Britton P, Day NE. Screening with breast ultrasound in a population at moderate risk due to family history. J Med Screen. 2001;8(2):106-9.
104. Hou MF, Chuang HY, Ou-Yang F, Wang CY, Huang CL, Fan HM, et al. Comparison of breast mammography, sonography and physical examination for screening women at high risk of breast cancer in taiwan. Ultrasound Med Biol. 2002;28(4):415-20.
105. Podo F, Sardanelli F, Canese R, D'Agnolo G, Natali PG, Crecco M, et al. The Italian multi-centre project on evaluation of MRI and other imaging modalities in early detection of breast cancer in subjects at high genetic risk. J Exp Clin Cancer Res. 2002;21(3 Suppl):115-24.
106. Warner E, Plewes DB, Hill KA, Causer PA, Zubovits JT, Jong RA, et al. Surveillance of BRCA1 and BRCA2 mutation carriers with magnetic resonance imaging, ultrasound, mammography, and clinical breast examination. JAMA. 2004;292(11):1317-25.

107. Lehman CD, Isaacs C, Schnall MD, Pisano ED, Ascher SM, Weatherall PT, et al. Cancer yield of mammography, MR, and US in high-risk women: prospective multi-institution breast cancer screening study. Radiology. 2007;244(2):381-8.
108. Riedl CC, Ponhold L, Flory D, Weber M, Kroiss R, Wagner T, et al. Magnetic resonance imaging of the breast improves detection of invasive cancer, preinvasive cancer, and premalignant lesions during surveillance of women at high risk for breast cancer. Clin Cancer Res. 2007;13(20):6144-52.
109. Berg WA, Blume JD, Cormack JB, Mendelson EB, Lehrer D, Bohm-Velez M, et al. Combined screening with ultrasound and mammography vs mammography alone in women at elevated risk of breast cancer. JAMA. 2008;299(18):2151-63.
110. Weinstein SP, Localio AR, Conant EF, Rosen M, Thomas KM, Schnall MD. Multimodality screening of high-risk women: a prospective cohort study. J Clin Oncol. 2009;27(36):6124-8.
111. Sardanelli F, Podo F, Santoro F, Manoukian S, Bergonzi S, Trecate G, et al. Multicenter surveillance of women at high genetic breast cancer risk using mammography, ultrasonography, and contrast-enhanced magnetic resonance imaging (the high breast cancer risk italian 1 study): final results. Invest Radiol. 2011;46(2):94-105.
112. Leconte I, Feger C, Galant C, Berliere M, Berg BV, D'Hoore W, et al. Mammography and subsequent whole-breast sonography of nonpalpable breast cancers: the importance of radiologic breast density. AJR Am J Roentgenol. 2003;180(6):1675-9.
113. Protection EUHC. European guidelines for quality assurance in breast cancer screening and diagnosis. Fourth ed. Perry NB, M.; de Wolf, C.; Törnberg, S.; Holland, R.; von Karsa, L., editor. Lyxembourg: Office for Official Publications of the European Communities; 2006.
114. Abdullah N, Mesurolle B, El-Khoury M, Kao E. Breast imaging reporting and data system lexicon for US: interobserver agreement for assessment of breast masses. Radiology. 2009;252(3):665-72.
115. Park CS, Lee JH, Yim HW, Kang BJ, Kim HS, Jung JI, et al. Observer agreement using the ACR Breast Imaging Reporting and Data System (BI-RADS)-ultrasound, First Edition (2003). Korean J Radiol. 2007;8(5):397-402.
116. Lazarus E, Mainiero MB, Schepps B, Koelliker SL, Livingston LS. BI-RADS lexicon for US and mammography: interobserver variability and positive predictive value. Radiology. 2006;239(2):385-91.
117. Yu J, Park A, Morris E, Liberman L, Borgen PI, King TA. MRI screening in a clinic population with a family history of breast cancer. Ann Surg Oncol. 2008;15(2):452-61.
118. Lapierre-Combes M, Rousset J, Combes E, Chinelatto S, Dupre PF, Andre V. [Retrospective study conducted in northern Finistere about the role of breast MRI in normal breast screening, experience in 51 patients]. Gynecol Obstet Fertil. 2009;37(5):401-9.
119. Elmore L, Margenthaler JA. Use of breast MRI surveillance in women at high risk for breast cancer: a single-institutional experience. Ann Surg Oncol. 2010;17 Suppl 3:263-7.
120. Shah P, Rosen M, Stopfer J, Siegfried J, Kaltman R, Mason B, et al. Prospective study of breast MRI in BRCA1 and BRCA2 mutation carriers: effect of mutation status on cancer incidence. Breast Cancer Res Treat. 2009;118(3):539-46.
121. Price J, Chen SW. Screening for breast cancer with MRI: recent experience from the Australian Capital Territory. J Med Imaging Radiat Oncol. 2009;53(1):69-80.
122. Abramovici G, Mainiero MB. Screening breast MR imaging: comparison of interpretation of baseline and annual follow-up studies. Radiology. 2011;259(1):85-91.
123. Peters G, Anderson J, Longman G, Thomson J, Taylor D, Bennett M, et al. Magnetic resonance findings in women at high risk for developing breast cancer: an Australian feasibility study. J Med Imaging Radiat Oncol. 2008;52(1):29-35.
124. Kerlikowske K, Grady D, Barclay J, Sickles EA, Ernster V. Effect of age, breast density, and family history on the sensitivity of first screening mammography. JAMA. 1996;276(1):33-8.
125. Robson M, Offit K. Clinical practice. Management of an inherited predisposition to breast cancer. N Engl J Med. 2007;357(2):154-62.

126. Jansen-van der Weide MC, Greuter MJ, Jansen L, Oosterwijk JC, Pijnappel RM, de Bock GH. Exposure to low-dose radiation and the risk of breast cancer among women with a familial or genetic predisposition: a meta-analysis. Eur Radiol. 2010;20(11):2547-56.
127. CBO. Mammacarcinoom. CentraalBegeleidingsOrgaan; 2008. Richtlijn Available from: www.cbo.nl
128. Key J, Hodgson S, Omar RZ, Jensen TK, Thompson SG, Boobis AR, et al. Meta-analysis of studies of alcohol and breast cancer with consideration of the methodological issues. Cancer Causes Control. 2006;17(6):759-70.
129. Duijm LEM, Groenewoud JH, Fracheboud J, van Ineveld BM, Roumen RMH, de Koning HJ. Introduction of additional double reading of mammograms by radiographers: effects on a biennial screening programme outcome. European Journal of Cancer. 2008;44(9):1223-8.
130. Bennett RL, Blanks RG, Moss SM. Does the accuracy of single reading with CAD (computer-aided detection) compare with that of double reading?: A review of the literature. Clin Radiol. 2006;61(12):1023-8.
131. Ciatto S, Ambrogetti D, Bonardi R, Catarzi S, Risso G, Rosselli Del Turco M, et al. Second reading of screening mammograms increases cancer detection and recall rates. Results in the Florence screening programme. J Med Screen. 2005;12(2):103-6.
132. Duijm LE, Groenewoud JH, Hendriks JH, de Koning HJ. Independent double reading of screening mammograms in The Netherlands: effect of arbitration following reader disagreements. Radiology. 2004;231(2):564-70.
133. Liston JC, Dall BJ. Can the NHS Breast Screening Programme afford not to double read screening mammograms? Clinical radiology. 2003;58(6):474-7.
134. Gilbert FJ, Astley SM, McGee MA, Gillan MG, Boggis CR, Griffiths PM, et al. Single reading with computer-aided detection and double reading of screening mammograms in the United Kingdom National Breast Screening Program. Radiology. 2006;241(1):47-53.
135. Mucci B, Athey G, Scarisbrick G. Double read screening mammograms: The use of a third reader to arbitrate on disagreements. BREAST. 1999;8(1):37-9.
136. Georgian-Smith D, Moore RH, Halpern E, Yeh ED, Rafferty EA, D'Alessandro HA, et al. Blinded comparison of computer-aided detection with human second reading in screening mammography. AJR. 2007;American Journal of Roentgenology. 189(5):1135-41.
137. Gromet M. Comparison of computer-aided detection to double reading of screening mammograms: review of 231,221 mammograms. AJR. American journal of roentgenology. 2008;190(4):854-9.
138. Destounis SV, DiNitto P, Logan-Young W, Bonaccio E, Zuley ML, Willison KM. Can computer-aided detection with double reading of screening mammograms help decrease the false-negative rate? Initial experience. Radiology. 2004;232(2):578-84.
139. Taylor P, Given-Wilson R, Champness J, Potts HW, Johnston K. Assessing the impact of CAD on the sensitivity and specificity of film readers. Clin Radiol. 2004;59(12):1099-105.
140. Yang SK, Moon WK, Cho N, Park JS, Cha JH, Kim SM, et al. Screening mammography-detected cancers: sensitivity of a computer-aided detection system applied to full-field digital mammograms. Radiology. 2007;244(1):104-11.
141. Skaane P, Kshirsagar A, Stapleton S, Young K, Castellino RA. Effect of computer-aided detection on independent double reading of paired screen-film and full-field digital screening mammograms. AJR. American journal of roentgenology. 2007;188(2):377-84.
142. Sala M, Comas M, Macia F, Martinez J, Casamitjana M, Castells X. Implementation of digital mammography in a population-based breast cancer screening program: effect of screening round on recall rate and cancer detection. Radiology. 2009;252(1):31-9.
143. Hagen AI, Kvistad KA, Maehle L, Holmen MM, Aase H, Styr B, et al. Sensitivity of MRI versus conventional screening in the diagnosis of BRCA-associated breast cancer in a national prospective series. Breast. 2007;16(4):367-74.