

# The role of ultrasonography in diagnostic breast imaging

## Citation for published version (APA):

Flobbe, K. (2003). The role of ultrasonography in diagnostic breast imaging. Maastricht: Universitaire Pers Maastricht.

## Document status and date:

Published: 01/01/2003

## Document Version:

Publisher's PDF, also known as Version of record

## Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

## General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

[www.umlib.nl/taverne-license](http://www.umlib.nl/taverne-license)

## Take down policy

If you believe that this document breaches copyright please contact us at:

[repository@maastrichtuniversity.nl](mailto:repository@maastrichtuniversity.nl)

providing details and we will investigate your claim.

## Summary

Routine diagnostic breast imaging in the radiology departments of a hospital consists of clinical examination, mammography, and ultrasonography when indicated. This thesis focuses on the diagnostic performance of breast imaging in diagnosing breast cancer and more specifically on the role of ultrasonography herein.

Chapter 2 discusses a prospective study that was performed in 1994, assessing the diagnostic performance of breast imaging in a non-screening population. A consecutive set of 1944 breast examinations in 1890 patients was studied, consisting of palpation and mammography, supplemented with ultrasonography when clinically indicated. In total, 3816 breasts were examined. Pathology results after a follow up of one year were used as reference standard for the presence of breast cancer and demonstrated 118 malignancies in 115 patients (6.1% prevalence).

With a cut-off point between benign and probably benign results, a sensitivity of 89% and a specificity of 98% was found for radiological breast imaging. There were 13 false-negative and 69 false-positive results. Logistic regression and ROC curve analysis demonstrated that adding the imaging results to the information available from palpation and the age of the patient significantly improved the diagnostic performance. Combining all information a sensitivity of 100% could be reached hypothetically, with a specificity of around 60%.

This study concludes that radiological breast imaging has an important role in the diagnosis of breast cancer in daily clinical practice. The results on the prevalence of breast cancer and the performance of breast imaging were comparable to other studies in national and international radiology departments.

In this study, additional ultrasonography was performed in 28% of all patients referred for diagnostic breast imaging. Current practice shows that the use of ultrasonography is increasing, reporting over 35% of patients undergoing ultrasonography as an adjunct to mammography in 1999. This is partly the result of the wider availability and accessibility, the broadening clinical indications and the belief of physicians that is less likely to be blamed for doing more testing than for doing too little. However, this practice leads to the question what the diagnostic performance is of additional ultrasonography and whether ultrasonography could be restricted to subgroups of patients who benefit most.

The study discussed in chapter 2 was not able to answer these questions. First, as ultrasonography was performed on clinical indication, as decided by the individual radiologist, the study population who had undergone breast ultrasonography was a selected one. Secondly, one definitive radiological diagnosis was given, in which mammography results were integrated with ultrasonography results. This made it impossible to analyze the impact of incremental ultrasonography.

Following, in chapter 3 a systematic literature review was performed in which all publications in MEDLINE and EMBASE between 1990 and 2000, assessing the role of ultrasonography as an adjunct to mammography, were analyzed. Diagnostic odds ratios for mammography and ultrasonography were calculated as a combined measure for sensitivity and specificity and were compared between the different studies. 22 studies were included, showing a large variety of indications for ultrasonography and variations in the diagnostic performance of mammography and ultrasonography. The study showed that little evidence-based support was available for the well-recognized value of ultrasonography as an adjunct to mammography in

clinical practice. The method of selecting the study population and the method of ultrasonography interpretation, with or without prior knowledge of mammography results, strongly influenced the performance of mammography and ultrasonography relative to each other.

These results emphasized the need for a prospective study in which systematic ultrasonography is performed in a non-selected consecutive patient population to identify the subgroups of patients in which the highest diagnostic yield would be obtained. In order to facilitate and optimize the logistic planning of breast imaging early in the diagnostic work-up, it was decided to focus mainly on the reasons of referral when defining subgroups of patients.

In chapter 4, the results are presented of a large prospective study, named the MAMMOED study (MAMmacarcinoom-studie Maastricht: Optimalisatie Echografische Diagnostiek) that was performed at Maastricht University Hospital between October 1999 and August 2000, as part of the Cost Containment Research Program of the Dutch Health Care Insurance Board. Aims of the study were to assess the diagnostic value of ultrasonography as an adjunct to mammography and clinical examination in the detection of breast cancer and the feasibility of selecting subgroups of patients who benefit most.

In 3835 breasts in 2020 consecutive patients referred for breast imaging, bilateral clinical examination, mammography and ultrasonography were performed. Pathology results from biopsy or surgical excision during 12 months follow up were used as gold standard for the presence of breast cancer and reported 129 malignancies in 127 patients (6.3% prevalence).

Breast ultrasonography detected 8 extra malignancies and correctly downgraded 332 positive results on mammography and clinical examination to a negative final diagnosis.

ROC curve analysis showed that there was a significant improvement in diagnostic value by adding ultrasonography to mammography and clinical examination. Subgroups of patients were defined according to indication for referral and age. The diagnostic yield improved significantly in patients referred for palpable breast masses, especially in patients younger than 50 years old, or patients referred from the national breast cancer screening program. The increase in area under the ROC curve was not statistically significant in patients referred for breast symptomatology and follow up of a prior breast malignancy. Due to the small numbers and low prevalence of breast cancer, the diagnostic benefit remained undefined for patients referred for follow up of a prior benign lesion, patients referred for reassurance, or when ultrasonography was performed as screening examination of the contralateral breast.

The study concluded that systematic application of ultrasonography significantly improved the overall diagnostic yield, and that the diagnostic value increased most significantly in patients with palpable breast masses and patients referred with abnormal screening mammograms.

In the diagnostic work-up for breast disease, all information on the lesions and abnormalities found by the different tests is combined with the knowledge of patient characteristics from medical history into one overall diagnosis indicating the suspicion of breast cancer. The contribution of each diagnostic test in defining the overall diagnosis was studied by multivariate logistic regression analysis and reported in chapter 5. Additionally, the contribution of additional ultrasonography in the prediction of breast cancer was studied for different patient populations. Patient data

of the MAMMOED study were used and diagnostic variables were included from medical history, clinical examination, mammography and ultrasonography.

A prediction model for breast cancer was presented, which included the variables: age, menopause, specific breast symptoms, diagnosis of clinical examination, mammographic calcifications and the integrated diagnosis resulting from ultrasonography, mammography and clinical examination. Using this model, the probability of breast cancer can be calculated for individual patients by combining the results of the different diagnostic tests. The "worst-case" scenario consisted of postmenopausal women older than 50 years, having specific breast symptoms such as skin retraction or nipple discharge, a suspicious clinical examination, suspicious calcifications on the mammogram and a suspicious final diagnosis of clinical examination, mammography and ultrasonography together, resulting in a probability of breast cancer of 0.997.

ROC curve analysis of the generated probabilities for each category of variables, showed that each step in the diagnostic work-up improved the diagnostic value of the prediction model for breast cancer. Adding results of ultrasonography to the other modalities significantly improved the predictive value of the model. Using all available information a sensitivity of 95% and specificity of 92% could be reached.

Stratified analysis according to reason for referral showed that additional ultrasonography significantly improved the predictive value of the model in patients referred for palpable masses ( $p=0.04$ ). Among patients without a palpable lump, ultrasonography contributed most to the diagnosis in the subgroup of premenopausal patients ( $p=0.39$ ).

It was concluded that, although the value of this prediction model in clinical practice may be argued, the model demonstrated a set of diagnostic variables, which are most important in the composition of a final imaging diagnosis.

In order to study the costs and effects of the use of additional ultrasonography in the evaluation of palpable breast masses, a decision analytical model was presented in chapter 6.

The model was based on the subgroup of patients from the MAMMOED study presenting with a palpable mass ( $n=522$ , 93 malignancies). Routine diagnostic work-up of palpable breast masses includes the performance of clinical examination, mammography and fine needle aspiration cytology (FNAC), combined as the "triple test".

In this model, 3 different diagnostic strategies of incorporating ultrasonography in the triple test were compared to the routine conventional triple test. In strategy 1, a cut-off between benign and probably benign ultrasonography results was used for further referral for FNAC; in strategy 2, this cut-off was shifted to a cut-off between normal and benign results; and in strategy 3, prior ultrasonography was only performed in patients with normal, benign results on mammography and clinical examination, whereas immediate FNAC was performed in patients with suspicious lesions. Outcome variables included the total costs of strategies and the expected number of life years.

All strategies reported similar life expectancies, namely 31.0 life years. Cost-minimization demonstrated that experimental strategy-3 was the least expensive strategy (€ 2946). Experimental strategy-2 was the most costly strategy (€ 3475). Compared to the conventional strategy of immediate FNAC (€ 3010), both ultrasonography strategy 1 and 3 were preferred.

The study concluded that a reduction of the total costs for the diagnosis and treatment of breast cancer could be achieved by integrating ultrasonography in the triple test.

These results confirmed the results found by ROC curve and logistic regression analysis and favor the use of additional ultrasonography in patients with a palpable mass.

The overall conclusion of this thesis was that systematic application of ultrasonography as an adjunct to mammography and clinical examination significantly improved the diagnostic performance. Integrating the results from clinical epidemiological methodology, such as ROC curves and logistic regression analysis, with the results from decision analytical modeling indicated that in patients with palpable breast masses additional ultrasonography results in more accurate diagnosis at lesser costs.

## Samenvatting

Het gangbare beeldvormende onderzoek van de mamma bestaat op de meeste radiologie afdelingen uit klinisch onderzoek en diagnostische mammografie, hetgeen kan worden aangevuld met echografisch onderzoek wanneer geïndiceerd. Dit proefschrift concentreert zich op het vermogen van deze onderzoeken om borstkanker te diagnosticeren en in het bijzonder de rol van de echografie hierbij.

Hoofdstuk 2 beschrijft een prospectieve studie uit 1994 naar het diagnostisch vermogen van beeldvormend onderzoek van de mammae in een poliklinische patiëntenpopulatie. Bij 1944 opeenvolgende diagnostische onderzoeken bij 1890 patiënten werden 3816 mammae onderzocht met behulp van palpatie en mammografie, aangevuld met echografie wanneer klinisch geïndiceerd.

Als gouden standaard voor de aanwezigheid van borstkanker werden de beschikbare pathologie uitslagen gedurende één jaar follow up gebruikt. Er werden 118 maligniteiten gevonden bij 115 patiënten ( $115/1890 = 6,1\%$  prevalentie). Bij een afkappunt tussen benigne en waarschijnlijk benigne testresultaten had het radiologisch beeldvormend onderzoek een sensitiviteit en specificiteit van respectievelijk 89% en 98%. Er werden 13 foutnegatieve en 69 foutpositieve resultaten gevonden.

Met behulp van logistische regressie en ROC curve analyse werd voorts aangetoond dat het diagnostisch vermogen significant verbeterde wanneer testresultaten van beeldvormend onderzoek werden toegevoegd aan resultaten van palpatie en leeftijd. Hypothetisch kon een sensitiviteit van 100% worden behaald bij een specificiteit van 60%, wanneer alle informatie werd gecombineerd. Er werd geconcludeerd dat radiologisch beeldvormend onderzoek een belangrijke rol speelt bij de diagnose van borstkanker in de dagelijkse praktijk.

De resultaten van dit onderzoek en de prevalentie van borstkanker waren vergelijkbaar met die uit andere studies door nationale en internationale radiologie groepen. Aanvullend echografisch onderzoek werd bij 28% van de patiëntenpopulatie uitgevoerd. De huidige praktijk laat echter zien dat deze toepassing van echografie is toegenomen; in 1999 werd op de afdeling radiologie van het Academisch Ziekenhuis Maastricht bij meer dan 35% van de patiënten die doorverwezen werd voor diagnostisch beeldvormend onderzoek van de mamma ook een echografie uitgevoerd. Gedeeltelijk is dit het resultaat van de toegenomen beschikbaarheid en toegankelijkheid van echografie, de verbrede klinische indicaties en het geloof van artsen dat het uitvoeren van teveel testen hen minder snel wordt kwalijk genomen dan het uitvoeren van te weinig testen. Desalniettemin leidt dit tot de vraag wat het diagnostisch vermogen is van aanvullende echografie in deze patiëntenpopulatie en of echografie beperkt kan worden tot subgroepen patiënten die het meeste baat hebben bij dit onderzoek.

De studie die in hoofdstuk 2 werd beschreven, kon deze vragen niet beantwoorden. Ten eerste werd de echografie hier op klinische indicatie uitgevoerd, zoals bepaald door de individuele radioloog, waardoor de groep patiënten die een echografie ondergingen een sterk geselecteerde populatie was.

Ten tweede werd één definitieve radiologische diagnose gesteld, waarin de testresultaten van mammografie geïntegreerd werden met die van de echografie. Dit maakt het onmogelijk om de invloed van aanvullende echografie te bestuderen.

Vervolgens werd in hoofdstuk 3 een systematisch literatuuronderzoek besproken, waarin alle publicaties uit MEDLINE en EMBASE tussen 1990 en 2000 over de rol van

echografie als aanvulling op mammografie werden geanalyseerd. Als gecombineerde maat voor de sensitiviteit en specificiteit werden diagnostische odds ratios berekend voor mammografie en echografie en vergeleken tussen de verschillende studies. In totaal werden 22 studies in het onderzoek meegenomen en lieten een grote verscheidenheid zien in gebruikte indicaties voor echografie en diagnostisch vermogen van mammografie en echografie.

De studie liet zien dat er maar weinig literatuur voorhanden is, die de reeds lang geaccepteerde waarde van echografie als aanvulling op mammografie in de klinische praktijk onderbouwt. In de geanalyseerde publicaties had de methode om de patiëntenpopulatie te selecteren en de methode van echografie beoordeling, namelijk met of zonder voorkennis van de resultaten van mammografie, een grote invloed op de prestatie van mammografie en echografie ten opzichte van elkaar.

De resultaten van dit literatuuronderzoek benadrukten de behoefte aan een prospectief onderzoek waarin systematische echografie wordt uitgevoerd in een niet-geselecteerde patiëntenpopulatie, teneinde subgroepen patiënten te identificeren met de grootste diagnostische opbrengst. Om de logistische planning van beeldvormend onderzoek van de mamma op de radiologie afdeling te kunnen vergemakkelijken en te optimaliseren, werd besloten om de subgroepen patiënten voornamelijk in te delen volgens de verwijfsreden.

In hoofdstuk 4 van dit proefschrift werden de resultaten gepresenteerd van een grote prospectieve studie, de MAMMOED studie genoemd, (MAMmacarcinoom-studie Maastricht: Optimalisatie Echografische Diagnostiek). Deze studie werd uitgevoerd in het Academisch Ziekenhuis Maastricht tussen oktober 1999 en oktober 2001 en maakt deel uit van het Doelmatigheidsonderzoek programma van het College voor Zorgverzekeraars.

Doelen van de studie waren het bepalen van de diagnostische waarde van echografie als aanvulling op mammografie en klinisch onderzoek in het diagnosticeren van borstkanker en de mogelijkheid om subgroepen patiënten te selecteren bij wie aanvullende echografie het meeste baat heeft.

In 3835 mammae van 2020 opeenvolgende poliklinische patiënten die door huisarts of specialist doorverwezen werden voor diagnostisch beeldvormend onderzoek, werd bilateraal klinisch onderzoek, mammografie en echografie uitgevoerd. Alle onderzoeken werden gescoord op een 5-punts schaal met oplopende graad van verdenking op borstkanker: 1= normaal, 2= benigne, 3= onzeker benigne, 4= verdacht maligne, 5= maligne.

De pathologie uitslagen van biopsieën en chirurgische excisies gedurende 12 maanden follow up werden gebruikt als gouden standaard voor de aanwezigheid van borstkanker. In totaal werden 129 maligniteiten gevonden in 127 patiënten ( $127/2020=6,3\%$  prevalentie).

Echografisch onderzoek detecteerde 8 extra maligniteiten en zorgde voor een correcte aanpassing van 332 verdachte diagnoses op mammografie en klinisch onderzoek in een negatieve einddiagnose. ROC curve analyse liet zien dat er een significante verbetering was van de diagnose, wanneer resultaten van echografie werden toegevoegd aan die van mammografie en klinisch onderzoek. Er werden subgroepen patiënten gedefinieerd volgens de verwijfsreden en leeftijd. Aanvullend echografisch onderzoek vergrootte de diagnostische opbrengst statistisch significant bij patiënten die doorverwezen werden vanwege een palpabele afwijking, vooral wanneer ouder dan 50 jaar en patiënten die werden doorverwezen vanuit het nationaal bevolkingsonderzoek voor borstkanker.

De toename in oppervlakte onder de ROC curve was niet statistisch significant voor patiënten die doorverwezen werden vanwege andere symptomatologie, zoals pijn of tepelvloed, of vanwege follow up van een eerdere maligniteit in de mamma. Tenslotte bleef, vanwege de kleine aantallen en de lage prevalentie borstkanker, de waarde van aanvullende echografie onduidelijk voor patiënten die doorverwezen werden vanwege een eerdere benigne afwijking, ter geruststelling, of wanneer een screeningsechografie werd uitgevoerd van de tegengestelde mamma.

De studie concludeerde dat systematische toepassing van aanvullende echografie in deze patiëntenpopulatie de uiteindelijke diagnostische opbrengst verbeterde, en dat dit effect het duidelijkst was voor patiënten die verwezen werden met een palpabele afwijking of met een afwijkend screeningsmammogram vanuit het bevolkingsonderzoek.

In de diagnostische work-up voor borstafwijkingen worden alle resultaten van de verschillende onderzoeken gecombineerd met patiëntkarakteristieken vanuit de anamnese tot een eindiagnose en graad van verdenking op borstkanker. De bijdrage van elke diagnostische test in de totstandkoming van deze eindiagnose werd bestudeerd met behulp van logistische regressie analyse en werd besproken in hoofdstuk 5. Tevens werd hier de bijdrage van echografie in de voorspelling van de aanwezigheid van borstkanker voor verschillende patiëntenpopulaties bestudeerd. Hiertoe werden van de patiënten uit de MAMMOED studie verschillende diagnostische variabelen uit de anamnese, klinisch onderzoek, mammografie en echografie gebruikt.

In een voorspellend model voor de aanwezigheid van borstkanker werden de volgende variabelen opgenomen: leeftijd, menopauze, specifieke borstklachten, diagnose uit klinisch onderzoek, mammografisch aangetoond kalk en de geïntegreerde eindiagnose van echografie, mammografie en klinisch onderzoek. Met behulp van dit model kan de waarschijnlijkheid op borstkanker worden geschat voor individuele patiënten door de testresultaten van de verschillende diagnostische onderzoeken te combineren. Het slechtste scenario bestond uit postmenopausale vrouwen ouder dan 50 jaar, met specifieke borstklachten, zoals huidintrekking of tepelvloed, een verdacht klinisch onderzoek, verdacht kalk op het mammogram en een verdachte eindiagnose van klinisch onderzoek, mammografie en echografie samen, en resulteerde in een geschatte kans op borstkanker van 0.997.

ROC curve analyse van de gegenereerde kansen voor elke categorie variabelen liet zien dat elke stap in de diagnostische work-up de voorspellende waarde van het model verbeterde. Het model verbeterde sterk wanneer de resultaten van echografie werden toegevoegd aan de resultaten van de andere modaliteiten. Wanneer alle informatie werd gebruikt, kon een sensitiviteit en specificiteit van respectievelijk 95% en 92% behaald worden.

Gestratificeerde analyse naar verwijssreden liet zien dat aanvullende echografie de voorspellende waarde van het model vooral verbeterde in patiënten die verwezen werden voor een palpabele afwijking ( $p=0.04$ ). Bij patiënten zonder een palpabele afwijking, droeg echografie het meest bij aan de diagnose in de groep premenopausale patiënten ( $p=0.39$ ).

Er werd geconcludeerd dat dit voorspellende model, alhoewel de waarde ervan voor de klinische praktijk kan worden betwist, een set diagnostische variabelen presenteert die belangrijk is bij de samenstelling van een beeldvormende eindiagnose.



Vervolgens werd in hoofdstuk 6 een besliskundig model gepresenteerd, dat de kosten en effecten bestudeert van de toepassing van aanvullende echografie bij de evaluatie van palpabele borstafwijkingen. Dit besliskundig model werd gebaseerd op de groep patiënten uit de MAMMOED studie die een palpabele afwijking hadden ( $n=522$ , 93 maligniteiten). Het gangbare beeldvormend onderzoek bij palpabele afwijkingen bestaat uit klinisch onderzoek, diagnostische mammografie en dunne naald cytologie (FNAC), samen de "triple test" genoemd. In dit besliskundige model werden 3 verschillende diagnostische strategieën waarbij echografisch onderzoek werd geïntegreerd in de triple test, vergeleken met deze routine conventionele methode. In strategie 1 werd een afkappunt tussen benigne en waarschijnlijk benigne echografie uitslagen gebruikt voor verdere doorverwijzing voor FNAC; in strategie 2 werd dit afkappunt verlegd tussen normale en benigne uitslagen; en in strategie 3 werd aanvullende echografie alleen uitgevoerd bij patiënten met normale, benigne resultaten op mammografie en klinisch onderzoek, terwijl directe FNAC uitgevoerd werd in patiënten met verdachte laesies. De uitkomstmaten waren de totale kosten van de strategieën en het verwachte aantal levensjaren.

Alle strategieën hadden een gemiddelde levensverwachting van 31,0 jaar. Kostenminimalisatie liet zien dat experimentele strategie 3 de goedkoopste was (€ 2946). Experimentele strategie 2 was de duurste strategie (€ 3475). Vergeleken met de conventionele strategie van directe FNAC (€ 3010), ging de voorkeur uit naar echografie strategieën 1 en 3.

De studie concludeerde dat een reductie van de totale kosten van de diagnose en behandeling van borstkanker bewerkstelligd kan worden wanneer echografie in de triple test zou worden opgenomen.

Deze resultaten bevestigden de resultaten van de ROC curve en logistische regressie analyse en moedigen het gebruik van echografie in patiënten met een palpabele afwijking aan.

Uit de resultaten van de verschillende studies die beschreven worden in dit proefschrift kan geconcludeerd worden dat systematische toepassing van echografie als aanvulling op mammografie en klinisch onderzoek in een poliklinische patiëntenpopulatie de diagnostische opbrengst significant verbeterde.

Het combineren van de resultaten van klinisch epidemiologische methodologie, zoals ROC curves en logistische regressie analyse, met de resultaten van besliskundig modelleren, wijzen erop dat het gebruik van aanvullende echografie bij patiënten met een palpabele afwijking een meer nauwkeurige diagnose kan bewerkstelligen tegen lagere kosten.