

Mammography Screening – as of 2013

Mammografie-Screening – Stand 2013

Authors

S. Heywang-Koebrunner¹, K. Bock², W. Heindel³, G. Hecht⁴, L. Regitz-Jedermann⁵, A. Hacker⁶, V. Kaeab-Sanyal⁷

Affiliations

The affiliations are listed at the end of the article.

Key words

- mamma
- breast cancer
- mammogram screening

Schlüsselwörter

- Mamma
- Mammakarzinom
- Mammografie-Screening



Deutschsprachige
Zusatzinformationen
online abrufbar unter:
[www.thieme-connect.de/
ejournals/toc/gebfra](http://www.thieme-connect.de/ejournals/toc/gebfra)

received 16. 5. 2013
revised 10. 7. 2013
accepted 9. 8. 2013

Bibliography

DOI [http://dx.doi.org/
10.1055/s-0033-1350880](http://dx.doi.org/10.1055/s-0033-1350880)
Geburtsh Frauenheilk 2013; 73:
1007–1016 © Georg Thieme
Verlag KG Stuttgart · New York ·
ISSN 0016-5751

Correspondence

**Prof. Sylvia
Heywang-Koebrunner**
Munich Mammography
Reference Centre
Sonnenstraße 29
80331 Munich
[sylvia.heywang@
referenzzentrum-muenchen.de](mailto:sylvia.heywang@referenzzentrum-muenchen.de)

Abstract

Introduction: Since 2008 the German Mammography Screening Programme has been available throughout Germany to all women aged between 50 and 69. The programme strictly follows the European Guidelines. There are controversial discussions in the media as well as in the specialised press.

Materials and Methods: Overview of the available data with regard to an evaluation of randomised studies and with regard to quality-assured screening programmes in accordance with EU Guidelines (including data from 18 screening countries).

Results: Positive effects of screening: reduction in mortality, less invasive treatment. Negative effects: False-positive diagnoses and biopsy recommendations, so-called overdiagnoses, radiation dose. Limits of screening: Interval carcinomas, incomplete reduction in mortality. A mathematical synopsis of the latest publications from the European screening programmes with the diagnosis rates in Germany determined from >4.6 million screening examinations produces the following: a total of 10 000 mammograms are created for 1000 women (P) taking part in the Mammography Screening Programme (each of whom undergoes 10 mammograms in 20 years). Overall, the risk of triggering breast cancer through a mammogram is very clearly below the annual natural risk of suffering from breast cancer. In the German screening, of these 1000 women, an average of 288 women are called back once in 20 years as a result of changes that are ultimately benign (<3% per cycle). Of these, 74 of the 288 women undergo a biopsy due to a benign change (false-positive biopsy recommendations, usually punch or vacuum biopsies). According to EUROSREEN, 71 carcinomas develop among participants (56 are discovered in the screening, 15 in the interval), and 67 carcinomas among non-participants (N-P) (in

Zusammenfassung

Einleitung: Seit 2008 steht das Deutsche Mammografie-Screening-Programm flächendeckend allen Frauen zwischen 50 und 69 Jahren zur Verfügung. Das Programm folgt streng den europäischen Leitlinien. In den Medien wie auch in der Fachpresse gibt es kontroverse Diskussionen.

Material und Methoden: Überblick über die Datenlage bez. der Bewertung der randomisierten Studien und bez. der Qualitätsgesicherten Screening-Programme gemäß EU-Leitlinien (inklusive Daten aus 18 Screening-Ländern).

Ergebnisse: Positive Wirkungen des Screenings: Mortalitätsreduktion, schonendere Behandlung. Negative Wirkungen: Falsch positive Befundungen und Biopsieempfehlungen, sog. Überdiagnosen, Strahlendosis. Grenzen des Screenings: Intervallkarzinome, nicht vollständige Mortalitätsreduktion. Eine rechnerische Synopse der aktuellsten Veröffentlichungen der europäischen Screening-Programme mit den Abklärungsraten in Deutschland ermittelt aus >4,6 Mio. Screening-Untersuchungen ergibt folgendes: Von 1000 Frauen (TN), die am Mammografie-Screening-Programm teilnehmen (jede davon erhält in 20 Jahren 10 Mammografien), werden insgesamt 10 000 Mammografien erstellt. Insgesamt liegt das Risiko, durch eine Mammografie Brustkrebs auszulösen, ganz deutlich unter dem jährlichen natürlichen Risiko, an Brustkrebs zu erkranken. Im deutschen Screening werden von diesen 1000 Frauen wegen schlussendlich benignen Veränderungen durchschnittlich 288 Frauen 1-mal in 20 Jahren wiedereinbestellt (<3% pro Runde). Davon erhalten 74 der 288 Frauen eine Biopsie wegen einer benignen Veränderung (falsch positive Biopsieempfehlungen, i.d.R. Stanz- oder Vakuumbiopsie). Bezogen auf diesen Zeitraum entstehen entsprechend EUROSREEN 71 Karzinome bei den Teilnehmerinnen (56 werden im Screening, 15 im Intervall entdeckt), bei Nichtteilnehmerinnen 67 Karzinome (56 werden im Screening, 15 im Intervall entdeckt), bei Nichtteilnehmerinnen 67 Karzinome (56 werden im Screening, 15 im Intervall entdeckt), bei Nichtteilnehmerinnen

some cases, several years later) during this period. The 4 additional diagnoses among the Ps are referred to as overdiagnoses, as they do not contribute to a reduction in mortality (these participants die beforehand from other causes of death). With regard to the carcinomas that concern the screening periods, 11 women out of 1000 die among the Ps; there are 19 deaths among the N-Ps (within the observation period plus follow-up period).

Discussion: The false-positive rate is unavoidable, but is far lower with mammography screening than with other methods. Overdiagnoses are to be expected with any early detection. All calculations require assumptions and are therefore highly discrepant. They have very low evidence levels. The radiation dose should not be an argument against screening when applied correctly due to the very low risk and significant benefits. Interval carcinomas indicate the limits of a mammography screening programme. False-negatives only represent a subset of the interval carcinomas and are not to be equated with them. There is a very high evidence level for a significant reduction in mortality through mammography screening. For the first time, an independent expert commission has confirmed the results of the randomised studies and the statement of the WHO from 2002 and their further validity. Participants can expect a reduction in mortality of 30%. Data from the current European screening programmes confirm a mortality reduction of 43%, corresponding to 8/19 saved lives among 71 women with breast cancer or 1000 asymptomatic Ps. Many additional Ps benefit from less invasive treatment due to the early detection.

Conclusions: As a result of the risk/benefit ratio, mammography screening should absolutely be recommended to asymptomatic women aged between 50–69. High importance is given to the provision of education for women by the treating gynaecologists as regards the opportunities for quality-assured early detection available to them in the healthcare system.

merinnen (N-TN) 67 Karzinome (z. T. mehrere Jahre später). Die 4 zusätzlichen Diagnosen bei den TN werden als Überdiagnosen bezeichnet, da sie nicht zur Mortalitätsreduktion beitragen (diese Teilnehmerinnen versterben vorher an anderen Todesursachen). Bezogen auf die Karzinome, welche die Screening-Periode betreffen, sterben bei den TN 11 Frauen von 1000, bei den Nicht-TN sind es 19 Todesfälle (innerhalb des Betrachtungszeitraums plus Nachbeobachtungszeit).

Diskussion: Die Falsch-positiv-Rate ist unvermeidbar, ist aber mit Mammografie-Screening weitaus geringer als mit anderen Methoden. Überdiagnosen sind bei jeglicher Früherkennung zu erwarten. Alle Berechnungen erfordern Annahmen und sind daher höchst diskrepant. Sie haben sehr niedrigen Evidenzlevel. Die Strahlendosis sollte bei korrekter Anwendung wegen des sehr geringen Risikos bei deutlichem Benefit kein Grund gegen Screening sein. Intervallkarzinome zeigen die Grenzen eines Mammografie-Screening-Programms auf. Falsch negative stellen nur eine Teilmenge der Intervallkarzinome dar und sind nicht mit ihnen gleichzusetzen. Für eine signifikante Mortalitätsreduktion durch Mammografie-Screening besteht ein sehr hoher Evidenzlevel. Erstmals hat eine unabhängige Expertenkommission die Ergebnisse der randomisierten Studien und das Statement der WHO von 2002 und deren weitere Gültigkeit bestätigt. Teilnehmerinnen können eine Mortalitätsreduktion um 30% erwarten. Daten aus den aktuellen europäischen Screening-Programmen belegen eine Mortalitätsreduktion von 43%, entsprechend 8/19 geretteten Leben bei 71 Frauen mit Mammakarzinomen bzw. 1000 asymptomatischen TN. Viele weitere TN profitieren aufgrund der Früherkennung von einer schonenderen Behandlung.

Schlussfolgerung: Aufgrund der Nutzen-Risiko-Relation ist Mammografie-Screening für asymptomatische Frauen zwischen 50–69 Jahren unbedingt zu empfehlen. Einer Aufklärung der Frauen über die für sie im Gesundheitssystem geschaffenen Möglichkeiten für qualitätsgesicherte Früherkennung durch die betreuenden Frauenärzte kommt hier eine hohe Bedeutung zu.

Introduction

Since it became possible to verify the efficacy of mammography screening with regard to mortality reduction using randomised studies approximately 30 years ago, mammography has been introduced as an early detection method in numerous countries. Quality-assured centrally organised mammography screening programmes have been introduced in particular in Europe, Australia and New Zealand. These organised screening programmes follow a defined quality control procedure, which – corresponding to EU Guidelines – should comprise all parts of screening, beginning with the invitation and creation of the mammograms through to determining a benign or malignant result [1]: centralised quality assurance for all of the mammography technology with continuous external monitoring of equipment and recording technology each working day, monthly and annually, minimum requirements for all medical and non-medical staff involved in screening, with regard to training and retention of professional skills, systematic diagnosis of all scans by 2 independent reviewers, consensus diagnosis of all abnormalities, regular participation at interdisciplinary conferences, standardised investigation chain, systematic documentation, and regular external verification of the quality of processes and results using performance parameters.

In contrast to this, to date only what is known as “grey” (opportunistic) screening is carried out in other countries, such as the USA, as well as Austria and Switzerland. Besides the considerably lower quality requirements, no systematic documentation takes place.

In Germany, a Mammography Screening Programme carried out in strict accordance with EU was established between 2005 and 2009 (period guidelines until complete coverage).

Using the comprehensive documentation, it is now already possible to determine and assess important quality parameters from the German programme. These enable the first assessments of efficacy to be made. However, more precise calculations into mortality reduction are naturally only possible 8–10 years after a screening programme is established.

The efficacy of organised mammography screening is now well documented by the results of randomised studies [2], as well as by epidemiological evaluations from 18 screening countries [3]. Despite the high level of evidence available, mammography screening is intensely debated among advocates and early detection critics.

Materials and Methods

This work is intended to provide an insight into the data currently available, show the evaluation of randomised studies, also with regard to quality-assured screening programmes in accordance with EU Guidelines (including data from 18 screening countries), explain the possibilities and limits of mammography screening, and report on the results already available from Germany. This is not a systematic review.

Results

Reducing mortality was defined as the most important goal of a mammography screening programme. Besides this main effect, there are other positive effects that result from an earlier detection of cancer. They relate above all to less invasive treatment options (more breast retention, avoidance of axillary dissection, reduction in chemotherapy and a better cosmetic result).

The fact that interval carcinomas and also advanced stages are not preventable indicates the limits of mammography screening. Potential side effects result from the low dose of radiation applied, from false-positive diagnoses, and possible overdiagnoses.

Reduction in mortality

The reduction in mortality through mammography screening was investigated in 8 randomised studies. These showed that a reduction in mortality of approximately 30% is to be expected on average for a woman taking part every 2 years over a period of 20 years.

Since the year 2000, the results of the randomised studies have been called into question by Goetzsche. He initially attempted to declare 6/8 randomised studies as invalid due to the cluster randomisation (a common and recognised randomisation technique) used in these because, in principle, age differences can occur between study and control groups when cluster randomisation (invitation/non-invitation to entire locations/cities) is used [4]. He therefore excluded 6 out of 8 randomised studies based on his own formalistic definition, although it was possible to show that the, in some cases opposing, average age differences between the studies (of just a few months in each case) had no effect on the overall result.

Of the remaining 2 studies, which Goetzsche continues to this day to assess as reliable according to his own definition, the study conducted in Canada is highly controversial due to significant deficiencies in mammography quality and due to possible severe flaws in randomisation [5–7]. During the study conducted in Malmö, the verifiable effect was reduced as there was a high cross-over level (high proportion of participants among the uninvited women and vice versa).

In further publications, Goetzsche [8] doubted the objectivity of the manner in which cause of death was determined in some of the studies. These claims have not been confirmed following re-evaluation of these studies by an independent WHO commission [9].

Despite the clear statement by the WHO in 2002, Goetzsche continues to insist on a selectively different assessment of the present randomised studies and also postulates a mortality reduction of just 15% between invited vs. uninvited women.

In 2012, an independent panel of experts in the United Kingdom was commissioned to re-evaluate the mortality reduction [2]. The independent panel re-evaluated both the data from the ran-

domised studies and the available meta-analyses and took into account arguments from screening advocates and opponents. It was clear that a mortality reduction of 20% can be assumed with regard to invited vs. uninvited women. As in the randomised studies only 70% of the invited women took part in the screened group, a mortality reduction of 20% relating to invited vs. uninvited women corresponds approximately to a mortality reduction of > 30% for actual participants [10]. Overall, the independent British expert commission thus confirms the results of the randomised studies and the statement by the WHO from 2002.

The British expert commission explicitly confirms that it can find no reason to exclude some of the randomised studies. It also takes the view that the effects of improved treatment and effects of diagnosis are very likely independent of each other, so the proven mortality reduction achieved through mammography screening is to be expected in addition to the treatment improvements made since. Reference is also made to the simultaneous improvement in mammography technology since the randomised studies were concluded (30 years ago).

Naturally, no further randomised studies have been possible since state screening programmes were introduced. This means that other study types are required in order to be able to assess the effect among participants vs. non-participants. Institutions which supervise the screening programmes in the screening countries and have assessed the primary data have explicitly indicated the need for a correct statistical methodology [11].

Simple trend analyses are considered unsuitable as cases of breast cancer that already occurred prior to the start of the screening programme and the deaths resulting from these cannot be correctly separated out. This can lead to significant underestimates of mortality reduction. Furthermore, additional external influences cannot be sufficiently identified or taken into account.

Incidence-based mortality studies and case-control studies appear suitable [11, 12]; however, an under- and overestimate of the effect is possible both for mortality studies and case-control studies. An adequate follow-up period is of the utmost importance. As breast cancer was discovered many years earlier in some cases in the study group, it is essential for calculating mortality reduction that the tumours discovered significantly later in some cases in the control group and their later follow-up are recorded. If one takes into account that tumours become symptomatic following a period of time that differs greatly from individual to individual (example DCIS), but can then progress in very different ways, a follow-up period of at least 10 years is considered indispensable.

Additional possible influencing factors are also to be taken into account for all the study types. These relate to regional differences, time-based trends, relevant age differences, and the influence of diet and medication that can occur between the study group and the selected control groups.

A systematic review, which analysed studies from 18 screening countries with regard to the correct statistical methodology and included those studies that fulfilled the aforementioned requirements, showed an average mortality reduction for actual participants of 43% (38–48%) using the data from 18 screening countries [13].

Studies that recently reported an incomprehensibly low mortality reduction [14, 15] indicated the aforementioned problems, e.g. no segregation of the carcinomas that occurred prior to screening, as well as, in the final case, a follow-up period of just 2 years (1).

Table 1 Most important results: Performance parameters with regard to the detection of breast cancer. Performance parameters affecting the detection of breast cancer: all parameters correspond to the requirements of the EU Guidelines [1,24].

Parameter	Reference range EU Guidelines, 4th edition		Results screening Germany 2008–2009	
	First examinations	Subsequent examinations	First examinations	Subsequent examinations
4. Breast cancer detection rate (multiple of the regional background incidence)	$\geq 3 \times \text{IR}$	$\geq 1.5 \times \text{IR}$	Regionally fluctuating (1.7–4.6)	Regionally fluctuating (1.4–4.0)
5. Breast cancer detection rate (cases per 1 000)	approx. 7.5/1 000	approx. 3.8/1 000	8.2‰	5.6‰
6. Proportion of in-situ carcinomas	10% > 15%	10% > 15%	19.5%	19.8%
7. Proportion of invasive carcinomas ≤ 10 mm	n. a. $\geq 25\%$	25% $\geq 30\%$	30.0%	34.9%
8. Proportion of invasive carcinomas < 15 mm	$\geq 50\%$	$\geq 50\%$	51.3%	57.2%
9. Proportion of carcinomas not involving the lymph nodes (N–)	n. a. > 70%	$\geq 75\%$	74.8%	79.0%
10. Proportion of carcinomas in UICC stage II+	n. a. < 30%	$\leq 25\%$	28.9%	24.2%

Table 2 Most important results: Performance parameters with regard to investigation. All performance parameters affecting investigation fulfil the requirements of the EU Guidelines 4th edition [1, 24].

Parameter	Reference range EU Guidelines, 4th edition		Results screening Germany 2008–2009	
	First examinations	Subsequent examinations	First examinations	Subsequent examinations
3. Call-back rate	5–7%	3–5%	6.1%	3.0%
11. Proportion of preoperatively detected carcinomas	$\geq 90\%$ (3rd ed. > 70%)		91.7%	
12. PPV I	n. a.		14.8%	
13. PPV II	n. a.		49.4%	

Data on mortality reduction are to be expected from the German programme at the earliest 10 years after the national screening programme is established nationwide. The necessary data protection regulations are currently still being finalised for this purpose across the country.

The process and results parameters available to date that influence mortality (detection rates, stage distribution of the discovered carcinomas) enable an effect comparable with that of other programmes to be expected for Germany. **Table 1** and **Table 2** show the most important results obtained in the years 2008–2009 and the requirements in accordance with EU Guidelines. **Fig. 1** offers an overview of the stage distribution of the carcinomas discovered in the screening programme compared with the stage distribution within the same age groups before the screening programme started. **Fig. 2** shows the carcinomas that do not involve the lymph nodes in the target population and among the screening participants.

Less invasive treatment

If one combines the more favourable stage distribution of the carcinomas among screening participants with the current approach in line with the guidelines, it is clear that benefits of early detection in no way only apply to women whose lives can be saved. Naturally, better cosmetic results, a higher rate of breast-conserving therapy (BCT), fewer axillary dissections and less chemotherapy are to be expected when smaller carcinomas are discovered.

Screening opponents also report here on a lack of benefits or even of disadvantages (e.g. higher mastectomy rates) [8, 16].

The reported disadvantages relate in one case to very old data (prior to systematic introduction of BCT) [8], in another case [16], these calculations may have arisen from a counteracting effect caused by the simultaneous introduction of screening and a parallel change in therapeutic standards (modern treatments particularly benefitted the carcinomas discovered later in the control group).

Contrary to the calculations by screening opponents, responsible institutions in various countries were able to demonstrate the following benefits (to be expected from the changed stage distribution) for screening participants:

- ▶ A reduction in mastectomies in favour of a higher BCT rate
- ▶ A higher rate of sentinel lymph node operations with a reduction in axillary dissections [17–19]

There is still the absence of a Cancer Register comparison in Germany, but an incidence-based trend analysis from Bavaria is now available for the first time. This compares treatment trends from 2000 to 2008 for various age groups. These show the following initial trends for the age group of 50- to 69-year-olds (at participation in screening, which rose during this period from 0 to approx. 30%) in comparison with the other age groups: mastectomy rates falling more significantly, more pronounced fall in adjuvant chemotherapies, more pronounced fall in axillary dissections [20].

Radiation dose

As radiation can trigger cancer, it is particularly important that the radiation dose is optimised for screening, where predominantly healthy women are regularly examined with X-rays. How-

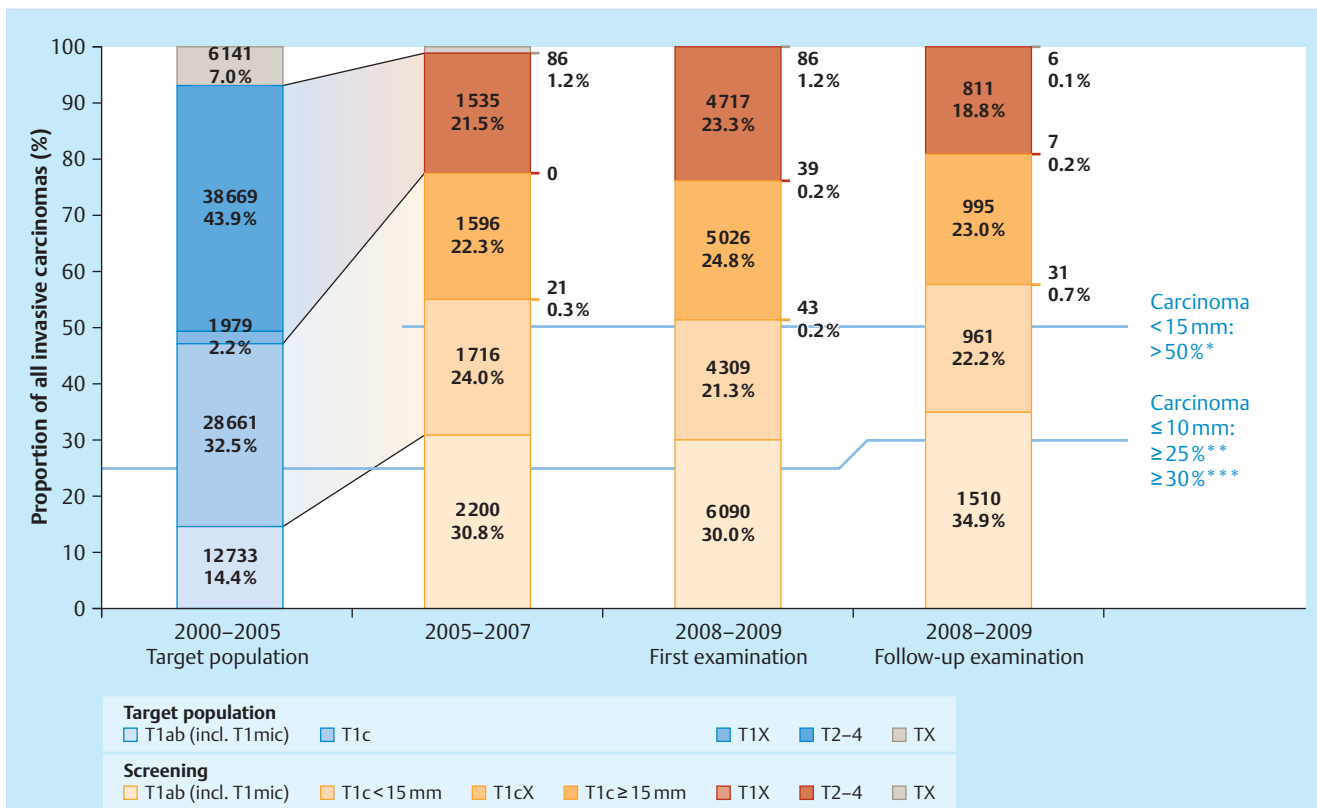


Fig. 1 Simple T-stage distribution of invasive breast cancers in the screening (2005–2007, 2008–2009) and in the target population (2000–2005).

There is a noticeable increase in small invasive carcinomas in the screened population [24].

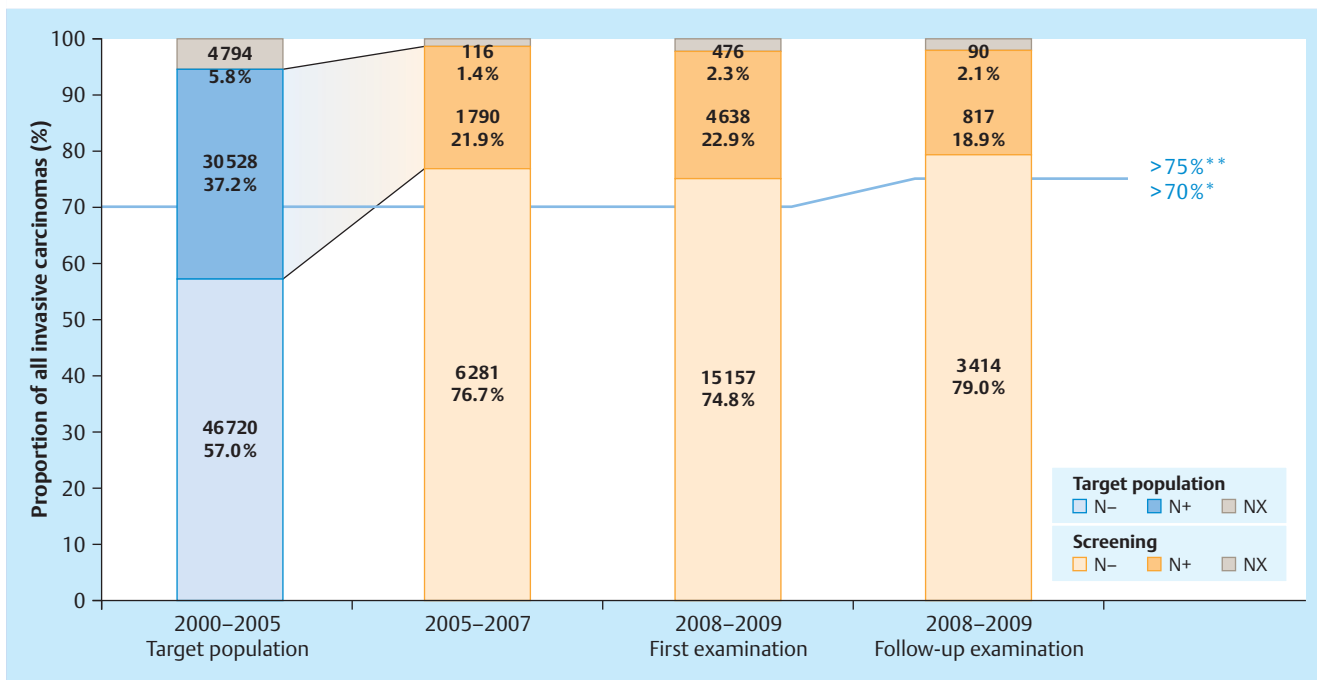


Fig. 2 Proportion of lymph node-negative carcinomas in the screening (2005–2007, 2008–2009) and in the target population (2000–2005). The effect of the screening can already be seen at the beginning of the pro-

gramme: an increase in node-negative carcinomas with a simultaneous reduction in node-positive carcinomas in the screened women [24].

ever, the image quality and dose are closely linked. In order to achieve optimal image quality with a minimal required dose and thus ensure that the smallest changes are detected, apart from using modern technology, optimally coordinated equipment, parameters and positioning, the continuous training and testing of diagnosticians, and the optimisation of the entire screening chain are crucial. The most important technical components are monitored daily online and annually in person during the screening. According to early cancer detection guidelines and German X-ray regulations, what are known as IGeL mammograms (individual healthcare services not featured on the list of services covered by statutory health insurance companies), for which all the provided quality assurance for the screening can be circumvented, are not permissible for early detection in asymptomatic women. A mean parenchyma dose of significantly below 4 mGy is now assumed within the quality assurance provided for the screening. (The parenchyma dose required on an individual basis varies considerably depending on breast size and compression.) Overall, (according to current calculations) the risk of triggering breast cancer through a mammogram is very clearly (factor of 50–100) below the annual natural risk of suffering from breast cancer as early as from the 40th year of life [21,22]. If one assumes that mortality can be reduced through quality-assured mammography screening, which naturally also applies to rare cases of triggering cancer through mammography, the calculated benefit is significantly greater than the risk. This is the reason why mammography screening is approved and recommended in Western countries. According to the S3 Guidelines [23], the radiation dose applied during a screening mammogram cannot be considered an argument against screening mammography.

False-positive rate

The German Mammography Screening Programme is conducted purely in the form of mammography screening according to EU Guidelines. If an abnormality is detected as part of independent double reading and the subsequent consensus conference, the doctor responsible for the programme (DRP) invites the woman for further tests (called an imaging assessment). During most imaging assessments, malignancy can be excluded from the very first diagnosis appointment. If this is not possible, a histological assessment follows, which is carried out by highly experienced examiners as gently as possible using minimally invasive methods in over 90% of cases. When assessing the programme, every invitation for investigation counts as a positive finding and is designated as “false-positive” if the finding is ultimately discovered to be benign. The investigations conducted in Germany are already being collected prospectively and are thus clear [24]. For the years 2008–2009 (• Tables 1 and 2), 6.1% of women were invited for further investigation at the first examinations, which corresponds to a “false-positive rate” of 5.3% with 8.2 carcinomas discovered among 1000 screened women. At subsequent examinations (women who take part in screening repeatedly) an average of 3% of women were invited for diagnosis with 5.6/1000 discovered carcinomas, corresponding to a “false-positive rate” of 2.44%.

Histological assessments (usually minimally invasive investigations) were conducted in the screening in 2008–2009 for 1.5% of all women (first and subsequent examinations)¹ whereby a ma-

lignant tumour was found in 7.6/1000 women¹. This corresponds to a rate of 0.74% false-positive biopsy recommendations.

Screening opponents calculate the number of “false-positive findings” over 20 years of screening based on far higher investigation rates in the USA and talk of so-called “false-positive rates” of up to 50% in the screening. This does not apply to the German programme.

In addition to the rate of false-positive biopsy recommendations, the false-positive rate (relating to the recommendation for further investigation) in the Mammography screening is far lower than for other imaging methods, at 5.3% in the first and 2.44% in the subsequent screening cycles [25].

If one then calculates false-positive rates relating to the 20-year “screening life” of a woman in the German Mammography screening, this indicates, using the results gathered prospectively for 2008–2009 [24], that < 300 out of 1000 women on average would be invited for additional imaging once within 20 years due to an ultimately benign finding. A histological assessment is required to determine a benign finding in 74 of these women on average. At the same time, 76 breast cancers are found in the screening during this period.

In comparison with the pooled results from other screening countries [26], the call-back rate for imaging investigations in Germany is somewhat higher on average, with a comparable rate of histological examinations. Of particularly favourable note is the considerably lower proportion of open biopsies, which is explained by the consistent introduction of modern minimally invasive techniques in the German programme.

Interval carcinomas

Interval carcinomas are defined as carcinomas that occur within 24 months following a screening mammogram without pathological findings (i.e. in the screening interval). They are not to be equated with overlooked findings, as these only constitute a proportion of interval carcinomas (see below) [29]. However, they indicate the limits of a screening programme.

Breast cancers can develop at any time (and not just every 2 years); the screening measures only take place at set intervals (every 2 years). There are no methods (either mammography or other imaging procedures) that have a 100% prediction probability for 2 years, meaning that interval carcinomas are unavoidable throughout the world and with the use of all screening methods. Interval carcinomas in mammography screening are comprised as follows:

1. Carcinomas that develop in the interval and thus did not exist (“no sign”) at the time of screening.
2. Carcinomas that exist at the time of screening but are not identifiable (“no sign”) using the Mammography method (even retrospectively).
3. Carcinomas that are prospectively not identifiable or cannot be differentiated from benign changes with sufficient certainty, but their existence or development in a corresponding location can be suggested retrospectively based on an unspecific mammographic change (“minimal sign”).
4. Carcinomas that are overlooked or incorrectly assessed (even by 2 experienced examiners with independent readings) (“missed” or “misinterpreted”).

A distinction between the aforementioned groups is in no way trivial, as there are naturally blurred transitions. As mammography is difficult to interpret (like other mamma-diagnostic methods) due to the countless standard variants and various benign changes, widely different assessments can be made in some cases

¹ Data from the German Mammography Screening Programme.

– depending on the prior information (knowledge of location and appearance of the carcinoma subsequently diagnosed vs. blind diagnosis mixed with unremarkable cases) [26]. Where there is a lack of blinding and a retrospective assessment, the prospective diagnosability of a finding can sometimes be significantly overestimated, particularly by less experienced assessors when the localisation of a finding is known to them.

For patients and investigators from outside this field, it is also sometimes difficult to understand that, although mammography generally offers good sensitivity for small carcinomas, it is sometimes not possible to diagnose larger carcinomas using mammography. This is explained by the fact that, depending on the mammary glands and carcinoma type, carcinomas that do not contain any micro-calcifications can be hidden by denser glandular tissue or benign changes, or imitate these changes. This can particularly apply to diffusely growing carcinomas and carcinomas that develop in pre-existing asymmetries (which frequently occur as normal findings).

Overall, based on the results of European screening programmes, one assumes that approx. 25% of all carcinomas that occur or develop among screening participants within 2 years are diagnosed in the interval [1,27]. Carcinomas that are actually overlooked may only constitute a small proportion of all interval carcinomas (below 5–10%) with independent double reading in this regard. No precise figures are yet available for the German Mammography Screening Programme due to the data protection regulations still undergoing clarification.

An initial fully anonymised Cancer Register comparison for the years 2005–2007 (first screening cycle) from North Rhine-Westphalia [28] indicated a proportion of 22% interval carcinomas. This corresponds to a mammography sensitivity of 78% for the first cycle based on the whole (!) 2-year interval. For a woman with an unremarkable screening finding, the probability that *no* breast cancer will also be found in the interval during the following 2 years is at > 98.8% (negative prediction value).

In other words, one could also say that an interval carcinoma must be expected in 15 out of 1000 women with unremarkable mammography screening findings (first cycle).

Due to the anonymisation currently required under data protection law, it was still not possible to conduct analyses of individual cases. Owing to the naturally lower breast cancer detection rate in the subsequent screening cycles, a somewhat lower overall sensitivity is to be expected for future calculations. Programme sensitivities between 67 and 84% are reported from established European screening programmes [27].

These results, like (see above) the other surrogate parameters, are indicative of the good quality of the German Mammography Screening Programme by international comparison. This result also includes the possibilities of errors (counteracting in each case), as there is no individual case comparison and there are remaining gaps in the Cancer Register.

Overdiagnoses

Overdiagnosis is defined as the diagnosis of breast cancer which would not have been discovered during the woman's lifetime without screening. Overdiagnoses are thus unwanted screening results. In the biological sense, additionally diagnosed breast cancer in the screening group is just as much real breast cancer as any breast cancer discovered outside of the Screening Programme, the early detection of which can save lives.

Overdiagnoses are created in that breast cancer is sometimes identified much earlier in the screened group than in the

unscreened group (called lead-time bias). This early diagnosis then becomes overdiagnosis when the breast cancer, for instance in the case of a low-grade DCIS, develops so slowly that the patient dies before the breast cancer would have become symptomatic (without screening).

As neither the remaining lifetime of the woman nor the precise course of the illness without discovery can be predicted on an individual basis, overdiagnoses can never be determined in an individual woman with certainty.

The calculation of overdiagnoses and how they differ from early diagnoses is extremely difficult and subject to considerable uncertainties, as assumptions are always required here. Assumptions are always required for such calculations. Depending on the calculation method, there are extreme deviations in the calculation of expected overdiagnoses. Estimates of between 0 and 50% of all breast cancers diagnosed in the screening have been published [30–34].

The possibilities of errors may be present for all estimation procedures [11,30,35].

Calculations of extremely high overdiagnoses [31,32] can largely be explained methodologically, e.g. through an inadequate follow-up period, insufficient consideration of the length-time bias in the statistical approach, selection of unsuitable control groups, no consideration of external trends). The EUROSCREEN Group analysed the literature on this topic, paying particular attention to the correct methodology, and estimated the rate of overdiagnoses as 1–10%, 6% on average [30].

Reliable estimates regarding the frequency of overdiagnoses may come to between 5 and 20% of all cancer diagnoses [2,30,33,34]. Overall, it should be emphasised that, due to the strong influence of the calculation methodology and the controversial and discrepant results in the literature, the existing evidence should be classed as very low with regard to the calculation of overdiagnoses. The scientific data currently available on this subject are thus clear and are inconsistent with the views of screening opponents.

Overview and absolute figures

Percentage information, both from screening opponents and screening advocates, is used. This can vary significantly depending on the reference basis and lead to relative distortions with a changing reference basis, producing a false image.

An overview is provided below of the absolute frequencies of positive and negative effects to be expected from screening (► **Table 3**). It is important to understand that, although breast cancer can affect one in 9 women in their lives (lifetime risk approx. 11%), the annual risk (annual!) is low at 3/1000. This shows the need for regular examinations of many asymptomatic women if one wishes to detect breast cancer early or in good time.

Below, 1000 women who take part in the screening over 20 years (P) are compared with 1000 women who do not take part (N-P). The participants undergo a mammogram every 2 years (= 10,000 screening mammograms in 20 years between 50–69).

As a result of a benign finding¹, 288 women on average are invited for diagnosis *once* in 20 years (28.8/1000 per cycle or almost 3%), of which 74 (7.4 per cycle²) have to undergo a histological investigation once in 20 years due to a benign finding.

² The figures from the German screening may currently be calculated too highly as this figure for 2008–2009 is comprised primarily of the more frequent histological investigations in the first screening cycle.

Table 3 Overview with regard to the numerical data to be expected in the screening (absolute values for 20 years of screening among 1 000 participants). The values come from the review by the EUROSCREEN Working Group [3] and include the median values in the German Screening Programme with regard to diagnoses. 19 women² die without screening in the control group; 11 (10 to 12) women² die in the screened group. This corresponds to 8 lives saved among 1 000 screening participants or a mortality reduction of approx. 43%.

	With screening (P)	Without screening (N-P)	Difference
Number of women	1 000	1 000	
Mammograms/20 years	10 000	unknown	
Call-backs for any diagnosis due to benign changes	288 (incl. biopsies)	unknown	
of which:			
▶ Biopsy due to benign change	74	unknown	
▶ discovered carcinomas (screening period and follow-up)	71	67	4 (- 14)
▶ Deaths due to carcinomas (relating to screening period)	11	19	7 to 9

According to the calculations by the EUROSCREEN Group, an average of 71 breast cancers are found in Ps (approx. 56 in the screening and 15 in the interval), while there are an average of 67 breast cancers in the control group (N-Ps) (taking into account an adequate follow-up period). The difference of 4 carcinomas thus corresponds to the overdiagnoses. (These are actual breast cancers that would not be detected during the woman's lifetime without screening. According to calculations by the UK panel, it could be up to 14³.)

A relevant proportion of the 56 women with breast cancer discovered in the screening may benefit from earlier discovery (more breast retention, fewer axillary dissections, less chemotherapy due to discovery at an earlier stage). There are no model calculations available here.

19 women² die without screening in the control group; 11 (10 to 12) women² die in the screened group. This corresponds to 8 lives saved among 1,000 screening participants or a mortality reduction of approx. 43%.

Discussion

Since its introduction between 2005 and 2009, mammography screening has become increasingly established in Germany. The outcome parameters of the German Mammography Screening Programme fulfil the requirements of the EU Guidelines and also show very good results (▶ Tables 1 and 2) by international comparison.

In view of the astoundingly high rate of carcinomas discovered (despite opportunistic grey screening, which exists in parallel), not only can it be argued that there is a generally slightly higher diagnosis rate using imaging investigations compared with other screening countries, but this also lies within the scope of the limit values specified by the EU Guidelines. The biopsy rate calculated for the German programme appears high with regard to the whole programme, but is distorted by the high number of first examinations, for which higher biopsy rates must be applied. The gratifyingly high proportion of minimally invasive biopsies among the histological investigations in the German programme is to be noted. This may be due to the early inclusion of vacuum biopsy, as well as to the introduction of digital screening, for both of which Germany proved to be a pioneer in Europe. Data on

mortality reduction or overdiagnoses are not to be expected yet, as coverage was only achieved in 2009. According to current knowledge (based on stage distribution and detection rates), it is to be expected that the German programme can achieve comparable effects to those seen in other quality-assured screening programmes in Europe.

In general, both positive effects and possible side effects are to be expected from screening programmes (as is the case for other medical measures).

Following countless attacks against the Mammography Screening Programme, an independent British panel has since confirmed that the mortality reduction calculated using randomised studies and specified by WHO (approx. 30% for the actual participant) continues to be valid. The EUROSCREEN Working Group was able to calculate a mortality reduction of approx. 43% for the current quality-assured screening programmes using the data from 18 countries (▶ Table 3). All the data available so far indicate that the programme is highly effective, including in Germany.

A mortality reduction of 43% corresponds to approx. 8 lives saved per 1000 screening participants. (It should be taken into account here that only 70 breast cancers occur among 1000 participants within the 20 screening years between 50–69, of which approx. 19 are fatal without screening.)

Information from screening opponents, who claim a “benefit” among “just 1/1000” women or a similarly low efficacy, does not correspond to the data resulting from all of the meta-analyses and randomised studies [2], and relates to other periods of time or another age group (e.g. 10 years of screening among under-40s, etc.) The advantages of earlier treatment are also ignored.

The radiation dose, which is often stated as the primary disadvantage, should still not be an argument against screening when applied correctly due to the very low risk and significant benefits [21–23]. Warnings should be issued about so-called IGeL mammograms (with questionable application in some cases, e.g. due to low-dose mammography or a “wellbeing factor”) due to the lack of quality assurance and thus also uncertain effects.

If breast cancer is to be discovered at an early stage, further investigations with a certain number of ultimately benign changes (as for almost all medical tests) are unavoidable. Call-backs are necessary during the screening for this reason. Fortunately, quality-assured mammography screening has a very high specificity, so only approx. 30/1000 women need to be called back per screening cycle. Based on the screening period, this means that almost 300 women are called back *once* (i.e. in 10 screening cycles) within 20 years due to an ultimately benign finding. Quality-assured investigations in mammography screening take place in as minimally invasive a manner as possible (mostly without biopsies,

³ The calculations result from 2 randomised studies with a follow-up period of 6 years. An incorrect increase in this calculation can also be produced when former screening participants have also undergone a grey screening.

otherwise usually with minimally invasive biopsy techniques). Approximately one in every two needle biopsies demonstrates a malignant tumour. If one contrasts these disadvantages with the benefit of saving lives and the less invasive treatment on earlier diagnosis, the stress of investigation is medically low and highly justifiable. A close collaboration between the screening unit and the treating gynaecologists and GPs should absolutely be strived for in order to bypass any possible short-term, highly individual mental stress experienced by the patient between the invitation and investigation.

There is actually no other diagnostic method with such high specificity as the Mammography Screening Programme. If one compares ultrasound and MRI when used for the early detection of breast cancer with mammography, it is initially to be noted that there is a lack of prospective studies among asymptomatic women without an increased risk. Among risk collectives it can be seen that, depending on the risk constellation, sensitivity can be increased using supplementary methods. However, specificity falls considerably when other methods are added. With mammography alone, significantly fewer invasive investigations (> factor 3) are required, in particular [25, 38]. Recommendations for short-term check-ups (where stated in the literature) are, in some cases, more than 10 times higher for sonographic and MRI investigations than for mammography [38, 39].

A further unwanted effect of every screening investigation that is relevant when applying sensitive test procedures (i.e. procedures that can identify cancer long before it is palpable) is what is known as overdiagnosis. Overdiagnoses are thus to be expected with any early detection method (including grey screening), but, in principle, cannot be determined due to a lack of data. They occur when breast cancer is discovered very early using imaging (or another test) and/or grows very slowly, so the woman would not have known about it during her lifetime without screening. An excessively high number of overdiagnoses is to be avoided as it does not contribute to a reduction in mortality, yet the patient is placed under unnecessary stress by the knowledge of her breast cancer and the treatment. Unfortunately, all estimates are tainted by significant calculatory fluctuation margins due to the uncertain state of data. The current discussion regarding overdiagnoses [31], which are even incorrectly shown as wrongly diagnosed breast cancer in some cases, is scientifically incomprehensible and hardly defensible medically in terms of the patient in light of the very low evidence level that exists.

In view of the benefits, the number of overdiagnoses ascertained according to reliable calculations appears to be absolutely justifiable in medical terms. It is important to understand that it is not possible to predict for any individual woman whether the very early diagnosis of breast cancer benefits her (by saving her life and through better treatment options) or could be an overdiagnosis. Discovery and appropriate treatment therefore remain sensible and important. The “damage” of an overdiagnosis is actually closely linked to the treatment that follows it. Chemotherapy and axillary dissections are usually rare, especially in the case of early breast cancers. However, with increasingly earlier discovery, it is particularly important that treatment is adjusted optimally to suit the individual risk of the patient.

Despite the high sensitivity of mammography, interval carcinomas are unavoidable. Through appropriate investigation, it should be ensured that clinical findings undergo a further investigation, even after a screening mammogram without pathological findings.

Although it is known that additional breast cancer cases can be found with sonography or MRI, pure sonography or MRI-based screening is currently not possible due to their far lower specificity and the as yet insufficient quality assurance for screening.

The first encouraging results have been reported on the prospective application of tomosynthesis, a mammographic method whereby tomographic images of the breast are produced [36, 37]. Regardless of the discussions generated by screening critics (these are actually opponents of any early detection measures), quality-assured mammography screening proves to be a sensible and important measure. Any side effects and limits are to be taken into account and minimised. It should be explicitly noted that any side effects of screening apply at least equally to grey screening, although the effects and side effects cannot be verified for this, nor can the quality appropriate for early detection. Although acceptance of screening is high among screening participants [24], the national participation rate should be further increased from the recently published 54% of women. Gynaecologists, in particular, can perform important clarification work here.

Future research should be concentrated on possible and sensible amendments, especially to reduce interval carcinomas and late stages. The optimal adjustment of treatment is important for small breast cancers discovered at an especially early stage. A reduction in stressful treatments should be looked at here using suitable studies.

Conflict of Interest



All authors are affiliated to the German Mammography Screening Program.

Affiliations

- ¹ Munich Mammography Reference Centre, Munich
- ² Southwest Reference Centre, Marburg
- ³ Münster Reference Centre, Münster
- ⁴ North Reference Centre, Oldenburg
- ⁵ Berlin Reference Centre, Berlin
- ⁶ Munich Reference Centre, Munich
- ⁷ Mammography Cooperation Community, Berlin

Literatur

- 1 Perry N, Broeders M, De Wolf C et al., eds. European Guidelines for Quality Assurance in Mammography Screening. 4th ed. Luxembourg: Office for Official Publications of the European Communities; 2006
- 2 Independent UK Panel on Breast Cancer Screening. The benefits and harms of breast cancer screening: an independent review. *Lancet* 2012; 380: 1778–1786
- 3 EUROSCREEN Working Group. Summary of the evidence of breast cancer service screening outcomes in Europe and first estimate of the benefit and harm balance sheet. *J Med Screen* 2012; 19 (Suppl. 1): 5–13
- 4 Göttsche PC, Olsen O. Is screening for breast cancer with mammography justifiable? *Lancet* 2000; 355: 129–134
- 5 Baines CJ, Miller AB, Kopans DB et al. Canadian National Breast Screening Study: assessment of technical quality by external review. *Am J Roentgenol* 1990; 155: 743–747
- 6 Tarone RE. The excess of patients with advanced breast cancers in young women screened with mammography in the Canadian National Breast Screening Study. *Cancer* 1995; 75: 997–1003
- 7 Kopans DB. Why the critics of screening mammography are wrong. *Diagnostic Imaging* 2009; 31: 1–5
- 8 Göttsche PC, Nielsen M. Screening for breast cancer with mammography. *Cochrane Database Syst Rev* 2009; 4: CD001877
- 9 International Agency for Research on Cancer (WHO). Press release no. 139 19-03-02. Mammography screening can reduce deaths from breast cancer. 2002. www.iarc.fr/en/media-centre/pr/2002/pr139.html; last access: 15.10.2013

- 10 Brenner H, Heywang-Köbrunner S, Becker N. Re: [answer to Gigerenzer G, Mata J, Frank R] Public knowledge of benefits of breast and prostate cancer screening in Europe. *J Natl Cancer Inst* 2010;102: 356
- 11 Hackshaw A. The benefits and harms of mammographic screening for breast cancer: building the evidence base using service screening programmes. *J Med Screen* 2012; 19 (Suppl. 1): 1–2
- 12 Broeders M, Moss S, Nyström N et al.; EUROSCREEN Working Group. The impact of mammographic screening on breast cancer mortality in Europe: a review of observational studies. *J Med Screen* 2012; 19 (Suppl. 1): 14–25
- 13 Broeders M, Moss S, Nyström L et al.; EUROSCREEN Working Group. The impact of mammographic screening on breast cancer mortality in Europe: a review of trend studies. *J Med Screen* 2012; 19 (Suppl. 1): 26–32
- 14 Jørgensen KJ, Zahl PH, Gøtzsche PC. Breast cancer mortality in organised mammography screening in Denmark: comparative study. *BMJ* 2010; DOI: 10.1136/bmj.c1241
- 15 Kalager M, Zelen M, Langmark F et al. Effect of screening mammography on breast cancer mortality in Norway. *NEJM* 2010; 363: 1203–1210
- 16 Suhrke P, Mæhlen J, Schlichting E et al. Effect of mammography screening on surgical treatment for breast cancer in Norway: comparative analysis of cancer registry data. *BMJ* 2011; 343: d4692
- 17 Samnakay N, Tinning J, Ives A et al. Rates for mastectomy are lower in women attending a breast-screening programme. *ANZ J Surg* 2005; 75: 936–939
- 18 Zorzi M, Puliti D, Vettorazzi M et al. Mastectomy rates are decreasing in the era of service screening: a population-based study in Italy (1997–2001). *Br J Cancer* 2006; 95: 1265–1268
- 19 Dillon MF, Hill AD, Quinn CM et al. Surgical intervention in screen-detected patients versus symptomatic patients with breast cancer. *J Med Screen* 2004; 11: 130–134; Erratum in: *J Med Screen* 2004; 11: 211
- 20 Schrodi S, Braisch U, Schenkirsch G et al. Veränderungen der Therapie des Mammakarzinoms in Bayern seit Einführung des Mammographie-Screenings. Eine Analyse bayerischer Krebsregisterdaten der Jahre 2000 bis 2008. *Gesundheitswesen* 2013; [Epub ahead of print]
- 21 Yaffe MJ, Mainprize JG. Risk of radiation-induced breast cancer from mammographic screening. *Radiology* 2011; 258: 98–105
- 22 Hendrick RE. Radiation doses and cancer risks from breast imaging studies. *Radiology* 2010; 257: 246–253
- 23 Albert U. Stufe-3-Leitlinie Brustkrebs-Früherkennung in Deutschland. München, Wien, New York: W. Zuckschwerdt Verlag; 2008
- 24 Kooperationsgemeinschaft Mammographie. Evaluationsbericht 2008–2009. Ergebnisse des Mammographie-Screening-Programms in Deutschland. http://www.mammo-programm.de/cms_upload/datenpool/qualibericht_2012_final_web.pdf; last access: 15.10.2013
- 25 Heywang-Köbrunner S, Schreer I, Heindel W et al. Bildgebung für die Brustkrebsfrüherkennung. *Dtsch Arztebl* 2008; 105: 541–547
- 26 Hofvind S, Skaane P, Vitak B et al. Influence of review design on percentages of missed interval breast cancers: retrospective study of interval cancers in a population-based screening program. *Radiology* 2005; 237: 437–443
- 27 Törnberg S, Kemetli L, Ascunce N et al. A pooled analysis of interval cancer rates in six European countries. *Eur J Cancer Prev* 2010; 19: 87–93
- 28 Heidinger O, Batzler WU, Krieg V et al. Häufigkeit von Intervallkarzinomen im deutschen Mammographie-Screening-Programm: Auswertungen des Epidemiologischen Krebsregisters Nordrhein-Westfalen. *Dtsch Arztebl Int* 2012; 109: 781–787
- 29 Bock K, Hecht G, Heindel W et al. Krebsregisterdaten – Bedeutung von Intervallkarzinomen im Mammografie-Screening-Programm. *Senologie* 2013; 10: 12–14
- 30 Puliti D, Duffy SW, Miccinesi G et al.; EUROSCREEN Working Group. Overdiagnosis in mammographic screening for breast cancer in Europe: a literature review. *J Med Screen* 2012; 19 (Suppl. 1): 42–56
- 31 Jørgensen KJ, Gøtzsche PC. Overdiagnosis in publicly organised mammography screening programmes: systematic review of incidence trends. *BMJ* 2009; 339: b2587
- 32 Zahl PH, Gøtzsche PC, Mæhlen J. Natural history of breast cancers detected in the Swedish mammography screening programme: a cohort study. *Lancet Oncol* 2011; 12: 1118–1124
- 33 Njor SH, Olsen AH, Blichert-Toft M et al. Overdiagnosis in screening mammography in Denmark: population based cohort study. *BMJ* 2013; 346: f1064
- 34 Falk RS, Hofvind S, Skaane P et al. Overdiagnosis among women attending a population-based mammography screening program. *Int J Cancer* 2013; 133: 705–712
- 35 Duffy SW, Lynge E, Jonsson H et al. Complexities in the estimation of overdiagnosis in breast cancer screening. *Br J Cancer* 2008; 99: 1176–1178
- 36 Skaane P, Bandos AI, Gullien R et al. Prospective trial comparing full-field digital mammography (FFDM) versus combined FFDM and tomosynthesis in a population-based screening programme using independent double reading with arbitration. *Eur Radiol* 2013; 23: 2061–2071
- 37 Ciatto S, Houssami N, Bernardi D et al. Integration of 3D digital mammography with tomosynthesis for population breast-cancer screening (STORM): a prospective comparison study. *Lancet Oncol* 2013; 14: 583–589
- 38 Berg WA, Blume JD, Cormack JB et al. Combined screening with ultrasound and mammography vs. mammography alone in women at elevated risk of breast cancer. ACRIN 6666 Investigators. *JAMA* 2008; 299: 2151–2163; Erratum in: *JAMA* 2010; 303: 1482
- 39 Kuhl CK, Schrading S, Weigel S et al. The “EVA” trial: evaluation of the efficacy of diagnostic methods (mammography, ultrasound, MRI) in the secondary and tertiary prevention of familial breast cancer. Preliminary results after the first half of the study period. *Rofo* 2005; 177: 818–827