MARIA JOSÉ AFONSO TEODÓSIO BENTO

EVALUATION OF BREAST CANCER SCREENING IN THE NORTH OF PORTUGAL

Tese de Candidatura ao grau de Doutor em Ciências Médicas submetida ao Instituto de Ciências Biomédicas Abel Salazar da Universidade do Porto.

Orientador – Doutor Vítor Rodrigues

Categoria – Professor Associado

Afiliação – Faculdade de Medicina da Universidade de Coimbra

Co-orientador – Doutor Guilherme Gonçalves

Categoria – Professor Associado com Agregação

Afiliação – Instituto de Ciências Biomédicas Abel Salazar da Universidade do Porto

De acordo com o disposto no n.º 1 do artigo 34.º do Decreto-Lei n.º 74/2006, publicado em Diário da República, 1.ª série, n.º 60 de 24 de Março de 2006, e republicado pelo Decreto-Lei n.º 115/2013, publicado em Diário da República, 1.ª série, n.º 151 de 7 de Agosto de 2013, que procede à terceira alteração ao Decreto-Lei n.º 74/2006, de 24 de março de 2006, constam nesta tese os artigos já publicados e os submetidos para publicação, que a seguir se discriminam:

I. Castro C, Bento MJ, Lunet N, Campos P. Assessing the completeness of cancer registration using suboptimal death certificate information. Eur J Cancer Prev 2012; 21(5):478-9.

II. Crocetti E, Caldarella A, Ferretti S, Ardanaz E, Arveux P, Bara S, Barrios E, Bento MJ, Bordoni A, Buzzoni C, Candela G, Colombani F, Delafosse P, Federico M, Francart J, Giacomin A, Grosclaude P, Guizard AV, Izarzugaza I, Konzelmann I, La Rosa F, Lapotre B, Leone N, Ligier K, Mangone L, Marcos-Gragera R, Martinez R, Michelena MJ, Michiara M, Miranda A, Molinié F, Mugarza-Gomez C, Paci E, Piffer S, Puig-Vives M, Sacchettini C, Sánchez MJ, Traina A, Tretarre B, Tumino R, Van Vaerenbergh E, Velten M, Woronoff AS. Consistency and inconsistency in testing biomarkers in breast cancer. A GRELL study in cut-off variability in the Romance language countries. Breast. 2013:22(4):476-81.

III. Giordano L, von Karsa L, Tomatis M, Majek O, de Wolf C, Lancucki L, Hofvind S, Nyström L, Segnan N, Ponti A; Eunice Working Group, Van Hal G, Martens P, Májek O, Danes J, von Euler-Chelpin M, Aasmaa A, Anttila A, Becker N, Péntek Z, Budai A, Mádai S, Fitzpatrick P, Mooney T, Zappa M, Ventura L, Scharpantgen A, Hofvind S, Seroczynski P, Morais A, Rodrigues V, Bento MJ, Gomes de Carvalho J, Natal C, Prieto M, Sánchez-Contador Escudero C, Zubizarreta Alberti R, Fernández Llanes SB, Ascunce N, Ederra Sanza M, Sarriugarte Irigoien G, Salas Trejo D, Ibáñez Cabanell J, Wiege M, Ohlsson G, Törnberg S, Korzeniewska M, de Wolf C, Fracheboud J, Patnick J J, Lancucki L, Ducarroz S, Suonio E. Mammographic screening programmes in Europe: organization, coverage and participation. J Med Screen. 2012;19 Suppl 1:72-82.

IV. Hofvind S, Ponti A, Patnick J, Ascunce N, Njor S, Broeders M, Giordano L, Frigerio A, Törnberg S; EUNICE Project and Euroscreen Working Groups, Van Hal G, Martens P, Májek O, Danes J, von Euler-Chelpin M, Aasmaa A, Anttila A, Becker N, Péntek Z, Budai A, Mádai S, Fitzpatrick P, Mooney T, Zappa M, Ventura L, Scharpantgen A, Hofvind S, Seroczynski P, Morais A, Rodrigues V, Bento MJ, Gomes de Carvalho J, Natal C, Prieto M, Sánchez-Contador Escudero C, Zubizarreta Alberti R, Fernández Llanes SB, Ascunce N, Ederra Sanza M, Sarriugarte Irigoien G, Salas Trejo D, Ibáñez Cabanell J, Wiege M, Ohlsson G, Törnberg S, Korzeniewska M, de Wolf C, Fracheboud J, Patnick J, Lancucki L, Ducarroz S, Suonio E. False-positive results in mammographic screening for breast cancer in Europe: a literature review and survey of service screening programmes. J Med Screen. 2012;19 Suppl 1:57-66.

V. Bento MJ, Gonçalves G, Aguiar A, Castro C, Veloso V, Rodrigues V. Performance indicators evaluation of the population-based breast cancer screening programme in Northern Portugal using the European Guidelines (submitted).

VI. Bento MJ, Gonçalves G, Aguiar A, Antunes L, Veloso V, Rodrigues V. Clinicopathological differences between interval and screen-detected breast cancers diagnosed within a screening programme in the North of Portugal. J Med Screen 2014;21(2):104-109.

VII. Bento MJ, Gonçalves G, Aguiar A, Antunes L, Castro C, Veloso V, Rodrigues V. Clinicopathological characteristics of invasive breast cancers diagnosed in participants, non-participants and not invited to the organized population-based Breast Cancer Screening, in the North of Portugal (submitted).

Contents

Abstract

Abstract

The aim of a population-based breast cancer screening programme (BCSP) is to lower the burden of the disease in a population. As a public health intervention it needs to be evaluated and when specific mortality analysis cannot be used, the evaluation of performance and impact indicators is an alternative, as they can be evaluated shortly after the screening implementation. Performance analyses include statistics of key monitoring data of cancer screening (as coverage rate, participation rate, and recall rate) and early impact indicators include the analysis of the clinicopathological characteristics of the detected breast cancers and interval cancer rate, among others.

The aim of the research reported in this thesis was to evaluate the population-based BCSP implemented in the Northern Region of Portugal.

The existence of a population-based cancer registry with high completeness is an important pre-condition for an accurate evaluation of many indicators of the screening programme. A specific study was conducted to assess the completeness of the case ascertainment at the North Region Cancer Registry (RORENO). Results obtained warranted the conclusion that RORENO is a valuable source of information on the new cases of breast cancer diagnosed in the population.

Another important issue in the evaluation of screening is related to the molecular biomarkers profile of the breast cancer cases. A study was conducted in Romance language countries, and although there was high prevalence of biomarker testing, the variability of categorical labelling stressed the need of a more extensive use of the existing guidelines.

A first insight on the performance of the BCSP of the Northern Region was done in the framework of two international collaborative studies, within the project of European Network for Information on Cancer (EUNICE). Parameters as organization, coverage, participation rate and false-positive rate were studied. The BCSP showed similar coverage by invitation, higher coverage by examination and a lower rate of further assessment, compared to the overall results of the participating screening programmes.

The monitoring of the BCSP of the Northern Region through comparison of the performance and impact indicators with the standard European Guidelines assessed the quality of screening and provided means to predict the mortality outcome. The main results indicated that the BCSP was being highly accepted by the population, that it was detecting the expected number of invasive breast cancers in the prevalent and subsequent rounds and was able to identify small size breast cancers among the participants. Most of the performance and impact indicators evaluated were consistent with the desirable levels of the European Guidelines and with other international mammography screening programmes.

Interval breast cancers are an important indicator of the quality of mammography as well as of the probable impact of the screening programmes on breast cancer. In this research,

the detection rates of interval cancers during the first and in the second year of the screening interval were in accordance with the desirable levels of the European Guidelines; the clinicopathological characteristics of the interval cancers showed higher size, higher grade and less frequent oestrogen receptor positivity in these cancers when compared to screen-detected cancers. This pattern of more aggressive characteristics found in interval cancers was also described in the majority of the international studies.

From a public health perspective it was necessary to evaluate the impact of the screening programme on the whole population (not only the screened women) and on routine health-care settings. The assessment of the screening experience of the women resident in the district of Bragança was compared to a contemporaneous population of women resident in Vila Real and not exposed to the organized screening programme. It was noteworthy that in screen-detected breast cancers the maximum dimension of the invasive tumour was smaller, and significantly different from the cancers detected in non-participant or not invited women to the organized screening.

The main conclusions from this thesis are:

- The organized population-based BCSP implemented in the Northern Region of Portugal from 2000 to 2009 provided a high quality service and it is foreseeable a mortality reduction due to breast cancer among the population covered by the programme.

- The programme should be expanded to cover all eligible women in the Northern Region.

- Assessment is a never ending process; the work produced in this study is in progress.

Future research should evaluate the impact of the Northern Region BCSP in the mortality of breast cancer. Cohort and case-referent studies can be used to fulfil this objective. Other possible studies should address the prognostic factors associated with screen-detected cancers and its impact on treatment and survival rates, and the effectiveness of breast cancer screening in younger women.

Resumo

Resumo

O objetivo de um programa de rastreio do cancro da mama de base populacional (PRCM) é reduzir a carga da doença na população. Como intervenção de saúde pública precisa de ser avaliada e, quando a análise da mortalidade específica não pode ser utilizada, a avaliação de indicadores de desempenho e de impacto é uma alternativa, uma vez que podem ser avaliados logo após a implementação do rastreio. Análises de desempenho incluem estatísticas de dados fundamentais de monitorização do rastreio do cancro (como taxa de cobertura, taxa de participação e taxa de chamada para aferição) e os indicadores iniciais de impacto incluem a análise das características clínico-patológicas dos cancros da mama detetados e taxa de cancro de intervalo, entre outros.

O objetivo da investigação nesta tese foi avaliar o PRCM de base populacional implementado na Região Norte de Portugal.

A existência de um registo de cancro de base populacional com elevada exaustividade na deteção de casos é uma pré-condição importante para uma avaliação precisa de muitos indicadores do programa de rastreio. Um estudo específico foi realizado para avaliar a exaustividade na identificação dos casos no Registo Oncológico Regional do Norte (RORENO). Os resultados obtidos suportam a conclusão de que o RORENO é uma fonte valiosa de informações sobre os novos casos de cancro da mama diagnosticados na população.

Outra questão importante na avaliação do rastreio está relacionada com o perfil dos biomarcadores moleculares dos casos de cancro de mama. Um estudo foi realizado nos países de língua Latina, e embora tenha havido elevada prevalência de biomarcadores testados, a variabilidade na categorização dos resultados revelou a necessidade de uma utilização mais ampla das orientações existentes.

A primeira análise sobre o desempenho do PRCM da Região Norte foi feita no âmbito de dois estudos colaborativos internacionais, enquadrados no projeto da Rede Europeia de Informação sobre o Cancro (EUNICE). Foram estudados parâmetros como organização, cobertura, taxa de participação e taxa de falsos-positivos. Os resultados do PRCM revelaram, em relação aos resultados globais dos programas de rastreio participantes, uma taxa de cobertura por convite semelhante, maior cobertura por exame e uma menor taxa de aferição complementar.

A monitorização do PRCM da Região Norte, através da comparação dos indicadores de desempenho e de impacto com as Normas Europeias, avaliou a qualidade do rastreio e forneceu informação sobre um previsível impacto na mortalidade. Os resultados principais indicaram que o PRCM teve grande aceitação na população, detetou o número esperado de cancros invasores da mama nas voltas de rastreio prevalente e subsequente e entre as mulheres participantes, identificou cancros da mama de tamanho pequeno. A maioria dos

indicadores de desempenho e de impacto avaliada foi concordante com os níveis desejáveis das Normas Europeias e com outros programas internacionais de rastreio por mamografia.

Os cancros da mama de Intervalo são um importante indicador da qualidade da mamografia, bem como do impacto provável dos programas de rastreio sobre o cancro da mama. Neste estudo, as taxas de deteção de cancros durante o primeiro e o segundo ano do intervalo entre voltas de rastreio estavam de acordo com os níveis desejáveis das Normas Europeias; as características clínico-patológicas dos cancros de intervalo revelaram que estes tinham maior tamanho, grau mais elevado e positividade dos recetores de estrogénio menos frequente quando comparados com os cancros detetados por rastreio. Este padrão de características mais agressivas encontrado nos cancros de intervalo também foi descrito na maior parte dos estudos internacionais.

Numa perspetiva de saúde pública, foi necessário avaliar o impacto do programa de rastreio em toda a população (não só nas mulheres rastreadas) e em condições de rotina dos cuidados de saúde. A experiência de rastreio das mulheres residentes no distrito de Bragança foi comparada com a de uma população contemporânea de mulheres residentes em Vila Real e não expostas ao programa de rastreio organizado. De notar, que nos cancros da mama detetados por rastreio a dimensão máxima do tumor invasivo foi menor, e significativamente diferente da dimensão dos cancros detetados em mulheres não-participantes ou não convidadas para o rastreio organizado.

As principais conclusões desta tese são:

- O PRCM organizado e de base populacional implementado na Região Norte de Portugal de 2000 a 2009 prestou um serviço de alta qualidade e é previsível uma redução da mortalidade por cancro da mama entre a população abrangida.

- O programa deverá ser expandido para abranger todas as mulheres elegíveis na Região Norte.

- A avaliação é um processo que nunca termina; o trabalho produzido neste estudo continua em andamento.

Investigações futuras deverão avaliar o impacto do PRCM da Região Norte na mortalidade por cancro de mama. Os estudos de coortes e “case-referent” podem ser utilizados para atingir este objetivo. Outros estudos possíveis deverão abordar os fatores de prognóstico associados a cancros detetados no rastreio e impacto sobre o tratamento e as taxas de sobrevivência, bem como a efetividade do rastreio do cancro da mama em mulheres mais jovens.

Abbreviations

ARS-N	Administração Regional de Saúde do Norte
BCSP	Breast Cancer Screening Programme
CI	Confidence interval
CrI	Credible interval for meta-analysis results
DCIS	Ductal carcinoma *in situ*
ER	Oestrogen receptor
EUNICE	European Network for Information on Cancer
GRELL	Grupo de Registos e de Epidemiologia nos Países de Língua Latina
HER2	Human epidermal growth factor type 2 receptor
IARC	International Agency for Research on Cancer
IBM	Incidence based mortality
LPCC	Liga Portuguesa Contra o Cancro
NNI	Number needed to be invited
NNS	Number needed to be screened
OECD	Organisation for Economic Co-operation and Development
OR	Odds ratio
PR	Progesterone receptor
RR	Relative risk
RORENO	Registo Oncológico Regional do Norte
WHO	World Health Organization

I – Background

I – Background

1. Epidemiology and clinicopathogical characteristics of breast cancer

1.1 Incidence and mortality worldwide and in Portugal

Breast cancer is the most common cancer in women in many regions of the world, including Australia, the western part of Asia, North Africa, Western Europe, North America and parts of South America. In 2012, it is estimated that 1,67 million new cases had occurred, representing a standardized (world standard population) rate of 42.3/10^5, a quarter of all cancers diagnosed in women. The highest rates were reported in Switzerland, the white population of the USA, Italy and other European countries and the lowest were in Africa.[1-4] There are large differences between the incidence rates recorded in developed countries (except Japan) all above 80/10^5 and the rates found in most developing countries, with rates below 40/10^5. Breast cancer is also the leading cause of death in women worldwide. In 2012 around 522,000 women were estimated to have died from this cause. The distribution pattern of mortality is similar to the distribution of incidence,[1-3] however, the range in mortality rates between developed regions and those in developing countries, is less than that of incidence (between 6-20/10^5).[3]

In Portugal, the number of new cases of breast cancer in 2008 was estimated at 5,333 which corresponded to 27.6% of all cancers diagnosed in women, or just over 1 in 4 cases of cancer were attributed to breast cancer. The age-standardized incidence rate stood at 60.0/10^5 (World Standard Population).[3] The estimated prevalence of breast cancer cases at 5 years was 21,272.[5]

Regarding mortality from this cause, in Portugal 1,661 women died in 2010 and the absolute numbers have increased since 1955, mainly due to increased life expectancy and aging of the female population, and changes in lifestyle.[6] However, the risk of dying from breast cancer for women aged 35-74 years increased 1.55%/year between 1955 and 1992, and changed -2.20%/year from 1992 to 2002.[7]

1.2 Incidence of breast cancer in the Northern Region of Portugal

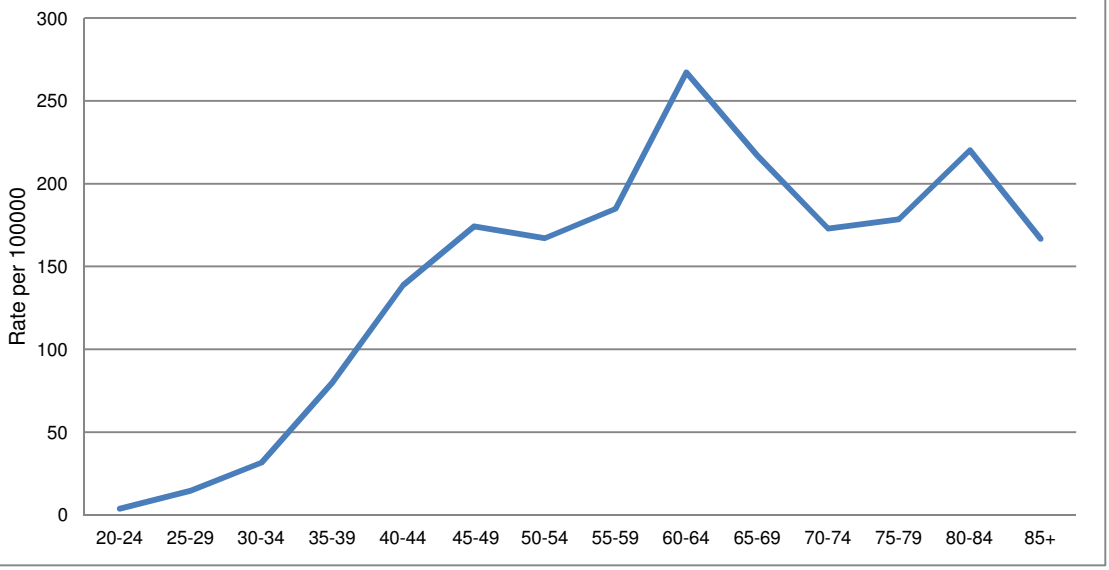
According to the North Region Cancer Registry of Portugal (*Registo Oncológico Regional do Norte* – RORENO) 1,764 new invasive breast cancer cases were diagnosed in 2008, corresponding to an incidence rate of 103.6/10^5. This represented more than a quarter of all invasive cancers diagnosed amongst women that year.[8]

The age-standardized incidence rate (World Standard Population) in the North Region was 66.4/10^5, a higher value than the one estimated for the whole country.[3]

1.3 Risk factors for breast cancer

Just being a woman and getting older are the main risk factors to have breast cancer.[9] Men can also develop breast cancer, but it represents approximately 1% of the total frequency in women.[10] Breast cancer rises sharply with age.[9] According to data from RORENO, in 2008, the overall incidence rate of breast cancer was low at younger ages ($3.8/10^5$ in age group 20-24 years), after which incidence rates more than double in the next 5-year age groups till age 44 (Figure 1). In the following ages, the increase was attenuated and a peak rate was achieved in the 60-64 age-group, with an incidence of $267.3/10^5$.

Figure 1 Breast cancer incidence in women resident in the Northern Region of Portugal, in 2008, by age-group



Family history is a major risk factor, especially if the family member was diagnosed with the disease at a young age. A woman whose first-degree relative had breast cancer before age 40, has 6 times more risk of developing breast cancer before age 40, compared with a women of the same age but with no family history.[9]

The BRCA1 and BRCA 2 genes have been related to the occurrence of familial breast cancer, manifesting in premenopausal women, but on the whole they don´t contribute to more than 10% of cases of breast cancer.

Other risk factors were linked to some benign breast pathologies, breast density, life style, diet, reproductive and hormonal factors.

Although many risk factors are already identified, only a few can be changed or prevented. Besides, most women who do get breast cancer don't have any significant risk factors (other than being a woman and growing older).

1.4 Histopathology and prognosis

In general, there are three categories of breast abnormalities: benign conditions, in-situ and invasive cancer.[11]

Benign conditions are associated with a risk for breast cancer ranging from one- to fivefold, depending on the degree of epithelial proliferation and atypia.

Breast cancer is probably an heterogeneous group of diseases with more than one natural history.[11] More than 95% of breast cancers originate from the epithelial elements of the mammary gland, particularly from the cells of the terminal ductal lobular units of the breast.[12] The lobular carcinoma *in situ* is associated with an increased risk for invasive breast cancer but is usually an incidental finding and is not generally detected by mammography. Data on the natural history of ductal carcinoma *in situ* (DCIS) are limited, but it is likely that high-grade carcinomas are associated with a significantly higher risk for development of invasive carcinoma than low-grade DCIS. Weather DCIS is an obligate precursor to invasive ductal cancer, or if both entities derive from a common progenitor cell line is unclear.[13]

In relation to invasive carcinomas, 75% to 80% are infiltrating ductal carcinomas and 5-10% lobular carcinomas. Other types of breast cancer include the medullary, mucinous, tubular and other less frequent tumours.[12]

The prognosis of a patient with breast cancer is associated with time-dependent variables (tumour size, presence and extent of lymph nodes metastasis, and distant metastasis, the three variables defining the TNM stage[14]) and is also associated with variables related to the biology of the individual tumour, as the histological grade[15] tumour type, hormone receptor status, among others.[11] These features of breast cancers are fundamental for predicting response to treatment and overall outcome.[12]

Tumour size is one of the strongest predictors of breast cancer behaviour and studies have shown that screening leads to the detection of smaller size tumours.[16-18] The 5-year survival for patients with node-negative disease is 82.8%, but there is a direct relationship between the number of involved axillary nodes and the risk for distant recurrence and shorter survival.[18-20] Grade utility for staging and prognostication became more relevant in the screening era as a higher proportion of tumours are T1N0M0 at diagnosis.[15] The 10-year survival for ductal carcinomas was 76% for women with grade 1 carcinoma and 39% for those with grade 3 tumours.[21]

Biological markers became essential for prognosis definition and therapy. Patients with carcinomas with oestrogen (ER) and/or progesterone receptor (PR) positive have a better survival than hormone receptor negative tumours with a 5-year overall survival (all stages) of 83% in the ER+/PR+ group versus 69% in the double negatives.[22]

1.5 Treatment and survival

Treatment of breast cancer commonly encompasses a combination of treatments as surgery, chemotherapy, radiation therapy, and hormone therapy. It depends on factors such as stage of the disease at diagnosis, histological grade, age, co-morbidities and the women´s preferences. Some pathological characteristics as ER and PR status and the human epidermal growth factor type 2 receptor (HER2) influence the use of target therapies.[23] Breast conserving surgery is the treatment of choice for the majority of small sized tumours.[24] Locally advanced breast cancers are treated with a combined modality including upfront chemotherapy, surgery and radiation.[24]

In the last decades, significant improvements were registered in the survival rates for women with breast cancer.[4] Five-year survival rates over 80% were verified in North America, Europe and Australia. Nowadays, women with breast cancer have a higher survival rate than for most other types of cancer.[4,25] The 10-year survival rate for breast cancer in most western populations reached 70%.[26]

2. Breast cancer screening

2.1 Principles of breast cancer screening

The objective of screening for a disease is to discover those among the apparently well who are in fact suffering from the disease.[27] In 1968, Wilson and Junger[27] established general principles of screening for the World Health Organization (WHO), that are still valid today.[28] These principles of screening can be summarized[28] as:

- Screening should be directed towards an important health problem
- There should be a simple, safe, precise and validated screening test
- Treatment started at an early stage should be of more benefit than treatment initiated later
- There should be evidence that the screening test is effective in reducing mortality and morbidity
- The benefit of screening should outweigh the physical and psychological harm caused by the test, diagnostic procedures and treatment
- The opportunity cost of the screening programme should be economically balanced in relation to expenditure on medical care as a whole
- There should be a plan for managing and monitoring the screening programme and an agreed set of quality assurance standards
- Potential screening participants should receive adequate information about benefits and disadvantages of participation

In general, breast cancer screening is consistent with these principles:

- Breast cancer is the commonest cause of death from cancer in women in many regions of the world.
- Breast cancer is a progressive pathology and its development can be hindered by early detection.[29,30] The importance of early detection is evident upon examination of the strong association between stage at diagnosis and survival.[30,31]
- Systematic examination with mammography has the potential to lower breast cancer mortality rates in approximately 20% in women invited to screening[29,32] and to reduce the burden of the disease in the population.[33] Mammography screening is the only screening method that has proven to be effective.[34]
- Sensitivity of the mammographic examination to detect malignant lesions, as reported by IARC in 2002,[11] ranged from 52% to 82% and specificity was higher than 90%; the predictive positive value ranged from 2%-22%, with most studies reporting this value as 12%.[11]
- There is sufficient evidence[13] that inviting women 50-69 years of age to screening reduces their mortality from breast cancer.
- Treatment is considered to be more effective if cancer is detected at earlier stages.[35]
- After randomized trials screening programmes for women aged 50-69 at a 2- or 3-year interval are expected to be cost-effective in high-incidence countries with well organized programmes.[11]
- The benefits of breast cancer screening on mortality reduction outweigh the harms associated with screening, especially the risk of overdiagnosis.[29]

Current data indicate that mammographic screening may not be effective in all age groups and the ratio of benefit to cost varies significantly with age.[35,36] The sensitivity of mammography is inversely proportional to breast density,[37,38] and younger women tend to have higher density.[36]

Also the natural history of the disease still has many unanswered questions, namely the malignant potential of DCIS to become invasive, ie, to consider DCIS as a marker of malignancy requiring active treatment or a benign condition of no clinical significance.[29]

2.2 Organization and components of breast cancer screening

According to International Agency for Research on Cancer (IARC) an organized cancer screening has 6 characteristics:[39] a written policy specifying the target age categories, the method of screening and interval; a defined target population; a management team that is responsible for overseeing facilities where screening occurs and for ensuring that the target population is screened; a clear decision structure and responsibility for health care

management; a quality assurance structure and a method for identifying cancer occurrence in the target population.[39]

Population-based breast cancer screening by mammographic examination is a multi-step process.[40] Hakama[41] proposed four main components of cancer screening programmes that can be applied to breast cancer screening:

1. Population component

- definition of target population

- identification of individuals: unique personal identifiers are required to compile the full information of a woman over multistep screening episodes, and to link this information to other data sources in health-care.[42]

- measures to achieve sufficient coverage and attendance: use of an individual letter of invitation to screen and reminders to attend was found to increase access and attendance.[39] Invitation gives each eligible person an equal chance to benefit from screening and therefore reduces health inequalities.[43]

2. Test execution

- test facilities for mammographic examination and analysis

- quality control programme for obtaining mammography and its analysis: this component extends from the technical quality control of the radiologic equipment, procedures, and the radiologists performance. Mammographic examination of the breast with two views is likely to increase sensitivity by approximately 20%.[41] Double reading of the mammography by two experienced radiologists compared to single reading increases the detection rate by 10% but lowers specificity.[41] In case of discrepancy, consensus or arbitration by a third reader should be decided.

3. Clinical components

- facilities for diagnosis, treatment and follow-up of patients with screen-detected disease: a rapid referral and diagnostic evaluation by multi-disciplinary team to avoid unnecessary delay should be provided

4. Coordination

- a referral system linking the screen, screening unit and clinical facility

- monitoring, quality control and evaluation of the programme; follow-up of incidence and mortality in the entire target population, and for both attenders and non-attenders.

Major organizational considerations are the ages at which the programme starts and stops and the interval at which the test is applied.[41,44] Women 50-69 years of age seem to be the ones that benefit most from screening mammography.[45,46] Data are limited regarding the effects of screening mammography in women that are 70 years of age or older[32], and there is uncertainty over the value of screening women between ages of 40 to 49.[32,47] Most of the

screening programmes have adopted a 2-yearly screening because of the high interval cancer rates seen in the third year in trials.[11]

Breast cancer screening can be conducted outside the organized programme when it is known as “opportunistic screening”.[39] This type of screening refers to activities that involve referral to mammography facilities by clinicians and self-referral by women themselves.[43] Compared to an organized screening programme, the opportunistic screening usually results in increased costs as there is no inbuilt mechanism to prevent unnecessarily frequent screens and uncontrolled adverse effects.[44]

2.3 Potential bias in breast cancer screening

There are three main types of bias that can suggest benefit from screening when there is none. These biases are the self-selection, lead-time, and length bias.[36]

The self-selection bias occurs when a group of individuals comes forward to be screened.[36] Volunteers usually are more health-conscious and they are more likely to have a better outcome than the general population.[36,48,49]

Lead-time bias is related to the period of time between the detection of the cancer by screening and the time when the cancer would have been diagnosed clinically.[28] It constitutes an artificial addition to the survival time of screen detected cancer cases.[49,50] Although a women diagnosed through screening may spend more time aware of the existence of her breast cancer, the date of her death might well remain unaltered.[48]

The length bias occurs when slow-growing, less aggressive cancers are detected during screening.[36] The probability of a cancer being detected at screening depends on the length of time the lesion is detectable in a preclinical phase, the so called sojourn time.[11] The probability of a cancer to be screen-detected is greater when sojourn time is longer. Within a screening programme an unwarranted proportion of cancers detected will have a longer sojourn time and probably a better prognosis.[11,36,50]

2.4 Quality guidelines in breast cancer screening

According with the Council of the European Union recommendation launched in 2003, screening for cancer should be based on a well-organized population-based approach using systematic quality assurance at all appropriate levels.[51] Implementation of breast cancer screening of high quality has the potential to not only lower the burden of disease in the population attending screening but also on the quality and effectiveness of symptomatic, ie, usual care.[43]

Evaluation of breast cancer screening programmes involves analyses of performance and impact. Performance analyses include statistics of key monitoring data of cancer

screening.[24,42] Reduction in disease-specific mortality, being the primary purpose of screening, is the outcome of choice for studies of effectiveness.

2.4.1 Performance indicators

Performance of the screening programmes should be continuously monitored and compared with short-term indicators and standards. Comprehensive multidisciplinary guidelines for quality assurance in breast cancer screening and diagnosis were developed by experts and published by the European Commission. The first edition was published in 1993[52] and the fourth and latest edition was available in 2006.[24,43,52-55] The publication of "European Guidelines for Quality Assurance in Breast Cancer Screening and Diagnosis"[24] became an internationally recognized reference for best practice screening and diagnosis and later also a multidisciplinary management of breast cancer.[43]

There is a multitude of possible process indicators. Some of these indicators that are of importance epidemiologically, are listed in table 1, together with the acceptable and desirable levels, as defined by the European guidelines.[24] For reasons of comparability and in accordance with European policy, data should be reported separately for the 50-69 age group.[24]

Table 1 – Performance indicators of a breast screening programme and acceptable and desirable levels

Performance indicator		Acceptable level	Desirable level
Participation rate		> 70 %	> 75%
Recall rate	Initial screening	< 7%	< 5%
	Subsequent screening	< 5%	< 3%
Benign to malignant biopsy ratio		≤ 1:2	≤ 1:4
Elegible women reinvited within the specified screening interval		> 98%	100%

2.4.2 Impact indicators

The best impact indicator for an organized breast cancer screening programme is mortality due to breast cancer. This measure does not suffer from important bias as lead-time or length bias that can seriously affect other kind of measures, as survival analysis: screened women could have better survival only because diagnosis was performed earlier and not because of an improved prognosis.

The first studies on the impact of mammography screening were the randomized controlled trials.[29,56] After the trials, the impact of the screening service was evaluated by comparison of the breast cancer mortality trends, either by geographically regions or over

time. Breast screening has also been evaluated by comparison of breast cancer mortality in women invited/screened to not invited/unscreened women.

Because the use of mortality to evaluate breast cancer screening takes more than 10 years before an effect can be expected,[57] an attractive alternative is to identify early surrogate indicators.[24] Table 2 contains some of these indicators and the acceptable and desirable levels according to the European Guidelines.[24]

Table 2 – Early surrogate indicators of the impact of a breast screening programme and acceptable and desirable levels*

Surrogate indicator		Acceptable level	Desirable level
Interval cancer rate/background incidence rate (%)	0-11 months 12-23 months	30%	< 30%
Detection rate	Initial screening	3xIR	> 3xIR
	Subsequent screening	1.5xIR	> 1.5xIR
Stage II+/total cancers SD	Initial screening	NA	< 30%
	Subsequent screening	25%	< 25%
Invasive cancers ≤ 10 mm/total invasive cancers SD (%)	Initial screening	NA	≥ 25%
	Subsequent screening	≥ 25%	≥ 30%
Invasive cancers/total cancers SD (%)		90%	80-90%
Node-negative cancers/total invasive cancers SD (%)	Initial screening	NA	> 70%
	Subsequent screening	75%	> 75%

IR - background incidence; NA – not applicable; SD – screen-detected; *adapted from "European guidelines for quality assurance in breast cancer screening and diagnosis"[24]

2.5 Monitoring and evaluation studies

2.5.1 Performance indicators and early surrogate indicators of mortality

Within the framework of the European Network for Information on Cancer (EUNICE) screening activity of 18 European countries and 26 programmes was evaluated, mostly for the year 2005. A wide geographical variation in the coverage by examination was evident, and half the programmes indicated an acceptable (> 70%, according to the EU guidelines) uptake; however, differences between programmes were more than three-fold. Lower examination coverage was found mostly in screening programmes recently implemented and when rollout was not yet completed.[58,59]

2.5.2 Randomized controlled trials

Randomized controlled trials provide the most reliable evidence on the effects of screening.[29] Since 1963, several trials have been undertaken to determine the effectiveness

of screening using mammography in decreasing breast cancer mortality.[29,32] These studies were conducted mainly in North America, Scandinavia and United Kingdom.

Some of these trials were used in a recent meta-analysis[29] made by The Independent United Kingdom Panel on Breast Cancer Screening, to calculate the Relative Risk (RR) of breast cancer mortality in screened women compared to a control group. Authors concluded that for women aged 50-70 years invited for screening every 3 years to undergo mammography, a 20% reduction in breast cancer mortality at ages 55-79 was likely (assuming that women who began screening at 50 would gain no benefit in the first five years, but that the mortality reduction would continue for 10 years after screening ended). Additionally, they calculated the number of women needed to be invited (NNI) for screening for 20 years at age 50 to prevent one breast cancer death, and it was 235. The number of women needed to be screened (NNS) to prevent one breast cancer death, was 180.[29] Although other authors[60-62] had fiercely criticized some of these trials, the authors from that Independent Panel considered that problems and biases detected were unlikely to have had a major distorting effect on overall result.

Another systematic evidence review was published in 2009 to update recommendations from the United States Preventive Task Force.[32] For women aged 50-59 years, trials provided a pooled RR of 0.86 (credible interval - CrI, 0.75 to 0.99) and for women aged 60-69 years, the RR was 0.68 (CrI, 0.54-0.87). The NNI to prevent one breast cancer death was 1339 (CrI, 322-7455) for women aged 50-59, and the NNI for the older ones was 377 (CrI, 230-1050). Using the same scenario applied in the UK Independent Panel, the NNS to prevent one death over ages 55-69 years was 193.[63]

Other meta-analyses gave different results, mainly due to diverse methodological approaches, but in general, they all pointed towards a mortality reduction associated with breast cancer screening.[57,60,64]

2.5.3 Observational studies

After the results from the randomized controlled trials, many population-based and organized breast cancer screening programmes were launched in several countries.[65,66] From a public health perspective it was necessary to evaluate the impact of the screening programme in the whole population (not only the screened women) and in real life settings. The observational studies became the principal source of information on the impact of a breast cancer screening programme in a population.

As a whole, the observational evidence showed that breast cancer screening for women aged 50 to 69 years reduced mortality from breast cancer.[67]

Cohort studies

A cohort study published in 2012, compared attenders to the Florentine Screening Programme with no attenders. The estimated mortality reduction in attenders was 45% for women aged 50-59 years, and raised to 51% among 60-69 years old attenders.[68]

Incidence based mortality studies (IBM) compare breast cancer mortality in patients with breast cancer diagnosed during similar periods before (pre-screening period) and after (screening period) screening introduction.[69] In 2012 a systematic review was conducted and 20 studies on incidence-based mortality were included. The reported reduction in breast cancer mortality was 25% (RR 0.75, 95% CI 0.69-0.81) among invited women and 38% (RR 0.62, 95% CI 0.56-0.69) among those actually screened.[70]

Case-control studies

Case-control studies are a traditional tool to evaluate the effectiveness of screening.[71] In 2012 a published systematic review evaluated 8 case-control studies reported in the period of 2004 till 2012; breast cancer mortality reduction was 48% (odds ratio (OR) 0.52, 95% CI 0.42-0.65) for screened versus non-screened women and 31% (OR 0.69, 95% CI 0.57-0.83) for invited versus not invited women.[66] This type of observational studies have some reported potential biases such as the self-selection bias, which can alter the magnitude of the mortality reduction; this value also depends on factors as the attendance rate, the screening organization and in the quality of treatment, among others.[66,72]

Ecological studies

This kind of study is frequently done due to the common availability of the data but as it has important methodological constraints, conclusions should be carefully drawn.[65,71] Usually the mortality rates for breast cancer were compared before and after the starting of the screening programme, or the trends of breast cancer mortality were compared between regions with contrasting screening policies. The majority of studies suggested reductions in breast cancer mortality ranging from 28-36% as a result of mammographic screening or reductions from 1%-9% per year in studies reporting an annual percent change.[73] However, these ecological studies didn´t provide reliable evidence and were considered of limited value for screening assessment.[29,73]

2.6 Potential harms associated with breast cancer screening

2.6.1 Overdiagnosis (and overtreatment)

Mammography preferentially detects indolent tumours because they are detectable for longer periods. Length bias occurs because indolent tumours are less likely to be lethal.[74] Overdiagnosis is considered an extreme form of length bias and represents the major harm

of screening.[29] The consequences of overdiagnosis are substancial, as unnecessary treatment (including toxicity and treatment associated morbidity) of people with inconsequential disease, adverse psychological effects of labelling as a cancer patient, quality of life adversely affected, and the costs incurred.[29,36,75]

Overdiagnosed cancers tend to be more likely DCIS (possibly low/intermediate rather than high grade) and invasive tumours exhibiting grade 1 or 2 rather than grade 3.[29]

In an article on overdiagnosis, 13 observational studies reported estimates ranging from 1% till 10%.[76] Estimates were adjusted for the natural changes in breast cancer incidence and lead time bias.[76]

The UK panel considered that overdiagnosis was 19% during the screening period and 11% was the best estimate for a screened population above the long-term expected incidence in the absence of screening.[29] Authors recognized the uncertainty on this estimation,[77] as they were able to show that by varying assumptions and regression methods, overdiagnosis estimates varied from trivial to the alarming.[29]

Overdiagnosis was also reported to vary with the age of the women and screening round: younger cases showed lower values,[78] and prevalent screening registered higher rates of overdiagnosis compared to subsequent screening.[76]

Jørgensen and Gotzsche in 2009 published a work reporting an estimated rate of overdiagnosis of 52%;[79] a possible reason for such result was an underestimation of the expected incidence of breast cancer in the screening period and lack of fully adjustment for lead time.[63,76,80] In studies that have individual data and/or take in consideration the underlying incidence trends and lead time, reported overdiagnosis ranges between 0-10%.[63]

2.6.2 Interval cancers

According to the European Guidelines, interval cancers are those that occur after a negative mammogram and before the scheduled date for the next examination.[24]

Interval cancers are an important indicator of the quality of mammography as well as of the likely impact of the screening programmes on breast cancer.[81,82] As they are unlikely to be eliminated,[83,84] they should be reduced and the European Guidelines established a limit of 30% proportional incidence in the first year after a negative examination and 50% in the second year.[24]

European Guidelines recommend reviewing the interval cancers as an essential part of routine radiological audit and proposed the following categories for classification of interval cancers: true interval, occult, minimal signs, false negative and unclassifiable[24] False negative cases should not exceed 20% of the total number of interval cancers.[24]

Mammography is difficult to interpret. In order to classify the mammography, retrospective reviews have been performed but their design strongly influences the number

of interval cancers that are classified as missed.[83-85] In a study in Norway, the missed interval-detected cancers ranged from 1% to 36%, depending on the methodology.[86] For this reason some authors questioned the usefulness of retrospective reviews to separate the interval cancer categories,[87,88] while others considered reviews an important and necessary tool for continuous education of radiologists and a way to improve the quality of radiological skills in breast cancer screening[38,89-91]

Interval cancers are related with the programme sensitivity and with the length of the inter-screening interval.[92-94] In a pooled analysis of eight breast cancer screening programmes in six European countries,[95] that shared common performance aspects as the age-group of the participating women (mostly from 50 to 69 years), two-view mammography, independent double reading, 24 months interval and linkage to cancer registries with high completeness, the reported sensitivities varied between 67% and 84% (72%, in total). The overall interval cancer rate (including invasive and *in situ* cancers) was 18.5/100000, with large differences between programmes: 8.4/10000 the lowest rate till 21.3/10000, the highest value. These differences were reported to depend on (besides technical skills and the sensitivity of the test) the intensity of opportunistic screening, the recall rate and background incidence rate.

Other factors such as younger age,[96,97] increased mammographic breast density,[36,82,97-99] use of hormone replacement therapy,[36,96] lobular histology[94,100,101] were associated with higher difficulty to detect cancer in the mammography, lowering sensitivity and increasing incidence of interval cancers.[96,99] Also, fast-growing tumours with shorter sojourn time than the screening interval will be often interval-detected cancers.[102]

The predominance of interval cancers of poor prognosis compared to screen-detected cancers[87,94,101,103] may indicate failure of the screening programme to detect cancers at an earlier stage. Studies conducted to evaluate the prognostic characteristics of interval cancers compared to breast cancers detected in women who did not participate in the screening revealed contradictory results.[87,92,94,100,104-107]

Survival in women with interval breast cancer is worse than in screen-detected cases.[88,102,103,108] Compared to women not participating or not invited to screening, survival of women with interval cancer cases showed conflicting results, with some studies finding significant differences in favour of interval cancers,[102] while others didn´t find any difference.[88]

2.6.3 False-positive results

A false-positive test is considered the most common (though not most serious) adverse effect of mammographic screening. [48,109] Psychological distress, harmful diagnostic follow-up

and economic burden for the women and healthcare system, are some reasons of concern after a false-positive result.[45]

The estimated cumulative risk of a false-positive recall for assessment in women aged 50-69 years undergoing 10 biennial screening tests varied between 8% and 21% (pooled estimate 19.7%).[110] It´s inversely related to age, as younger ages have a tendency for more dense breasts and lower incidence of the disease.[45]

A study in Spain that included 762,506 women aged 45 to 69 years, observed an increased risk of cancer detection in women with a previous false-positive test in mammographic screening.[111]

2.6.4 Other harm factors associated with breast cancer screening

Risk of cancer associated with radiation, pain during procedures, anxiety and other psychological distress associated with a false-positive screening result[32,36] are some additional side effects related to screening with mammography.

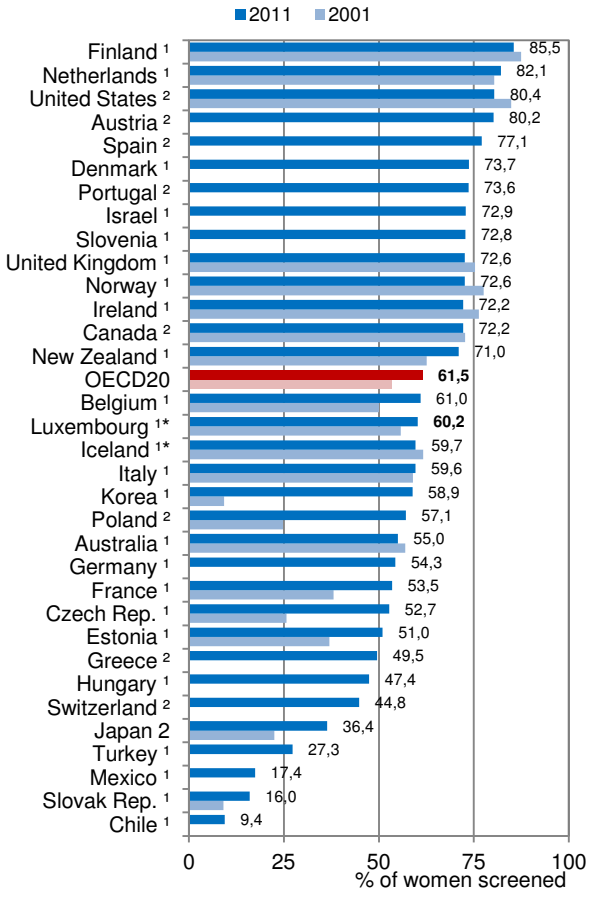
2.7 Organized breast cancer screening programmes in the world

Screening programmes are organized regionally or nationally, and most of them target women 50 to 69 years old, with a 2-year interval between screening tests.[41] According to the Organisation for Economic Co-operation and Development (OECD)[112] screening rates varied widely across countries in 2011, ranging from less than 10% in Chile to over 80% in Finland, the Netherlands, the Unites States and Austria - Fig 2. Some countries that had high screening rates in the past, experienced some reductions over the last decade, including Finland, the United States, the United Kingdom, Norway, Ireland and Canada.

2.8 Breast cancer screening programmes in Portugal

In Portugal, the organized population-based breast cancer screening was initiated in 1990 in the Centre Region, conducted by the Portuguese Cancer League (*Liga Portuguesa Contra o Cancro* – LPCC); women aged 45 years or over were invited to participate.[113] Screening procedures included single-view mammograms, centralized and independently read by two radiologists, with a final reading by a third independent and experienced radiologist, in case of discrepancy. Positive results were assessed at the Portuguese Oncology Institute of Coimbra and all diagnoses and cancer treatment procedures followed standardized therapeutic protocols. Double view mammograms were gradually introduced until the late ‘90s, first for age group 45-49 at initial screening, then for all age groups and afterwards, subsequent screening was also included; it was performed in every mobile and fixed unit. The age-group for invitation was settled at 45-69 years.

Fig. 2 - Mammography screening in women aged 50- 69, 2001 to 2011 (or nearest year)



1. Programme. 2. Survey. * Three-year average.

Information on data for Israel: http://dx.doi.org/10.1787/888932315602.

Source: OECD Health Statistics 2013, http://dx.doi.org/10.1787/health-data-en.

In Portugal the organized BCSP was implemented throughout the country in a staggered way and over a long period of time. After the start in the Centre Region in 1990, it was implemented in the Southern Region in 1997, Northern Region and Madeira in 1999, Algarve in 2005[114] and in Açores the programme started in 2009.[115] Since the beginning of the screening programme, quality assurance was considered a priority, in accordance with the European Guidelines.[24]

In the Northern Region of Portugal the population-based BCSP is conducted by the north branch of LPCC and coordinated by the North Regional Health Administration (*Administração Regional de Saúde do Norte* – ARS-N);[116] it started in October 1999 in one municipality and it gradually expanded to 65 municipalities by the end of 2012 (76% of the

Region). In that year, 90659 women were submitted to a mammography, which represents an overall participation rate of 61%.[117] The screening programme is organized with eight mobile units and one fixed facility (in 2012). The database/information system from BCSP is settled at the LPCC and stores data elements on demographic characteristics, screening services and results, diagnostic tests and cancer information.

The programme includes a structured individual invitation to all resident women aged 45-69 years using the lists of users enrolled in the Health Centres. From the beginning, screening procedures included bilateral mammography with two-view (craniocaudal, mediolateral oblique), centralized and independently read by two radiologists, with a final reading by a third independent and experienced radiologist, in case of discrepancy. The rating scale applied to classify the radiological findings follows the European Guidelines;[116] this classification is widely used in European screening programmes.[24] Assessment of screen positive mammography is carried out by a multidisciplinary team (radiologist, surgeon and pathologist) at a dedicated clinical setting outside the hospital, and cancer treatment is established according to standardized hospital therapeutic protocols. Screen film mammography was used till mid-2007, and thereafter, computer-aided mammography was performed in all screening units. The screening interval is 24 months.

In the Northern Region, opportunistic breast cancer screening coexists with the organized programme. The precise estimative of the extent of opportunistic screening in the Region is unknown, but according to ARS-N in 2009, 54% of the women aged 45 to 69 years, using the primary health centres had undergone a mammography.[115,118] In 2007 and 2008, 16% of the women that were referred to breast screening outside the organized programme had to repeat the mammography and 59% had an additional ultrasound examination.[119]

2.9 Importance of population-based cancer registries in the planning and evaluation of breast cancer screening programmes

The main objective of a breast cancer screening programme is to decrease mortality from this cause. However, as this benefit can only be assessed after 7-10 years of programme operation, short term indicators were developed that allow an earlier assessment of the impact of screening.[81] In order to perform this evaluation, it is necessary a close collaboration between the screening programme and the population-based cancer registry. As it is clear from the considerations made in the following paragraphs, quality of a cancer registry is important for an appropriate evaluation of a screening programme.[120]

The cancer registry covers the entire female population of a given region, providing information on all women either they accept or not to participate in breast cancer screening; population-based cancer registries have an important role in the evaluation of the impact of screening on the entire target population.[24,92]

The key functions of a population-based cancer registry in the assessment of breast cancer screening programmes are the estimated breast cancer incidence, identification of interval cancers, monitoring of tumour characteristics and comparison of survival rates between groups.[81,92]

Population-based cancer registries provide information on breast cancer incidence rates in a certain region, supporting cancer control planning activities. One of the most important contributions of a cancer registry to the screening evaluation is to provide reliable data for the analysis of incidence trend of breast cancer after the implementation of a screening programme and the comparison between the observed and the expected incidence rates.[81,121]

The cancer registry is essential for identifying interval cancers.[92] It is necessary to ensure that the registry has a high case ascertainment and cases are not missed, or bias might be introduced. Due to time-lag in the recording of cancer cases in a registry, some underestimation of interval cancer rates in the most recent years is unavoidable. However, it is recommended to use only this source of information in order to guarantee the comparability of data, namely stage or other prognostic information.[81,122]

Tumour size, type, grade and lymph node status, have the greatest impact on the prognosis of breast cancer. It is very important to compare the stage of tumours diagnosed over a screening cycle and those diagnosed in an unscreened population as stage is regarded as one of the early indicators of the effectiveness of cancer screening.[92]
Unfortunately, stage comparison of tumours reported from different institutions and over time has been found to vary significantly so the use of histological size as a proxy to stage is recommended.[81] In order to confirm the effectiveness of the screen to detect small, early stage lesions, the different categories of detection should be considered as screen-detected cancers, cancers in non-attenders, interval cancers, cancers in lapsed attenders and cancers in women not yet invited.[92,104]

2.9.1 North Region Cancer Registry of Portugal (RORENO)

RORENO is a population-based cancer registry, established in 1988 by governmental initiative and located at the Portuguese Oncology Institute, in Porto. It constitutes the main source of information on cancer burden for the Northern Portugal and it covers the area of five districts: Braga, Bragança, Porto, Viana do Castelo e Vila Real. In 2008, the estimated number of residents in the area was 3,294,709.

New cases are collected from the public and private hospitals and pathology laboratories. All invasive cancers are registered. Information on the patient demographic data and on cancer is included in the registry computerised database. Cancer variables include date of

diagnosis, topography and morphology, behaviour and grade, stage and treatment. RORENO uses passive and active follow-up activities to obtain survival.

In order to achieve high quality of cancer reporting and completeness, some routine audits are performed as well as computer checks of consistency, warning programmes, training courses for registrars and medical staff, and especially designed research.

In 2008, the number of received notifications was 26,703 which corresponded to 16,935 new cancer cases diagnosed in residents in the Northern Region.

II – Aims

II – Aims

Every public health intervention should be evaluated.[123] Breast cancer screening is a public health intervention planned to lower the burden of the disease in a population.[43] As explained before performance and impact indicators are used to evaluate a population-based Breast Cancer Screening Programme (BCSP). Though specific mortality is considered the best impact indicator, it was explained that often it cannot be used. Thus, early surrogate impact and performance indicators are often used as an alternative.

The aim of the research reported in this thesis was to evaluate the population-based BCSP implemented in the Northern Region of Portugal. For that purpose, some available surrogate impact and performance indicators were used. Furthermore, some preconditions had to be analysed, in order to assess the validity of data used to build the above mentioned indicators. Thus the specific objectives of this research were:

1 – To evaluate the completeness of cancer registry (RORENO) and biomarkers classification, preconditions necessary for the assessment of population-based breast screening programmes.

2 – To evaluate performance indicators of the BCSP in comparison with other population-based BCSP within the European Network for Information on Cancer project.

3 – To evaluate the BCSP, using the standards of performance and impact indicators, recommended by the European Guidelines.

4 – To compare the clinicopathological characteristics of the interval breast cancers with the screen-detected cancers.

5 – To evaluate the clinicopathological characteristics of breast cancers detected in a population invited to an organized BCSP compared to the tumour characteristics of a non-invited population

III – Results

1. Important preconditions for breast cancer screening evaluation

III – Results

1. Important preconditions for breast cancer screening evaluation

As it was explained in point 2.9 (Background) the quality of a cancer registry is important for an appropriate evaluation of a screening programme.[120]

Completeness is one of the key points of that quality and a specific study was conducted to assess it.

Cancer cases with high survival, as breast cancer, are considered to be more easily identified by the registry than cases with high fatality rates. For this reason, it was decided to evaluate the completeness of RORENO using gastric cancer, a cancer with low survival and more likely to remain undiagnosed or untreated. The study was conducted in 2012, and included all gastric cancer cases diagnosed during 2001-2006 in the district of Porto. Three different quantitative methods were used: capture-recapture, death certificate and mortality/incidence, and the flow-method. Results provided by the three methods were similar and the overall estimates for completeness of gastric cancer registration ranged between 82.9% and 96.4% (**paper I**)

Another important issue in the evaluation of screening is related to the molecular biomarkers profile reported in breast cancer cases and its prognostic and predictive significance (as reported in point 1.4 of the Background section). Standardization of methods and results is required especially when comparing with previous studies.

In the framework of the Association for Cancer Registries and Epidemiology in Romance Language Countries (Grupo de Registos e de Epidemiologia nos Países de Língua Latina – GRELL) a study was conducted to evaluate methods for testing and cut-offs of oestrogen (ER) and progesterone (PR) receptors, human epidermal growth factor type 2 receptor (HER2) and proliferation index. Analysis across and within countries were performed (**paper II**).

A questionnaire was employed to collect data. RORENO was invited and participated with a sample of breast cancer cases diagnosed in 2007. Data was retrieved from the pathology reports.

This study highlighted the high prevalence of biomarkers testing reported by the 34 participating cancer registries; nevertheless, it was of concern the variability of categorical labelling of ER/PR, HER2 or markers of proliferation activity, found among and within countries.

This raised some questions about the positivity labelling of biomarkers in breast cancers, consequent prognosis definition and treatment modality given to women with breast cancer.

Letter to the Editor

European Journal of Cancer Prevention 2012, 21:478–479

Assessing the completeness of cancer registration using suboptimal death certificate information

Clara Castro[a,c,d], Maria José Bento[a], Nuno Lunet[b,c] and Pedro Campos[d],
[a]North Region Cancer Registry (RORENO), [b]Department of Clinical Epidemiology, Predictive Medicine and Public Health, Faculty of Medicine, University of Porto, [c]Institute of Public Health of the University of Porto (ISPUP) and [d]LIAAD INESC Porto, L.A. and FEP, Porto, Portugal

Correspondence to Dr Clara Castro, Portuguese Oncology Institute, Porto. Rua Dr António Bernardino de Almeida, 4200-072 Porto, Portugal
Tel: +351 225 084 067; e-mail: clara.castro@ipoporto.min-saude.pt

Received 16 September 2011 **Accepted** 21 November 2011

The completeness of cancer registration depends on the probability of cases being identified by the registries at different time after diagnosis (Montanaro *et al.*, 2006), and therefore varies with the type of cancer, and between and within registries (Shin *et al.*, 2009). The understanding of the cancer-specific patterns of completeness in each registry is essential for a proper interpretation of the measures of cancer burden, and for a valid appraisal of its variation with time and across registries.

As cases of cancer with a high survival may be more easily identified by the registries, those with high case fatality rates are more likely to remain undiagnosed or untreated (Regal and Hook, 1991), providing a good model to address the ability of cancer registries to achieve high levels of completeness. The survival of gastric cancer patients is low (Crew and Neugut, 2006), and the northern region of Portugal [North Region Cancer Registry (RORENO), 2010] presents the highest gastric cancer incidence and mortality rates in Western Europe (Lunet *et al.*, 2004). Therefore, we estimated the completeness of gastric cancer registration in the north of Portugal (specifically in the district of Porto), during the time period 2001–2006. We compared the estimates obtained from three different quantitative methods – 'capture–recapture' (Robles *et al.*, 1988), 'death certificate (DC) and mortality/incidence' (Ajiki *et al.*, 1998; Parkin *et al.*, 1994; Parkin and Bray, 2009) and the 'flow method' (Bullard *et al.*, 2000; Montanaro *et al.*, 2006). As DCs are not routinely available to cancer registries in Portugal, and no information on all the potential causes mentioned in the DC as immediate or contributory causes or significant conditions contributing to death could be obtained, we further addressed the implications of using suboptimal DC information (underlying cause of death instead of a mention of cancer in any field of the DC).

We analysed data from the RORENO, on the incident cases of gastric cancer ($n = 3787$) occurring in the district of Porto between 2001 and 2006. The vital status at the end of 2006 was assessed by linkage with the National Health System Database, which includes up to date information on the citizens' vital status, but no information on the cause of death. A list of the underlying cause of death of all the individuals dying between 2001 and 2006, and living in Porto at the time of their death, was obtained from the Northern Health Sub-Region (ARS-Norte) for manual inspection.

Completeness of registration was obtained by estimating the number of gastric cancer cases that were not registered by RORENO, including those that were alive by the end of 2006 as well as those who were dead before the end of 2006 but with gastric cancer not mentioned as the underlying cause of death. We considered two alternative scenarios in data analysis: first, assuming that the information on the underlying cause of death allowed the identification of all DC with a mention of gastric cancer, which is expected to yield a lower bound estimate of the completeness; second, assuming that all gastric cancer patients registered before death and known to have died had a mention of gastric cancer in the DC, which is expected to yield an upper bound estimate of the completeness.

The overall estimates for completeness of gastric cancer registration ranged between 82.9 and 96.4% under the worst and best case scenarios, respectively (Table 1). The three quantitative methods yielded similar conclusions, with the most precise estimates being obtained with the capture–recapture method. The flow method showed that no meaningful improvements in completeness can be expected in this setting after a period of 3 years since diagnosis.

The present work is the first application of quantitative methods in the assessment of completeness of cancer registration in Portugal, and provides valuable information to support improvement of the procedures towards a more effective case ascertainment. We were able to show the applicability of these methods in a setting with suboptimal DC information and the impact of the lack of direct access to complete information from DC by the cancer registry. However, we could only provide lower and upper bound estimates of completeness, and this can hardly be considered as an option to monitor this quality indicator. Furthermore, the collection of the necessary information was too labour intensive; therefore, the routine use of the quantitative methods is not an option until a timely and efficient use of DC information for the purpose of cancer registration is possible.

0959-8278

DOI: 10.1097/CEJ.0b013e32834f811c

Table 1 Completeness estimates[a], assuming that the information on the underlying cause of death identifies all death certificates with a mention of gastric cancer (first scenario) and that all gastric cancer patients registered before death and known to have died have a mention of gastric cancer in the death certificate (second scenario)

	Men		Women	
Method	Completeness (%)	95% CI[b]	Completeness (%)	95% CI[b]
First scenario				
Capture–recapture				
Model M_t[c]	81.5	(78.3; 84.2)	84.5	(82.0; 87.4)
DC and M/I ratio				
Ajiki *et al.* (1998)	81.4	–	84.7	–
Parkin *et al.* (1994)	85.4	–	88.0	–
Flow method				
1 year	71.5	(57.7; 83.5)	83.6	(58.2; 98.1)
2 years	76.4	(63.0; 87.5)	85.1	(63.2; 97.9)
3 years	77.2	(64.5; 87.7)	85.8	(67.8; 97.1)
4 years	77.8	(65.7; 87.9)	89.5	(71.7; 98.9)
5 years	78.0	(66.1; 87.9)	89.6	(73.2; 98.6)
Second scenario				
Capture–recapture				
Model M_t[c]	96.7	(95.8; 97.6)	94.4	(92.9; 96.7)
DC and M/I ratio				
Ajiki *et al.* (1998)	96.7	–	94.3	–
Parkin *et al.* (1994)	97.1	–	95.2	–
Flow method				
1 year	82.3	(69.9; 91.9)	88.7	(67.2; 99.2)
2 years	88.8	(78.9; 95.8)	91.2	(74.1; 99.4)
3 years	90.3	(81.3; 96.5)	92.7	(77.5; 99.6)
4 years	91.6	(83.4; 97.1)	96.6	(81.2; 100.0)
5 years	92.1	(84.5; 97.3)	96.9	(82.6; 100.0)

CI, confidence interval; DC, death certificate; M/I, mortality/incidence.
[a]Death certificates (DC) stating gastric cancer as the underlying cause of death were identified for 1885 individuals, from which 348 were Death Certificate Only cases. Among the cases registered by North Region Cancer Registry (RORENO) before the inclusion of information from DC, 2024 were reported dead before the end of 2006 in the National Health System Database, from which 1182 had a DC with gastric cancer as the underlying cause of death.
[b]95% confidence interval.
[c]Under the assumption that RORENO and DC sources are not independent, by using a log-linear model with a temporal effect to overcome the dependency between sources.

Despite these limitations, our results quantify the potential for improvement in cancer registration completeness in this setting, and define the minimum lag to be respected between diagnosis and the publication of valid incidence estimates.

Acknowledgements

The authors wish to thank the Studies and Planning Department from the ARS-Norte for providing access to the DC information, and Dr Fabio Montanaro and Dr David Robinson, for making Stata ado-files available for the flow method and for the valuable comments on its application.

Conflicts of interest

There are no conflicts of interest.

References

Ajiki W, Tsukuma H, Oshima A (1998). Index for evaluating completeness of registration in population-based cancer registries and estimation of registration rate at the Osaka Cancer Registry between 1966 and 1992 using this index. *Nippon Koshu Eisei Zasshi* **45**:1011–1017.

Bullard J, Coleman MP, Robinson D, Lutz JM, Bell J, Peto J (2000). Completeness of cancer registration: a new method for routine use. *Br J Cancer* **82**: 1111–1116.

Crew KD, Neugut AI (2006). Epidemiology of gastric cancer. *World J Gastroenterol* **12**:354–362.

Lunet N, Pina F, Barros H (2004). Regional trends in Portuguese gastric cancer mortality (1984–1999). *Eur J Cancer Prev* **13**:271–275.

Montanaro F, Robinson D, Bordoni A, Lutz JM (2006). A modification to the flow method to estimate completeness in cancer registries with delayed registration. *J Public Health (Oxf)* **28**:274–277.

North Region Cancer Registry (RORENO) (2010). *Registo Oncológico Regional do Norte 2006*. Porto, Portugal: Instituto Português de Oncologia do Porto.

Parkin DM, Bray F (2009). Evaluation of data quality in the cancer registry: principles and methods Part II. Completeness. *Eur J Cancer* **45**:756–764.

Parkin DM, Chen VW, Ferlay J, Galceran J, Storm HH, Whelan SL (1994). *Comparability and Quality Control in Cancer Registration (IARC Technical Reports No 19)*. Lyon: IARC.

Regal RR, Hook EB (1991). The effects of model selection on confidence intervals for the size of a closed population. *Stat Med* **10**:717–721.

Robles SC, Marrett LD, Clarke EA, Risch HA (1988). An application of capture–recapture methods to the estimation of completeness of cancer registration. *J Clin Epidemiol* **41**:495–501.

Shin HR, Curado MP, Ferlay J, Heanue M, Edwards B, Storm H (2009). Comparability and quality of data. In: Curado MP, Edwards B, Shin HR, Storm H, Ferlay J, Heanue M, Boyle. P, editors. *Cancer Incidence in Five Continents*. Vol. IX. (IARC Scientific Publication, No. 160). Lyon, France: IARC and IACR. pp. 67–94.

The Breast 22 (2013) 476–481

Contents lists available at SciVerse ScienceDirect

The Breast

journal homepage: www.elsevier.com/brst

Original article

Consistency and inconsistency in testing biomarkers in breast cancer. A GRELL study in cut-off variability in the Romance language countries☆

Emanuele Crocetti [a], Adele Caldarella [a], Stefano Ferretti [b], Eva Ardanaz [c], Patrick Arveux [d], Simona Bara [e], Enrique Barrios [f], Maria J. Bento [g], Andrea Bordoni [h], Carlotta Buzzoni [a,*], Giuseppina Candela [i], Françoise Colombani [j], Patricia Delafosse [k], Massimo Federico [l], Julie Francart [m], Adriano Giacomin [n], Pascale Grosclaude [o], Anne V. Guizard [p], Isabel Izarzugaza [q], Isabelle Konzelmann [r], Francesco La Rosa [s], Benedicte Lapotre [e], Nathalie Leone [t], Karine Ligier [u], Lucia Mangone [v], R. Marcos-Gragera [w], Ruth Martinez [q], Maria J. Michelena [x], Maria Michiara [y], Ana Miranda [z], Florence Molinié [aa], Conception Mugarza-Gomez [x], Eugenio Paci [a], Silvano Piffer [ab], Montserrat Puig-Vives [w], Claudio Sacchettini [a], Maria J. Sánchez [ac], Adele Traina [ad], Brigitte Tretarre [ae], Rosario Tumino [af], Elke Van Vaerenbergh [m], Michel Velten [ag], Anne S. Woronoff [ah]

[a] *Tuscany Cancer Registry, Clinical and Descriptive Epidemiology Unit, ISPO Via delle Oblate 2, 50141 Florence, Italy*
[b] *Ferrara Cancer Registry, Italy*
[c] *Navarra Cancer Registry, Navarra Public Health Institute, Pamplona, Spain*
[d] *Breast and Gyno Côte-d'Or Cancer Registry, France*
[e] *Somme Cancer Registry, France*
[f] *Uruguay Cancer Registry, Uruguay*
[g] *North region Cancer Registry of Portugal (RORENO), Portugal*
[h] *Ticino Cancer Registry, Switzerland*
[i] *Trapani Cancer Registry, Italy*
[j] *Registre général des cancers de la Gironde, France*
[k] *Isère Cancer Registry, France*
[l] *Modena Cancer Registry, Italy*
[m] *Belgium Cancer Registry, Belgium*
[n] *Biella Cancer Registry, Italy*
[o] *Tarn Cancer Registry, France*
[p] *Calvados Cancer Registry, France*
[q] *Basque Country Cancer Registry, Spain*
[r] *Registre valaisan des tumeurs, Valais, Switzerland*
[s] *Umbria Cancer Registry, Italy*
[t] *Limousin Cancer Registry, France*
[u] *Registre general des cancers de Lille et de sa region, Lille, France*
[v] *Reggio Emilia Cancer Registry, Italy*
[w] *Epidemiology Unit and Girona Cancer Registry, Oncology Coordination Plan, Department of Health and Catalan Institute of Oncology, Girona, Spain*
[x] *Onkologikoa Cancer Registry, Spain*
[y] *Parma Cancer Registry, Italy*
[z] *Lisbon Cancer Registry, Portugal*
[aa] *Registre des tumeurs de Loire-Atlantique et Vendée, France*
[ab] *Trento Cancer Registry, Italy*
[ac] *Granada Cancer Registry, Andalusia School of Public Health, Granada and CIBER Epidemiologia y Salud Pública (CIBERESP), Spain*
[ad] *Palermo Breast Cancer Registry, Italy*
[ae] *Hérault Cancer Registry, France*
[af] *Cancer Registry and Histopathology Unit, "Civile M.P. Arezzo" Hospital, ASP 7 Ragusa, Italy*
[ag] *Bas-Rhin Cancer Registry, France*
[ah] *Doubs Cancer Registry, France*

☆ Association for cancer registration and epidemiology in Romance language countries.

* Corresponding author. Tel.: +39 (0)557972518; fax: +39 (0)557972535.
E-mail address: c.buzzoni@ispo.toscana.it (C. Buzzoni).

0960-9776/$ – see front matter
http://dx.doi.org/10.1016/j.breast.2013.04.012

ARTICLE INFO

Article history:
Received 20 November 2012
Received in revised form
27 February 2013
Accepted 3 April 2013

Keywords:
Biomarkers
Breast cancer
HER2
Oestrogen receptor
Progesterone receptor
Proliferation index

ABSTRACT

Purpose: Biological markers are crucial factors in order to differentiate female breast cancers and to determine the right therapy. This study aims at evaluating whether testing for biomarkers for female breast cancer has similar frequency and characteristics across and within countries.
Methods: Population-based cancer registries of the Association for cancer registration and epidemiology in Romance language countries (GRELL) were asked to complete a questionnaire on biomarkers testing. The data collected referred to invasive female breast cancer cases diagnosed between 2004 and 2009. The investigation focused on 1) the overexpression and amplification of the human epidermal growth factor receptor 2 oncogene (HER2); 2) the expression of oestrogen (ER) and progesterone (PgR) receptors; and 3) the proliferation index (PI). Weighted percentages, the heterogeneity among and within countries, and the correlation between responses and calendar years were evaluated. The study was based on 19,644 breast cancers.
Results: Overall, 85.9% of the cases were tested for HER2, 91.8% for both ER and PgR, and 74.1% for proliferative markers. For HER2 and ER−PgR, the frequency of testing increased from 2004 to 2009. Testing varied among countries (HER2 from 82.0% to 95.9%, ER−PgR from 89.3% to 98.9%, PI from 10% to 92%) and also within the same country (e.g. HER2 in Italy from 51% to 99%) as well as within single cancer registries. The most relevant differences were in the scores for positive/negative/not clearly defined HER2 (e.g. HER2 was defined positive if IHC 3+ in 21/33 registries), and in the cut-off of positive cells for ER/PgR (from >0% to >30%) and PI positivity (from >0% to >20%).
Conclusions: Biological markers are widely tested in the Romance language countries; however, the parameters defining their positivity may vary, raising concerns about homogeneity in breast cancer classification and treatment.

Introduction

In the last decade, studies based on variation's patterns in gene expression derived from c-DNA microarrays have produced a new classification for breast cancer with sound predictive and prognostic value.[1–4] New breast cancer subgroups have been identified on the basis of 1) distinct molecular features of mammary epithelial biology corresponding to phenotypes with different expression of oestrogen (ER) and progesterone (PgR) receptors, 2) overexpression and amplification of human epidermal growth factor receptor 2 oncogene (HER2), and 3) proliferation index (PI). The assessment of luminal and basal cell has helped to establish two main categories, and several clusters of different lesions have been proposed, e.g. luminal A (ER+ and/or PgR+, HER2−, PI < 14), luminal B HER2− (ER+ and/or PgR+, HER2−, PI ≥ 14), luminal B HER2+ (ER+ and/or PgR+, HER2+), basal/triple negative (ER−, PgR−, HER2−) and HER2+ (ER−, PgR−, HER2+).[5] The identification of different types and subtypes is in the making.[6–8] These types of breast cancers, stage by stage, show different prognoses[4,7,9–15] and different sensitivities to treatment.[16–19]

Since the introduction of Trastuzumab in 1998, two aspects of the assessment of HER2 status have been debated — the significance of score 2+ in the Hercep Test (Dako Denmark A/S, Glostrup, Denmark), and the use of immunohistochemistry or in situ hybridization.[20] In 2007, the ASCO/CAP guidelines defined IHC 3+ as HER2 positive (FISH more than 6.0 HER2 gene copies per nucleus or FISH ratio of more than 2.2), IHC 0 &1+ as negative (FISH less than 4.0 HER2 copies per nucleus or FISH ratio less than 1.8) and IHC 2+ as not clearly defined (FISH ratio 1.8–2.2 or HER2 gene copy 4.0–6.0).[21] More recently, the ASCO/CAP guidelines addressed the issue of ER/PgR immunohistochemistry testing, and set the cut-off for invasive tumour cell positivity: positive if ≥1%, and negative if <1%.[9]

Data on biomarker expression are usually available in pathology reports. These reports are one of the basic sources for population-based cancer registries (CRs).[22] In addition to the traditional data garnered by CRs (date of incidence, site, morphology and the behavior of tumours), CRs have also collected for some years the stage at diagnosis, the type of therapy,[23] and, many of them, also the ER/PgR status[24] and information on treatment modalities. The collection of HER2 data has started only recently,[25] however, population-based studies on breast cancer biomarker classification are increasingly available.[10,26] Some of them address also survival.[14,27,28]

Recent papers show evidence of the clinical relevance of biomarkers in the standardization of their testing: the definition of positive or negative results for HER2[7,29,30] and, for hormonal receptors, the standard for ER/PgR positivity[9] and the issue of intensity or threshold.[31] Without standardization, therefore, the determination of the cancer type, and the selection of the therapy, might vary from one laboratory to the other. This study aims at carrying out a survey on the variability in biomarker testing for invasive female breast cancer in national or regional CRs included in the GRELL (Association for cancer registration and epidemiology in Romance language countries: www.grell-network.org).

Material and methods

We employed an ad hoc questionnaire to collect information on HER2, ER, PgR and PI. The 31 questions included in the questionnaire gathered data about the testing's frequency, the kind of test, the kind of report (numerical scale, proportion of positive, categorical variables [e.g. negative, positive, etc.]), the cut-offs used for positivity, negativity and ambiguous results, and, for HER2, the type of analysis (immunohistochemistry − IHC and/or in situ hybridization − ISH) and the percentage of patients treated with specific monoclonal antibodies (Trastuzumab). The study specifically addressed biomarkers for incident invasive breast cancers in women.

We sent the questionnaire to the CRs in the GRELL area.

The participating CRs were asked to check the information on biomarkers in the pathology reports that they collect when garnering incidence data. Each CR was able to provide information for all the incident cases in one or more years starting from 2004 or for a simple random sample of cases. We chose the lower time bound on the assumption that HER2 testing would have been unusual before 2004. For each CR, we weighted the proportion of patients tested on the number of cases. We evaluated the

correlation between frequency of testing and calendar year with the Pearson coefficient, and tested its statistical significance with the method proposed by Snedecor and Cochran.[32] We assessed the heterogeneity of responses among and within countries by means of a logistic regression considering the number of not tested and tested women, and adjusting for the period of incidence if there was a statistically significant correlation ($p < 0.05$) between frequency of testing and calendar year.

The evaluation of heterogeneity within countries was limited to those countries in which there were responses from 3 or more CRs (for fewer than 3 it would not have been possible to distinguish the role of calendar year from that of the CR).

Results

The questionnaire was sent to 71 CRs in 9 countries. Overall 34 (47.9%) CRs from 7 countries (Belgium, France, Italy, Portugal, Spain, Switzerland, and Uruguay) responded. We drew responses from the garnered data for 19,644 female invasive breast cancers incident between 2004 and 2009. The countries, the CRs, the number of cases and the period relevant to the results are presented in Table 1. Table 2 summarizes the main findings on HER2, oestrogen and progesterone receptors, and proliferation testing index. More specific data on such biomarkers by country are reported in the following three sections and presented in Table 3.

Table 1
Collaborative GRELL study on biomarker testing for incident female invasive breast cancer. Participating population-based cancer registries, analysed years and total number of cases considered.

Country	Population-based cancer Registry	Data from year/s of incidence	N. of cases
Belgium	Belgian	2008	249
France	Bas-Rhin	2008	50
	Calvados	2008	50
	Côte-d'Or	2008	50
	Doubs	2007	50
	Gironde	2008	50
	Hérault	2008	50
	Isère	2008	50
	Lille et région	2008	50
	Limousin	2006	50
	Loire-Atlantique et Vendée	2007	100
	Manche	2008	50
	Somme	2006	50
	Tarn	2007	50
Italy	Biella	2004–2006	544
	Ferrara	2007	394
	Modena	2004–2008	2924
	Palermo-Female breast	2004–2007	2800
	Parma	2004–2007	1276
	Ragusa	2004–2005	330
	Reggio-Emilia	2004–2005	800
	Trapani	2005	230
	Trento	2009	350
	Tuscany	2005	820
	Umbria	2004–2006	1879
Portugal	Lisbon	2008	2480
	RORENO	2007	1100
Spain	Basque	2007	150
	Girona	2005	292
	Granada	2005	292
	Navarra	2006	290
Switzerland	Valais	2008	234
	Ticino	2005–2009	290
Uruguay	National cancer registry	2009	1220
Total			19,644

Table 2
Collaborative GRELL study on biomarker testing for incident female invasive breast cancer. Proportions of cancer registries where women were tested for HER2, ER and PgR, PI and some used cut-offs. Overall and country-specific results.

	Proportion (by countries) of CR in which the test was performed	Number of CR (and country) where test was performed
HER2	85.9%[a]	
HER2 test used in ≥90% of woman in:	France 76.9%	France 10
	Spain 25%	Spain 1
	Italy 18.2%	Italy 2
	Switzerland 50%	Switzerland 1
	Belgium 100%	Belgium 1
	Portugal 50%	Portugal 1
HER2 IHC positive 3+	France 69.2%	France 9
	Spain 50%	Spain 2
	Italy 70%	Italy 7
	Switzerland 100%	Switzerland 2
	Belgium 0%	Belgium 0
	Portugal 0%	Portugal 0
	Uruguay 100%	Uruguay 1
Both ER and PgR	91.8%[a]	
ER+ cut-off equal to ≥1	5.3%	1 (Italy)
ER+ cut-off below <1	10.5%	2 (Italy, Spain)
ER+ cut-off above ≥1	89.5%	16 (Italy, France, Spain, Portugal, Switzerland)
PI	74.1%[a]	
either HER or ER/PR or PI	100%	In all countries and CRs

CR = Cancer registry.
[a] We weighted the proportion of patients tested on the number of cases of each CR.

HER2

Overall 85.9% female invasive breast cancers were tested for HER2. The proportion varied from 82.0% to 93.9% among countries (this information was not available for Uruguay). The proportion of tested women was 82.0% in Italy (range 51.0–99.0%), 92.2% in France (72.3–100%), 93.9% in Portugal (80.0–100%), 92.3% in Switzerland (82.5–94.0%), 89.8% in Spain (88.0–92.0%) and 95.9% in Belgium. HER2 test was used in ≥90% of the woman in the areas of 10/13 CR in France, 1/4 in Spain, 2/11 in Italy, 1/2 in Switzerland, in the Belgian CR, and in 1/2 CR in Portugal.

The frequency of testing increased with calendar year ($p < 0.001$). The frequency of testing, adjusted for period of time, was significantly different across countries ($p < 0.001$) ranging from 82% in Italy to 93.9% in Portugal, but also among CRs in Italy and France.

For 6 out of 33 CRs (information was not available for 1 CR) it was not possible to identify in the pathology reports the antibodies used for IHC testing and, moreover, for 16 CRs the information was available only for some of the pathology departments in the area. The results of IHC tests were presented in different ways among and within countries (percentage of positive cells, categories [negative, positive, not clearly defined], and score [0+…3+]), but they were also available in all CRs as a score. Out of 32 answering CRs some 23 responded that they kept HER2 test results in their database. Our investigation found that the test used for performing in situ hybridization was, most frequently, FISH (28 of 32 CRs with available information), followed by CISH, SISH and multiple tests. The laboratories used 'IHC only' for 83.6% of the women (range 0–100%). This proportion was 76.1% in Italy (0–97.1%), 89.8% in France (44.8–100%), 80.3% in Spain (62.0–94.0%), 83.3% in Switzerland (83.0–84.5%), 100% in Portugal and 71.1% in Belgium. There was a statistically significant variability across countries ($p < 0.001$).

Table 3
Collaborative GRELL study on biomarker testing for incident female invasive breast cancer. Weighted proportion of women tested for HER2, ER, PgR and Ki-67, number of population-based cancer registries within each country using official cut-off for HER2 IHC positive, negative and not clearly defined results, cut-off for the percentage of positive cell for ER receptor and cut-off for Ki-67 positivity.

Country	HER2 tested % (range)	Number of registries for which HER2 IHC positive is 3+	Number of registries for which HER2 IHC not clearly defined is 2+	Number of registries for which HER2 IHC negative is 0–1+	ER and PgR tested (%)	Cut-off % positive cells for ER Positivity (>)	% Tested proliferation activity	Cut-off % positive cells for PI (>)
Italy	82.0 (51.0–99.0) [a] $p < 0.001$	7/10	7/10	8/10	89.3 (75.5–100) [a] $p < 0.001$	0,1,5,9,10,30	81.7 (50.0–100) [a] $p < 0.001$	10,20
Belgium	95.9	0/1	1/1	0/1	95.5	n.a.	47.3	n.a.
France	92.2 (72.3–100) [a] $p < 0.001$	9/13	7/13	9/13	97.3 (87–100) [a] $p = 0.10$	9,10,15	33.5 (9.0–100) [a] $p < 0.001$	15,20
Portugal	93.9 (80.0–100)	0/2	1/2	1/2	92.8 (90.0–99)	10	10	
Spain	89.8 (88.0–92.0) [a] $p = 0.38$	2/4	4/4	4/4	98.9 (92.6–98.0) [a] $p = 0.006$	0,10	49.7 (34.7–79.0) [a] $p < 0.001$	0
Switzerland	92.3 (82.5–94.0)	2/2	2/2	2/2	95.5 (95.0–98.3)	4	92.0	n.a.
Uruguay	n.a.	1/1	1/1	1/1	100	n.a.	n.a.	n.a.
Among countries	[b] $p < 0.001$				[b] $p < 0.001$		[b] $p < 0.001$	

n.a. = not available.
[a] Heterogeneity within country adjusted for period of incidence.
[b] Heterogeneity among countries adjusted for period of incidence.

In contrast, 'ISH only' was used on average for only 4.7% of laboratories (rather unusual, except in one Italian CR where it was used for 80% of the women). The majority of those which used both tests (IHC and ISH) stated that ISH was used in cases of an uncertain IHC result (score 2+).

With regard to the IHC scores considered to be positive, 3 out of 33 CRs did not have such information available, and 9 CRs showed worrisome results – in 7 CRs scores 2+ & 3+ were both considered positive (in Italy, France, Spain and Portugal) and in 2 CRs (Portugal and Belgium) the score for positivity varied in different pathology laboratories. A certain amount of variability was also present among countries for the most recent years, 2008–2009, with varying scores in different laboratories for Portugal and Belgium.

Also IHC scores considered to be negative varied among and within countries. In one Italian CR 0, 1+ & 2+ were considered negative, while there were different scores in different laboratories in four CRs (in Italy, France, Portugal and Belgium). And again for recent years, 2008–2009, there were varying scores in different laboratories in France, Portugal and Belgium.

The IHC scores considered to be not clearly defined were 1+ & 2+ in 2 CRs (Italy and France), with different scores (1+ & 2+, or 2+) in two CRs (Italy and Portugal). Two other CRs did not have clearly defined results (France). There was variability in recent years, too.

Trastuzumab was the specific treatment for about 61.7% of the women positive for HER2 and for 0.2% of those negative for HER2 (in 2 Italian CRs).

Oestrogen and progesterone receptors

Overall, 91.8% of cases were tested for both ER and PgR. The proportion varied from 89.3% to 100% among countries. The frequency of testing increased with calendar year ($p = 0.0029$). Such testing, adjusted for period of time, was significantly different among countries ($p < 0.001$) and also among CRs in Italy and Spain. Only one French CR stated that 'ER only' or 'PgR only' tests were performed while for the other CRs single receptor testing was rather unusual. The identification of antibodies used for testing ER/PgR was possible in the majority of the areas although different antibodies were usually used in different pathology departments.

Also the reports from different laboratories showed great variability — the most common report showed the proportion of positive cells. The positivity cut-off, the minimum percentage of positive cells needed for defining the positivity of receptors, varied very much both among and within countries. It ranged from >0% to >30%. It varied from >0% to >30% in Italy, to >9% to >15% in France, from >0% to >10% in Spain, and it was >4% in Switzerland and >10% in Portugal. For Belgium and Uruguay cut-off were not reported in the questionnaire. There was a slight, but not statistically significant, correlation between the cut-off and the calendar year ($p = 0.082$). Moreover, several CRs responded that the cut-offs for positivity varied within their area among the different laboratories (2 of 8 CRs in Italy, 2 of 8 in France, 2 of 3 in Spain, 2 of 2 in Portugal, 1 of 1 in Belgium).

Assuming ≥1 the cut-off for ER positivity, only 1 Italian CR reports this level in its area, 2 CRs (1 Italian and 1 Spanish) had even a lower limit for positivity (>0), while all the others showed higher cut-offs, Table 2.

Biomarkers for proliferation index

Overall 74.1% of cases were tested for PI. The proportion varied from 10.0% in Portugal to 92.0% in Switzerland. The index was tested mainly (78%) by Ki-67/MIB-1 antibodies. There was no statistical correlation between frequency of testing and calendar year ($p = 0.40$). No information was available for Uruguay and for 2 French, 1 Spanish, 1 Suisse and 1 Portuguese CRs. The proportion of tested women was 81.7% in Italy (range 50–100%), 33.5% in France (9–100%), 49.7% in Spain (34.7–79.0%), 92% in Switzerland, 10% in Portugal and 47.3% in Belgium. There was a statistically significant heterogeneity in the frequency of testing among countries and within Italy, France and Spain. Heterogeneity could not be tested in the other countries.

The most frequent PI report was based on the proportion of positive cells (24 of 30 CRs) but quite frequently (in 6 out of 30 CRs) different forms of reporting results were used in different departments of the same area. Cut-offs for positivity were not used in the areas of 17 of 29 CRs (in Italy, France, Spain and Switzerland), while they were used only in some of the pathology departments of the area, more specifically in 6 out of 29 CRs (in Italy, France, Belgium and Portugal). When the proportion of positive cells was used for defining positivity, different cut-offs were adopted varying from >0% to >20%. Cut-offs were >10% or >20% in Italy, >15% or >20% in France and >0% in Spain.

Finally, 21 CR answered that also other biomarkers were tested, among them the most frequent were CA 15-3, carcinoembryonic antigen, cytokeratin 5/6 and E-Cadherin.

Overall, in all the areas of all CR, women were tested for either HER2 or ER/PgR or KI.

Discussion

Female invasive breast cancer biomarkers were widely tested in patients living in the analysed areas of Belgium, France, Italy, Portugal, Spain, Switzerland and Uruguay. Between 2004 and 2009, the frequency of testing increased for both HER2 and ER/PgR. However, the frequency of testing differed among countries, even taking into account the years of incidence. Moreover, such frequency varied within country for ER/PgR in Italy and Spain, for HER2 in Italy and France, and for proliferation activity in Italy, France and Spain. This latter variability may be due to the lack during the last decade of specific guidelines regarding Ki-67 among the routinarely biological markers. New genetic tests have recently documented the prognostic role of proliferation genes, including Ki-67.[33]

The frequency, methodology, type and thresholds of measures of testing were based on the information that CRs collected from the Pathology Departments. As a consequence, completeness and interpretation might be affected. Some pathology reports may have been misplaced between the CR and a pathology department (either within the CR's area or, for some patients, outside the CR's area). Therefore, the frequency of biomarker testing may have been underestimated, and some of the differences detected among and within countries may, in fact, stand for differences in the data collection's completeness, rather than in the extent of biomarker testing. Interpretation bias is especially related to those pathology reports expressing uncertainty.[34] Moreover, the different periods of time included in the study by the CRs, could influence the use of in situ hybridization in the analysis of HER2 since the use of FISH was later incorporated in the clinical practice for those cases with not clearly defined results. Also, differences in cancer patients' age in the analysed areas may have contributed to some differences in the frequency of testing.

Finally, only half of the invited cancer registries participated in the study. However, this study did not aim at being geographically representative. The target was to find, in the thirty-four areas of the seven selected countries, and in this period, a variability in the frequency and especially in the cut-offs for positivity of several biomarkers for breast cancer.

Even taking into account the potential biases, the high prevalence of biomarker testing is a positive result that should ensure the best treatment for women living in the GRELL countries.[16–18] Also some of the characteristics of the testing procedures were reassuring about the general good quality of standard operating practice: hormonal receptors were generally tested together and given as a proportion of positive cells, KI-67 was given as a proportion of positive cells and, for HER2, the first test was usually IHC, given as a score, while ISH (mainly FISH) was used to confirm not clearly defined results, that were usually, but not always, score 2+. Moreover, the CRs answered that almost all the women treated with Trastuzumab had tested positive for HER2 (but without specifying the use of IHC 3+ score for positivity).

On the other hand, our findings of the variability of categorical labelling are less encouraging. Our results show that the procedure with which an invasive breast cancer patient was labelled positive for HER2, ER/PgR or markers of proliferation activity, varied among and within countries, with a certain variability also in recent years. The crucial point is the huge variability within countries in the cut-offs for HER2 positivity, negativity and not clearly defined results. It was also true for ER/PgR that were considered positive for proportions of positive cells ranging from >0% to >30%, sometimes with different cut-offs in the same areas (Italy, France and Spain). The positivity cut-offs for KI-67 ranged from >0% to >20%. The variability may arise from updated recommendations. For example, thresholds for ER+ of about 10% and 1% were recommended in 2003[18] and 2010,[9] respectively, but this may not explain why different cut-offs were adopted in the same area and period as reported for HER2, ER/PgR and Ki-67.

Nowadays, biological markers have a fundamental role in defining the prognoses of breast cancers and in indicating treatment.[5] In general terms, testing positive for hormonal receptors indicates hormonal treatment endocrine therapy (tamoxifen/aromatase inhibitor), while testing negative for hormonal receptors indicates, when appropriate, chemotherapy. To test positive for the HER2 oncogene, signals sensitivity to the specific biological therapy – Trastuzumab.[19]

Conclusion

In the analysed period, the criteria for female breast cancer biomarker positivity varied among countries, within countries and even within the area of a CR. Therefore, the biological labelling and treatment of female patients may be laboratory dependant and not cancer dependant. The composition of guidelines for female breast cancer biomarker testing[9,21] has, in recent years, addressed the issue, but the guidelines' national and international extent of application has not been studied. The same treatment should, theoretically, be provided to all cancers of the same type. This did not seem to be the case for breast cancers diagnosed in the GRELL countries. This issue is extremely relevant from the perspective of public health bridging, as it does, epidemiologists and clinicians, but if women were to start inquiring whether they did receive the right treatment, the same issue might also carry legal implications for clinicians and pathologists.

Role of the funding source

There are no funding sources for the manuscript.

Conflict of interest statement

We declare that we have no conflict of interest.

References

1. Perou CM, Sørlie T, Eisen MB, van de Rijn M, Jeffrey SS, Rees CA, et al. Molecular portraits of human breast tumours. *Nature* 2000;**406**:747–52.
2. Sørlie T, Perou CM, Tibshirani R, Aas T, Geisler S, Johnsen H, et al. Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *Proc Natl Acad Sci* 2001;**98**:10869–74.
3. Sørlie T, Tibshirani R, Parker J, Hastie T, Marron JS, Nobel A, et al. Repeat observation of breast cancer tumor subtypes in independent gene expression data sets. *Proc Natl Acad Sci* 2003;**100**:8418–23.
4. Yang XR, Sherman ME, Rimm DL, Lissowska J, Brinton LA, Peplonska B, et al. Differences in risk factors for breast cancer molecular subtypes in a population-based study. *Cancer Epidemiol Biomarkers Prev* 2007;**16**:439–43.
5. Goldhirsch A, Wood WC, Coates AS, Gelber RD, Thürlimann B, Senn HJ, et al. Strategies for subtypes-dealing with the diversity of breast cancer: highlights of the St. Gallen International Expert Consensus on the Primary Therapy of Early Breast Cancer 2011. *Ann Oncol* 2011;**22**:1736–47.
6. Cheang MC, Voduc D, Bajdik C, Leung S, McKinney S, Chia SK, et al. Basal-like breast cancer defined by five biomarkers has superior prognostic value than triple negative phenotype. *Clin Cancer Res* 2008;**14**:1368–76.
7. Rahka EA, Reis-Filho JS, Ellis IO. Combinatorial biomarker expression in breast cancer. *Breast Cancer Res Treat* 2010;**120**:293–308.
8. Chong LY, Cheok PY, Tan WJ, Thike AA, Allen G, Ang MK, et al. Keratin 15, transcobalamin I and homeobox gene Hox-B13 expression in breast phylloides tumors: novel markers in biological classification. *Breast Cancer Res Treat* 2011 May 15. [Epub ahead of print].

9. Hammond MEH, Hayes DF, Dowsett M, Allred DC, Hagerty KL, Badve S, et al. American Society of Clinical Oncologist/College of American Pathologists guideline recommendations for immunohistochemical testing of estrogen and progesterone receptors in breast cancer. *Arch Pathol Lab Med* 2010;**134**:E1–16.
10. Kwan ML, Kushi LH, Weltzien E, et al. Epidemiology of breast cancer subtypes in two prospective cohort studies of breast cancer survivors. *Breast Cancer Res* 2009;**11**:R31.
11. Fernandes RCM, Bevilacqua JLB, Soares IC, Maringı B, Kutner SE, Fulton RS, et al. Coordinated expression of ER, PR and HER2 define different prognostic subtypes among poorly differentiated breast carcinomas. *Histopathology* 2009;**55**: 346–55.
12. Brown M, Tsodikov A, Bauer KR, Parise CA, Caggiano V. The role of human epidermal growth factor receptor 2 in the survival of women with estrogen and progesterone receptor-negative, invasive breast cancer. *Cancer* 2008;**112**: 737–47.
13. Millar EKA, Graham PH, O'Toole SA, McNeil CM, Browne L, Morey AL, et al. Prediction of local recurrence, distant metastases, and death after breast-conserving therapy in early-stage invasive breast cancer using a five-biomarker panel. *J Clin Oncol* 2009;**27**:4701–8.
14. Spitale A, Mazzola P, Soldini D, Mazzucchelli L, Bordoni A. Breast cancer classification according to immunohistochemical markers: clinicopathologic features and short-term survival analysis in a population-based from the South of Switzerland. *Ann Oncol* 2009;**20**:628–35.
15. Cancello G, Maisonneuve P, Rotmensz N, Viale G, Mastropasqua MG, Pruneri G, et al. Prognosis in women with small (T1mic, T1a, T1b) node-negative operable breast cancer by immunohistochemically selected subtypes. *Breast Cancer Res Treat* 2011;**127**:713–20.
16. Engelsman E, Persijn JP, Korsten CB, Cleton FJ. Oestrogen receptor in human breast cancer tissue and response to endocrine therapy. *Br Med J* 1973;**2**: 750–2.
17. Barghava R, Beriwal S, Dabbs DJ, Ozbek U, Soran A, Johnson RR, et al. Immunohistochemical surrogate markers of breast cancer molecular classes predicts response to neoadjuvant chemotherapy. *Cancer* 2010;**116**:1431–9.
18. Goldhirsch A, Wood WC, Gelber RD, Coates AS, Thürlimann B, Senn SH. Meeting highlights: updated international expert consensus on the primary therapy of early breast cancer. *J Clin Oncol* 2003;**21**:3357–65.
19. Romond EH, Perez EA, Bryant J, Suman VJ, Geyer Jr CE, Davidson NE, et al. Trastuzumab plus adjuvant chemotherapy for operable HER2-positive breast cancer. *N Engl J Med* 2005;**353**:1673–84.
20. Atkinson R, Mollerup J, Laenkholm AV, et al. Effects of change in cutoff values for human epidermal growth factor receptor 2 status by immunohistochemistry and fluorescence in situ hybridization. *Arch Pathol Lab Med* 2011;**135**:1010–6.
21. Wolff AC, Hammond EH, Schwartz JN, Hagerty KL, Allred DC, Cote RJ, et al. American Society of Clinical Oncology/College of American Pathologists guidelines recommendations for human epidermal growth factor receptor 2 testing in breast cancer. *J Clin Oncol* 2007;**25**:118–45.
22. Shanmugaratnam K. Introduction. In: Jensen OM, Parkin DM, MacLennan R, Muir CS, Skeet RG, editors. *Cancer registration: principles and methods. IARC scientific publication*, vol. 95. Lyon: IARC Press; 1991. p. 1–2.
23. Sant M, Allemani C, Capocaccia R, et al. Stage at diagnosis is a key explanation of differences in breast cancer survival across Europe. *Int J Cancer* 2003;**106**: 416–22.
24. Allemani C, Storm H, Voogd AC, Hakulinen T, Aareleid T, Coebergh JW, et al. Variation in 'standard care' for breast cancer across Europe: a EUROCARE-3 high resolution study. *Eur J Cancer* 2010;**46**:1528–36.
25. Reichman ME, Altekruse S, Li CL, Chen VW, Deapen D, Potts M, et al. Feasibility study for collection of HER2 data by NCI Surveillance, Epidemiology and End Results (SEER) Program Central Cancer Registries. *Cancer Epidemiol Biomarkers Prev* 2010;**19**:144–7.
26. Bauer KR, Brown M, Cress RD, Parise CA, Caggiano V. Descriptive analysis of estrogen receptor (ER)-negative, progesterone receptor (PR)-negative, and HER2-negative invasive breast cancer, the so-called triple-negative phenotype. A population-based study from the California Cancer Registry. *Cancer* 2007;**109**:1721–8.
27. Carey LA, Perou CM, Livasy CA, Dressler LG, Cowan D, Conway K, et al. Race, breast cancer subtypes, and survival in the Carolina Breast Cancer Study. *JAMA* 2006;**295**:2492–502.
28. Caldarella A, Crocetti E, Bianchi S, Vezzosi V, Urso C, Biancalini M, et al. Female breast cancer status according to ER, PR and HER2 expression: a population based analysis. *Pathol Oncol Res* 2011;**17**:753–8.
29. Sauter G, Lee J, Bartlett MS, Slamon DJ, Press MF. Guidelines for human epidermal growth factor receptor 2 testing: biologic and methodologic considerations. *J Clin Oncol* 2009;**27**:1323–33.
30. Gown AM. Current issues in ER and HER2 testing by IHC in breast cancer. *Mod Pathol* 2008;**21**:S8–15.
31. Welsh AW, Moeder CB, Kumar S, Gershkovich P, Alarid ET, Harigopal M, et al. Standardization of estrogen receptor measurement in breast cancer suggests false-negative results are a function of threshold intensity rather than percentage of positive cells. *J Clin Oncol* 2011;**29**:2978–84.
32. Snedecor GW, Cochran WG. *Statistical methods*. 8th ed. Ames, IA: Iowa State University Press; 1989177–95.
33. Yerushalmi R, Woods R, Ravdin PM, Hayes MM, Gelmon KA. Ki67 in breast cancer: prognostic and predictive potential. *Lancet Oncol* 2010;**11**:174–83.
34. Galloway M, Taiyeb T. The interpretation of phrases used to describe uncertainty in pathology reports. *Pathol Res Int* 2011. Article ID:656079.

III – Results

2. International comparison of breast cancer screening programmes

III – Results

2. International comparison of breast cancer screening programmes

As it was described in point 2.5 of the Background section, international collaborative work has the potential to improve and optimize screening programmes.[124,125] Efforts have been done to standardize data collection and definitions in order to enable meaningful comparisons across countries.[124]

Recently, the Portuguese screening programmes from the Northern and Centre regions were invited to join two international studies by participation in the project of European Network for Information on Cancer (EUNICE), co-funded by the European Union. These two studies were described in two published **papers** (**III and IV**).

The first study involved 10 national and 16 regional screening programmes covering topics as organization, coverage and participation rates. From results there was an evident wide agreement between programmes on the mammography as the screening test, the target age range 50-69 and the screening interval (two years). Differences were more notorious at the organization level, volume concentration of services and the size of target populations. The screening programme from the Northern Region contributed with 32122 examinations registered in 2005. Compared to the other programmes, this BCSP showed similar coverage by invitation (80.2% versus 79.3%) and higher coverage by examination (54.0% versus 48.2%). The highest values of coverage by examination were found in Spain (Navarra - 92.1% and Valencia - 73.9%) and Sweden (Västmanland – 82.5%).

Concerning the problem of the false-positive cases (point 2.6.3. of the Background section) in the second study mentioned above, a literature review was performed on the false-positive rate reported by four original research papers; additionally, data collected in the EUNICE project from 20 European breast cancer screening programmes were used to calculate cross-sectional rates of further assessment, with and without needle biopsy and surgery, and the positive predictive value. For this study, the BCSP of the Northern Region contributed with data on 12299 initial and 12709 subsequent screening mammographic examinations, performed in 2005.

The overall further assessment rate was 9.3% in initial and 4.0% in subsequent screenings, in women aged 50-69. For BCSP in the Northern Region those values were 7.1% and 1.9%, respectively. Surgical intervention to clarify previous findings using less invasive techniques in BCSP was 0.7% in initial and 0.3% in subsequent examinations, lower values than the ones verified for the overall screening programmes studied (1.0% and 0.7%, respectively). Compared to the European guidelines,[24] in the Northern Region BCSP the rate of further assessment at initial screening was very close to the acceptable value (<7%) while the subsequent screening value met the desirable level (< 3%).

ORIGINAL ARTICLE

Mammographic screening programmes in Europe: organization, coverage and participation

Livia Giordano, Lawrence von Karsa, Mariano Tomatis, Ondrej Majek, Chris de Wolf, Lesz Lancucki, Solveig Hofvind, Lennarth Nyström, Nereo Segnan, Antonio Ponti and The Eunice Working Group (Eunice Working Group members are listed at the end of the paper)

J Med Screen 2012;**19 Suppl 1**:72–82
DOI: 10.1258/jms.2012.012085

See end of article for authors' affiliations

Correspondence to: Dr Livia Giordano, Unit of Epidemiology – CPO Piemonte, Via S. Francesco da Paola 31–10123 Turin, Italy; livia.giordano@cpo.it

Accepted for publication 25 June 2012

Objectives To summarize participation and coverage rates in population mammographic screening programmes for breast cancer in Europe.

Methods We used the European Network for Information on Cancer (EUNICE), a web-based data warehouse (*EUNICE Breast Cancer Screening Monitoring, EBCSM*) for breast cancer screening, to obtain information on programme characteristics, coverage and participation from its initial application in 10 national and 16 regional programmes in 18 European countries.

Results The total population targeted by the screening programme services covered in the report comprised 26.9 million women predominantly aged 50–69. Most of the collected data relates to 2005, 2006 and/or 2007. The average participation rate across all programmes was 53.4% (range 19.4–88.9% of personally invited); or 66.4% excluding Poland, a large programme that initiated personal invitations in 2007. Thirteen of the 26 programmes achieved the European Union benchmark of acceptable participation (>70%), nine achieved the desirable level (>75%). Despite considerable invitation coverage across all programmes (79.3%, range 50.9–115.2%) only 48.2% (range 28.4–92.1%) of the target population were actually screened. The overall invitation and examination coverage excluding Poland was 70.9% and 50.3%, respectively.

Conclusions The results demonstrate the feasibility of European-wide screening monitoring using the EBCSM data warehouse, although further efforts to refine the system and to harmonize standards and data collection practices will be required, to fully integrate all European countries. The more than three-fold difference in the examination coverage should be taken into account in the evaluation of service screening programmes.

INTRODUCTION

Monitoring early indicators of effectiveness of mammographic breast cancer screening is essential to ensure the quality of all procedures, to optimize the use of resources and ultimately to produce an observable reduction in breast cancer mortality. The fourth edition of the European Guidelines for quality assurance in breast cancer screening and diagnosis defines several performance parameters and indicators that should be monitored in any screening programme, and recommends standards.[1,2] These performance targets address the entire range of activities in screening for and diagnosis of breast cancer, including invitation of the target population, performance of the screening examination, assessment, diagnosis and treatment.

In Europe, most programmes for breast cancer screening have developed their own screening information systems for running day-to-day operations, managing quality, monitoring and evaluating services, and for preparing information for organizations such as local government or ministries, with no explicit priority on promoting an exchange of information between programmes in different countries.

The European Network for Information on Cancer (EUNICE) was a project co-funded by the European Union. A key aim of the project was to produce a monitoring tool capable of calculating a selection of key performance parameters and early impact indicators from the European Guidelines, which could be used to compare screening programmes across Europe on a regular basis. The user-friendly tool facilitates monitoring of screening activity in a standardized format. It enables the uniform, automatic calculation of pre-defined indicators for benchmarking and for comparison between programmes.

This paper describes the design of the tool, with selected programme characteristics, coverage and participation obtained from its initial application in 10 national and 16 regional breast cancer screening programmes in 18 European countries.

METHODS

In 2007 a web-based data warehouse (*EUNICE Breast cancer screening monitoring, EBCSM*) was developed for collection of aggregated data on implementation and performance of breast cancer screening programmes in Europe (www.qtweb.it/eunice).

The database is accessible to authorized users only. It allows data uploading and verification, calculation of screening indicators with standardized algorithms and formats, and comparison between programmes and with benchmarks. The parameters and indicators are shown for the entire age-range and for 5-year age groups. They are generated in the following eight modules:

- Coverage and participation rates
- Number of mammograms performed
- Further assessment, including needle biopsies performed (fine needle aspirations and core biopsies)
- Outcome of further assessment (e.g. number of referrals to surgery)
- Outcome of surgical referral (e.g. number of cancers, benign lesions, ductal carcinoma in situ)
- Number of invasive cancers detected, total and by (TNM) stage
- Size of invasive cancers (1–10 11–20, >20 mm)
- Type of surgery for invasive cancer (number of breast-conserving surgeries, mastectomies)

Each module has two sections, one for routine indicators generated from a minimum set of parameters (standard), and one for optional, more differentiated indicators based on additional parameters (extended). For example, the standard section of the 'participation and coverage' module generates the participation rate by age, whereas the extended version shows the participation rate by age separately for women who were invited to attend screening for the first time.

The online data collection instrument also has a general section, with a questionnaire format, that includes items on programme characteristics, such as the policy on the number of mammographic views or double reading of screening mammograms. The questionnaire and the operational definition of the indicators, as well as a documentation manual (www.qtweb.it/eunice), were prepared by the Eunice Working Group, based on the fourth edition of the European Guidelines.[1,2] Representatives of breast cancer screening programmes from all 27 European Member States plus Norway and Switzerland were invited to join the Group. Two pan-European meetings (Brno, Czech Republic, 2006 and Budapest, Hungary, 2008) were organized to agree the design of the data warehouse, study procedures and data collection.

A survey was then conducted using the EBCSM. The previously identified reference persons from these 29 European countries were asked to provide aggregated data describing service screening activity in the reference year 2005, and supplemental information on programme characteristics in the reference year 2007. Completion of all standard sections was requested, plus the extended sections, where feasible. Checks for internal consistency and completeness were performed on the data received, and detected errors in classification or data entry were corrected. Missing data were reported to participants and completed where possible.

The main outcome measures we report here are coverage and participation, and key organizational and policy characteristics of the programmes. Coverage is defined as the extent to which the screening programme covers the eligible population within the appropriate interval in a given period by invitation (invitation coverage) and the extent to which the screening programme covers the eligible population with screening tests (examination coverage). In practice, coverage has been calculated as the annual number of invitations (or tests) divided by the annual target population, which in turn is represented by the total target population divided by the screening interval in years. Participation is defined as the proportion of women attending screening of those personally invited.

To provide a more detailed picture of the organization of screening services, the extent of invitations and tests performed per screening unit and mammography machine in 2007 were estimated, using data from programme organization in 2007 and the programme performance in earlier years (in most cases 2005 and/or 2006).

RESULTS

Eighteen of the 29 European countries provided aggregated data and information on programme characteristics (Figure 1). National data was provided by 10 countries: Czech Republic, Estonia, Finland, Hungary, Italy, Luxembourg, Norway, Poland, the Netherlands and the United Kingdom (Figure 1). The eight other countries provided data limited to 16 regional programmes: Belgium (Flanders), Denmark (Copenhagen), Germany (pilot projects), Portugal (Centre and North), Republic of Ireland (East), Spain (Asturias, Baleares, Galicia, Navarra, Pais Vasco, Valencia), Sweden (Södermanland, Stockholm, Västmanland) and Switzerland (Fribourg). The results are presented for 26 national or regional programmes. Although the UK also provided national data, we include here only those related to England, as these data are more complete. Performance data from the reference year 2005 were provided by 24 programmes, 10 of which included data from one or two additional reference years. The data from one programme (Germany, pilot projects) referred to the years 2001–2004. The data from Poland referred only to the year 2007.

Policies and organization

Basic information on the programmes is shown in Table 1. Most programmes began in the late 1980s or early 1990s. Exceptions were Belgium (Flanders), Czech Republic, Estonia, Germany (pilot projects in Bremen, Weser-Ems and Wiesbaden), Hungary, Poland, Republic of Ireland and Switzerland (Fribourg), which started more recently.

Women were targeted from age 50 in 17 programmes, while a lower target age was applied in nine programmes. The target age specified in the European Union policy on cancer screening[3] (50–69 years) was adopted in eight programmes, though three others used 50–70 years. In addition to age, gender and geographical area, some

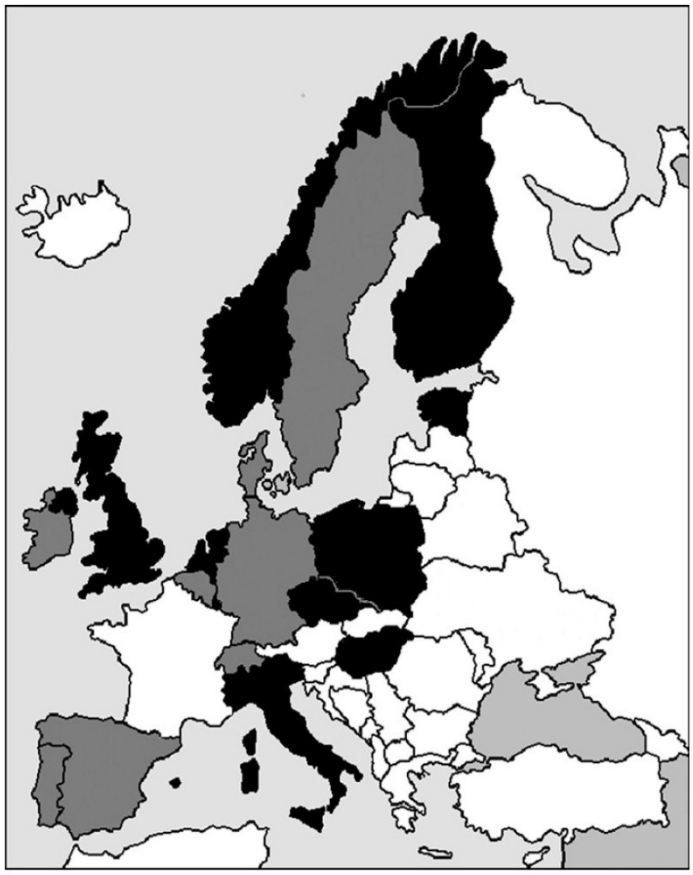


Figure 1 Countries represented in *EUNICE Breast Cancer Screening Monitoring* survey by type of data provided. National data (black): Czech Republic, Estonia, Finland, Hungary, Italy, Luxemburg, Norway, Poland, The Netherlands, United Kingdom; Regional data (grey): Belgium (Flanders), Denmark (Copenhagen), Germany (pilot projects), Portugal (North, Centre) Republic of Ireland (East), Spain (Asturias, Baleares, Galicia, Navarra, Pais Vasco, Valencia), Sweden (Södermanland, Stockholm, Västmanland), Switzerland (Fribourg), Regional and national data: Hungary, Italy, United Kingdom

programmes applied other eligibility criteria, such as exclusion of women with previous breast lesions, previous mastectomies, breast implants, pregnancy or terminal illness. All eligible women received an individual invitation letter, except in the Czech Republic, where women were referred by general practitioners or gynaecologists. Personal invitations were introduced in the Czech Republic on a pilot basis in 2007.

All programmes reported the use of two-view mammography for the initial screening examination; nine programmes used only a single-view at subsequent screening for all or selected groups of women (one missing value). Screening mammograms were read by two independent radiologists in all but five programmes. Mammography was the only screening test performed in 25 programmes. In Hungary, clinical breast examination (BCE) was also used. The screening interval was two years in all programmes except for the United Kingdom (England) where the maximum interval was three years.

In all but two programmes, further assessment was performed on recall. In Spain (Valencia) and the Czech Republic it was possible to perform assessment on the same day as the screening examination. With the exception of the Netherlands, women were recalled for further assessment in units dedicated full or part-time to the screening programme.

Table 2 presents estimates of the annual number of screening tests and the average annual numbers of tests per screening unit, and per mammography machine in 2007 based on the reported programme characteristics in 2007 and volume of tests reported by most programmes in earlier years, assuming that there was no change in volume over time. For most programmes, performance data were provided for the years 2005 and/or 2006. On average, 19 programmes performed less than 10,000 tests per screening unit per year. Ten programmes performed less than 5,000 screening mammograms per machine per year.

Table 1 Breast cancer screening programme features by country or region in 26 European programmes (2007)

Country or region	Start (year)	Target age (years)	Interval (months)	Intermediate mammograms: After screening (Yes/No)	Intermediate mammograms: After further assessment (Yes/No)	Mammography views at screening (N)*	Double reading (Yes/No)
Belgium, Flanders	2001	50–69	24	Yes	No	2	Yes
Czech Republic	2002	45–69	24	Yes	Yes	2	Yes†
Denmark, Copenhagen	1992	50–69	24	No	No	2/1	Yes
Estonia	2002	50–59	24	No	Yes	2	Yes
Finland	1989	50–69	24	NA	NA	NA	Yes
Germany, pilot projects	2001	50–70	24	No	Yes	2	Yes
Hungary	2002	45–65	24	Yes	Yes	2	Yes
Italy	1990	50–69	24	Yes	Yes	2/1	Yes
Luxembourg	1992	50–69	24	Yes	No	2	Yes
Norway	1996	50–69	24	No	No	2	Yes
Poland	2007	50–69	24	Yes	Yes	2	No
Portugal, centre	1990	45–69	24	No	Yes	2	Yes
Portugal, north	1999	45–69	24	No	Yes	2	Yes
Republic of Ireland (East)	2000	50–64	24	No	Yes	2	Yes
Spain, Asturias	1991	50–69	24	No	Yes	2/1	No
Spain, Baleares	1990	50–64	24	Yes	Yes	2	Yes
Spain, Galicia	1992	50–66	24	Yes	Yes	2	Yes
Spain, Navarra	1990	45–69	24	Yes	Yes	2	No
Spain, Pais Vasco	1990	50–64	24	Yes	Yes	2	No
Spain, Valencia	1992	45–69	24	Yes	Yes	2/1	Yes
Sweden, Södermanland	1990	40–74	24	NA	NA	2/1	No
Sweden, Stockholm	1989	40–69	24	NA	NA	2/1	Yes
Sweden, Västmanland	1986	40–69	24	NA	NA	2/1	Yes
Switzerland, Fribourg	2004	50–70	24	No	Yes	2/1	Yes
The Netherlands	1988	50–75	24	NA	NA	2/1	Yes
UK, England	1988	50–70	36	No	Yes	2	Yes

NA, not available
*2/1: two views at first screening, one at subsequent screening
†Performed not in all but in most screened women

Table 2 also shows the availability of full-field digital mammography (FFDM) in 2007. In the 16 programmes using digital mammography, only two relatively small programmes (Estonia and Switzerland, Fribourg) were equipped essentially only with FFDM machines (100% and 95% respectively). In the other programmes FFDM accounted for less than 20% of mammography machines.

Information on breast cancer screening data management and monitoring in 2007 is presented in Table 3, including website addresses from which further information and reports can be obtained. Regional and national monitoring was implemented for 16 programmes in nine countries. Four programmes in three countries used only regional monitoring systems. Monitoring was established only at the national level in five programmes. Most of the programmes used either individual ($n = 8$) or mixed individual and aggregated data ($n = 14$) for monitoring; three programmes used only aggregated data.

Coverage and participation

All 26 programmes completed the standard section of the coverage and participation module. The results shown for most programmes refer to the age group 50–69 years (Table 4). The total number of women in the target population of the 26 programmes used to calculate coverage rates was 26.9 million, and the total number of invitations used to calculate coverage by invitation in the 26 programmes was 13.9 million.

The total number of screening tests in predominantly 50–69-year-old women reported in the 26 programmes and shown in Table 4 (9.16 million) was 7.6% less than the total number performed in women of all ages (9.92 million, Table 2). About 20% of the tests reported in the study were for initial (prevalent) screening, with substantial respective volumes of initial screening reported for the Czech Republic, Estonia and Poland (data not shown).

Coverage by invitation

The coverage by invitation ranged from 50.9% in Italy, to 115.2% in Poland, the latter exceeding 100% as more than 50% of the target population were invited in a single year (i.e. exceeding 100% coverage for a two-year programme) (Table 4). The invitation coverage in Poland was inflated in 2007 due to the initiation of personal invitation in the screening programme in that year. The overall coverage by invitation across all 26 programmes was 79.3%. Excluding Poland, the overall coverage by invitation was 70.9%.

Participation rates

The participation rate varied from 19.4% in Poland to 88.9% in Navarra (Spain) (Table 5). Half of the programmes (13 out

Table 2 Breast cancer screening facilities and estimated annual volume of activity in 26 European programmes (2007)

Country or region	Period (year)	Exams in the period (N)*	Screening tests (1 year average) (N)*	Screening units (N)	Tests per screening unit (1 year average) (N)†	Mammography machines (N)	Tests per machine (1 year average) (N)†	Digital mammography (Y/N)	Digital mammography Proportion of machines (%)
Belgium, Flanders	2005	134,356	134,356	172	781	200	672	Yes	17.4
Czech Republic	2005–2007	1,122,472	374,157	58‡	6451	79‡	4736	Yes	10.3
Denmark, Copenhagen	2005	16,897	16,897	1	16,897	3	5632	No	
Estonia	2005–2006	41,068	20,534	5	4107	5	4107	Yes	100
Finland	2005	211,183	211,183	NA	NA	50	4224	Yes**	NA
Germany	2001–2004§	80,388	20,097	4	5024	5	4019	Yes††	NA
Hungary	2005–2007	658,218	219,406	45	4876	NA	NA	Yes	–
Italy	2005	1,072,357	1,072,357	123	8718	NA	NA	Yes	NA
Luxembourg	2004–2005	28,017	14,009	9	1557	NA	NA	No	–
Norway	2005–2006	370,778	185,389	30	6180	30	6180	Yes	7.7
Poland	2007	935,416	935,416	348	2688	356	2628	Yes	2.0
Portugal, centre	2005	73,182	73,182	10	7318	10	7318	Yes	100
Portugal, north	2005	32,122	32,122	4	8031	4	8031	No	–
Republic of Ireland (East)	2005	59,960	59,960	8	7495	13	4612	Yes	7.8
Spain, Asturias	2005	40,136	40,136	8	5017	10	4014	No	–
Spain, Baleares	2005	13,018	13,018	6	2170	6	2170	Yes	15
Spain, Galicia	2005–2006	172,341	86,170	13	6628	13	6603	No	–
Spain, Navarra	2005–2006	74,087	37,044	2	18,522	3	12,348	No	–
Spain, Pais Vasco	2005	76,229	76,229	9	8470	13	5864	No	–
Spain, Valencia	2005–2006	418,542	209,271	23	9099	23	9099	Yes	9.0
Sweden, Södermanland	2005	21,222	21,222	1	21,222	2	10,611	No	–
Sweden, Stockholm	2005	76,371	76,371	6	12,729	NA	NA	Yes	NA
Sweden, Västmanland	2005	19,617	19,617	2	9809	2	9809	No	–
Switzerland, Fribourg	2005	6886	6886	8	861	8	861	Yes	95
The Netherlands	2005	890,837	890,837	65	13,705	65	13,705	Yes	4.0
UK, England	2005–2007†‡	3,269,375	1,634,688	83	19,695	202	8093	Yes	3.5
Total		9,915,075	6,480,554§§	1043		1102			

NA, not available
*Estimates for 2007 based on tests performed in all age groups in the period indicated in second column
†Estimates for 2007 based on units/machines reported for 2007 and tests performed in all age groups in the period indicated in second column
‡Data refer to 2006
§July 2001–September 2004 in Bremen and Weisbaden, May 2002–September 2004 in Weser Ems
**Only in subsequent routine screening, not in the pilot projects
††Since2007
‡‡July 2005–June 2007
§§Total does not add up due to rounding of country data to the nearest digit

Table 3 Breast cancer screening data managing and monitoring in 26 European programmes (2007)

Country or region	Regional monitoring (Y/N)	National monitoring (Y/N)	Web site references
Belgium, Flanders	Yes (i)	No	www.zorg-en-gezondheid.be/ziektes/vlaams-bevolkingsonderzoek-naar-borstkanker/
Czech Republic	No	Yes (i)	www.mamo.cz/index-en.php
Denmark, Copenhagen	Yes (i)	No	www.cancer.dk/international/english/Screening+breast+cancer+english/
Estonia	No	Yes (a)	www.cancer.ee/?op=body&id=123
Finland	NA	Yes (i)	www.cancer.fi/syoparekisteri/en/mass-screening-registry/breast_cancer_screening/
Germany	Yes (i)	Yes (i)	www.mammo-programm.de
Hungary	Yes (a)	Yes (a)	NA
Italy	Yes (m)	Yes (a)	www.gisma.it; www.osservatorionazionalescreening.it
Luxembourg	No	Yes (i)	www.mammographie.public.lu/
Norway	Yes (m)	Yes (i)	www.kreftregisteret.no
Poland	No	Yes (m)	NA
Portugal, centre	Yes (i)	No	www.ligacontracancro.pt
Portugal, north	Yes (i)	No	www.ligacontracancro.pt
Republic of Ireland (East)	No	Yes (m)	www.cancerscreening.ie
Spain, Asturias	Yes (i)	Yes (a)	www.cribadocancer.es
Spain, Baleares	Yes (i)	Yes (a)	www.cribadocancer.es
Spain, Galicia	Yes (i)	Yes (a)	www.cribadocancer.es
Spain, Navarra	Yes (NA)	Yes (a)	www.cribadocancer.es
Spain, Pais Vasco	Yes (m)	Yes (a)	www.cribadocancer.es
Spain, Valencia	Yes (m)	Yes (a)	www.cribadocancer.es
Sweden, Södermanland	Yes (m)	Yes (a)	NA
Sweden, Stockholm	Yes (i)	Yes (a)	NA
Sweden, Västmanland	Yes (a)	Yes (a)	NA
Switzerland, Fribourg	Yes (i)	Yes (a)	http://www.liguecancer-fr.ch/fr/; http://www.fgdcs.ch/accueil/index.php
The Netherlands	Yes (i)	Yes (a)	http://www.bevolkingsonderzoekborstkanker.nl/
UK, England	Yes (m)	Yes (a)	www.cancerscreening.nhs.uk

NA, not available
(i) = individual data
(a) = aggregated data
(m) = mixed (individual and aggregated) data

of 26) achieved the acceptable level of participation recommended in the European Union Guidelines (>70%). Nine programmes achieved the higher desirable level specified in the European Union Guidelines (>75%). None of the six programmes that started more recently met the acceptable European Union target; in one of these (Czech Republic) the participation rate was not reported because women access the programme without a personal invitation letter.

The participation rate calculated across all 25 programmes sending personal invitations was 53.4% (Table 5). Excluding Poland, the average participation rate was 66.4%.

Except for two regions in Spain (Baleares and Valencia), the programmes that provided data permitting separate calculation of participation after invitation to attend screening for the first time revealed lower participation rates for initial screening compared with the overall participation rate (initial and subsequent screening combined). The differences ranged from 3 to 33 percentage points (see Table 5, footnote).

Coverage by examination

The coverage by examination ranged from 28.4% in Italy to 92.1% in Navarra, Spain (Table 4). The overall coverage by examination calculated across all 26 programmes was 48.2%.

DISCUSSION

A limited number of publications have presented data on the characteristics and performance of breast cancer screening programmes across Europe and internationally.[4–10] These reports have been instrumental in demonstrating the need for uniform standards of reporting to improve the exchange of information and experience between programmes. The fourth edition of the European Guidelines for quality assurance in breast cancer screening and diagnosis recommends a comprehensive set of performance parameters and indicators for monitoring and evaluating any population-based breast cancer screening programme, but does not provide a means of routinely collecting the requisite data and uniformly generating indicators. The EUNICE data warehouse addresses this important need. This report demonstrates the feasibility of the EBCSM module on coverage and participation, and provides an overview of programme organization in 26 screening programmes in Europe. The results are relevant to the current discussion

Table 4 Coverage by invitation and examination in 26 European breast screening programmes

Country or region	Period (year)	Target population of women (N)*	Personal invitations (N)*	Examinations (N)*	Annual target population (N)*,†	Annual invitations (N)*,†	Annual examinations (N)*,†	Invitation coverage (%)*,†	Examination coverage (%)*,†
[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]
Belgium, Flanders	2005	717,856	295,150	134,356	358,928	295,150	134,356	82.2	37.4
Czech Republic‡	2005–2007	1,351,706	NA	831,778	675,853	0	277,259		41.0
Denmark, Copenhagen	2005	54,351	17,559	12,678	27,176	17,559	12,678	64.6	46.7
Estonia (50–59)	2005–2006	96,236	75,372	37,667	48,118	37,686	18,834	78.3	39.1
Finland	2005	688,778	242,796	211,183	344,389	242,796	211,183	70.5	61.3
Germany, pilot projects§	2001–2004	153,667	152,371	80,388	76,834	50,571	26,680	65.8	34.7
Hungary (45–65)	2005–2007	1,512,324	1,721,707	658,218	756,162	573,902	219,406	75.9	29.0
Italy	2005	7,240,570	1,843,119	1,027,964	3,620,285	1,843,119	1,027,964	50.9	28.4
Luxembourg	2004–2005	47,870	44,958	28,017	23,935	22,479	14,009	93.9	58.5
Norway	2005–2006	513,998	484,030	370,778	256,999	242,015	185,389	94.2	72.1
Poland	2007	4,747,156	2,734,513	935,416	2,373,578	2,734,513	935,416	115.2	39.4
Portugal, Centre	2005	193,378	94,131	58,447	96,689	94,131	58,447	97.4	60.4
Portugal, North (45–69)	2005	115,308	46,261	31,123	57,654	46,261	31,123	80.2	54.0
Republic of Ireland, East (50–64)	2005–2006	179,910	157,105	123,011	89,955	78,553	61,506	87.3	68.4
Spain, Asturias	2005	134,280	54,905	40,136	67,140	54,905	40,136	81.8	59.8
Spain, Baleares (50–64)	2005	70,642	19,198	13,018	35,321	19,198	13,018	54.4	36.9
Spain, Galicia (50–66)	2005–2006	258,375	218,542	172,341	129,188	109,271	86,171	84.6	66.7
Spain, Navarra	2005–2006	59,574	61,716	54,873	29,787	30,858	27,437	103.6	92.1
Spain, Pais Vasco (50–64)	2005	206,452	98,044	74,636	103,226	98,044	74,636	95.0	72.3
Spain, Valencia	2005–2006	433,368	441,758	320,268	216,684	220,879	160,134	101.9	73.9
Sweden, Södermanland	2005	34,325	14,516	12,192	17,163	14,516	12,192	84.6	71.0
Sweden, Stockholm	2005	215,440	102,887	71,972	107,720	102,887	71,972	95.5	66.8
Sweden, Västmanland	2005	29,424	13,779	12,138	14,712	13,779	12,138	93.7	82.5
Switzerland, Fribourg	2005	29,480	13,073	6886	14,740	13,073	6886	88.7	46.7
The Netherlands	2005	1,861,200	881,862	730,263	930,600	881,862	730,263	94.8	78.5
UK, England**	2005–2007	5,989,817	4,088,143	3,114,205	1,996,606	2,044,072	1,557,103	102.4	78.0
Total		26,935,485	13,917,495	9,163,952	12,469,440††	9,882,079††	6,006,334††	79.3	48.2

NA, not available

*The data and rates in columns [3] to [10] refer to the reference period indicated in column [2] and the age range 50–69 years, unless a different age range is indicated in column [1]

†Calculations: ANNUAL target population [6] = [3]/no. of years in screening interval reported in Table 1; Annual invitations [7] = [4]/no. of years in [2]; Annual examinations [8] = [5]/no. of years in [2]; Invitation coverage [9] = [7]/[6]; Examination coverage [10] = [8]/[6]

‡A pilot study was conducted in 2007–2008 with personal invitation of non-participant women

§Average rates shown are based on data from the first screening round, 07/2001–09/2004 in Bremen and Weisbaden, 05/2002–09/2004 in Weser Ems. Cumulative rates for round 1 are 99.2% (invitation coverage) and 52.3% (examination coverage)

¶July 2005–June 2007

††Totals do not add up due to rounding of country data to the nearest digit

Table 5 Participation rates in 25 population-based European breast screening programmes

Country or region	Period (year)	Personal invitations (*N*)*	Examinations (*N*)*,†	Annual invitations (*N*)*,‡	Annual examinations (*N*)*,†,‡	Participation (%)*,‡
[1]	[2]	[3]	[4]	[5]	[6]	[7]
Belgium, Flanders	2005	295,150	111,794	295,150	111,794	37.9
Denmark, Copenhagen	2005	17,559	12,989	17,559	12,989	74.0
Estonia (50–59)	2005–2006	75,372	37,667	37,686	18,834	50.0
Finland	2005	242,796	211,183	242,796	211,183	87.0
Germany, pilot projects§	2001–2004	152,371	80,388	50,571	26,680	52.8
Hungary (45–65)	2005–2007	1,721,707	658,218	573,902	219,406	38.2
Italy	2005	1,843,119	1,044,338	1,843,119	1,044,338	56.7**
Luxembourg	2004–2005	44,958	28,017	22,479	14,009	62.3
Norway	2005–2006	484,030	370,778	242,015	185,389	76.6
Poland	2007	2,734,513	530,300	2,734,513	530,300	19.4
Portugal, Centre	2005	94,131	58,447	94,131	58,447	62.1**
Portugal, North (45–69)	2005	46,261	31,123	46,261	31,123	67.3**
Republic of Ireland, East (50–64)	2005–2006	157,105	123,011	78,553	61,506	78.3
Spain, Asturias	2005	54,905	40,136	54,905	40,136	73.1
Spain, Baleares (50–64)	2005	19,198	13,018	19,198	13,018	67.8**
Spain, Galicia (50–66)	2005–2006	218,542	172,341	109,271	86,171	78.9**
Spain, Navarra	2005–2006	61,716	54,873	30,858	27,437	88.9**
Spain, Pais Vasco (50–64)	2005	98,044	74,636	98,044	74,636	76.1
Spain, Valencia	2005–2006	441,758	320,268	220,879	160,134	72.5**
Sweden, Södermanland	2005	14,516	12,192	14,516	12,192	84.0
Sweden, Stockholm	2005	102,887	71,972	102,887	71,972	70.0
Sweden, Västmanland	2005	13,779	12,138	13,779	12,138	88.1
Switzerland, Fribourg	2005	13,073	5790	13,073	5790	44.3
The Netherlands	2005	881,862	728,151	881,862	728,151	82.6**
UK, England††	2005–2007	4,088,143	3,032,433	2,044,072	1,516,217	74.2**
Total		13,917,495	7,836,201	9,882,079‡‡	5,273,987‡‡	53.4

NA, not available

*The data and rates in columns [3] to [7] refer to the reference period indicated in column [2] and the age range 50–69 years, unless a different age range is indicated in column [1]

†No. of screening examinations reported in column [4] differs from respective number reported in Table 4 for programmes that used invitation cohort method (Belgium, Flanders; Denmark, Copenhagen; Italy; Poland; Switzerland, Fribourg; The Netherlands; UK, England). For explanation see Methods section

‡Calculations: Annual invitations [5] = [3]/no. of years in [2]; Annual examinations [6] = [4]/no. of years in [2]; Participation [7] = [6]/[5]

§Average rates shown are based on data from the first screening round, 07/2001–09/2004 in Weser-Ems

**Participation after invitation to attend screening for the first time: Italy: 40.7%; Portugal centre: 36.4%; Spain, Baleares 72.0%, Galicia: 76.2%, Navarra: 56.4%, Valencia: 80.3%; The Netherlands: 78.7%; UK England: 69.4%

††July 2005–June 2007

‡‡Totals do not add up due to rounding of country data to the nearest digit

of the impact of breast cancer service screening programmes, and to the current preparations to update the first report on the implementation of cancer screening programmes in the European Union.[11]

All of the programmes that participated in this survey are involved in the European Cancer Network (ECN), into which the former European Cancer Screening networks have been consolidated. In addition to collaboration in EUNICE, key projects in the ECN and the former European Union cancer screening networks have been the development and updating of the European Guidelines for quality assurance in breast, cervical and colorectal cancer screening and reporting on the implementation of cancer screening programmes in the European Union.[1,2,11–14]

Although 10 of the 26 European Union Member States with breast screening programmes are not represented in the current survey, the present results are consistent with the findings in the first report on implementation of cancer screening programmes in the European Union. This applies, for example, to the wide consensus in the European Union that breast cancer screening should be conducted in organized, population-based programmes, with personal invitations to each individual in the target population. The results presented here also show that despite wide agreement in Europe on additional policy aspects recommended by the Council of the European Union, such as the screening test (mammography), the target age range (50–69 years) and the screening interval (two years), there are still potentially significant differences in the way screening programmes are organized, particularly with regard to the volume and concentration of services and the size of target populations and screening programmes.

Professionals require a sufficient volume of tests to develop and maintain specialized skills in screening. The larger the testing volume of a screening unit, the shorter the time that will be required to accumulate sufficient data to reliably determine performance indicators, such as the rates of referral to surgery and detection of breast cancer, or the benign-to-malignant biopsy ratio. Delays in detecting potential problems necessarily also delay the time until corrective action can be taken. The present survey reveals a high, 27-fold variation between programmes in the estimated yearly number of examinations per screening unit, and a 20-fold variation in the estimated number of tests performed per mammography machine. This wide variation suggests that programmes with lower unit volumes may

require additional efforts and resources to achieve and maintain appropriate quality.

Very large programmes, with millions of eligible participants, must also ensure the same high level of quality across a large number of screening units and mammography machines. Significant, sustainable resources for uniform, timely reporting of appropriate performance parameters and indicators are essential to reliably and promptly detect differences between screening units that may require further investigation and action.

Decision-makers, programme coordinators and scientists should be aware of the substantial differences in Europe in the extent to which target populations are actually exposed to screening. There is a nearly two-fold difference in the invitation coverage across the 25 programmes included in the survey that routinely sent personal invitations. Furthermore, there is a more than three-fold difference in the examination coverage in the 26 programmes included in the survey (Table 4). The low examination coverage in some programmes may be attributed, to a large extent, to the exclusion criteria in the case of the Hungarian programme (women with a mammogram in the previous 24 months were not eligible to attend), and the incomplete rollout of the very large screening programme in Italy and the pilot projects in Germany during the respective reference periods.

Low examination coverage should not, however, be misinterpreted as a reason to interrupt screening activities of appropriate quality, particularly in the rollout phase of programmes, because potentially high coverage in some regions will be masked by little or no coverage in regions that have not yet initiated or completed rollout. In general, the measureable impact of screening on a target population should be greater in a programme with higher examination coverage. In practice, the relationship between examination coverage and impact may not be proportional, but the importance of the degree to which a target population is actually exposed to the intended screening examination should not be overlooked when evaluating the impact of screening. The lower the examination coverage, the more difficult it will be to distinguish the impact of screening from other trends affecting the burden of breast cancer in the population, particularly when methods of analysis are used that do not distinguish carefully between those women who are exposed to the screening test, and those who are not.

Given the importance of maximizing the benefit of screening, while minimizing the negative effects, professionals responsible for the implementation of breast cancer screening programmes should make every reasonable effort to ensure that the screening examination is available to all eligible women. As pointed out in the European Guidelines, effective communication is crucial to the overall success of these activities.[1,2] Even if 90% of the target population is invited to screening, and the participation rate reaches the 'acceptable ' target recommended in the European Union Guidelines (>70%), 3–4 out of 10 women will not have the mammographic examination during a given round of screening. Careful attention should therefore be paid not only to effective communication, enabling women to make an informed choice about attending screening, but also to technical and administrative aspects which ensure that all eligible women are reliably invited.

The present results show less variation between programmes in invitation coverage than in participation rates. The pronounced differences in participation rates underline the fact that the areas and target populations served by programmes may differ substantially with regard to the health-care environment and the characteristics of the target population. Breast cancer awareness among women and the extent of opportunistic screening can strongly affect participation in population-based screening programmes. Low participation rates should also stimulate careful examination of organizational procedures. For example they may result from preselection of the invited population to only include women with previous tests or invitations. The comparatively low participation rates (<60%) in more recently established breast screening programmes are consistent with previous experience in the ECN, particularly in the initial rounds of programmes in areas with significant opportunistic screening activity. The potential impact of opportunistic screening on participation rates is also quite relevant in some Italian areas, especially in younger women.[15] Facilitating the switch to organized screening when eligible women seek an appointment for mammography outside the programme can help to improve participation. A point which is often overlooked is that organized programmes are usually subject to rigorous quality assurance, whereas opportunistic screening activities may not be. This illustrates the importance of informing general practitioners and office-based gynaecologists and radiologists about the programme, and involving them in communication with women.[1,2]

While this overview of European breast screening programmes provides a useful snapshot of key aspects relevant to monitoring and evaluation, there are also limitations. Data were collected predominantly for the years 2005–2007, but significant changes in policies or performance of some programmes may have occurred subsequently. The use of digital technology, for example, which is now established in a number of programmes, was not yet widespread. The information provided in the report is derived from aggregated data. Though aggregated data permit some useful conclusions they are limited regarding the depth of analysis.

CONCLUSIONS

The feasibility of a web-based data warehouse (*EBCSM*) for standardized data collection, analysis and benchmarking of screening programmes in different countries has been demonstrated. The quality of data collection, and the validity and reliability of the information generated with the EBCSM database, is likely to improve if this resource is used on a regular basis to monitor regional and national programme performance, and to compare results between countries. The results presented here show a substantial difference in the extent to which eligible women are offered and participate in screening in European programmes; this should be taken into account when evaluating the impact of screening.

EUNICE WORKING GROUP

The following authors were responsible for data collection from participating screening programmes or for data verification:

Belgium: Van Hal G[1], Martens P[2]; Czech Republic: Májek O[3], Danes J[4]; Denmark: von Euler-Chelpin M[5]; Estonia: Aasmaa A[6]; Finland: Anttila A[7]; Germany: Becker N[8]; Hungary: Péntek Z[9], Budai A[9]; Mádai S[10]; Ireland: Fitzpatrick P,[16] Mooney T[16]; Italy: Zappa M[11], Ventura L[11]; Luxembourg: Scharpantgen A[12]; Norway: Hofvind S[13]; Poland: Seroczynski P[14]; Portugal: Morais A[15], Rodrigues V[15], Bento M J[15], Gomes de Carvalho J[15†]; Spain: Natal C[17], Prieto M[17], Sánchez-Contador Escudero C[18], Zubizarreta Alberti R[19], Fernández Llanes S B[19], Ascunce N[20], Ederra Sanza M[20], Sarriugarte Irigoien G[21], Salas Trejo D[22], Ibáñez Cabanell J[22]; Sweden: Wiege M[23]; Ohlsson G[23], Törnberg S[24], Korzeniewska M[25]; Switzerland: de Wolf C[26]; The Netherlands: Fracheboud J[27]; United Kingdom: Patnick J[28], Lancucki L[28]; IARC: Ducarroz S.[29] , Suonio E.[29]

† Dr J Gomes de Carvalho now deceased

Affiliations:

[1]Voorzitter Consortium, Centrum voor Kankerpreventie Antwerpen, Belgium

[2]Vzw Vroegtijdige Opsporing van Borstklierkanker, Brugge, Belgium

[3]Institute of Biostatistics and Analyses, Masaryk University, Brno, Czech Republic

[4]Department of Radiology, First Faculty of Medicine, Charles University in Prague, Czech Republic

[5]Centre for Epidemiology and Screening, University of Copenhagen, Denmark

[6]Epidemiology and Biostatistics Department-National Institute for Health Development, Tallinn, Estonia

[7]Mass Screening Registry and Finnish Cancer Registry, Helsinki, Finland

[8]German Cancer Research Center, Division of Cancer Epidemiology, Heidelberg, Germany

[9]National Mammography Screening Evaluation Committee (NMSEC), Budapest, Hungary

[10]MaMMa Healthcare Co., Budapest, Hungary

[11]National Centre of Screening Monitoring, Florence, Italy

[12]Programme Mammographie, Direction de la Santé, Luxembourg

[13]The Cancer Registry of Norway, Oslo, Norway

[14]Asseco Poland S.A.

[15]Liga Portuguesa Contra o Cancro, Lisbon, Portugal

[16]BreastCheck, National Cancer Screening Service, Dublin, Ireland

[17]Subdirección de Gestión Clínica y Calidad. Servicio de Salud del Principado de Asturias, Spain

[18]Dirección General de Salud Pública y Participation. Conselleria de Salut i Consum, Baleares, Spain

[19]Dirección General de Innovación y Gestión de la Salud Pública, Consellería de Sanidade, Galicia, Spain

[20]Instituto de Salud Pública, Breast Cancer Screening Program, Navarra, Spain

[21]Programa Detección Precoz Cáncer de Mama, Direccion Territorial de Sanidad de Bizkaia, Bilbao, Spain

[22]General Directorate Research and Public Health & Centre for Public Health Research (CSISP), Valencia, Spain

[23]Mälarsjukhuset, Eskilstuna, Sweden

[24]Cancer Screening Unit, Stockholm, Sweden

[25]Brostenheten Västerås Central Hospital, Västmanland, Sweden

[26]Swiss federation of breast cancer screening, Bern, Switzerland

[27]Department of Public Health, NETB, Erasmus University Medical Centre, Rotterdam, The Netherlands

[28]NHS Cancer Screening Programmes, Sheffield, UK

[29]Quality Assurance Group, Early Detection and Prevention Section, International Agency for Research on Cancer, Lyon, France

. . . . . .

Authors' affiliations
Livia Giordano, MD MPH Epidemiologist, Epidemiology Unit, CPO Piemonte, AOU S. Giovanni Battista, Turin, Italy
Lawrence von Karsa, MD, Coordinator European Cancer Network (ECN) for Screening and Prevention, Quality Assurance Group, Section of Early Detection and Prevention, International, Agency for Research on Cancer, Lyon, France
Mariano Tomatis, Statistician, Epidemiology Unit, CPO Piemonte, AOU S. Giovanni Battista, Turin, Italy
Ondrej Majek, MSc Biostatistician, Institute of Biostatistics and Analyses, Masaryk University, Brno, Czech Republic
Chris de Wolf, MD MPH, in charge of the breast cancer screening program, Fribourg, Swiss Federation of Breast Cancer Screening, Bern, Switzerland
Lesz Lancucki, Statistics Consultant, NHS Cancer Screening Programmes, Sheffield, UK
Solveig Hofvind, Researcher, Department of Research, Cancer Registry of Norway, Oslo, Norway
Lennarth Nyström, Associate Professor of Epidemiology, Department of Public Health and Clinical Medicine, Umeå University, Umeå, Sweden
Nereo Segnan, MD MSc, Head of Cancer Epidemiology Unit, Department of Cancer Screening, CPO Piemonte, AOU S. Giovanni Battista, Turin, Italy
Antonio Ponti, MD MPH Epidemiologist, Epidemiology Unit, CPO Piemonte, AOU S. Giovanni Battista, Turin, Italy
The Eunice Working Group (Eunice Working Group members are listed at the end of the paper)

ACKNOWLEDGEMENTS

Financial support for this publication was provided by the Italian 'National Centre for Screening Monitoring' (www.osservatorionazionalescreening.it) within the EUROSCREEN Project. The EUNICE breast cancer screening survey was financially supported by the European Commission (grant agreements No. 2004114, European Network for Information on Cancer, and No. 2006322, European Cancer Network). Financial support from the European Union Public Health Programme (project no. 2006322, European Cooperation on Development and Implementation of Cancer Screening and Prevention Guidelines [ECCG-ECN]) was also provided for the preparation of this report. Many thanks to Roberta Castagno and Denise Casella (CPO Piemonte) for their assistance during the project and the preparation of this report. The authors thank Tracy Lignini (IARC) for assistance in the preparation of the manuscript. A special acknowledgment to Adam Svobodnik, formerly at Masarick University, Brno (Czech Republic), for providing great support to the project and hosting, with Sylvia Madai, MaMMa Healthcare,

Budapest (Hungary), the project meetings. Finally, thanks to Ahti Anttila for his encouragement and for contributing to refining and dissemination of survey methods and results within the Eurocourse project.

REFERENCES

1 European Commission. *European guidelines for quality assurance in breast cancer screening and diagnosis*. 4th edn. Perry N, Broeders M, de Wolf C, Törnberg S, Holland R, von Karsa L, eds. Luxembourg: Office for Official Publications of the European Communities, 2006

2 Perry N, Broeders M, de Wolf C, Törnberg S, Holland R, von Karsa L. European guidelines for quality assurance in breast cancer screening and diagnosis. Fourth edition – Summary document. *Ann Oncol* 2008;**19**:614–22

3 Council of the European Union. Council Recommendation of 2 December 2003 on cancer screening (2003/878/EC). *Off J Eur Union* 2003;**L 327**:34–8

4 Broeders MJ, Scharpantgen A, Ascunce N, *et al*. European Breast Cancer Network. Comparison of early performance indicators for screening projects within the European Breast Cancer Network: 1989–2000. *Eur J Cancer Prev* 2005;**14**:107–16

5 Shapiro S, Coleman EA, Broeders M, *et al*. Breast cancer screening programmes in 22 countries: current policies, administration and guidelines. International Breast Cancer Screening Network (IBSN) and the European Network of Pilot Projects for Breast Cancer Screening. *Int J Epidemiol* 1998;**27**:735–42

6 Lynge E, Olsen AH, Fracheboud J, Patnick J. Reporting of performance indicators of mammography screening in Europe. *Eur J Cancer Prev* 2003;**12**:213–22

7 Klabunde CN, Sancho-Garnier H, Taplin S, Thoresen S, Ohuchi N, Ballard-Barbash R. Quality assurance in follow-up and initial treatment for screening mammography programs in 22 countries. *Int J Qual Health Care* 2002;**14**:449–61

8 Yankaskas BC, Klabunde CN, Ancelle-Park R, *et al*. International comparison of performance measures for screening mammography: can it be done? *J Med Screen* 2004;**11**:187–93

9 Klabunde CN, Ballard-Barbash R, for the International Breast Cancer Screening Network. Evaluating population-based screening mammography programs internationally. *Semin Breast Dis* 2007;**10**:102–7

10 Dowling EC, Klabunde C, Patnick J, Ballard-Barbash R, International Cancer Screening Network (ICSN). Breast and cervical cancer screening programme implementation in 16 countries. *J Med Screen* 2010;**17**:139–46

11 von Karsa L, Anttila A, Ronco G, *et al*. *Cancer Screening in the European Union*. Report on the implementation of the Council Recommendation on cancer screening – First Report. European Commission, Luxembourg, 2008

12 European Commission. *European guidelines for quality assurance in cervical cancer screening*. 2nd edn. Arbyn M, Anttila A, Jordan J, *et al.*, eds. Luxembourg: Office for Official Publications of the European Communities, 2008

13 Arbyn M, Anttila A, Jordan J, *et al*. European guidelines for quality assurance in cervical cancer screening. Second edition – summary document. *Ann Oncol* 2010;**21**:448–58

14 European Commission. *European guidelines for quality assurance in colorectal cancer screening and diagnosis*. 1st edn. Segnan N, Patnick J, von Karsa L, eds. Luxembourg: Publications Office of the European Union, 2010

15 The Italian cancer screening activity in 2008. Italian Health Ministry – National Centre for prevention and Disease Control (CCM) – National Centre for Screening Monitoring (I programmi di screening in Italia 2008. Ministero del lavoro e della salute delle politiche sociali, Centro Nazionale per la prevenzione e il controllo delle malattie (CCM), Osservatorio Nazionale Screening. Zadig editor, Rome, February 2009

ORIGINAL ARTICLE

False-positive results in mammographic screening for breast cancer in Europe: a literature review and survey of service screening programmes

Solveig Hofvind, Antonio Ponti, Julietta Patnick, Nieves Ascunce, Sisse Njor, Mireille Broeders, Livia Giordano, Alfonso Frigerio and Sven Törnberg The EUNICE Project and Euroscreen Working Groups (Members of the EUNICE Project and Euroscreen Working Groups listed at end of paper)

J Med Screen 2012;**19 Suppl 1**:57–66
DOI: 10.1258/jms.2012.012083

See end of article for authors' affiliations

Correspondence to: Solveig Hofvind, Department of Research, Cancer Registry of Norway, Box 5313 Majorstua, 0403, Oslo, Norway; Solveig.Hofvind@kreftregisteret.no

Accepted for publication 22 June 2012

Objective To estimate the cumulative risk of a false-positive screening result in European mammographic screening programmes, and examine the rates and procedures of further assessment.
Methods A literature review was conducted to identify studies of the cumulative risk of a false-positive result in European screening programmes (390,000 women). We then examined aggregate data, cross-sectional information about further assessment procedures among women with positive results in 20 mammographic screening programmes from 17 countries (1.7 million initial screens, 5.9 million subsequent screens), collected by the European Network for Information on Cancer project (EUNICE).
Results The estimated cumulative risk of a false-positive screening result in women aged 50–69 undergoing 10 biennial screening tests varied from 8% to 21% in the three studies examined (pooled estimate 19.7%). The cumulative risk of an invasive procedure with benign outcome ranged from 1.8% to 6.3% (pooled estimate 2.9%). The risk of undergoing surgical intervention with benign outcome was 0.9% (one study only). From the EUNICE project, the proportions of all screening examinations in the programmes resulting in needle biopsy were 2.2% and 1.1% for initial and subsequent screens, respectively, though the rates differed between countries; the corresponding rates of surgical interventions among women without breast cancer were 0.19% and 0.07%.
Conclusion The specific investigative procedures following a recall should be considered when examining the cumulative risk of a false-positive screening result. Most women with a positive screening test undergo a non-invasive assessment procedure. Only a small proportion of recalled women undergo needle biopsy, and even fewer undergo surgical intervention.

INTRODUCTION

Mammographic screening aims to detect breast cancer at an early stage in women without symptoms. Suspicious findings on screening mammography need further investigation. The decision to recall a woman for further assessment should be a consensus decision made between two radiologists or through arbitration of a third radiologist, and is an essential part of the screening process.

Women who have further assessments and are diagnosed as not having breast cancer are referred to as having a 'false-positive screening result', which is considered an adverse effect of mammographic screening. The likelihood of having breast cancer will be higher in women who are recalled for further assessment compared with those who are not, and the likelihood is even higher in women who have a needle biopsy (fine needle aspiration cytology, FNAC, or core needle biopsy, CNB) compared with those who do not. It is therefore not surprising that a recall for further assessment could be stressful.

The psychological impact of being recalled for further assessment in mammographic screening has been reported in several studies. Different psychological measurement scales and quality-of-life assessments have been used[1–5] and reviews have concluded that false-positive screening results did not cause anxiety and distress at a general level, but rather breast cancer specific distress, anxiety and apprehension.[6,7] The negative impact on a woman's wellbeing is seen during the examination period, in the period between screening and further examination, and after further examination, regardless of the final outcome. The adverse effect among women found not to have breast cancer has been

reported to be transient,[4,6,8] but occasionally can be seen more than a year after the screening test.[3,9]

The psychological harm of a false-positive screening result could lead to a lower participation in future screening rounds, but results from studies have been inconclusive.[4,7,10–13] A recent meta-analysis from four European countries based on data from more than 340,000 women showed that a false-positive screening result did not have any material influence on the overall participation during subsequent rounds.[13]

The risk of a false-positive screening result is strongly positively correlated with the recall rate. This rate is influenced by the training and experience of the radiologist, by the image quality and number of views, by other factors related to screening (e.g. prevalence or incidence screen, screening interval, single versus double reading and screening technique) and characteristics of the women (e.g. age, screening history, use of hormone therapy, breast density, previous invasive procedure and familial breast cancer).[14–19]

To optimize the balance between benefits and harms in mammographic screening, it is important to keep the false-positive rate low without missing breast cancers.[20] We reviewed European studies that estimate the cumulative risk of a recall for further assessment among women without breast cancer, associated with participation in biennial screening mammography starting at ages 50–51 and continuing to ages 68–69. This age group was chosen to be in line with the European guidelines for quality assurance in breast cancer screening and diagnosis.[20] We stratified the cumulative risk by whether invasive procedures (needle biopsy and/or surgery) were used. In an additional analysis, we present cross-sectional observed rates of further assessment, with and without needle biopsy, and surgery, and the positive predictive value (PPV) of the screening test for several European countries, based on data collected as a part of the European Network for Information on Cancer project (EUNICE).[21]

METHODS

Cumulative risk of a false-positive screening result

Procedures for further assessment following a recall differ between countries and centres, but usually include additional mammographic imaging (additional views and magnification), ultrasound and clinical breast examination as the first step. If the finding on the screening mammograms is not resolved after this step, a needle biopsy (CNB or FNAC) is usually performed to obtain a histological or cytological diagnosis. If the biopsy is positive for cancer, surgical treatment is considered. If no histological or cytological diagnosis is available using needle biopsy, women are usually referred for surgical intervention (surgical biopsy and/or surgical treatment). Preoperative breast MRI is being performed with increasing frequency, particularly after an unsuccessful needle biopsy. These three steps (imaging, clinical assessment and biopsy) represent a typical patient flow associated with procedures for further assessment following a recall, and are crucial for understanding the differences in the described extent and psychological consequences of further assessments with a benign outcome. The number of false-positive screening results includes recalls for further assessment with a final diagnosis that is not breast cancer, regardless of which procedures were performed.

We defined the cumulative risk of having a false-positive screening result as the risk of being recalled for further assessment at least once during 10 biennial screens performed from age 50 to 51 until 68 to 69, among women without a diagnosis of breast cancer. The cumulative risk of an invasive procedure with benign outcome was defined as the risk of having at least one needle biopsy or surgical intervention performed during the same time span, again among women without a diagnosis of breast cancer.

We conducted a literature review based on articles in PubMed with a title, abstract or keywords including a reference to both false-positive results and screening mammography which were published in 1995–2011 (the search was performed in October 2011). The search terms were (false-positive OR abnormal OR benign) AND (breast cancer OR mammog*). We also manually searched the reference sections of relevant papers. We aimed to identify reports of the cumulative risk of a false-positive screening result in women aged 50–69. After searching on title and abstract, and the whole paper if it appeared relevant, four original research papers[10,22–24] and one research letter[25] were considered relevant for our analysis. The Spanish Cumulative False-Positive Risk Group has published several studies on this topic, based on the same women.[18,24,26,27] We chose the first published study, of more than 1.5 million women, including 251,275 aged 50–51 and followed for six screening rounds.[24] The 8502 women in the study by Castells *et al.*[22] were also included in Salas *et al.*[24] and the study by Castells *et al.* was therefore excluded from the pooled analyses and from our review results.

We estimated both the unweighted average of the cumulative risk of a false-positive screening result after 10 screening examinations and the average weighted by study size. The cumulative risk of women undergoing needle biopsy is reported only in two studies.[10,24] The cumulative risk of surgical intervention was given only in the study by Hofvind *et al.*[10] As a consequence of the limited numbers of studies covering needle biopsy and surgical intervention, the average rates, both weighted and unweighted, should be considered with care.

Performance parameters in European mammographic screening programmes

We augmented the cumulative estimates with cross-sectional estimates from population-based mammography programmes in Europe. Aggregate data regarding performance parameters achieved in European mammographic population-based screening programmes were collected within the EUNICE project in 2008–2009. The EUNICE project is described in detail elsewhere in this Supplement.[21] We included data from 20 national or regional screening programmes (from 17 countries), based on 7,658,586 screening examinations performed between 2005 and 2007. Compared with the report on screening coverage and participation,[21] seven programmes (two national and five regional) were excluded because of missing

information on key variables concerning further assessment or because data from initial and subsequent examinations could not be separated. One regional programme (MaMMa Network, Budapest) was added.

Information about further assessment procedures following a positive screening test (i.e. needle biopsy and surgical intervention) were calculated as proportions of screened women. The PPV of the screening test was calculated as the number of screen-detected breast cancers (including ductal carcinoma *in situ*) divided by the number of all positive screening tests (among those with and without cancer). The benign-to-malignant (B/M) biopsy ratio, as well as the rate of surgical intervention includes all procedures performed as a consequence of a positive screening test. Time trends by age groups were analysed using Poisson regression. Spearman's rank correlation co-efficient (rho) was used for examining correlations between outcome parameters. Time trends and correlation coefficients were only investigated for subsequent screening tests. In order to ensure that the assumption of linearity was reasonable, the correlation between the rate of further assessment and 1/PPV of the screening test was plotted, the latter of which is interpretable as the number of women recalled per cancer detected.

RESULTS

Cumulative risk of a false-positive screening result

Table 1 describes the studies included, and their respective estimates. Together, they are based on 390,000 screened women.

The study by Hofvind *et al.*[10] was based on women in the first three screening rounds in the pilot study of the Norwegian Breast Cancer Screening Programme, conducted between 1996 and 2001. The cumulative risk of having one or more false-positive screening results was estimated for women aged 50–51, who are expected to be screened 10 times until age 68–69. The authors assumed independence between the outcomes of subsequent screening tests which may overestimate the cumulative risk. To extrapolate to a follow-up time of 20 years, they further assumed that the risk of having a false-positive screening result from each of the fourth to the 10th screening tests was the same as that observed in women in corresponding age groups in the third screening round. It was estimated that women aged 50–51 who participate in 10 biennial screening tests have a 20.8% cumulative risk of a false-positive screening result. The cumulative risks of having FNAC or CNB were estimated to be 3.9% and 1.5%, respectively. The cumulative risk of having a surgical intervention (both surgical biopsy and final surgical treatment) with benign outcome was 0.9%. This was the only study that presented the cumulative risk of surgical intervention with benign outcome.

The largest study (more than 250,000 women) came from Spain.[24] The cumulative risk of a false-positive screening result was estimated for women aged 50–51 at the start of screening. The database included women who were screened at least once in eight of the 17 administrative regions of the mammographic screening programme. The estimates were obtained from a regression model (discrete time hazard model), after adjustment for possible confounding factors (screening unit and calendar period). The overall cumulative risk of having a false-positive result was 20.4%, and 1.8% for an invasive procedure (fine or thick needle biopsy, biopsy and/or other invasive tests) with benign outcome.

The study by Njor *et al.*[23] was based on two areas within Denmark. The study used the same estimation model as used by Hofvind *et al.* but Njor *et al.* first analysed whether there was independence between the outcomes of subsequent screens. The hypothesis of independence between the outcomes of subsequent screens was accepted, and the estimated cumulative risk of having a false-positive screening result was 15.8% for women residing in the Copenhagen area and 8.1% for women residing in the Fyn area. The risks were 22.6% for Copenhagen and 9.9% for Fyn when crude proportions of participants with a false-positive test from each of the first five invitation rounds were used.[23]

Puliti *et al.*[25] estimated the cumulative risk of a false-positive screening test at 15.2% in 28,500 women who had participated in seven screening rounds in Florence, Italy (Table 1). The cumulative risk of a needle biopsy with benign outcome was 1.8%.

We derived pooled estimates of the cumulative risks using those studies that estimated the risk over 10 years, that is all except the study by Puliti *et al.*[25] Thus, the pooled estimates were based on 364,991 screenees. The weighted pooled estimate of the cumulative risk of having at least one false-positive screening result over 10 biennial screening sessions in women aged 50–51 without breast cancer was 19.7% based on three studies.[10,23,24] The unweighted average was 16.3%. Both are consistent with the empirical result from Italy (15.2% after seven screening rounds). The weighted pooled cumulative risk of having a false-positive screening test that led to an invasive procedure was 2.9% based on two studies (unweighted average 4.1%).[10,24] The estimated cumulative risk of a false-positive screening result without an invasive procedure was therefore 16.8%.

Performance parameters collected in the EUNICE project

The EUNICE project is described by Giordano *et al.*[21] in this supplement of the journal. Data from 20 mammographic screening programmes in 17 European countries illustrate that the rate of further assessment varied between 2.2% and 15.6% for initial screening tests, and between 1.2% and 10.5% for subsequent screening tests in women aged 50–69 (Table 2). The overall rates were 9.3% and 4.0%, respectively, for initial and subsequent screening tests. The rate of needle biopsy varied between 0.8% and 3.3% for initial screening tests and between 0.3% and 1.5% for subsequent screening tests respectively (2.2% and 1.1% overall). Surgery was performed in an average of 1.0% of the initial screens (range: 0.4–1.4%) and 0.73% (range: 0.3–1.1%) of the subsequent screens. The variation in further assessment rate is reflected in the PPV which differed substantially in programmes that provided data: 4.9–24.2% in initial screening tests and 6.8–49.5% in subsequent tests. The rate of surgical intervention with benign outcome was only 0.19% in initial and 0.07% in subsequent screening,

Table 1 Summary of studies performed in Europe to estimate cumulative risk of a further assessment without cancer diagnosed during a given screening period

Authors and year of publication	Country	Study group (n)	Age group (years)	Study period (years)	Screening interval (year)	Views (#)	Recall with negative outcome (range %)	Adjusted for characteristics		Observed screening rounds (#)	Screening rounds in estimation (#)	Cumulative risk of further assessments among women without cancer diagnosed	
								In the protocol	For the individual			All proc (%)	Invasive proc (%)
Hofvind *et al.* 2004[10]	Norway	83,416*	50–69	1996–2001	Two	Two	Ini‡: 2.7–4.4, Sub§: 1.6–2.6	No	No	3; values for 4th–10th rounds based on 3rd round	10	20.8	FNAC: 3.9, CNB: 1.5, SI: 0.9, total: 6.3
Njor *et al.* 2007[23]	Denmark	21,261*, 9039†	50–69	1993–2003, 1991–2001	Two	Pre: two, Sub: one; two if dense breast	Ini: 1.8, Sub: 0.7–0.9, Ini: 5.7, Sub: 0.9–3.3	No	No	5; values for 6th–10th rounds based on 5th round	10, 10	8.1, 15.8	NA, NA
Salas *et al.* 2011[24]	Spain	251,275	50–69	1990–2006	Two	One or two	Ini: 7.3**, Sub: 3.7**	Yes	Yes	6; values for 7th and 10th round based on 6th round	10	20.4 (20.0–20.8)	1.8 (1.7–1.9)
Puliti *et al.* 2011[25]	Italy	28,500*	50–69	1991–2006	Two	Two	Ini: 3.1, Sub: 1.4–2.6	No	No	7; empirical data	7	15.2	IP: 1.8
Weighted pooled estimates			50–69		Two						10	19.7[10,23,24]	All 2.9[10,24]

All programmes performed independent double reading and assumed independence between subsequent screening tests
FNAC: fine needle aspiration cytology, CNB: core needle biopsy, SI: surgical intervention, IP: invasive procedure (fine or thick needle biopsy, biopsy and/or other invasive tests)
*Fyn
†Copenhagen
‡Ini: Initial screening round
§Sub: subsequent screening rounds
**Crude rates

Table 2 Number of screening examinations in women aged 50–69 who were screened within 2005–2007 in European mammographic screening programmes, rate (%) of further assessment (FA), preoperative needle biopsy and surgical intervention (all and with benign outcome, including all surgical procedures performed as a consequence of a positive screening test) among those screened, positive predictive value (PPV) of the screening test and benign-to-malignant (B/M) ratio of surgical biopsies , by screening history (initial and subsequent screening test) and area

	Initial screening test								Subsequent screening test							
					Surgical intervention							Surgical intervention				
	Period	Screening exams (*n*)	FA (%)	Needle biopsy (%)	All (%)	Ben* (%)	PPV* (%)	B/M* (ratio)	Screening exams (*n*)	FA (%)	Needle biopsy (%)	All (%)	Ben* (%)	PPV* (%)	B/M* (ratio)
Belgium	2005	47,104	6.2	–	0.8	0.13	10.7	0.20	87,252	3.5	–	0.5	0.06	14.0	0.12
Flanders															
Czech Republic	2005–2006	256,425	15.6	1.1	0.7	–	–	–	234,900	10.5	0.6	0.4	–	–	–
Denmark	2005	2910	4.3	–	1.3	0.27	24.2	0.27	10,079	2.0	–	1.1	0.07	49.5	0.07
Copenhagen															
Estonia†	2005–2006	20,555	2.2	0.9	0.4	0.06	15.6	0.18	–	–	–	–	–	–	–
Finland	2005	–	–	–	–	–	–	–	211,183‡	2.7	–	0.6	0.09	17.1	0.20
Hungary	2005	15,631	7.0	2.2	0.8	0.12	7.3	0.23	27,060	5.2	1.5	0.6	0.05	8.2	0.12
Budapest§															
Italy	2005	170,427	7.2	1.2	0.8	0.16	7.5	0.29	576,207	4.4	0.9	0.6	0.10	11.2	0.21
Luxembourg	2004–2005	5094	14.1	–	0.9	0.22	4.9	0.31	22,923	5.3	–	0.7	0.06	12.4	0.09
Norway	2005–2006	76,058	5.7	2.1	0.7	0.12	9.7	0.22	283,184	3.1	1.1	0.6	0.06	18.0	0.12
Poland	2007	403,596	10.6	–	1.3	–	–	–	531,820	7.1	–	0.8	–	–	–
Portugal centre	2005	13,841	10.2	–	0.8	–	–	–	44,606	2.4	–	0.4	–	–	–
Portugal north	2005	12,299	7.1	–	0.7	–	–	–	12,709	1.9	–	0.3	–	–	–
Republic of	2005–2006	38,170	5.6	–	1.0	0.25	12.9	0.35	84,841	2.0	–	0.5	0.06	21.7	0.13
Ireland**															
Spain Galicia††	2005–2006	30,969	4.5	0.8	0.5	0.15	6.8	0.47	141,372	1.3	0.3	0.3	0.04	17.1	0.18
Spain Navarra	2005–2006	734	13.9	2.2	1.4	0.41	6.9	0.43	54,139	3.8	0.6	0.5	0.04	10.8	0.11
Spain Valencia	2005–2006	15,826	12.6	1.7	1.0	0.26	5.7	0.36	304,442	5.3	0.6	0.4	0.06	6.8	0.16
Sweden Stockholm	2005	8102	4.8	–	0.7	0.17	10.2	0.35	63,870	2.6	–	0.8	0.10	25.4	0.14
Switzerland	2005	5790	7.4	–	1.3	0.50	10.3	0.66	–	–	–	–	–	–	–
Fribourg															
The Netherlands	2005	62,025	2.7	1.2	–	–	19.9	–	668,238	1.2	0.7	–	–	36.5	–
UK England‡‡	2005–2007	531,870	8.3	3.3	1.0	0.21	9.7	0.26	2,582,335	3.7	1.5	0.9	0.07	21.6	0.09
Total – all screening programmes (range)		1,717,426 (2.2–15.6)	9.3 (2.2–15.6)	2.2 (0.8–3.3)	1.0 (0.4–1.4)	0.19 (0.06–0.50)	9.6 (4.9–24.2)	0.27 (0.18–0.66)	5,941,160 (2.2–15.6)	4.0 (1.2–10.5)	1.1 (0.3–1.5)	0.73 (0.3–1.1)	0.07 (0.04–0.10)	18.6 (6.8–49.5)	0.11 (0.07–0.21)

*Results are shown for screening programmes with <10% missing information on postoperative diagnosis
†50–59 years old
‡Including initial tests, estimated to be <10% of total number of screening tests
§50–65 years old; screening test includes breast clinical examination
**East Republic of Ireland; 50–64 years old
††50–66 years old
‡‡July 2005–June 2007

while the B/M ratio of surgical biopsy was 0.27 and 0.11, respectively, for initial and subsequent screening. There was less variation in the B/M ratio between programmes. False-positive rates can be calculated by subtraction of cancer detection rates from assessment rates. Therefore, in estimating the overall false-positive rates, we excluded those programmes not providing detection rates. We also excluded results for UK England, which were based on a three-year interval. After these exclusions, the overall false-positive rate was 5.4% in initial screens (6.0% further assessment rate, 0.55% detection rate) and 2.5% in subsequent screens (3.0% further assessment rate, 0.46% detection rate).

The European Union (EU) guidelines indicate that the desirable rate of further assessment should be <5% for initial and <3% for subsequent screens, but acceptable levels are <7% and <5%, respectively.[20] The EUNICE data show that several (but not all) European programmes meet these targets. For initial screening 42% (8/19) were acceptable and 26% (5/19) were desirable, while for subsequent screening 72% (13/18) were acceptable and 44% (8/18) desirable (Table 2).

The PPV of the screening test and the rates of needle biopsy and surgery (for diagnosis and treatment of screen-detected breast cancer) increased with increasing age, while the B/M ratio decreased with age (Table 3). In subsequent screening, the rate of further assessment was positively correlated with 1/PPV as might be expected, while no correlation was observed between the further assessment rate and needle biopsy, surgical intervention or surgical intervention with benign outcome (Figure 1). The rates of needle biopsy presented in Figure 1 and Table 2 are on average 24% and 28%, respectively, of the reported further assessment rates for initial and subsequent screens.

DISCUSSION

In mammography service screening, a biopsy is required to verify the presence of breast cancer, but it is important to keep the biopsy rate as low as possible in women not found to have cancer (false-positives), using the least invasive techniques. Inherent to a false-positive screening test is the psychological stress, assumed to be related to the fear of having breast cancer and the subsequent risk of death. However, the stress is mainly transient and related to the time between when the call is received and the suspicious finding is resolved.[4,6] If a diagnosis of breast cancer is excluded with additional imaging and clinical examination only, the anxiety will usually begin to subside as soon as the outcome is communicated. Performance of a needle biopsy or a surgical biopsy and the subsequent waiting time for the result prolongs the stress,[3] which might be related to an element of unpreparedness. Women with symptoms are probably more prepared to undergo further examinations than those who are recalled due to a positive screening mammogram. Fixed protocols for further assessment procedures communicated to the women and avoidance of short-term follow up (early re-screens), as recommended in the European guidelines,[20] could reduce the psychological stress related to recalls. Specialist screening units with dedicated and well-trained staff (clinicians, administrative personnel, breast nurses and radiographers) equipped with good communication skills are more likely to achieve these aims than non-dedicated clinical practices.[28]

Using data from three published observational studies,[10,23,24] the estimated cumulative risk of having a recall with additional imaging but a benign result is 20% (ranging between 8% and 21%), and the risk of having an invasive procedure (i.e. needle biopsy or surgical intervention) is 3% (range 2–6%), among women aged 50–51 who attend 10 screening tests over two decades. Only one study reported the cumulative risk of a false-positive screening test resulting in surgical intervention (risk of 0.9%). The results were consistent with a fourth study[25] based on seven screening tests. The pooled estimates should be interpreted with caution due to the small number of studies. The largest study[24] has the longest follow-up, while the Norwegian study has the shortest. A substantial proportion of FNAC (3.9%) performed on cysts in the start-up period of the programme might explain the higher needle biopsy rate in Norway compared with the large study in Spain.

From the EUNICE data, the overall false-positive rate was 5.4% in initial screens (6.0% further assessment rate, 0.55% detection rate) and 2.5% in subsequent screens (3.0% further assessment rate, 0.46% detection rate), consistent

Table 3 Pooled results from European Network for Information on Cancer project (EUNICE) showing rates (%) of further assessment, preoperative needle biopsy and surgical intervention (all and with benign outcome), positive predictive value (PPV) of the screening test, and benign-to-malignant (B/M) ratio of surgical biopsies, by age groups in subsequent screened women aged 50–69 in European screening programmes between 2005 and 2007

	Age (years)					
Procedure	50–54	55–59	60–64	65–69	Total	Linear trend for proportions*
Number of women with the actual information and screened	1,152,188	1,972,213	1,577,939	1,238,820	5,941,160	
All further assessment (%)	4.8	3.8	3.7	4.0	4.0	$P = 0.460$
Needle biopsy (%)	0.89	1.00	1.18	1.50	1.14	$P = 0.031$
Surgical intervention (all) (%)	0.52	0.63	0.78	0.99	0.73	$P = 0.012$
Surgical intervention with benign outcome (%)	0.09	0.07	0.07	0.07	0.07	$P = 0.490$
PPV of the screening test (%)	10.9	17.1	21.3	24.4	18.6	$P = 0.013$
B/M ratio of surgical biopsies	0.22	0.12	0.10	0.08	0.11	$P = 0.070$

*Poisson distribution

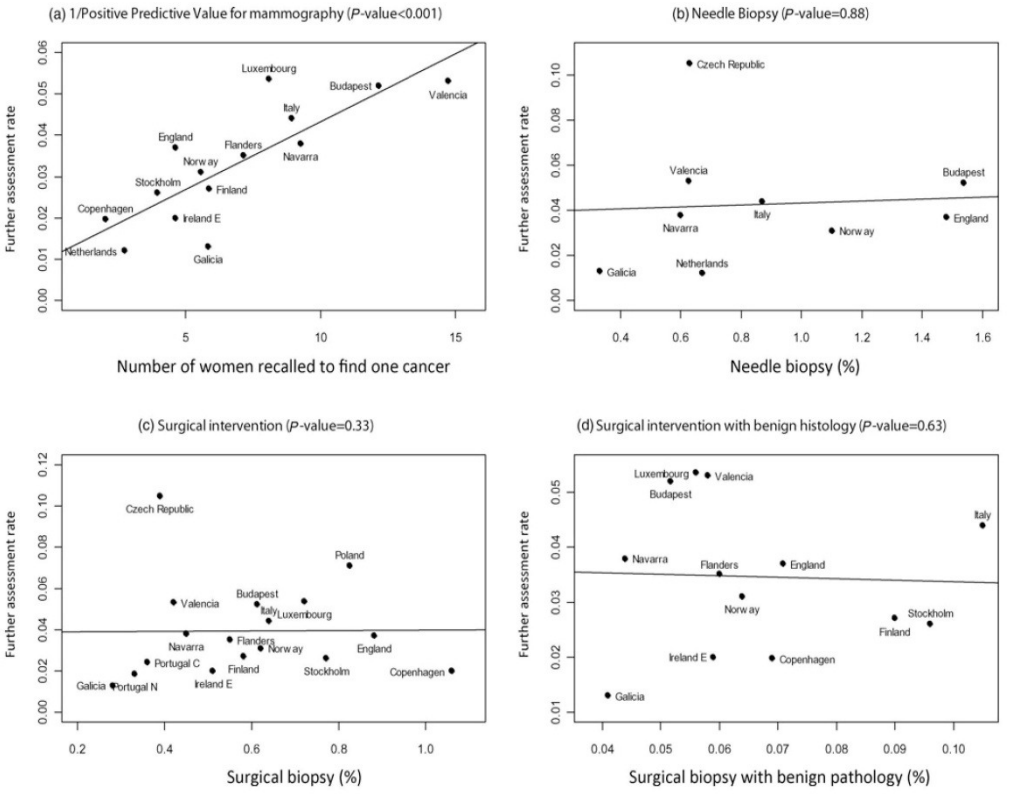


Figure 1 Correlation between further assessment rate and (a) 1/positive predictive value for the screening test (Spearman rank correlation co-efficient rho = 0.83), (b) needle biopsy (Spearman rank correlation co-efficient = 0.07, (c) surgical intervention (including all surgical procedures performed as a consequence of a positive screening test) (Spearman rank correlation co-efficient = 0.25) and (d) Surgical intervention with benign outcome (Spearman rank correlation co-efficient = −0.15) in subsequent screened women aged 50–69 years, within 2005–2007 in European screening programmes (EUNICE collaborative project)

with the results from the three observational studies[10,23,24] and one research letter.[25] Because of the three-year screening interval in the UK programme, England was excluded from this analysis along with programmes that did not provide information on detection rate. It should be noted that the observational studies are based on cohorts of women followed from age 50–51 on their first screen and attending successive screens, while the EUNICE survey represents a cross section of screens in a particular time period including participants with both regular and irregular attendance. The EUNICE data also show that the proportion of surgical intervention with benign outcome (Table 2) was well below the European desirable target of 0.25[20] and declined compared with results in the early years of service screening.[29] This target can probably be lowered in the next edition of the European Guidelines, due to improvements in imaging quality and a general shift from FNAC to larger CNB, which provides more diagnostic material.

The EUNICE data illustrate that the rates of further assessment and needle biopsy and surgery vary considerably between screening programmes in Europe. The personal characteristics of women screened might explain some of the substantial variation in further assessment rate and utilization of assessment procedures in European screening programmes. A recent study from Spain showed that the cumulative risk of a false-positive screening result varied from 8% to 51% in women with low-risk and high-risk profiles, respectively.[18] Differences in other background conditions might also explain some of the observed variation.

The finding of a correlation between further assessment rate and PPV of the screening test in the EUNICE programmes (Figure 1) was expected.[30] However, the lack of correlation between further assessment rate and needle biopsy, and with surgical intervention with benign histology was somewhat surprising, and means that a screening programme with high further assessment rates does not necessarily have high rates of invasive assessment. Figure 1 should be interpreted with caution given the relatively low number of observations, particularly for needle biopsy, for which information was provided by only nine out of the 18 countries. Additional analyses also showed that the *P* value and rho did not differ substantially in Figures 1a, c and d when only the subset of 12 countries providing information on all relevant parameters was included in the analyses.

It should also be noted that, as expected, the assessment rate and the cancer detection rate were also positively associated (data available, not shown). The variability of detection rate among programmes was considerably lower than that of either the rate of further assessment or PPV, suggesting that it is possible to improve the detection rate without necessarily influencing the assessment rate.

An increase in PPV by age (Table 3) has been reported from other studies.[29,31] This is reasonable, given the relatively higher breast density and lower incidence of disease at younger ages. Somewhat surprisingly, the further assessment rate did not decrease significantly by age, contrary to the findings of others.[32] However, when we re-analysed excluding Ireland and Galicia, where women aged 65–69 were not invited, and England, which only recently initiated screening in this age group, the trend by age reached borderline significance (50–54: 5.4%; 55–59: 4.3%; 60–64: 4.0%; and 65–69: 3.8%, $P = 0.090$).

Continuous quality assurance and further improvements to optimize the screening process are needed to offer women in all European settings the quality achieved in programmes with the best results. Improvement in specificity over time has been demonstrated[31] and may have an impact on future results.[21] Analysis of the rate of further assessment should be performed in relation to the corresponding PPV of the screening test. It would arguably be preferable to have a recall rate of 4% and a PPV of 20%, implying a cancer detection rate of 0.8%, rather than a recall rate of 3% and a PPV of 15%, which would yield a cancer detection rate of 0.45%.

Key strengths of the EUNICE database include the large number of individuals for whom aggregate data are available, the large proportion of EU countries represented, the extensive feedback to project contributors, and data checking performed.[21] There are several limitations to the present analysis of the EUNICE data. Given the aggregate format and relatively limited detail of the currently available data, it does not permit a precise estimation of cumulative false-positive rates. Differences in screening protocols, health-care systems and reporting systems in European countries may have affected the accuracy and completeness of data collection. An example of the way different screening protocols may limit the comparison and interpretation of data is that two out of the 20 programmes perform further assessment on the same day as screening, so that those with suspicious features on the mammograms undergo further investigation and receive a diagnosis on the same day as the screen.[33] The expected avoidance of short-term anxiety associated with recall to assessment is accompanied, however, by an increase in further assessment rates. The reported correlations between the parameters calculated from the EUNICE data should be interpreted with care, as they are based on a relatively small number of observations (screening programmes).

Estimating the cumulative risk of a false-positive screening result is challenging due to a number of methodological issues, including the varying further assessment rates both within and between screening programmes. Different recall procedures may result from differences in screening regimens[26,27] and in the characteristics of women attending screening (age, use of hormone treatment and previous breast biopsy).[18,24] All the studies shown in Table 1 estimated the cumulative risk after 10 screening examinations, but none of the studies actually followed the women for 10 screening rounds. The cumulative risk is therefore based on estimated recall rates for four to seven screening rounds. This means that the risks of a false-positive test result might be overestimated in the studies cited in our review. An assumption of independence between screening results in subsequent screening rounds is discussed in several methodological studies, and different models for estimating the cumulative risk of false-positive screening results have been developed.[34–36] The assumption of independence between recalls in subsequent screening rounds was confirmed in Hofvind *et al.*[10] and Njor *et al.*[23]

In studies in the USA, the recall rate in a given round of screening has been reported to be as high as 15%[15,37–39] and the cumulative risk of a false-positive screening result, with and without invasive procedures, is therefore considerably higher than the estimates in the European studies. Screening readers in most European screening programmes perform a specified volume of screening and diagnostic mammograms, in compliance with the EU guidelines.[20] Adherence to the EU guideline recommendations is suggested as being the main reason for the lower recall rate, and therefore the lower cumulative risk of a false-positive screening test in European programmes compared with the results reported in the USA.[10,23,24,38]

CONCLUSIONS

Previously published estimates of the cumulative risk of a recall for further assessment among women aged 50–69 without breast cancer were reviewed. Over a period of two decades with biennial screening these varied from 8% to 21% in Europe (average 20%), and the cumulative risk of an invasive procedure was 3%. Performance monitoring data collected from several population-based service screening programmes in Europe are mostly within accepted levels according to the European guidelines, but outliers were identified. Continued quality assurance is required to offer women high quality mammographic screening.

EUNICE WORKING GROUP

The following authors were responsible for data collection from participating screening programmes or for data verification:

Belgium: Van Hal G[1], Martens P[2]; Czech Republic: Májek O[3], Danes J[4]; Denmark: von Euler-Chelpin M[5]; Estonia: Aasmaa A[6]; Finland: Anttila A[7]; Germany: Becker N[8]; Hungary: Péntek Z[9], Budai A[9]; Mádai S[10]; Ireland: Fitzpatrick P[16], Mooney T[16]; Italy: Zappa M[11], Ventura L[11]; Luxembourg: Scharpantgen A[12]; Norway: Hofvind S[13]; Poland: Seroczynski P[14]; Portugal: Morais A[15], Rodrigues V[15], Bento MJ[15], Gomes de Carvalho J[15†]; Spain: Natal C[17], Prieto M[17], Sánchez-Contador Escudero C[18], Zubizarreta Alberti R[19], Fernández Llanes SB[19], Ascunce N[20], Ederra Sanza M[20], Sarriugarte Irigoien G[21], Salas Trejo D[22]; Ibáñez Cabanell J[22]; Sweden: Wiege M[23]; Ohlsson G[23], Törnberg S[24], Korzeniewska M[25]; Switzerland: de Wolf C[26]; The Netherlands: Fracheboud

J[27]; United Kingdom: Patnick J[28], Lancucki L[28]; IARC: Ducarroz S[29], Suonio E[29]

†Dr J Gomes de Carvalho now deceased.

Affiliations:

[1]Voorzitter Consortium, Centrum voor Kankerpreventie Antwerpen, Belgium

[2]Vzw Vroegtijdige Opsporing van Borstklierkanker, Brugge, Belgium

[3]Institute of Biostatistics and Analyses, Masaryk University, Brno, Czech Republic

[4]Department of Radiology, First Faculty of Medicine, Charles University in Prague, Czech Republic

[5]Centre for Epidemiology and Screening, University of Copenhagen, Denmark

[6]Epidemiology and Biostatistics Department – National Institute for Health Development, Tallinn, Estonia

[7]Mass Screening Registry and Finnish Cancer Registry, Helsinki, Finland

[8]German Cancer Research Center, Division of Cancer Epidemiology, Heidelberg, Germany

[9]National Mammography Screening Evaluation Committee (NMSEC), Budapest, Hungary

[10]MaMMa Healthcare Co., Budapest, Hungary

[11]National Centre of Screening Monitoring, Florence, Italy

[12]Programme Mammographie, Direction de la Santé, Luxembourg

[13]The Cancer Registry of Norway, Oslo, Norway

[14]Asseco Poland S.A.

[15]Liga Portuguesa Contra o Cancro, Lisbon, Portugal

[16]BreastCheck, National Cancer Screening Service, Dublin, Ireland

[17]Subdirección de Gestión Clínica y Calidad, Servicio de Salud del Principado de Asturias, Spain

[18]Dirección General de Salud Pública y Participation, Conselleria de Salut i Consum, Baleares, Spain

[19]Dirección General de Innovación y Gestión de la Salud Pública, Consellería de Sanidade, Galicia, Spain

[20]Instituto de Salud Pública, Breast Cancer Screening Program, Navarra, Spain

[21]Programa Detección Precoz Cáncer de Mama, Direccion Territorial de Sanidad de Bizkaia, Bilbao – Spain

[22]General Directorate Research and Public Health & Centre for Public Health Research (CSISP), Valencia, Spain

[23]Mälarsjukhuset, Eskilstuna, Sweden

[24]Cancer Screening Unit, Stockholm, Sweden

[25]Brostenheten Västerås Central Hospital, Västmanland, Sweden

[26]Swiss Federation of Breast Cancer Screening, Bern, Switzerland

[27]Department of Public Health, NETB, Erasmus University Medical Centre, Rotterdam, The Netherlands

[28]NHS Cancer Screening Programmes, Sheffield, UK

[29]Quality Assurance Group, Early Detection and Prevention Section, International Agency for Research on Cancer, Lyon, France

EUROSCREEN WORKING GROUP

Coordinators:

E Paci, M Broeders, S Hofvind, S W Duffy

Members:

Ancelle-Park, R (F)[1], Armaroli P (I)[2], Ascunce N (E)[3], Bisanti, L (I)[4], Bellisario C (I)[2], Broeders M (NL)[5], Cogo C (I)[6], De Koning H (NL)[7], Duffy S W (UK)[8], Frigerio A (I)[2], Giordano L (I)[2], Hofvind S (N)[9], Jonsson H (S)[10], Lynge E (DK)[11], Massat N (UK)[8], Miccinesi G (I)[12], Moss S (UK)[8], Naldoni C (I)[13], Njor S (DK)[11], Nyström L (S)[14], Paap E (NL)[5], Paci E (I)[12], Patnick J (UK)[15], Ponti A (I)[2], Puliti D (I)[12], Segnan N (I)[2], Von Karsa L (D)[16], Törnberg S (S)[17], Zappa M (I)[12], Zorzi M (I)[6]

Affiliations:

[1]Ministère du travail de l'emploi et de la santé, Paris, France

[2]CPO-Piedmont, Turin, Italy

[3]Navarra Breast Cancer Screening Programme, Pamplona, Spain

[4]S.C. Epidemiologia, ASL di Milano, Italy

[5]Radboud University Nijmegen Medical Centre & National Expert and Training Centre for Breast Cancer Screening, Nijmegen, The Netherlands

[6]Veneto Tumor Registry, Padua, Italy

[7]Erasmus MC, Dept. Of Public Health Rotterdam, The Netherlands

[8]Wolfson Institute of Preventive Medicine, Queen Mary University of London, UK

[9]Cancer Registry of Norway, Research Department and Oslo and Akershus University College of Applied Science, Oslo, Norway

[10]Department of Radiation Sciences, Oncology, Umeå University, Umeå, Sweden

[11]Centre for Epidemiology and Screening, Universtiy of Copenhagen, Denmark

[12]ISPO Cancer Research and Prevention Institute, Florence, Italy

[13]Regional Cancer Screening Center, Emilia-Romagna Region, Bologna, Italy

[14]Department of Public Health and Clinical Medicine, Division of Epidemiology and Global Health, Umeå University, Umeå, Sweden

[15]NHS Cancer Screening Programmes and Oxford University, UK

[16]International Agency for Research on Cancer, Lyon, France

[17]Stockholm Cancer Screening, Stockholm, Sweden

Authors' affiliations

Solveig Hofvind, Researcher, Department of Research, Cancer Registry of Norway, Oslo, Norway

Antonio Ponti, MD MPH, Epidemiologist, Epidemiology Unit, CPO Piemonte, AOU S. Giovanni Battista, Turin, Italy

Julietta Patnick, Director, NHS Cancer Screening Programmes, Sheffield, UK

Nieves Ascunce, Public Health Doctor, Navarra Breast Cancer Screening Programme. Spanish Cancer Screening Network, Public Health Institute, Pamplona, Spain

Sisse Njor, Post Doc, Centre for Epidemiology and Screening, University of Copenhagen, Copenhagen, Denmark

Mireille Broeders, Senior Epidemiologist, Department of Epidemiology, Biostatistics and HTA, Radboud University Nijmegen Medical Centre, and National Expert and Training Centre for Breast Cancer Screening, Nijmegen, The Netherlands

Livia Giordano, MD MPH, Epidemiologist, Epidemiology Unit, CPO Piemonte, AOU S. Giovanni Battista, Turin, Italy

Alfonso Frigerio, Radiologist, Regional Reference Centre for Breast Cancer Screening, AOU S. Giovanni Battista, Turin, Italy
Sven Törnberg, Oncologist and Director, Cancer Screening Unit, Oncologic Centre S3:00, Karolinska University Hospital, Stockholm, Sweden
The EUNICE Project and Euroscreen Working Groups (Members of the EUNICE Project and Euroscreen Working Groups listed at end of paper)

ACKNOWLEDGEMENTS

The EUNICE breast cancer screening survey was financially supported by the European Commission (grant agreements no. 2004114, European Network for Information on Cancer, and no. 2006322, European Cancer Network). Financial support from the EU Public Health Programme (project no. 2006322, European Cooperation on development and implementation of Cancer screening and prevention Guidelines [ECCG-ECN]) was also provided for the preparation of this report. The contribution of Mariano Tomatis (CPO Piemonte) in designing the EUNICE data warehouse and in data management and analysis is gratefully acknowledged. Financial support was provided by the National Monitoring Italian Centre (ONS) to host the EUROSCREEN meetings in Florence in November 2010 and in March 2011 and the supplement publication and the National Expert and Training Centre for Breast Cancer Screening, Nijmegen, the Netherlands to host a meeting of the EUROSCREEN mortality working group in July 2011.

REFERENCES

1 Rimer BK, Bluman LG. The psychosocial consequences of mammography. *J Natl Cancer Inst Monogr* 1997;**22**:131–8
2 Scaf-Klomp W, Sanderman R, van de Wiel HB, Otter R, van den Heuvel WJ. Distressed or relieved? Psychological side effects of breast cancer screening in The Netherlands. *J Epidemiol Community Health* 1997;**51**:705–10
3 Brett J, Austoker J. Women who are recalled for further investigation for breast screening: Psychological consequences 3 years after recall and factors affecting re-attendance. *J Public Health Med* 2001;**23**:292–300
4 Brett J, Bankhead C, Henderson B, Watson E, Austoker J. The psychological impact of mammographic screening. A systematic review. *Psychooncology* 2005;**14**:917–38
5 van der Steeg AF, Keyzer-Dekker CM, De Vries J, Roukema JA. Effect of abnormal screening mammogram on quality of life. *Br J Surg* 2011;**98**:537–42
6 Brewer NT, Salz T, Lillie SE. Systematic review: The long-term effects of false-positive mammograms. *Ann Intern Med* 2007;**146**:502–10
7 Salz T, Richman AR, Brewer NT. Meta-analyses of the effect of false-positive mammograms on generic and specific psychosocial outcomes. *Psychooncology* 2010;**19**:1026–34
8 Barton MB, Morley DS, Moore S, *et al.* Decreasing women's anxieties after abnormal mammograms: A controlled trial. *J Natl Cancer Inst* 2004;**96**:529–38
9 Lampic C, Thurfjell E, Sjödén PO. The influence of a false-positive mammogram on a woman's subsequent behaviour for detecting breast cancer. *Eur J Cancer* 2003;**39**:1730–7
10 Hofvind S, Thoresen S, Tretli S. The cumulative risk of a false-positive recall in the Norwegian Breast Cancer Screening Program. *Cancer* 2004;**101**:1501–7
11 Andersen SB, Vejborg I, von Euler-Chelpin M. Participation behaviour following a false positive test in the Copenhagen mammography screening programme. *Acta Oncol* 2008;**47**:550–5
12 Alamo-Junquera D, Murta-Nascimento C, Macià F, *et al.* Effect of false-positive results on reattendance at breast cancer screening programmes in Spain. *Eur J Public Health* 2012;**22**:404–8
13 Salz T, DeFrank JT, Brewer NT. False positive mammograms in Europe: Do they affect reattendance? *Breast Cancer Res Treat* 2011;**127**:229–31
14 Fracheboud J, de Koning HJ, Boer R, *et al.* Nationwide breast cancer screening programme fully implemented in The Netherlands. *Breast* 2001;**10**:6–11
15 Smith-Bindman R, Ballard-Barbash R, Miglioretti DL, Patnick J, Kerlikowske K. Comparing the performance of mammography screening in the USA and the UK. *J Med Screen* 2005;**12**:50–4
16 Hofvind S, Moller B, Thoresen S, Ursin G. Use of hormone therapy and risk of breast cancer detected at screening and between mammographic screens. *Int J Cancer* 2006;**118**:3112–7
17 Njor SH, Hallas J, Schwartz W, Lynge E, Pedersen AT. Type of hormone therapy and risk of misclassification at mammography screening. *Menopause* 2011;**18**:171–7
18 Román R, Sala M, Salas D, Ascunce N, Zubizarreta R, Castells X, Cumulative False Positive Risk Group. Effect of protocol-related variables and women's characteristics on the cumulative false-positive risk in breast cancer screening. *Ann Oncol* 2012;**23**:104–11
19 Otten JD, Karssemeijer N, Hendriks JH, *et al.* Effect of recall rate on earlier screen detection of breast cancers based on the Dutch performance indicators. *J Natl Cancer Inst* 2005;**97**:748–54
20 Perry N, Broeders M, de Wolf C, Törnberg S, Holland R, von Karsa L. *European guidelines for quality assurance in breast cancer screening and diagnosis.* 4th edn. Luxembourg: Office for Official Publications of the European Communities, 2006
21 Giordano L, von Karsa L, Tomatis M, *et al.* M. Mammographic screening programmes in Europe: Organisation, coverage and participation. *J Med Screen* 2012;This issue
22 Castells X, Molins E, Macià F. Cumulative false positive recall rate and association with participant related factors in a population based breast cancer screening programme. *J Epidemiol Community Health* 2006;**60**:316–21
23 Njor SH, Olsen AH, Schwartz W, Vejborg I, Lynge E. Predicting the risk of a false-positive test for women following a mammography screening programme. *J Med Screen* 2007;**14**:94–7
24 Salas D, Ibáñez J, Román R, *et al.* Effect of start age of breast cancer screening mammography on the risk of false-positive results. *Prev Med* 2011;**53**:76–81
25 Puliti D, Miccinesi G, Zappa M. More on screening mammography. *N Engl J Med* 2011;**364**:284–5
26 Ascunce N, Ederra M, Delfrade J, *et al.* Impact of intermediate mammography assessment on the likelihood of false-positive results in breast cancer screening programmes. *Eur Radiol* 2012;**22**:331–40
27 Alberdi RZ, Llanes AB, Ortega RA, *et al.* Effect of radiologist experience on the risk of false-positive results in breast cancer screening programs. *Eur Radiol* 2011;**21**:2083–90
28 Giordano L, Cogo C, Patnick J, Paci E. Communicating the balance sheet. *J Med Screen* 2012;This issue
29 Broeders MJ, Scharpantgen A, Ascunce N, *et al.* Comparison of early performance indicators for screening projects within the European Breast Cancer Network: 1989–2000. *Eur J Cancer Prev* 2005;**14**:107–16
30 Lynge E, Olsen AH, Fracheboud J, Patnick J. Reporting of performance indicators of mammography screening in Europe. *Eur J Cancer Prev* 2003;**12**:213–22
31 Dean PB, Pamilo M. Screening mammography in Finland – 1.5 million examinations with 97 percent specificity. Mammography Working Group, Radiological Society of Finland. *Acta Oncol* 1999;**38**(Suppl. 13):47–54
32 Yankaskas BC, Klabunde CN, Ancelle-Park R, *et al.* International comparison of performance measures for screening mammography: Can it be done? *J Med Screen* 2004;**11**:187–93
33 Majek O, Danes J, Skovajsova M, *et al.* Breast cancer screening in the Czech Republic: Time trends in performance indicators during the first seven years of the organised programme. *BMC Public Health* 2011;**11**:288
34 Baker SG, Erwin D, Kramer BS. Estimating the cumulative risk of false positive cancer screenings. *BMC Med Res Methodol* 2003;**3**:11
35 Xu JL, Fagerstrom RM, Prorok PC, Kramer BS. Estimating the cumulative risk of a false-positive test in a repeated screening program. *Biometrics* 2004;**60**:651–60
36 Hubbard RA, Miglioretti DL, Smith RA. Modelling the cumulative risk of a false-positive screening test. *Stat Methods Med Res* 2010;**19**:429–49
37 Elmore JG, Barton MB, Moceri VM, Polk S, Arena PJ, Fletcher SW. Ten-year risk of false positive screening mammograms and clinical breast examinations. *N Engl J Med* 1998;**338**:1089–96
38 Hofvind S, Vacek PM, Skelly J, Weaver DL, Geller BM. Comparing screening mammography for early breast cancer detection in Vermont and Norway. *J Natl Cancer Inst* 2008;**100**:1082–91
39 Hubbard RA, Kerlikowske K, Flowers CI, Yankaskas BC, Zhu W, Miglioretti DL. Cumulative probability of false-positive recall or biopsy recommendation after 10 years of screening mammography: A cohort study. *Ann Intern Med* 2011;**18**:481–92

III – Results

3. Performance indicators evaluation of the population-based breast cancer screening programme in Northern Portugal using the European Guidelines

III – Results

3. Performance indicators evaluation of the population-based breast cancer screening programme in Northern Portugal using the European Guidelines

The monitoring of the screening programme through comparison of the performance and impact indicators with the standard European Guidelines assess the quality of screening and provide means to predict the mortality outcome in the early years,[17] as it was referred in points 2.4.1 and 2.4.2 of the Background section.

The objective of **paper V** was to evaluate the first 10 years of operation of the population-based BCSP implemented in the Northern Region of Portugal. A number of performance and impact indicators were chosen and they were compared with the desirable and acceptable levels reported in the European Guidelines.[24] The selected indicators were also reported in most of the international publications on this subject.

The determination of the background incidence for the period 2000-2009 in those municipalities covered by the screening programme was based on the trend observed in 1995-1999 for the districts where the screening was introduced. Data from RORENO was used to compute observed incidence rates and to ascertain interval invasive breast cancers. Except for the recall rate, most of the performance indicators evaluated were consistent with the desirable levels of the European Guidelines: the screening programme was highly accepted by the population, it was detecting the expected number of invasive breast cancers, and it was able to detect small size breast cancers among the participants. Although not sufficient, these results are considered necessary for the expected future reduction in mortality.[17,49]

Title:

Performance indicators evaluation of the population-based breast cancer screening programme in Northern Portugal using the European Guidelines

Authors names and affiliations:

Maria José Bento[a,b], MD MSc

[a]Institute of Biomedical Sciences Abel Salazar, University of Porto,

[b]Epidemiology Unit, Portuguese Oncology Institute,

Guilherme Gonçalves[a], MD PhD

[a]Institute of Biomedical Sciences Abel Salazar, University of Porto,

Ana Aguiar[c], MD

[c]Portuguese Cancer League – North Branch, Porto

Clara Castro[b], MSc

[b]Epidemiology Unit, Portuguese Oncology Institute,

Vitor Veloso[c], MD

[c]Portuguese Cancer League – North Branch, Porto

Vítor Rodrigues, MD PhD[e,f]

[e]Faculty of Medicine, University of Coimbra,

[f]Portuguese Cancer League – Centre Branch, Coimbra

Corresponding author and requests for reprints

Maria José Bento, MD MSc

Institute of Biomedical Sciences Abel Salazar, University of Porto,

Rua de Jorge Viterbo, nº 228, 4050-313 Porto, Portugal

mjbento@icbas.up.pt

Telephone: 00 351 225084067

Fax: 00 351 225084004

Abstract

Objective: To evaluate the first 10 years of operation of the population-based breast cancer screening programme implemented in the Northern Region of Portugal, using selected recommended standard performance indicators.

Methods: Data from women aged 50-69 screened with two-view mammography, biennially, in the period 2000-2009, were included. Main performance indicators were compared with the recommended levels of the European Guidelines.

Results: A total of 202,039 screening examinations were performed, 71731 (35.5%) in the prevalent screening and 130,308 (64.5%) in the subsequent screening. Coverage rate by examination reached 74.3% of the target population, in the last period evaluated. Recall rates were 8.1% and 2.4% and cancer detection rates were 4.4/1000 and 2.9/1000 respectively, for prevalent and subsequent screenings The breast detection rate of invasive cancer, expressed as a multiple of the background expected incidence was 3.1 in prevalent screen and 2.2 in subsequent screen. The incidence of invasive interval cancers met the desirable recommended levels both the first and second years since last screening examination, in the prevalent and subsequent screenings. Invasive tumours <15 mm were 50.4% and 53.8% of the invasive cancers detected in prevalent and subsequent screenings. Less favourable size, grading and biomarkers expression were found in interval cancers compared to screen-detected cancers.

Conclusions: Breast cancer screening programme in the Northern Region of Portugal was well accepted by the population. Most of the performance indicators were consistent with the desirable levels of the European Guidelines, which indicate an effective screening programme. Future research should verify the consistency of some of these results by using updated information from a larger population.

Keywords: breast cancer; performance indicators; population-based screening; mammography

1. Introduction

The main objective of breast cancer screening is to reduce mortality due to the disease. That is achieved by identifying breast cancer at a stage when it is more curable and the probability of the disease being disseminated is smaller [1-4]. Because screening programmes must be operating for many years before breast cancer mortality reduction can be verified, some early surrogate measures are considered valuable indicators of future changes in mortality rates [5]. In 2006, the European Union published the 4th edition of the European Guidelines on Quality Assurance in Breast Cancer Screening and Diagnosis [6] that have been used in the evaluation of specific European programmes [7, 8]. A good screening programme is a complex organisation with multiple steps encompassing the entire screening process and, to maintain a high-quality service, it requires continuous supervision and regular reporting of rigorous scientific studies [4, 9, 10]. These studies, also contribute with evidence-based knowledge on the risks and benefits of implementing breast cancer screening programmes in populations with diverse health systems and economic constraints [10].

In Portugal, the organized population-based breast cancer screening was initiated in 1990 in the Centre Region, conducted by the Portuguese Cancer League (*Liga Portuguesa Contra o Cancro* – LPCC) a private, non-profit organization; in the Northern Region an identical programme started in October 1999 [11, 12]. The programme is financed by the National Health Service (NHS). It is co-ordinated by the North Regional Health Administration (of the NHS) and operated by the north branch of LPCC. The organized programme coexists with opportunistic breast cancer screening; data on opportunistic screening is very limited [13].

Since the beginning of the programme, quality assurance was considered a priority, in accordance with the European Guidelines [6]. The objective of this study was to evaluate the breast cancer screening programme during the first 10 years of operation, in the Northern Region of Portugal, using some of the standard performance indicators recommended [6].

2. Material and methods

The population-based breast cancer screening implemented from October 1999 to 2009 in the Northern Region of Portugal started in one municipality and gradually expanded to 43 municipalities (mostly rural communities) by the end of 2009. The programme included a personal invitation sent by mail with a pre-booked appointment to resident women aged 45-69 years. Most Portuguese people are registered in their local Health Centers (from NHS); thus the lists of users enrolled in these Health Centers are very good proxies of a non-existent computerized residents databases [13]. Names and addresses of women registered in the local Health Centres were provided to LPCC, who managed the invitation process. Women not registered in these lists were also invited to participate through advertising, contacts with health professionals and community stakeholders. Mammography was offered free of charge and was performed at one of six mobile units or at one fixed facility. From the beginning screening procedures included bilateral mammography with two-view (craniocaudal, mediolateral oblique), centralized and independently read by two radiologists, with a final reading by a third independent and experienced radiologist, in case of discrepancy. Women with an abnormal screen mammography were recalled and reassessment was carried out by a multidisciplinary team (radiologist, surgeon and pathologist) at a dedicated clinical setting outside the hospital. Cancer treatment was established according to standardized hospital therapeutic protocols. Screen film mammography was used till mid-2007, and thereafter, computed radiography was performed in all screening units. The screening interval was 24 months.

Data collected within this programme, from the invitation process till the follow-up of cancer cases, were actively gathered and entered in a database centralized at LPCC. Information on individual women was checked for accuracy and completeness, before being introduced in the computer database. Data were organized by prevalent and subsequent screening; the latter includes regular and irregular screening [6]. For comparability with the European Guidelines, analysis shown here was restricted to women aged 50-69. Women enrolled in 1999 were not included in this analysis due to very low figures.

Coverage rate by examination was estimated using the ratio between the number of examinations within the organized screening programme and the number of eligible women during a two-year period [8]. Denominator was derived from the census data provided by the office of Statistics Portugal. In the last period of the study, the methodology to personally invite women to attend screening was definitively established and the participation rate [6] was also calculated.

Expected incidence rate in the absence of screening for the period 2000-2009 was defined as the predicted incidence rate of invasive breast cancer based on the trends observed in 1995-1999, from the districts where cancer screening was implemented. Using the database from the population-based North Region Cancer Registry (*Registo Oncológico Regional do Norte* – RORENO), a Poisson regression model was used to calculate the breast cancer incidence for women aged 50-69 and it was estimated as 1.23/1000, with an annual percent increase of 2.1% [6, 14].

Interval cancer (IC) was defined as breast cancer diagnosed in a woman who had a screening test, with/without further assessment, which was negative either before the next screening invitation or within a time period equal to a screen interval for women who have reached the upper age limit [6, 15]. To evaluate interval cancers, screening data were linked to the RORENO database but their ascertainment was limited to women participating in the screening programme from 2000-2007 as the Cancer Registry only had complete information on breast cancer incidence till 2009. The interval cancer rate was calculated as the number of interval cancers divided by the total number of screens within a specified time. To evaluate the proportion of interval cancers related to the background (expected) incidence, the estimates were made for the years 2000-2007 (estimated rate 1.20/1000). Interval cancers were divided in two groups according to time (in months) since screening examination [6]: 0-11 and 12-23. Radiological review of last screening/assessment imaging and diagnostic mammography of the interval cancer was not performed. Information on the maximum dimension of invasive tumour and expression of biomarkers as estrogen (ER) and

progesterone (PR) receptors and human epidermal growth factor receptor 2 (HER2) were gathered from the pathology reports.

Results are given in numbers, proportions and rates. Calculated parameters are displayed by two-year periods, from 2000 to 2009. As the implementation of the programme was gradual, not all the municipalities contributed with the same number of screening rounds. In the 2000-2001 period of evaluation 8 municipalities were covered, and in the next 4 periods 19, 32, 36 and 43 municipalities were included, respectively. Performance indicators used in this evaluation were: coverage rate by examination, recall rate, cancer detection rate, positive predictive value (PPV), ratio benign/malignant, interval cancer rate, tumour maximum dimension and tumour grade, which were calculated using the recommended standard definitions [6].

3. Results

From 2000 to 2009, a total of 202,039 screening examinations were performed, 71731 (35.5%) in the prevalent screening and 130,308 (64.5%) in the subsequent screening. By the end of 2009, 43 municipalities were covered by the screening programme, corresponding to 50% of the total number of municipalities in the Northern Region but comprised only 17% of the women aged 50-69 and resident in this region.

The coverage rate by examination was 47.2% at the start of the programme and in the next four periods of screening its values were 65.6%, 67.7%, 72.0% and 74.3%. Participation rate in the period 2008/2009 was 74.5%.

Results from the evaluation of the performance indicators according to prevalent and subsequent screening are shown in Tables 1 and 2.

The recall rate was 8.1% in prevalent screening, ranging from 6.3% in 2000/2001 to 12.5% in 2008/2009. In subsequent screening the recall rate was 2.4% with little variation along the 10 years. At initial screening, a total of 312 breast cancers were diagnosed corresponding to a detection rate of 4.4/1000 participants. In the subsequent screening a total of 374 cancers were detected (detection rate of 2.9/1000). The breast detection rate of invasive cancer, expressed as a multiple of the background expected incidence was 3.1 in

prevalent screen (3.7/1.2) and 2.2 (2.6/1.2) in subsequent screen (desirable levels in the European Guidelines are >3 for prevalent screen and >1.5 for subsequent screen).

The recall rates decreased with age and the detection rates showed the opposite pattern, both in prevalent and subsequent screenings – fig 1 and 2. In the prevalent screening, the increase in recall rates as the programme progressed was followed by a parallel increase in the rate of cancer detection, except for the 2008/2009 period (table 1). In subsequent screen, no patterns of tendencies were observed (table 2).

The invasive tumours in the prevalent and subsequent screening comprised 85.9% and 89.6% of the malignancies, respectively. At initial screening, the detection rate for invasive tumours increased over time except for the last period considered but ductal carcinoma *in situ* (DCIS) was relatively stable; for subsequent screening there was little variation in the invasive and DCIS rates.

The positive predictive value (PPV) was smaller at the prevalent screening compared with subsequent screening (5.6% vs.12.5%). With exception of the first period considered in both types of screening, time trends showed that as long as recall rate was getting higher, the PPV value was getting lower.

The ratio of benign lesions to malignant lesions surgically removed met the desirable recommendation from the European Guidelines, in both types of screening. In the prevalent screening smaller values were found in more recent years and for subsequent screening consistent low ratios were verified in all periods of evaluation.

A total of 112 invasive interval cancers were identified in the period 2000-2007, representing 21.8% (51 cases) of the screen detected invasive breast cancers in the prevalent and 29.3% (61 cases) in the subsequent screens. For prevalent screen, the interval cancer rate as a proportion of the expected breast cancer incidence was 18.3% within the first 0-11 months and 49.2% for the 12-23 months; for subsequent screen, these values were 15.0% and 49.6%, respectively (data not shown). The European Guidelines desirable level for this indicator was < 30% within the first 12 months and < 50% for the second 12 months.

Pathological characterization and biomarkers expression of the detected cancers in prevalent and subsequent screening are presented in tables 1, 2 and 3.

Overall, 34.8% of invasive cancers detected in the prevalent screening were less than 11 mm, 50.4% less than 15 mm, 85.5% were histology grade I or II. Biomarker status evaluation classified 87.1% of the cancers as ER positive and 75.3% as PR positive. HER2 overexpression was identified in 13.2% of the 106 cases searched for. The triple negative phenotype was verified in 6.6% cases (7/106).

For subsequent screening, 28.6% of invasive cancers were equal or less than 10 mm, 53.8% less than 15 mm, 77.3% were grade I or II. The overall ER positive status was 87.8%, PR positive 79.0% and HER2 overexpression was identified in 14.7% (34/232) cases. The triple negative phenotype was verified in 6.9% cases (16/232).

In invasive interval cancers, tumour size ≤10mm was verified in 7 cases out of 106 (6.6%), and 23.3% were <15mm (20/86). The proportion of histology grade I or II was 61.1% (66/108 cases). Biomarkers expression was less favourable in interval cancers (table 3), with a high proportion of triple negative of 20.4% (20 in 98 cases).

The mean size for invasive cancers in prevalent screening was slightly smaller than in subsequent screening (15.6 mm and 16.3 mm, respectively). For the interval cancers, the mean tumour size for 70 cases was 25.6 mm.

In screen-detected cancers, the proportions of invasive cancers ≤10mm didn´t show any tendency over time but it was very low in 2008/2009 in prevalent screening, compared with the previous years. For this indicator, subsequent screening had lower proportions than those verified in prevalent screening, except for the last period. However, these proportions values remained above the acceptable level (≥ 25%), according to the European Guidelines.

4. Discussion

Presently, breast cancer screening is under intense scrutiny. Although some authors consider that the benefits of screening overwhelm its disadvantages, others consider that the reduction in mortality was mainly due to advances in treatment and that overdiagnosis and overtreatment are major drawbacks [3, 16-22]. While the impact of screening services on

breast cancer mortality reduction takes many years before it can be evaluated, monitoring quality indicators should be an integral part of a breast cancer screening programme with a commitment on reduction of mortality rates of breast cancer in women [20, 23, 24].

This is the first comprehensive report on the population-based breast cancer screening programme implemented in the Northern Region of Portugal since it began in 1999. It covers the first ten years of operation of the programme and attempts to estimate its performance indicators and compare them with the European Guidelines. Some results must be carefully considered because of small numbers, but the various indicators evaluated and the chosen time frame allowed a deeper understanding of the screening process.

The long-term effect of screening is, to a large extent, dependent on coverage [25, 26]. In this evaluation, the coverage rate was low at start, but over the years the programme has been receiving increased acceptance and the coverage rate reached 74.3% of the eligible women in 2008/2009. In this last period the participation rate was very similar to the coverage rate which could somehow validate the information provided for the previous years, using only the data from the screening programme and from the population estimates of the years between census. Both the coverage and participation rates were compliant with the European Guidelines, largely exceeding the acceptable levels proposed (of over 70%).

The recall rate in the prevalent screening was higher than what is considered desirable by the European Guidelines but was lower than the reported overall recall rate of 9.3% from 20 screening programmes included in the European Network for Information on Cancer - EUNICE project [6, 8, 27]. The PPV were in the range of values provided by the aforementioned project, both in the prevalent and subsequent screenings [27]. Benign to malignant ratio values were well above the desirable level recommended in the Guidelines, since the assessment worked as an efficient classification routine: only a few women underwent surgical procedures before the diagnosis was confirmed [6, 28, 29].

The pattern of higher recall rates along with higher detection rates was verified, with the exception of period 2008/2009. Possible explanations for this exception include the result of the introduction of the computed radiography in the middle of 2007 or, most probably, the

consequence of the screening progression with a higher proportion of women entering the programme at younger ages [3, 6, 25, 30, 31]. For the subsequent screening, the values for these indicators were in accordance with the desirable level proposed by the European recommendations.

Recall rate was higher in younger age groups and the opposite pattern was shown with the detection rate, both indicators with higher values in the prevalent than subsequent screening, in accordance with other studies [7, 21, 26, 27].

Invasive breast cancer detection rate in relation to the expected rate in the absence of screening was in accordance with the highest recommended levels of the European Guidelines for the initial and subsequent screening. Interval cancer rates also met the desirable recommended levels, within both the first and the second years.

The proportion of DCIS in prevalent and subsequent screen was in accordance with the desirable level of the European Guidelines.

Tumour size is one of the strongest predictors of breast cancer behaviour and studies have shown that screening leads to the detection of smaller size tumours [32-37]. The proportions of small-sized tumours (size below 11mm or 15mm) were in accordance with the levels of the European Guidelines both in the prevalent and subsequent screenings.

Histological grading is a major prognostic factor and screen-detected cancers are likely to include a proportion of grade III tumours below 50% [6]. In our study, the proportion of grade III tumours was of 14.5% and 22.7% in the prevalent and subsequent screenings, respectively.

Less favourable size, grading and biomarker expression were found in interval cancers compared to screen-detected cancers [7, 38, 39]. The higher frequency of triple negative interval cancers compared with the screen-detected cancers was also identified in other studies, conveying a more aggressive behaviour and adverse prognosis to these tumours [38, 39].

In conclusion, in the Northern Region of Portugal the organization and implementation of a population-based breast cancer screening programme was well accepted by the

population. Most of the performance indicators were consistent with the desirable levels of the European Guidelines, which indicate an effective screening programme. These findings were obtained during a ten-year programme operating in a small-sized population and future research should verify the consistency of some of these results by using updated information from a larger population.

References

[1] Colditz G, Baer H,Tamimi R. Breast cancer. In: Schottenfeld D, Fraumeni J, editors. Cancer Epidemiology and Prevention. 3rd ed. Oxford: Oxford University Press; 2006. p. 995-1012.

[2] Bray F, McCarron P,Parkin DM. (2004). The changing global patterns of female breast cancer incidence and mortality. Breast Cancer Res. 6(6):229-239.

[3] Warner E. (2011). Breast-Cancer Screening. NEJM. 365(11):1025-1032.

[4] Blamey RW, Wilson AR,Patnick J. (2000). ABC of breast diseases: screening for breast cancer. BMJ. 321(7262):689-693.

[5] Schopper D,de Wolf C. (2009). How effective are breast cancer screening programmes by mammography? Review of the current evidence. Eur J Cancer. 45(11):1916-1923.

[6] Perry N, Broeders M, de Wolf C, Tornberg S, Holland R,von Karsa L, editors. European guidelines for quality assurance in breast cancer screening and diagnosis. 4th ed. Luxembourg: Office for Official Publications of the European Communities; 2006.

[7] Hofvind S, Geller B, Vacek PM, Thoresen S,Skaane P. (2007). Using the European guidelines to evaluate the Norwegian Breast Cancer Screening Program. Eur J Epidemiol. 22(7):447-455.

[8] Giordano L, von Karsa L, Tomatis M, Majek O, de Wolf C, Lancucki L, et al. (2012). Mammographic screening programmes in Europe: organization, coverage and participation. J Med Screen. 19 (Suppl 1):72-82.

[9] Perry N, Broeders M, de Wolf C, Tornberg S, Holland R,von Karsa L. (2008). European guidelines for quality assurance in breast cancer screening and diagnosis. Fourth edition--summary document. Ann Oncol. 19(4):614-622.

[10] Lynge E, Tornberg S, von Karsa L, Segnan N,van Delden JJ. (2012). Determinants of successful implementation of population-based cancer screening programmes. Eur J Cancer. 48(5):743-748.

[11] Klabunde CN,Ballard-Barbash R. (2007). Evaluating Population-Based Screening Mammography Programs Internationally. Semin Breast Dis. 10(2):102-107.

[12] Alves JG, Cruz DB, Rodrigues VL, Goncalves ML,Fernandes E. (1994). Breast cancer screening in the central region of Portugal. Eur J Cancer Prev. 3 (Suppl 1):49-53.

[13] Departamento de Estudos e Planeamento. Relatório de Actvidades 2009. Porto: Administração Regional de Saúde do Norte 2010.

[14] Castro C, Bento MJ, Lunet N,Campos P. (2012). Assessing the completeness of cancer registration using suboptimal death certificate information. Eur J Cancer Prev. 21(5):478-479.

[15] Tornberg S, Kemetli L, Ascunce N, Hofvind S, Anttila A, Seradour B, et al. (2010). A pooled analysis of interval cancer rates in six European countries. Eur J Cancer Prev. 19(2):87-93.

[16] Bock K, Borisch B, Cawson J, Damtjernhaug B, de Wolf C, Dean P, et al. (2011). Effect of population-based screening on breast cancer mortality. Lancet. 378(9805):1775-1776.

[17] Day NE. (2005). Overdiagnosis and breast cancer screening. Breast Cancer Res. 7(5):228-229.

[18] van Schoor G, Moss SM, Otten JD, Donders R, Paap E, den Heeten GJ, et al. (2011). Increasingly strong reduction in breast cancer mortality due to screening. Br J Cancer. 104(6):910-914.

[19] Autier P, Boniol M, Gavin A,Vatten LJ. (2011). Breast cancer mortality in neighbouring European countries with different levels of screening but similar access to treatment: trend analysis of WHO mortality database. BMJ. 343:d4411.

[20] Jorgensen KJ, Keen JD,Gotzsche PC. (2011). Is mammographic screening justifiable considering its substantial overdiagnosis rate and minor effect on mortality? Radiology. 260(3):621-627.

[21] Nelson HD, Tyne K, Naik A, Bougatsos C, Chan BK,Humphrey L. (2009). Screening for breast cancer: an update for the U.S. Preventive Services Task Force. Ann Intern Med. 151(10):727.

[22] Independent UK Panel on Breast Cancer Screening. (2012). The benefits and harms of breast cancer screening: an independent review. Lancet. 380(9855):1778-1786.

[23] Broeders M, Moss S, Nystrom L, Njor S, Jonsson H, Paap E, et al. (2012). The impact of mammographic screening on breast cancer mortality in Europe: a review of observational studies. J Med Screen. 19 (Suppl 1):14-25.

[24] Segnan N, Rosso S,Ponti A. (2012). Is the breast cancer mortality decrease in Sweden due to screening or treatment? Not the right question. J Natl Cancer Inst. 104(14):1040-1041.

[25] Lynge E, Olsen AH, Fracheboud J,Patnick J. (2003). Reporting of performance indicators of mammography screening in Europe. Eur J Cancer Prev. 12(3):213-222.

[26] Sarkeala T, Anttila A, Saarenmaa I,Hakama M. (2005). Validity of process indicators of screening for breast cancer to predict mortality reduction. J Med Screen. 12(1):33-37.

[27] Hofvind S, Ponti A, Patnick J, Ascunce N, Njor S, Broeders M, et al. (2012). False-positive results in mammographic screening for breast cancer in Europe: a literature review and survey of service screening programmes. J Med Screen. 19 (Suppl 1):57-66.

[28] Yankaskas BC, Klabunde CN, Ancelle-Park R, Renner G, Wang H, Fracheboud J, et al. (2004). International comparison of performance measures for screening mammography: can it be done? J Med Screen. 11(4):187-193.

[29] Sarkeala T, Anttila A, Forsman H, Luostarinen T, Saarenmaa I,Hakama M. (2004). Process indicators from ten centres in the Finnish breast cancer screening programme from 1991 to 2000. Eur J Cancer. 40(14):2116-2125.

[30] Moskowitz M. (1996). Retrospective reviews of breast cancer screening: what do we really learn from them? Radiology. 199(3):615-620.

[31] Vinnicombe S, Pinto Pereira SM, McCormack VA, Shiel S, Perry N,Dos Santos Silva IM. (2009). Full-field digital versus screen-film mammography: comparison within the UK breast screening program and systematic review of published data. Radiology. 251(2):347-358.

[32] Nagtegaal ID,Duffy SW. (2013). Reduction in rate of node metastases with breast screening: consistency of association with tumor size. Breast Cancer Res Treat. 137(3):653-663.

[33] Day NE, Williams DR,Khaw KT. (1989). Breast cancer screening programmes: the development of a monitoring and evaluation system. Br J Cancer. 59(6):954-958.

[34] Hofvind S, Wang H,Thoresen S. (2004). Do the results of the process indicators in the Norwegian Breast Cancer Screening Program predict future mortality reduction from breast cancer? Acta Oncol. 43(5):467-473.

[35] Hofvind S, Sorum R,Thoresen S. (2008). Incidence and tumor characteristics of breast cancer diagnosed before and after implementation of a population-based screening-program. Acta Oncol. 47(2):225-231.

[36] Vejborg I, Olsen AH, Jensen MB, Rank F, Tange UB,Lynge E. (2002). Early outcome of mammography screening in Copenhagen 1991-99. J Med Screen. 9(3):115-119.

[37] Broeders MJ, Scharpantgen A, Ascunce N, Gairard B, Olsen AH, Mantellini P, et al. (2005). Comparison of early performance indicators for screening projects within the European Breast Cancer Network: 1989-2000. Eur J Cancer Prev. 14(2):107-116.

[38] Domingo L, Sala M, Servitja S, Corominas JM, Ferrer F, Martinez J, et al. (2010). Phenotypic characterization and risk factors for interval breast cancers in a population-based breast cancer screening program in Barcelona, Spain. Cancer Causes Control. 21(8):1155-1164.

[39] Rayson D, Payne JI, Abdolell M, Barnes PJ, MacIntosh RF, Foley T, et al. (2011). Comparison of clinical-pathologic characteristics and outcomes of true interval and screen-detected invasive breast cancer among participants of a Canadian breast screening program: a nested case-control study. Clin Breast Cancer. 11(1):27-32.

Table 1 – Performance indicators and characteristics of cancers in prevalent screen, in women aged 50-69 years, by 2-year screening period and comparison with European guidelines

Parameters	2000-01	2002-03	2004-05	2006-07	2008-09	Total	European guidelines[a]
Screened women (n)	12910	15826	20573	13041	9381	71731	
Recall rate (%)	6.3	7.1	7.4	8.8	12.5	8.1	<7 / < 5
Women referred to hospital[b] (%)	0.5	0.7	0.7	0.8	0.7	0.7	
Screen cancers (n)	47	69	93	64	39	312	
Screen cancers (/1000)	3.6	4.4	4.6	5.0	4.6	4.4	
invasive (/1000)	2.9	3.7	4.0	4.4	3.6	3.7	
ductal *in situ* (/1000)	0.8	0.7	0.5	0.5	0.5	0.6	
Invasive (%)	78.7	84.1	88.2	89.1	87.2	85.9	90 / 80-90
Positive predictive value[c] (%)	5.8	6.2	6.3	5.9	3.8	5.6	
Ratio benign/malignant	0.23	0.33	0.35	0.16	0.03	0.24	≤1:2 / ≤1:4
Tumour diameter[d] (n)	35	54	80	55	32	256	
% ≤ 10 mm	31.4	40.7	33.8	41.8	18.8	34.8	NA / ≥25
% < 15 mm	45.7	51.9	47.5	58.2	46.9	50.4	50 / >50
% > 20 mm	34.3	24.1	27.5	16.4	15.6	23.8	
Tumour grade[d] (n)	36	52	76	52	31	247	
% grade 1	30.5	34.6	32.9	32.7	22.6	31.6	
% grade 2	55.6	46.2	56.6	51.9	61.3	53.9	
% grade 3	13.9	19.2	10.5	15.4	16.1	14.5	

[a]European Guideline, acceptable/desirable level; [b]number of women referred to hospital after a positive assessment as a proportion of the screened women; [c]number of cancers detected as a proportion of the women with a positive screening test
[d]Invasive tumours only; NA, not applicable

Table 2 – Performance indicators and characteristics of cancers in subsequent screen, in women aged 50-69 years, by 2-year screening period and comparison with European guidelines

Parameters	2002-03	2004-05	2006-07	2008-09	Total	European guidelines[a]
Screened women (n)	9550	23773	45643	51342	130308	
Recall rate (%)	2.2	2.0	2.4	2.6	2.4	<5 / <3
Women referred to hospital[b] (%)	0.4	0.3	0.4	0.4	0.4	
Screen cancers (n)	33	63	133	145	374	
Screen cancers (/1000)	3.5	2.7	2.9	2.9	2.9	
invasive (/1000)	3.2	2.5	2.6	2.5	2.6	
ductal *in situ* (/1000)	0.2	0.1	0.4	0.4	0.3	
Invasive (%)	93.9	95.2	88.0	87.6	89.6	90 / 80-90
Positive predictive value[c] (%)	15.7	13.8	12.5	11.5	12.5	
Ratio benign/malignant	0.15	0.13	0.12	0.08	0.11	≤1:2 / ≤1:4
Tumor diameter[d] (n)	30	58	112	125	325	
% ≤ 10 mm	33.3	29.3	28.6	27.2	28.6	≥25 / ≥30
% < 15 mm	53.3	60.3	50.9	53.6	53.8	50 / >50
% > 20 mm	30.0	15.5	20.5	21.6	20.9	
Tumour grade[d] (n)	31	57	114	124	326	
% grade 1	29.0	22.8	21.9	26.6	24.5	
% grade 2	51.6	52.6	57.9	48.4	52.8	
% grade 3	19.4	24.6	20.2	25.0	22.7	

[a]European Guideline, acceptable/desirable level; [b]number of women referred to hospital after a positive assessment as a proportion of the screened women; [c]number of cancers detected as a proportion of the women with a positive screening test; [d]Invasive tumours only; NA, not applicable

Table 3 – Biomarker expression in prevalent and subsequent screen-cancers and interval cancers

Parameters	Prevalent screen cancers	Subsequent screen cancers	Interval cancers
ER status (n)	263	329	107
% Positive	87.1	87.8	71.0
% Negative	12.9	12.2	29.0
PR status (n)	263	329	107
% Positive	75.3	79.0	60.7
% Negative	24.7	21.0	39.3
HER2 status (n)	106	232	98
% Positive	13.2	14.7	21.4
% Negative	86.8	85.3	78.6

Fig. 1 Recall rate (%) by five-year age group and type of screen

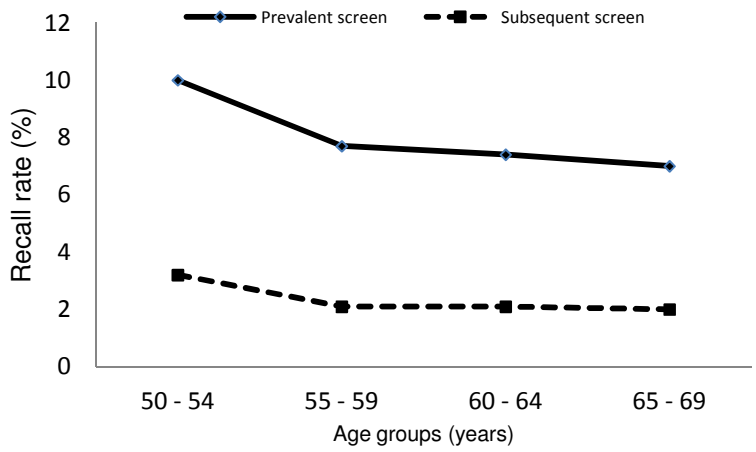
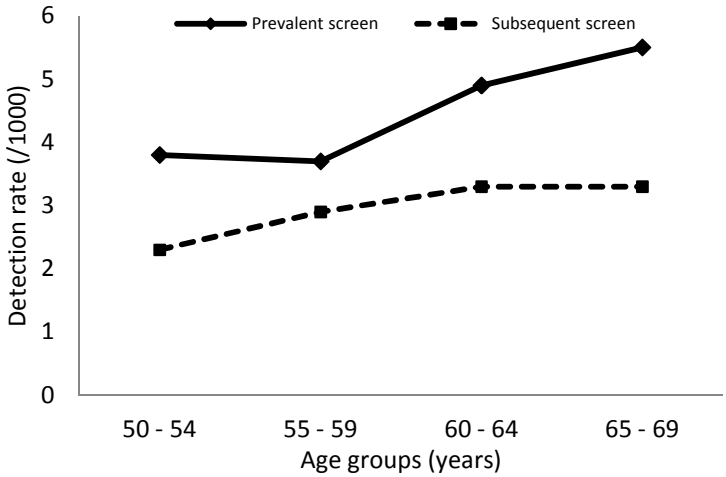


Fig.2 Breast cancer detection rate per 1000 by five-year age group and type of screen



III – Results

4. Clinicopathological differences between interval and screen-detected breast cancers diagnosed within a screening programme in Northern Portugal

III – Results

4. Clinicopathological differences between interval and screen-detected breast cancers diagnosed within a screening programme in Northern Portugal

As stated in the Background section (point 2.6.2) interval cancers are an important indicator of the quality of mammography as well as of the likely impact of the screening programmes on breast cancer.[81,82] The clinicopathological characterization of interval cancers is essential given its relation with the mortality from the disease.[91] This was the subject for **paper VI**.

Monitoring of interval cancers occurrence is a crucial part of the evaluation of a mammography screening programme[24] for the reason that it provides a mechanism to evaluate some of the technical processes involved in the screening, as performance and interpretation of the mammography, and it contributes to the evaluation of the impact mammography screening on breast cancer in the target population.[100,126]

The revision of screening and diagnostic mamography necessary to classify the interval cancers was not performed in this paper; some authors considered this task a necessary tool for continuous education of radiologists but of questionable usefulness the retrospective reviews to separate the interval cancer categories, as it is highly dependent on the adopted methodology.[85]

Due to the time lag between the diagnosis of the disease and registration at RORENO, the evaluation was restricted to the screening period of 2000 to 2007, with data from the registry completed till cancers diagnosed in 2009.

The objective of **paper VI** was to compare the clinicopathological characteristics of interval breast cancers to the screen-detected cancers. Of the results achieved it was noteworthy the higher size of interval cancers, the higher grade and less frequent oestrogen receptor positivity found in these cancers compared to screen-detected. This pattern is also described in the majority of the international studies.[94,100,127-132]

This study was only possible thanks to the extensive network of pathologists and hospital cancer registries that collaborated with RORENO in providing quality data.

Original Article

J Med Screen
2014, Vol. 21(2) 104–109

DOI: 10.1177/0969141314534406
msc.sagepub.com
SAGE

Clinicopathological differences between interval and screen-detected breast cancers diagnosed within a screening programme in Northern Portugal

Maria José Bento[1], Guilherme Gonçalves[2], Ana Aguiar[3], Luis Antunes[2], Vitor Veloso[3] and Vítor Rodrigues[4]

Abstract

Objective: To evaluate clinicopathological differences between screen-detected (SD) and interval (IC) breast cancers diagnosed in women enrolled in an organized breast screening programme in 2000–2007.

Setting: Breast Cancer Screening Programme of the north region of Portugal.

Methods: Using data from the screening programme and from the population-based North Region Cancer Registry, SD and IC were identified. Information on screening history, age, date of diagnosis, tumour size, histological type and grade, lymph node status, tumour stage, biomarkers, and treatment was obtained from the cancer registry and from clinical and pathological reports. Association between mode of detection and these clinicopathological characteristics was estimated by unconditional logistic regression.

Results: A total of 442 SD and 112 IC were identified in women aged 50–69. Compared with SD, IC were diagnosed in younger women (60.0 ± 5.8 years and 58.4 ± 6.0 years, respectively), were larger (tumour size >20 mm: 60.2% *versus* 25.1%), lobular (6.3% *versus* 16.1%), with a higher differentiation grade (grade 3: 17.7% *versus* 38.9%), had more lymph node metastases, more advanced stage, and oestrogen receptor (ER) negative (12.9% *versus* 29.0%) and progesterone negative, and HER2 positive. After multivariable analysis, compared with SD, IC were more likely to be larger than 20 mm, lobular, of grade 3 and negative for ER.

Conclusion: Our results are consistent with other studies. IC's have a more aggressive biology than SDs. Our findings did not show any unexpected pattern requiring changes to our screening procedures, but continuous identification and characterization of IC is advisable.

Keywords
interval cancer, organized screening, breast cancer

Date received: 20 December 2013; accepted: 15 April 2014

Introduction

Good practices in organized breast cancer screening include the identification and characterization of interval breast cancers (IC), those cancers that occur after a negative mammogram and before the scheduled date for the next examination.[1] In comparison with the evaluation of other routine performance indicators of screening, the identification of IC is a more complex task, relying on the availability of a population-based cancer registry with high coverage, operating continuously over time.[2–4] In the absence of a cancer registry, it is also possible to consult hospital data (including that from pathology units) in order to identify and characterize all breast cancers in the screened population.[2,5,6]

IC are related both to the sensitivity of the screening programme and the length of interval between screening rounds, and are expected events in any screening programme.[1,2,7,8] In general, compared with screen-detected cancers (SD), IC consist of larger tumours, with lymph node invasion, more advanced stage, and higher grade[9,10], but not all studies are consistent and some have shown different patterns.[5,11,12] Although there may be some controversy on the interpretation of the findings,

[1]Institute of Biomedical Sciences Abel Salazar, University of Porto, and Epidemiology Unit, Portuguese Oncology Institute, Porto, Portugal
[2]Institute of Biomedical Sciences Abel Salazar, University of Porto, Portugal
[3]Portuguese Cancer League – North Branch, Porto, Portugal
[4]Faculty of Medicine, University of Coimbra, and Portuguese Cancer League – Centre Branch, Coimbra, Portugal

Corresponding author:
Maria José Bento, Institute of Biomedical Sciences Abel Salazar, University of Porto, and Epidemiology Unit, Portuguese Oncology Institute, Porto, 4200-072, Portugal.
Email: mjbento@icbas.up.pt

IC are an heterogeneous group of tumours, the clinical and pathological features of which should be characterized in order to assess their relative contribution to the efficacy of breast cancer screening.[13–15]

The organized population-based breast cancer screening programme, conducted by the north branch of the Portuguese Cancer League (*Liga Portuguesa Contra o Cancro* – LPCC), was initiated in the Northern Region of Portugal in 1999. Preliminary results from the programme have been published, and considered satisfactory.[16]

The aim of this study was to evaluate the clinicopathological differences between SD and IC diagnosed in the northern region of Portugal, in women enrolled in the screening programme between 2000 and 2007, within the context of evaluation of the screening programme.

Methods

By the end of 2009 the screening programme had expanded to 43 municipalities, covering 17% of the 454666 eligible women resident in the Northern Region of Portugal.[17] Women aged between 45 and 69 were offered biennial screening with two-view mammography, with double independent reading of the examinations by trained radiologists and, in case of discrepancy, a final reading by a third independent and experienced radiologist.

SD was defined as a breast cancer diagnosed after a positive screening result. IC was defined as a breast cancer diagnosed in a woman who had a screening test, with or without further assessment, which was negative either before the next screening invitation or within a time period equal to a screen interval for women who have reached the upper age limit.[1] IC diagnosis occurred as a result of opportunistic screening examinations, or in the course of a medical examination sought by women due to symptoms (eg. a lump, nipple discharge, pain).

To identify SD cases in women with a positive mammogram, we conducted a search for diagnosed breast cancers from hospital clinical files, pathology reports, and the North Region Cancer Registry (*Registo Oncológico Regional do Norte* – RORENO), a population-based registry with high completeness.[18] Information on identified SD cases was added to the Breast Cancer Screening Programme database. To be included in this study, screening must have occurred in the period 2000–2007.

Information on IC was gathered by merging the RORENO and Breast Cancer Screening Programme databases.The woman's name, date of birth, and national health service number were used to link records. Inconsistencies and mismatching were checked manually. IC were those breast cancers registered in RORENO which matched with negative screening results in the previous two year period. This ascertainment was limited to women participating in the screening programme from 2000–2007. Because of the IC definition, breast cancers gathered might have been diagnosed in the period 2000–2009. As RORENO completeness is limited to invasive cancers, analysis was restricted to invasive SD and IC. Ductal carcinomas *in situ* were not considered. Following European Guidelines,[1] only women aged 50–69 were considered in the analysis.

The data collected was used to build a specific database for this study, including only cases of SD and IC. Information extracted from the screening programme database included variables such as date and round of screening, outcome of screening, woman's age at last screening examination, and residence. Information was also extracted from RORENO. To obtain additional data not routinely collected at the cancer registry, hospitals were contacted and clinical files and pathology reports were accessed. Overall, the following variables were collected: date of diagnosis, tumour size, histological type (using the International Classification of Diseases for Oncology-3rd edition;[19] 8500, 8521 coded as ductal; 8520, 8522, 8524 coded as lobular; 8211, 8480, 8510, 8530, 8540 coded as other), histological grade (according to Nottingham Grading System[20]), lymph node status, tumour stage (TNM classification - AJCC[21]), and first treatment (mastectomy, breast conserving surgery, chemotherapy). IC and SD tumour size were classified in two categories ($\leqslant 20$ mm or >20 mm; the cut-off between T1 and T2, T3, T4[21]). Information on biomarkers such as oestrogen (ER) and progesterone (PR) receptor status, and detection of overexpression and/or amplification of the human epidermal growth factor receptor 2 (HER2) were registered according to the pathology reports.

The two types of breast cancer (IC *versus* SD) were compared for each of these variables. To analyse the contribution of the different parameters to the risk of IC *versus* SD, an unconditional logistic regression was performed to calculate crude odds ratios (ORs) with corresponding 95% confidence intervals (CI-95%). T-test and χ^2 test were also used to compare differences between IC and SD, depending on the type of variables considered. After multivariable analysis (with all variables), the final model included only parameters significant to the risk of IC *versus* SD. HER2 status and triple negative assessment variables were not used in the multivariable analysis, due to the small number of cases. A $P < 0.05$ was considered to be statistically significant.

Results

In total, 554 invasive breast cancers were diagnosed in the study population, 442 (79.8%) classified as SD, and 112 (20.2%) as IC.

At last screening examination, women from the SD group were significantly older than the women with IC (mean age 60.0 ± 5.8 years and 58.4 ± 6.0 years, $p = 0.009$). IC were larger than SD (tumour size > 20 mm: 60.2% *versus* 25.1%). Results in table 1 are given as crude ORs, CIs and P values for each independent variable assessed.

The predominant histological type in both IC and SD cancers was ductal invasive, but the relative proportions

Table 1. Odds ratio of interval breast cancers relative to screen detected cancers according to clinicopathological characteristics.

Parameters	Interval cancers n = 112 (%*)	Screen-detected cancers n = 442 (%*)	Crude OR (95% CI)	P value†
Age group				
50–59	64 (57.1)	199 (45.0)	1	0.022
60–69	48 (42.9)	243 (55.0)	0.61 (0.40-0.93)	
Histology				
Ductal	84 (75.0)	371 (83.9)	1	0.004
Lobular	18 (16.1)	28 (6.3)	2.84 (1.50-5.37)	
Other	10 (8.9)	43 (9.7)	1.03 (0.50-2.13)	
Tumour grade				
grade 1	14 (13.0)	119 (28.4)	1	<0.001
grade 2	52 (48.1)	226 (53.9)	1.96 (1.04-3.67)	
grade 3	42 (38.9)	74 (17.7)	4.82 (2.47-9.44)	
Missing	4	23		
Tumour size				
⩽20 mm	43 (39.8)	329 (74.9)	1	<0.001
>20 mm	65 (60.2)	110 (25.1)	4.52 (2.91-7.03)	
Missing	4	3		
Lymph nodes				
node negative	48 (44.0)	303 (69.2)	1	<0.001
node positive	61 (56.0)	135 (30.8)	2.85 (1.86-4.38)	
Missing	3	4		
Stage				
I	32 (28.8)	260 (59.1)	1	<0.001
II	45 (40.5)	128 (29.1)	2.86 (1.73–4.71)	
III	26 (23.4)	49 (11.1)	4.31 (2.36–7.86)	
IV	8 (7.2)	3 (0.7)	21.67 (5.47–85.85)	
Missing	1	2		
ER status				
Positive	76 (71.0)	386 (88.9)	1	<0.001
Negative	31 (29.0)	48 (11.1)	3.28 (1.96–5.49)	
Missing	5	8		
PR status				
Positive	65 (60.7)	328 (75.6)	1	0.002
Negative	42 (39.3)	106 (24.4)	2.00 (1.28–3.12)	
Missing	5	8		
HER2 status				
Negative	77 (78.6)	171 (89.5)	1	0.011
Positive	21 (21.4)	20 (10.5)	2.33 (1.20–4.55)	
Missing	14	251		
Triple negative				
no	78 (79.6)	178 (93.2)	1	0.001
yes	20 (20.4)	13 (6.8)	3.51 (1.66–7.41)	
Missing	14	251		

*the percentages were calculated excluding those cancers with value unknown; †P value corresponds to χ^2 heterogeneity test; OR, odds ratio; CI, confidence interval; ER, oestrogen receptor; PR, progesterone receptor; HER2, epidermal growth factor receptor 2.

of ductal invasive and lobular were significantly different; lobular type was more likely to occur among IC (OR = 2.84, table 1). In this series of cases there were four inflammatory cancers, three in the IC group and one in the SD. In 27 cases, histological grade was not classified (table 1), either because the tumour had micro-invasion or because grade was not reported by the pathologist. In the remaining cancers, grade 3 was reported for 17.7% of the SD and for 38.9% of IC.

Women with IC presented more frequently than women with SD at diagnosis with lymph node invasion, and also with stage IV presentation (7.2% versus 0.7%). Evaluation of ER and PR status showed that SD more often had positive receptors than IC.

HER2 expression was evaluated in only 52.2% (n = 289) of the studied cancers; the proportion of HER2 positivity was higher in IC than in SD (21.4% and 10.5%, respectively) and the difference was significant. The phenotype triple negative for oestrogen, progesterone, and HER2 expression was more frequent in IC.

A multivariable logistic regression model was built to compute adjusted odds ratios of having IC (table 2). Only significant variables were considered in the final model: tumour size, histology, tumour grade, and ER status. Compared with SD, IC were more likely to be larger than 20 mm, be lobular, of grade 3 and negative for ER.

First treatment modality was significantly different between the two groups. Only 1.1% of SD cases had upfront chemotherapy compared with 13.8% of IC cases ($p < 0.001$). For those 525 women whose first treatment was surgery, breast conserving surgery was more commonly used in SD (261 in 431 cases, 60.6%) than IC (36 in 94 surgeries, 38.3%).

Table 2. Odds ratio of developing an interval cancer relative to screen detected cancer. Final model, after multivariable analysis.

Parameters	OR adjusted for covariates*	95% Confidence Interval
Tumour size		
⩽20 mm	1	
>20 mm	3.42	2.09–5.59
Histology		
Ductal	1	
Lobular	3.90	1.88–8.10
Other	1.23	0.55–2.74
Tumour grade		
grade 1	1	
grade 2	1.36	0.68–2.74
grade 3	2.58	1.17–5.71
ER status		
Positive	1	
Negative	2.50	1.34–4.66

*tumour size (categorical), histology, tumour grade, oestrogen receptors status.
OR, odds ratio; ER, oestrogen receptor.

Discussion

This is the first report on the clinicopathological characteristics of IC *versus* SD within this screening programme.

At screening examination, women with IC were younger than women with SD, as verified in other studies.[14,22–25] This has also been associated with higher prevalence of mammographic density in younger cases.[10,23,26]

Size of the breast cancer has been interpreted as reflecting the chronological age of the tumour.[14,27] In this study, IC were significantly larger than SD, in agreement with the findings from other studies.[10,12,28–30]

Lobular histological type was more frequent among IC than SD breast cancers. This has also been observed in some studies,[5,14,22,31] but others have found no difference in the histology type between IC and SD.[10,11,28] For the purposes of our analysis, "lobular" carcinoma and "mixed lobular and ductal" carcinoma were considered the same entity, as their characteristics and outcomes have been considered similar.[32] The growth patterns detected in lobular tumours make these more difficult to detect with mammography.[14]

Lymph node involvement in breast cancer has been considered to reflect both tumour chronology and biology.[27] In our study, after adjustment for the other variables, lymph node metastases were not significantly more frequent in IC than in SD. The small size of our studied population may explain this finding, but this lack of association has also been observed in other studies, after adjustment for size of the tumour.[7,14] This excess of lymph node metastases among IC depended not only on the chronological age and biology of the tumour, but also on other unexplained variables.[22,27]

Grade is an important indicator of biology and prognosis, considered by some as time independent,[13,33,34] but others have shown a phenotypic drift in cancer grade as tumour progresses.[35,36] With mammographic screening, a high proportion of tumours is stage T1N0M0 and grade is considered of particular utility in the evaluation of small tumours.[33] In our dataset the distribution of high grade tumours was more frequent in IC than SD, and this association was independent of the size of the tumour. This result is consistent with previous studies reporting a more aggressive behaviour of IC.[10,11,14,22,24,30,31]

This pattern of a more aggressive behaviour associated with IC was also expressed in the biomarkers results. The proportion of cases that were negative for RE or triple negative for RE, RP, and HER2, was higher in IC compared with SD.[7,24,37]

In the final model (after multivariable analysis), tumours larger than 20 mm, histological type lobular, high grade histology and RE negative were the independent predictors of the presence of interval cancers.

Primary treatment for women with IC was more likely to be extensive surgery, or systemic therapy, compared with SD cases. This is also in accordance with other published results, and reflects the more advanced stage at diagnosis for IC cases.[30]

Radiological review of last screening/assessment imaging and diagnostic mammography of the IC was not performed; this may be considered a limitation. However, although some authors have used a radiological category to distinguish between true interval and missed interval breast cancers,[28,38–41] others have not considered this assessment essential, and some have even thought it to be misleading.[11] Mammographic density has been considered a risk factor for IC,[23,26] but in this study the retrospective analysis of the mammographic pattern was not performed.

The sample size of our study may have hampered the detection of some associations of smaller magnitude. As mentioned in other studies, the existence of unregistered opportunistic screening can introduce some misclassification on the detection mode (SD *versus* IC).[29,42,43] In the northern region of Portugal, opportunistic screening coexists with the organized screening programme, but it is difficult to quantify with precision.[44]

We believe that limitations mentioned above do not seriously affect the internal validity of our study. Furthermore, the methodology used in data collection and management support the validity of the associations identified. Data were retrieved from the population-based breast cancer screening organized programme, ascertainment of IC cases was made through a population-based cancer registry (RORENO) to ensure complete capture of breast cancer cases in the population served by the screening programme, additional data not routinely collected at the cancer registry was actively sought in hospital clinical files and pathology reports, and this procedure was identical for SD and IC.

Variable by variable, these results are consistent with other studies conducted in different settings and with different screening methodologies. IC constitute a subset of breast cancers with a more aggressive biology than SD. They occur in any screening programme, but our findings did not show any unexpected pattern leading us to changes in the screening procedures. Continuous ongoing identification and characterization of IC, as we have carried out in this study, is advisable.

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Declaration of conflicting interests

Ana Aguiar is the Head of the Breast Cancer Screening Programme in the North Region of Portugal and Vítor Veloso is the President of the Portuguese League Against Cancer, North Branch. The remaining authors have no conflicts of interest to disclose.

Acknowledgements

The authors thank Manuel António Silva and Joana Bastos from the Centre Region Cancer Registry, Alice Coelho from São Sebastião Hospital, Isabel Carvalho from S. João Hospital and pathologists Vicente Gonçalves, Eduardo Ferreira and Fernando Pardal for their valuable contribution to data collection.

References

1. Perry N, Broeders M, de Wolf C, Tornberg S, von Holland R, Karsa L, eds *European guidelines for quality assurance in breast cancer screening and diagnosis*. 4th ed. Luxembourg: Office for Official Publications of the European Communities, 2006.
2. Stockton D, McCann J. Cancer registries in monitoring, evaluating and planning breast cancer screening programmes. In: Sankila R, Démaret E, Hakama M, Lynge E, Shouten LJ, Parkin DM, eds *Evaluation and Monitoring of Screening Programmes*. Luxembourg: Office for Official Publications of the European Communities, 2001:181–94.
3. Bennett RL, Sellars SJ, Moss SM. Interval cancers in the NHS breast cancer screening programme in England, Wales and Northern Ireland. *Br J Cancer* 2011;**104**:571–7.
4. Tornberg 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.
5. Baré M, Bonfill X, Andreu X on behalf of the Breast Cancer Screening Programme of Sabadell-Cerdanyola Research Group. Relationship between the method of detection and prognostic factors for breast cancer in a community with a screening programme. *J Med Screen* 2006;**13**:183–91.
6. Bare M, Sentis M, Galceran J, et al. Interval breast cancers in a community screening programme: frequency, radiological classification and prognostic factors. *Eur J Cancer Prev* 2008;**17**:414–21.
7. Domingo L, Sala M, Servitja S, et al. Phenotypic characterization and risk factors for interval breast cancers in a population-based breast cancer screening program in Barcelona, Spain. *Cancer Causes Control* 2010;**21**:1155–64.
8. Houssami N, Irwig L, Ciatto S. Radiological surveillance of interval breast cancers in screening programmes. *Lancet Oncol* 2006;**7**:259–65.
9. Burrell HC, Sibbering DM, Wilson AR, et al. Screening interval breast cancers: mammographic features and prognosis factors. *Radiology* 1996;**199**:811–7.
10. Kirsh VA, Chiarelli AM, Edwards SA, et al. Tumor characteristics associated with mammographic detection of breast cancer in the Ontario breast screening program. *J Natl Cancer Inst* 2011;**103**:942–50.
11. Gilliland FD, Joste N, Stauber PM, et al. Biologic characteristics of interval and screen-detected breast cancers. *J Natl Cancer Inst* 2000;**92**:743–9.
12. Caldarella A, Puliti D, Crocetti E, et al. Biological characteristics of interval cancers: a role for biomarkers in the breast cancer screening. *J Cancer Res Clin Oncol* 2013;**139**:181–5.
13. Webster LR, Bilous AM, Willis L, et al. Histopathologic indicators of breast cancer biology: insights from population mammographic screening. *Br J Cancer* 2005;**92**:1366–71.
14. Porter PL, El-Bastawissi AY, Mandelson MT, et al. Breast tumor characteristics as predictors of mammographic detection: comparison of interval- and screen-detected cancers. *J Natl Cancer Inst* 1999;**91**:2020–8.
15. Dawson SJ, Duffy SW, Blows FM, et al. Molecular characteristics of screen-detected vs symptomatic breast cancers and their impact on survival. *Br J Cancer* 2009;**101**:1338–44.

16. Giordano L, von Karsa L, Tomatis M, et al. Mammographic screening programmes in Europe: organization, coverage and participation. *J Med Screen* 2012;**19**(Suppl 1): 72–82.
17. Statistics Portugal. Annual estimates of resident population 2009 Lisboa2013 [Access Date: 2013/12/14]. Available from: http://www.ine.pt/
18. Castro C, Bento MJ, Lunet N, Campos P. Assessing the completeness of cancer registration using suboptimal death certificate information. *Eur J Cancer Prev* 2012;**21**:478–9.
19. World Health Organization. International Classification of Diseases for Oncology. 3rd ed. Fritz A, Percy C, Jack A, et al., editors. Geneva, 2000.
20. Elston CW, Ellis IO. Pathological prognostic factors in breast cancer. I. .The value of histological grade in breast cancer: experience from a large study with long-term follow-up. *Histopathology* 1991;**19**:403–10.
21. AJCC Cancer Staging Handbook. Greene FL, Page DL, Fleming ID, editors. New York, NY: Springer; 2002.
22. Bucchi L, Puliti D, Ravaioli A, et al. Breast screening: axillary lymph node status of interval cancers by interval year. *Breast* 2008;**17**:477–83.
23. Mandelson MT, Oestreicher N, Porter PL, et al. Breast density as a predictor of mammographic detection: comparison of interval- and screen-detected cancers. *J Natl Cancer Inst* 2000;**92**:1081–7.
24. Musolino A, Michiara M, Conti GM, et al. Human epidermal growth factor receptor 2 status and interval breast cancer in a population-based cancer registry study. *J Clin Oncol* 2012;**30**:2362–8.
25. Bucchi L, Ravaioli A, Foca F, Colamartini A, Falcini F, Naldoni C. Incidence of interval breast cancers after 650,000 negative mammographies in 13 Italian health districts. *J Med Screen* 2008;**15**:30–5.
26. Porter GJ, Evans AJ, Cornford EJ, et al. Influence of mammographic parenchymal pattern in screening-detected and interval invasive breast cancers on pathologic features, mammographic features, and patient survival. *AJR* 2007;**188**:676–83.
27. Smeets A, Ryckx A, Belmans A, et al. Impact of tumor chronology and tumor biology on lymph node metastasis in breast cancer. *Springerplus* 2013;**2**:480.
28. Raja MA, Hubbard A, Salman AR. Interval breast cancer: is it a different type of breast cancer? *Breast* 2001;**10**:100–8.
29. Domingo L, Jacobsen KK, von Euler-Chelpin M, et al. Seventeen-years overview of breast cancer inside and outside screening in Denmark. *Acta Oncol* 2013;**52**:48–56.
30. Olsson A, Borgquist S, Butt S, Zackrisson S, Landberg G, Manjer J. Tumour-related factors and prognosis in breast cancer detected by screening. *Br J Surg* 2012;**99**:78–87.
31. Crane CE, Luke CG, Rogers JM, Playford PE, Roder DM. An analysis of factors associated with interval as opposed to screen-detected breast cancers, including hormone therapy and mammographic density. *Breast* 2002;**11**:131–6.
32. Bharat A, Gao F, Margenthaler JA. Tumor characteristics and patient outcomes are similar between invasive lobular and mixed invasive ductal/lobular breast cancers but differ from pure invasive ductal breast cancers. *Am J Surg* 2009;**198**:516–9.
33. Rakha EA, Reis-Filho JS, Baehner F, et al. Breast cancer prognostic classification in the molecular era: the role of histological grade. *Breast Cancer Res* 2010;**12**:207.
34. Engstrom MJ, Opdahl S, Hagen AI, et al. Molecular subtypes, histopathological grade and survival in a historic cohort of breast cancer patients. *Breast Cancer Res Treat* 2013;**140**:463–73.
35. Anderson TJ, Waller M, Ellis IO, Bobrow L, Moss S. Influence of annual mammography from age 40 on breast cancer pathology. *Hum Pathol* 2004;**35**:1252–9.
36. Tabar L, Duffy SW, Vitak B, Chen HH, Prevost TC. The natural history of breast carcinoma: what have we learned from screening? *Cancer* 1999;**86**:449–62.
37. Nagtegaal ID, Allgood PC, Duffy SW, et al. Prognosis and pathology of screen-detected carcinomas: how different are they? *Cancer* 2011;**117**:1360–8.
38. Hofvind S, Geller B, Skaane P. Mammographic features and histopathological findings of interval breast cancers. *Acta Radiol* 2008;**49**:975–81.
39. Britton PD, McCann J, O'Driscoll D, Hunnam G, Warren RM. Interval cancer peer review in East Anglia: implications for monitoring doctors as well as the NHS breast screening programme. *Clin Radiol* 2001;**56**:44–9.
40. Hoff SR, Abrahamsen A-L, Samset JH, Vigeland E, Klepp O, Hofvind S. Breast Cancer: Missed Interval and Screening-detected Cancer at Full-Field Digital Mammography and Screen-Film Mammography— Results from a Retrospective Review. *Radiology* 2012;**264**:378–86.
41. Vitak B, Olsen KE, Manson JC, Arnesson LG, Stal O. Tumour characteristics and survival in patients with invasive interval breast cancer classified according to mammographic findings at the latest screening: a comparison of true interval and missed interval cancers. *Eur Radiol* 1999;**9**:460–9.
42. Heidinger O, Batzler WU, Krieg V, et al. The incidence of interval cancers in the German mammography screening program: results from the population-based cancer registry in North Rhine-Westphalia. *Dtsch Arztebl Int* 2012;**109**:781–7.
43. Bulliard JL, Sasieni P, Klabunde C, De Landtsheer JP, Yankaskas BC, Fracheboud J. Methodological issues in international comparison of interval breast cancers. *Int J Cancer* 2006;**119**:1158–63.
44. Departamento de Estudos e Planeamento. Relatório de Actividades 2009. Porto: Administração Regional de Saúde do Norte, 2010.

III – Results

5. Clinicopathological characteristics of invasive breast cancers diagnosed in participants, non-participants and not invited to the organized population-based Breast Cancer Screening, in the North of Portugal

III – Results

5. Clinicopathological characteristics of invasive breast cancers diagnosed in participants, non-participants and not invited to the organized population-based Breast Cancer Screening, in the North of Portugal

From the moment an organized screening programme has been introduced to a country or region it is necessary to study the effectiveness of the programme in routine health-care settings and observational studies become the main contributors of information on the impact of screening.[66] Many studies compared the breast cancer mortality among women invited and not invited to screening, or compared participants and non-participants.[120] Other studies compared early indicators of efficacy as size and stage of breast cancers diagnosed in invited and not invited populations (the control group) to the screening programme.[121,133] As it was referred in the Background section (1.4 Histopathology and prognosis) advanced breast cancer stage is related with higher mortality. It should be noticed that this population of not invited women can be subjected to more or less marked level of opportunistic screening rather than with no screening at all,[57,134] leading to a possible reduction in the BCSP reported benefits.

The phased implementation of the screening programme in the Northern Region allowed the assessment of the screening experience of the women resident in the district of Bragança compared to a contemporaneous population of women resident in Vila Real and not exposed to the organized screening programme. In **paper VII** the objective was to know how different the clinicopathological characteristics of the cancers were, depending on the modality of detection.

Although the size of the groups may hampered some associations, in screen-detected breast cancers the maximum dimension of the invasive tumour was smaller, and significantly different from the cancers detected in women non-participant or not invited to the organized screening.

The main limitation in this work derived from the impossibility to assign to the not invited population what was the modality of detection of the breast cancer: if symptomatic or through opportunistic screening.

Title:

Clinicopathological characteristics of invasive breast cancers diagnosed in participants, non-participants and not invited to the organized population-based Breast Cancer Screening, in the North of Portugal

Short title:

Clinicopathological characteristics of breast cancers in participants, non-participants and not invited to a screening programme

Authors names, degrees and affiliations:

- Maria José Bento, MD MSc, Institute of Biomedical Sciences Abel Salazar, University of Porto and Epidemiology Unit, Portuguese Oncology Institute, Porto, Portugal
- Guilherme Gonçalves, MD PhD, Institute of Biomedical Sciences Abel Salazar, University of Porto, Portugal
- Ana Aguiar, MD, Portuguese Cancer League – North Branch, Porto, Portugal
- Luis Antunes, MSc, Epidemiology Unit, Portuguese Oncology Institute, Porto, Portugal
- Clara Castro, MSc, Epidemiology Unit, Portuguese Oncology Institute, Porto, Portugal
- Vitor Veloso, MD, Portuguese Cancer League – North Branch, Porto, Portugal
- Vitor Rodrigues, MD PhD, Faculty of Medicine, University of Coimbra, and Portuguese Cancer League – Centre Branch, Coimbra, Portugal

Corresponding author and requests for reprints

Maria José Bento, MD MSc

Institute of Biomedical Sciences Abel Salazar, University of Porto,

Rua de Jorge Viterbo, nº 228, 4050-313 Porto, Portugal

mjbento@icbas.up.pt

Telephone: 00 351 225084067

Fax: 00 351 225084004

Abstract

Objective The aim of this study was to evaluate the clinical and pathological characteristics of the invasive breast cancers diagnosed in women participant in breast cancer screening programme, compared to cancers detected in non-participants and in not invited women.

Setting Breast Cancer Screening Programme (BCSP) of the north region of Portugal.

Methods Data was retrieved from the population-based North Region Cancer Registry and from the organized population-based BCSP, and records were matched to select the three groups for comparison.

Results In screening participants, 75.8% of invasive breast cancers were ≤ 20 mm, 67.7% had no axillary lymph nodes metastasis and 58.1% were stage I. These characteristics were significantly more favourable than those found in breast cancers detected in women non-participant or not invited. After multivariable analysis, size remained the only distinguishing characteristic of breast cancers detected within the screening programme compared to the other two studied groups. Breast cancers detected in screening participants were significantly smaller, which is consistent with findings by other authors.

Conclusion The more favourable prognostic characteristics of the breast cancers detected in a population exposed to screening (including interval cancers) indicate a possible mortality reduction in the future.

Keywords

Breast cancer; organized screening; non-participants; mammography; tumour size

Introduction

High-quality population-based breast cancer screening programmes, with periodic mammographic examination of asymptomatic women became an important tool in cancer control.[1,2] For logistic reasons the implementation of a new population-based screening programme in a certain country (or region) can take several years till it is fully implemented in all the geographical area considered; for that reason, during a certain time period, it happens that very similar neighbouring populations are being covered or not by the programme, creating an opportunity to compare likely outcomes between populations.[1]

Comparisons of characteristics of the cancers diagnosed in women invited or not to an organized screening programme, and the analysis of differences between screened-detected and symptomatic breast cancers, have been used as a further approach in the evaluation of screening programmes.[3-8]

The organized population-based Breast Cancer Screening Programme (BCSP) implemented in the Northern Region of Portugal, conducted by the north branch of the Portuguese Cancer League (*Liga Portuguesa Contra o Cancro* – LPCC) started in 1999 in one municipality and gradually expanded its coverage in the north region (5 districts and 68 municipalities). BCSP was implemented in the district of Bragança between 2003 and 2005 when full coverage was reached; in 2005, the estimated number of women aged 50-69 years living in the district was 19 554, representing 5.3% of the estimated 372 015 women of the same age living in the whole northern region. Bragança and Vila Real are neighbouring districts, with the same socioeconomical and cultural features and very close background breast cancer incidence.[9] In Vila Real the organized screening programme was only launched in 2009; in 2005, the estimated number of women aged 50-69 years living in this district was 27 644, representing 7.4% of the women of the same age group in the northern region.

We aimed at contributing to the assessment of BCSP. For that purpose, the specific objective of this study was to compare the characteristics of the invasive breast cancers

detected in populations with different screening exposure/participation status in our organized screening programme.

Methods

Briefly, the methods implemented at the BCSP were the following: every two years women aged between 45 and 69 years were sent a letter with an invitation for a two-view mammography examination at one of the mobile or fixed units. A blind-double reading was systematically performed at a dedicated centre by trained radiologists with a final reading by a third independent and experienced radiologist, in case of discrepancy. Since the beginning of the screening programme it has been operating in accordance with the European Guidelines[10] and preliminary results have been published.[11] A specific database with individual records for the screening procedures and results was created in 1999 (BCSP database).

Invasive breast cancers diagnosed in women resident in the northern region of Portugal have been registered since 1988, at the population-based North Region Cancer Registry (*Registo Oncológico Regional do Norte* – RORENO) which has high completeness.[12]

Data was retrieved from RORENO using the following criteria: invasive breast cancers diagnosed between 2003 and 2008, in women aged 50-69 years at diagnosis (to be in accordance with age group considered in the European Guidelines)[10] and resident in the districts of Vila Real and Bragança. Then, information on the screening history of breast cancers in women resident in Bragança was retrieved from the BCSP database. Variables as name, date of birth and national health number were used for matching. Similar to the "screening exposure"[4] and "participation"[6] status classifications used by other authors, the above described information was used to select three groups for comparison:

- women invited and participating in the screening, including screen-detected cancers and interval cancers (residents in Bragança) named *participants* in this analysis*;*

- women invited but not participating in screening, including women who never attended organized screening procedures, and those whose last participation had been more than 2 years before (residents in Bragança), named *non-participants;*
- women not invited to screening, which includes two subgroups: those resident in Vila Real district, who were not invited to screening in the study period, and women resident in Bragança district with breast cancer diagnosed prior to an invitation to participate in the screening programme, named *not invited*.

Data collected from the BCSP and RORENO databases included the patient date of birth, date and round of last mammography, outcome of screening, screening exposure/participation status (participants, non-participants, not invited), municipality of residence, date of diagnosis of breast cancer, age at diagnosis, tumour size in mm (with further division in 3 groups, according to the cut-offs of the European Guidelines)[10], histological type using the International Classification of Diseases for Oncology-3rd edition (8500, 8521 coded as ductal; 8520, 8522, 8524 coded as lobular; 8211, 8480, 8510, 8530, 8540 coded as other), histological grade according to Nottingham Grading System[13], lymph node status, tumour stage (TNM classification – AJCC[14]), first treatment (mastectomy, breast conserving surgery, chemotherapy). In cases with upfront chemotherapy, a clinical T and N were assigned. Information on biomarkers as oestrogen (ER) and progesterone (PR) receptors status, and detection of overexpression and/or amplification of the human epidermal growth factor receptor 2 (HER2) were registered according to the pathology reports.

Breast cancers detected in women participating in screening, in non-participants and in women not invited to screening were compared for each of the aforementioned variables. Comparisons were made pairwise. Proportions were compared using the Pearson χ^2 test or Fisher's exact test when χ^2 test was not applicable, and one-way analysis of variance was used to compare the means of the continuous variables.

Unconditional multivariable logistic regression was used to assess the association between screening exposure/participation status and clinicopathological characteristics of

breast cancer adjusted for possible confounder factors. Two models were tested for comparison of cancers detected in participants *versus* non-participants (including tumour size, lymph node status, grading, as covariates) and screening participants *versus* not invited (including tumour size, lymph node status, ER and PR expression, as covariates). Tumour size in the multivariable analysis was considered as ≤20 mm or > 20 mm. Since none of the screening participants had breast cancer with distant metastasis at diagnosis, this variable was not included in the multivariable models. HER2 was not used in this analysis, due to the small number of cases with this information. Differences were considered statistically significant for $P < 0.05$.

Results

Between 2003 and 2008, 125 breast cancer cases were detected in women participating in the programme (113 screen-detected and 12 interval cancers) and 57 breast cancers were diagnosed in non-participants, including 7 women with more than 2 years since last mammogram. In the same period, 314 cancers were detected in women not invited to screening, 278 were residents in Vila Real and 36 in Bragança.

The mean age of all (n= 496) selected women with breast cancer was 59.7±5.7 years, and there was no significant difference ($P = 0.56$) between the three groups. In table 1 are shown the main clinical and pathological characteristics of the three groups.

The predominant histological type was ductal and the proportions were very similar between groups.

In the group of screening participants, the proportion of cancers with maximum size ≤10 mm was 30.1%, < 15 mm was 52.1% and 75.8% were ≤ 20 mm. Compared to non-participants or not invited, screening participants had a significantly higher proportion of smaller breast tumours ($P < 0.001$ for the three cut-offs used).

When cancer dimensions were compared between non-participants and not invited groups, the proportion of breast cancers with a maximum dimension greater than 20 mm was significantly higher in the first group (66.7% compared to 49.5%, $P = 0.02$); when size cut-off

values used were of 10 and 15 mm, no significant differences were observed. For all the other variables, there were no significant differences between these two groups (non-participants and not invited).

Cancers detected in participants were found to be better differentiated than those detected in non-participants (P = 0.002); compared to not invited group, participants had lower grade tumours, though significance (P = 0.06) was slightly above the classical significance level.

The tumours in screening participants had less frequently lymph node metastasis than non-participants or not invited groups (P = 0.005 and P = 0.006, respectively). None of the cancers in participants had distant metastasis at diagnosis and it was significantly different from the 4.2% of the cancers with distant metastasis detected among the not invited (P = 0.02). In non-participants, 1.9% (one case) had distant metastasis at diagnosis and it was not significantly different from the group of participants.

Cancers in participants were more frequently found in an earlier stage than in each of the other two groups, with 58.1% of the cancers detected in stage I among participants (P < 0.001 for both comparisons, table 1). At diagnosis, 22.2% and 37.5% of the breast cancers diagnosed in non-participants and in the not invited group, respectively, were classified as stage I.

Cancers in participants showed significantly higher proportion of ER and PR positivity than cancers in the not invited group (P = 0.036 and P = 0.009, respectively) but a similar proportion when compared with breast cancers of non-participants. Although in this last group cancers were slightly more positive for the hormonal receptors than in the not invited group, the difference was not significant (P = 0.18).

Information on HER2 status was missing for almost half of the cancers in participants and not invited women. The association between HER2 status and exposure/participation was statistically significant: negative status was more frequent in the participant group compared to non-participants or to not invited (P = 0.015 and P = 0.024, respectively). There

were no significant differences according to exposure/participation status and the distribution of the triple-negative subtype.

In the multivariable analysis (table 2), 163 cases were included in the model for comparison of cancers in participants *versus* non-participants, and 370 cases for participants *versus* not invited group. Tumour size was the only significant variable in both final models. Smaller tumours had higher probability to be found in cancers diagnosed in the group of screening participants compared to the non-participant ($P<0.001$) or not invited ($P=0.002$) groups.

Information on treatment strategy was missing for 1.0% of participants, 12.3% of non-participants and 15.0% of not invited cases. When first treatment was surgery, the proportion of participants who underwent breast-conserving surgery was 57.4%, a value significantly higher compared to 34.1% of non-participants or 31.5% of not invited cancer cases ($P = 0.008$ and $P <0.001$, respectively). Chemotherapy as first treatment was recorded in 1.6% of participant women, which was significantly lower than 12.0% among non-participants and 10.9% in the not invited group ($P = 0.003$ and $P = 0.002$, respectively).

Discussion

In the evaluation of an organized screening programme, it is of paramount importance to describe the clinicopathological features of the cancers detected. In this study, we assessed these characteristics among breast cancers detected in a rolling population-based organized screening programme, comparing them to the breast cancers detected by usual practice or non-organized screening activities.

The results should be interpreted within the limitations imposed by the design of the study, the small sample size of the groups and missing values. Due to the small number of cancers in the participant group, we were not able to differentiate initial from subsequent screening round, which prevented a more in-deep analysis on the effect of length bias and overdiagnosis.[15] Also, some variables had a considerable amount of missing values.

In an initial analysis (univariate), breast cancers were significantly smaller among screened participants, less prone to the development of axillary metastasis and were found in an earlier stage, compared to breast cancers in women invited but not participant, or compared to the experience of breast cancer in a population not exposed to organized screening. Stage migration (down-staging) is an expected effect of screening.[6] This result is in agreement with other studies, either hospital-based or population-based, using comparison groups defined in a variety of ways, from cancers detected only by symptoms or opportunistic screening, cancers detected in populations not participating or not yet offered screening, among others.[4,6,7,16-18] Nevertheless, after multivariable analysis, size remained the only distinguishing characteristic of breast cancers detected within the screening programme compared to the other two studied groups. The small numbers in the multivariable analysis possibly hampered the disclosure of other significant associations. It is recognized that that expected benefit of early detection of breast cancer is not determined solely by tumour size but other variables as nodal status and grade are also significant.[19]

It was not surprising that conservative surgery was more frequently done in the screening participants, in which, detected cancers were smaller and with a higher proportion of stage I. Adoption of less harmful and more effective treatments in areas where organized screening has been implemented is a recognized benefit of screening programmes.[6,20-22]

Breast cancers detected in the not invited group had a significantly smaller dimension compared to cancers detected in women who didn´t participate or were less compliant with the organized screening programme. Several authors have raised this issue of the impact of opportunistic screening among populations without an organized screening service.[21,23-25] Opportunistic screening exists in the Northern Region of Portugal, though we have no precise estimates of its magnitude; furthermore, we were not able to assign individually, the participation in opportunistic screening for this group of women as it is recommended.[21] Not forgetting these limitations, it is legitimate to argue that opportunistic screening should have a stronger impact in the not invited group, as this was the only possibility for earlier diagnosis in this population, and a likely explanation for the differences in tumour size reported in this

study. Also, the implementation of a screening programme in a region has been considered to trigger cancer awareness in patients outside the screening programme, with a prognostic benefit in these.[18,26,27] The above explanations are plausible and eventually consistent with our findings and those published by other authors who reported a worse prognosis, as presenting larger dimensions for tumours detected in non-participants.[6,28-30]

The number of breast cancer cases with missing data on tumour size, nodal status and grading was greater in the not invited group than in the other two groups. However, it is unlikely that relevant selection bias had been introduced, since age and period of diagnosis (between 2003-2005 or 2006-2008) of the women with missing information did not differ from the age and period of diagnosis of the other women.

Reasons for non-participation can vary along the period of implementation of a screening programme.[18,31] In the beginning, most of the women not participating were not invited, but afterwards non-participation happens for other reasons such as worse accessibility and lower socioeconomic status;[6,18] this may lead to selection biases in this type of study.[18,24] We minimized the likelihood of this bias, since we were able to constitute more homogenous groups of not participant and non invited women to be compared.

Breast cancers detected in screening participants and non-participants or not invited women, were all diagnosed in the same time frame, close geographical location in the northern region and reflected the full experience of breast cancer incidence in the population. Thus, the possibility of bias due to improvement in cancer diagnosis and treatment in more recent years or bias due to selection of the cases was probably reduced.

Using the information from the organized population-based screening programme and matching it with information from a population-based cancer registry with high completeness, favours the validity of the reported associations. That is because it is more likely that we have got almost complete information on the clinicopathological characteristics of breast cancers in populations exposed and not exposed to an organized screening programme.

Breast cancers detected in screening participants were significantly smaller and tumour size is considered one of the strongest predictors of breast cancer behaviour.[32] As stated by

others, the more favourable prognostic characteristics of the breast cancers detected in a population exposed to screening (including interval cancers) indicate an eventual mortality reduction in the future, due to this cause.[6,21,33,34] Thus, though this is a limited descriptive study, its findings are consistent with an effective screening programme, which will have to be confirmed in future assessments.

Declaration of conflicting interests

Ana Aguiar is the Head of the Breast Cancer Screening Programme in the North Region of Portugal and Vítor Veloso is the President of the Portuguese Cancer League, North Branch. For the remaining authors there are no conflicts of interest to disclose.

References

1. Lynge E, Tornberg S, von Karsa L, Segnan N, van Delden JJ. Determinants of successful implementation of population-based cancer screening programmes. Eur J Cancer 2012;**48**:743-8.
2. Dijck J, Schouten L. Use of cancer registry data: prerequesites, limitations and solutions. In: Sankila R, Démaret E, Hakama M, Lynge E, Schouten L, Parkin DM, eds. *Evaluation and monitoring of screening programmes*. Brussels, Luxembourg: Europe Against Cancer Programme, 2000:43-57
3. Allgood PC, Duffy SW, Kearins O, et al. Explaining the difference in prognosis between screen-detected and symptomatic breast cancers. Br J Cancer 2011;**104**:1680-5.
4. Baré M, Bonfill X, Andreu X, on behalf of the Breast Cancer Screening Programme of Sabadell-Cerdanyola Research Group. Relationship between the method of detection and prognostic factors for breast cancer in a community with a screening programme. J Med Screen 2006;**13**:183-91.
5. Bucchi L, Puliti D, Ravaioli A, et al. Breast screening: axillary lymph node status of interval cancers by interval year. Breast 2008;**17**:477-83.
6. Hofvind S, Lee CI, Elmore JG. Stage-specific breast cancer incidence rates among participants and non-participants of a population-based mammographic screening program. Breast Cancer Res Treat 2012;**135**:291-9.
7. Nagtegaal ID, Duffy SW. Reduction in rate of node metastases with breast screening: consistency of association with tumor size. Breast Cancer Res Treat 2013;**137**:653-63.
8. Shen Y, Yang Y, Inoue LY, Munsell MF, Miller AB, Berry DA. Role of detection method in predicting breast cancer survival: analysis of randomized screening trials. J Natl Cancer Inst 2005;**97**:1195-203.
9. RORENO. Mapas - comparações efectuadas entre cada distrito e a restante região Porto2002 [Access Date: 2014/03/01]. Available from: http://www.roreno.com.pt/pt/estatisticas/mapas.html.

10. Perry N, Broeders M, de Wolf C, Tornberg S, Holland R, von Karsa L, editors. European guidelines for quality assurance in breast cancer screening and diagnosis. 4th ed. Office for Official Publications of the European Communities, Luxembourg; 2006.
11. Giordano L, von Karsa L, Tomatis M, et al. Mammographic screening programmes in Europe: organization, coverage and participation. J Med Screen 2012;**19 (Suppl 1)**:72-82.
12. Castro C, Bento MJ, Lunet N, Campos P. Assessing the completeness of cancer registration using suboptimal death certificate information. Eur J Cancer Prev 2012;**21**:478-9.
13. Elston CW, Ellis IO. Pathological prognostic factors in breast cancer. I. The value of histological grade in breast cancer: experience from a large study with long-term follow-up. Histopathology 1991;**19**:403-10.
14. AJCC Cancer Staging Handbook. Greene FL, Page DL, Fleming ID, editors. New York, NY: Springer; 2002.
15. Hakama M, Holli K, Isola J, et al. Aggressiveness of screen-detected breast cancers. Lancet 1995;**345**:221-4.
16. Bucchi L, Barchielli A, Ravaioli A, et al. Screen-detected vs clinical breast cancer: the advantage in the relative risk of lymph node metastases decreases with increasing tumour size. Br J Cancer 2005;**92**:156-61.
17. Burke JP, Power C, Gorey TF, Flanagan F, Kerin MJ, Kell MR. A comparative study of risk factors and prognostic features between symptomatic and screen detected breast cancer. Eur J Surg Oncol 2008;**34**:149-53.
18. Mook S, Van 't Veer LJ, Rutgers EJ, et al. Independent prognostic value of screen detection in invasive breast cancer. J Natl Cancer Inst 2011;**103**:585-97.
19. Narod SA. Tumour size predicts long-term survival among women with lymph node-positive breast cancer. Curr Oncol 2012;**19**:249-53.
20. Segnan N, Rosso S, Ponti A. Is the breast cancer mortality decrease in Sweden due to screening or treatment? Not the right question. J Natl Cancer Inst 2012;**104**:1040-1.

21. Bulliard JL, Ducros C, Jemelin C, Arzel B, Fioretta G, Levi F. Effectiveness of organised versus opportunistic mammography screening. Ann Oncol 2009;**20**:1199-202.
22. Berry DA, Cronin KA, Plevritis SK, et al. Effect of screening and adjuvant therapy on mortality from breast cancer. N Engl J Med 2005;**353**:1784-92.
23. Welch HG. Screening mammography--a long run for a short slide? N Engl J Med 2010;**363**:1276-8.
24. Hoff SR, Klepp O, Hofvind S. Asymptomatic breast cancer in non-participants of the national screening programme in Norway: a confounding factor in evaluation? J Med Screen 2012;**19**:177-83.
25. Vanier A, Leux C, Allioux C, Billon-Delacour S, Lombrail P, Molinie F. Are prognostic factors more favorable for breast cancer detected by organized screening than by opportunistic screening or clinical diagnosis? A study in Loire-Atlantique (France). Cancer Epidemiol 2013;**37**:683-7.
26. Kalager M, Haldorsen T, Bretthauer M, Hoff G, Thoresen SO, Adami HO. Improved breast cancer survival following introduction of an organized mammography screening program among both screened and unscreened women: a population-based cohort study. Breast Cancer Res 2009;**11**:R44.
27. Domingo L, Jacobsen KK, von Euler-Chelpin M, et al. Seventeen-years overview of breast cancer inside and outside screening in Denmark. Acta Oncol 2013;**52**:48-56.
28. Duffy SW, Tabar L, Fagerberg G, et al. Breast screening, prognostic factors and survival--results from the Swedish two county study. Br J Cancer 1991;**64**:1133-8.
29. McCann J, Stockton D, Day N. Breast cancer in East Anglia: the impact of the breast screening programme on stage at diagnosis. J Med Screen 1998;**5**:42-8.
30. Stockton D, McCann J. Cancer registries in monitoring, evaluating and planning breast cancer screening programmes. In: Sankila R, Démaret E, Hakama M, Lynge E, Shouten LJ, Parkin DM, eds. *Evaluation and Monitoring of Screening Programmes*. Luxembourg: Office for Official Publications of the European Communities, 2001:181-94

31. Nagtegaal ID, Allgood PC, Duffy SW, et al. Prognosis and pathology of screen-detected carcinomas: how different are they? Cancer 2011;**117**:1360-8.
32. Day NE, Williams DR, Khaw KT. Breast cancer screening programmes: the development of a monitoring and evaluation system. Br J Cancer 1989;**59**:954-8.
33. Hofvind S, Sorum R, Thoresen S. Incidence and tumor characteristics of breast cancer diagnosed before and after implementation of a population-based screening-program. Acta Oncol 2008;**47**:225-31.
34. Kim J, Lee S, Bae S, et al. Comparison between screen-detected and symptomatic breast cancers according to molecular subtypes. Breast Cancer Res Treat 2012;**131**:527-40.

Table 1 – Distribution of clinicopathological characteristics of invasive breast cancers diagnosed in women participant, in non-participant and not invited to the organized population-based Breast Cancer Screening Programme in 2003-2008

Clinicopathological characteristics		Exposure/participation status			Significance level		
Variable	value	Participants (P) n = 125 (%*)	Non-participants (NP) n = 57 (%*)	Not invited (NI) n = 314 (%*)	P value P/NP	P value P/NI	P value NP/NI
Histology							
	Ductal	109 (87.2)	49 (86.0)	271 (86.3)	0.27	0.57	0.63
	Lobular	13 (10.4)	4 (7.0)	29 (9.2)			
	Other	3 (2.4)	4 (7.0)	14 (4.5)			
Tumour size							
	≤10 mm	37 (30.1)	3 (5.6)	30 (11.2)	<0.001	<0.001	0.22
	>10	86 (69.9)	51 (94.4)	239 (88.8)			
	Missing	2	3	45			
Tumour size							
	<15 mm	63 (52.1)	8 (19.0)	55 (24.2)	<0.001	<0.001	0.47
	≥15 mm	58 (47.9)	34 (81.0)	172 (75.8)			
	Missing	4	3	87			
Tumour size							
	≤20 mm	94 (75.8)	18 (33.3)	142 (50.5)	<0.001	<0.001	0.02
	>20 mm	30 (24.2)	36 (66.7)	139 (49.5)			
	Missing	1	3	33			
Tumour grade							
	Grade 1	27 (23.1)	8 (16.3)	48 (19.4)	0.002	0.06	0.15
	Grade 2	75 (64.1)	23 (46.9)	141 (57.1)			
	Grade 3	15 (12.8)	18 (36.7)	58 (23.5)			
	Missing	8	8	67			
Lymph nodes							
	negative	84 (67.7)	24 (45.3)	143 (53.0)	0.005	0.006	0.31
	positive	40 (32.3)	29 (54.7)	127 (47.0)			
	Missing	1	4	44			
Distant metastasis							
	Negative	124 (100)	53 (98.1)	271 (95.8)	0.30	0.02	0.40
	Positive	0 (0)	1 (1.9)	12 (4.2)			
	Missing	1	3	31			
Stage							
	I	72 (58.1)	12 (22.2)	106 (37.5)	<0.001	<0.001	0.11
	II	40 (32.3)	23 (42.6)	94 (33.2)			
	III	12 (9.7)	18 (33.3)	71 (25.1)			
	IV	0	1 (1.9)	12 (4.2)			
	Missing	1	3	31			
ER status							
	Positive	108 (87.8)	43 (84.3)	206 (78.9)	0.54	0.036	0.38
	Negative	15 (12.2)	8 (15.7)	55 (21.1)			
	Missing	2	6	53			
PR status							
	Positive	96 (78.0)	38 (74.5)	167 (64.7)	0.61	0.009	0.18
	Negative	27 (22.0)	12 (25.5)	91 (35.3)			
	Missing	2	6	56			

HER2 status						
Negative	56 (87.5)	25 (67.6)	126 (73.7)	0.015	0.024	0.25
Positive	8 (12.5)	12 (32.4)	45 (26.3)			
Missing	61	20	143			
Triple negative						
no	57 (89.1)	35 (94.6)	151 (88.3)	0.35	0.87	0.26
yes	7 (10.9)	2 (5.4)	20 (11.7)			
Missing	61	20	143			

* the percents were calculated excluding those cancers with value unknown; P/NP, screen participants compared to non-participants; P/NI, screen participants compared to not invited; NP/NI, non-participants compared to not invited; ER, oestrogen receptor; PR, progesterone receptor; HER2, epidermal growth factor receptor 2

Table 2 – Multivariable logistic regression for the association between clinicopathological characteristics of breast cancer and mode of participation (screening participants *versus* non-participants or not invited)

Parameters	OR adjusted for covariates	95% Confidence Interval	P value
Participants/Non participants (n = 163)			
Tumour size			
≤20 mm	1		
>20 mm	4.36	2.00 – 9.71	<0.001
Lymph nodes			
negative	1		
positive	1.28	0.58 – 2.83	0.54
Tumour grade			
grade 1	1		
grade 2	0.78	0.28 – 2.19	0.64
grade 3	2.30	0.71 – 7.45	0.17
Participants/ Not invited (n = 370)			
Tumour size			
≤20 mm	1		
>20 mm	2.39	1.38 – 4.13	0.002
Lymph nodes			
negative	1		
positive	1.28	0.77 – 2.14	0.34
ER status			
Positive	1		
Negative	1.12	0.47 – 2.69	0.79
PR status			
Positive	1		
Negative	1.37	0.67 – 2.77	0.39

OR, odds ratio; ER, oestrogen receptor; PR, progesterone receptor

IV – General discussion and conclusions

IV – General discussion and conclusions

This thesis fulfils the desideratum of all public health interventions need to be assessed.[123] The research conducted within this thesis indicates that the population-based BCSP of the Northern Region of Portugal achieved high quality standards whether compared with the European Guidelines or compared to other European screening programmes, and should be expanded to cover all eligible women in the Northern Region.

This conclusion is supported by the valid studies included in this thesis. Strengths and limitations were extensively discussed in detail for each individual study.

Many indicators of the screening evaluation depend on the existence of a population-based cancer registry with high completeness; this was evaluated in **paper I**. Results obtained warranted the conclusion that RORENO is a valuable source of information on the new cases of breast cancer diagnosed in the female population. RORENO also participated in an international collaborative study to evaluate the criteria for the female breast cancer biomarker positivity used in different countries; this was addressed in **paper II** showing the need of a more extensive use of the existing guidelines.

Papers III and IV provided information on the characteristics and performance of breast cancer screening programmes across Europe. The BCSP of the Northern Region participated in this collaborative project of the EUNICE with data from 2005; it showed similar coverage by invitation, higher coverage by examination and a lower rate of further assessment, compared to the overall results of the participating screening programmes. These two papers demonstrated the feasibility and usefulness of comparing European-wide screening monitoring indicators. Furthermore, results concerning the BCSP in the Northern Region provided a first insight on the screening programme performance and were in line with the majority of the European screenings evaluated.

In **paper V**, performance and early surrogate impact indicators of the screening programme were evaluated by comparing them to the European Guidelines. The main results indicated that the screening programme conducted in the Northern Region of Portugal was highly accepted by the population, it was detecting the expected number of invasive breast cancers and in this group of tumours, it was able to detect small size breast cancers among the participants. These results are consistent with an effective screening programme and mortality reduction is to be expected in the future.[17,120]

In **papers IV** and **V** it was observed that the recall rate in the prevalent round exceeded the value recommended in the European Guidelines and the highest values were verified in the last years of evaluation. Nevertheless, this high recall rate was lower than values observed in other European programmes, as stated before (**paper IV**). High recall rate is positively correlated with false-positive results[110] that are considered a cause of needless psychological distress in addition to incurring in invasive investigations.[120] Possible

explanations include the introduction of the computed radiography in 2007, that increased sensitivity and the rate of false positives examinations,[135] the increasing number of new radiologists during roll-out phase of the programme,[136] and also, the increasing proportion of younger women entering the programme as long as the screening proceeds over the years.[24] The recall rate for subsequent screening was in accordance with the desirable level of the European Guidelines.

Monitoring of interval cancers occurrence is a crucial part of the evaluation of mammography screening programmes[24] for the reason that it provides a mechanism to evaluate some of the technical processes involved in the screening, as performance and interpretation of the mammography, and it contributes to the evaluation of the impact of mammography screening among the participants.[126] Interval cancers were assessed in **papers V and VI**. The main conclusions arising from these studies were: the rate of interval cancers during the first and in the second year were in accordance with the desirable levels of the European Guidelines and in relation to the clinicopathological presentation of the interval cancers, it was noteworthy the higher size, the higher grade and less frequent oestrogen receptor positivity found in these cancers compared to screen-detected cancers. This pattern of more aggressive characteristics found in interval cancers were also described in the majority of the international studies and would be expected in a well conducted screening programme.

After the evaluation of the main performance and impact indicators of the BCSP implemented in several municipalities in the Northern Region of Portugal, it was important to know how different these breast cancers were from those diagnosed in populations not covered by the organized screening programme. This was the objective of **paper VII**. From 2002 till 2007 the invasive breast cancers detected in women resident in Bragança (covered by the organized programme) were compared with those detected in women resident in Vila Real (not invited population). It was observed that invasive cancers detected within the organized programme (screen-detected and interval cancers) were significantly smaller than cancers detected among women not invited to the organized screening. These tumour characteristics favourably predict future mortality reduction.[137]

Strengths of this thesis result from the population-based approach for the evaluation of the BCSP in the Northern Region. The BCSP database provided information on individual invitations, further assessments and diagnosis for all population covered by the organized screening programme. This information was linked with the database from RORENO, a population-based cancer registry with high completeness, allowing for the evaluation of all breast cancers diagnosed inside (screen-detected, interval cancers and cancers in non-participants) and outside the screening programme.

This thesis has also recognized limitations. This work focused on a small proportion of the population covered by the organized screening (only 17% of the women aged 50-69 and residents in the Northern Region), precluding the generalization of the results to the whole Region as local populations may differ in terms of the risk of breast cancer, compliance, cancer awareness, among other influent factors.[57,134,138-140] The small number of diagnosed cancers prevented a more detailed research on subgroups; another limitation was the lack of exhaustiveness of some of the clinicopathological data (despite good collaboration of different entities) and the possible heterogeneity in the criteria for classification of some variables. Also the delay in registration of a case in RORENO didn't allow for the use of more recent data from the screening programme.

The main conclusions from this thesis are:

- The organized population-based screening programme implemented in the Northern Region of Portugal from 2000-2009 provided a high quality service and it is foreseeable a mortality reduction attributable to breast cancer in the population covered by the programme;
- The programme should be expanded to cover all eligible women in the Northern Region.
- Assessment is a never ending process; the work produced in the research included in this study is not a finished work. In order to proceed with a useful assessment some future research is proposed.

V – Future studies

V – Future studies

Screening programmes should ultimately be monitored in terms of deaths, the measure directly related to the purpose of screening.[120] The BCSP of the Northern Region is operating for almost 15 years and an effect in mortality is anticipated.

Some countries or areas compared time trends for breast cancer mortality before and after the implementation of screening or compared mortality trends in areas with and without screening.[73] However, several authors have raised arguments concerning the validity of such approaches to assess the effectiveness of the programmes.[29,73,141] Thus, though such studies may be done in the Northern Region of Portugal, probably they will provide a useful but not conclusive evaluation. Alternative approaches have been proposed as incidence-based mortality studies and case-control studies.[29]

Incidence-based mortality studies estimate the impact of screening by calculating the incidence-based breast cancer mortality in a population invited to screening compared to the incidence-based mortality in the absence of screening (control group).[57,66,70] The long phased implementation of the BCSP in the Northern Region allows for the comparison of contemporaneous invited and not yet invited women in relation to the breast mortality rates. Another approach estimates the expected breast cancer mortality rate in the absence of screening using historical data from a previous period to the screening implementation.[70,142]

A case-control study is an efficient method of combining screening information from a case series and a reference population where the cases originate from.[66] Data are collected from a cross section of cancer deaths cases and of the population from which the cases have emerged.[143] Some authors prefer to use the term case-referent study to the more commonly used term case-control study because the uptake of screening in the case group of breast cancer deaths is referred to the probability of having been screened in the population from which the cases originate.[143,144] The BCSP of the Northern Region has an information system that provides individual basic data on screening process, including data on invitations and rounds, diagnostic procedures and cancers detected. Linking this information with data from RORENO on breast cancer diagnosis and deaths (date and cause) in the target population and period of time, it will be possible to evaluate the BCSP in terms of mortality reduction, using this case-referent approach.

For this thesis a population-based cohort of women with breast cancer was studied regarding early surrogates of breast cancer mortality. Survival analysis comparing different modalities of cancer detection (within and outside organized screening) and prognostic variables will provide a deeper insight on the natural history of the disease and probably additional information for clinical practice.[145,146]

Another future research should address the effectiveness of breast cancer screening in younger women. The target age-group of the BCSP is women 45-69 years old. In this thesis,

no evaluation has been done in the group 45-49 years in terms of performance and impact indicators, in comparison to other international studies. Besides, the effectiveness of the screening programme in these younger ages has been under intense discussions among experts on screening.[13,46]

The methodology described in **papers III, IV, V** should be used routinely and an annual written report should be produced as a result from the continuous monitoring of the programme. Besides, new ways for reporting faster and significant epidemiological and clinical information on cancer diagnosis should be essayed, in order to provide up-to-date information to the screening programme; this information is vital to evaluate if a remedial action has to be triggered and/or to validate the effectiveness of the programme.[67]

Review of current evidence and controversy on the procedures, efficacy and cost-effectiveness of such screening programmes is also a never ending task, as it is illustrated in recent controversial discussions.[29,57,63,147-151]

VI – References

VI – References

1. Jemal A, Center MM, DeSantis C, Ward EM. Global patterns of cancer incidence and mortality rates and trends. Cancer Epidemiol Biomarkers Prev 2010;19:1893-907
2. Ferlay J, Shin HR, Bray F, Forman D, Mathers C, Parkin DM. Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008. Int J Cancer 2010;127:2893-917
3. Ferlay J, Shin H, Bray F, Forman D, Mathers C, Parkin D. GLOBOCAN 2008 v2.0, Cancer Incidence and Mortality Worldwide. Lyon France: International Agency for Research on Cancer; 2010.
4. Youlden DR, Cramb SM, Dunn NAM, Muller JM, Pyke CM, Baade PD. The descriptive epidemiology of female breast cancer: An international comparison of screening, incidence, survival and mortality. Cancer Epidemiol 2012;36:237-48
5. Bray F, Ren JS, Masuyer E, Ferlay J. Global estimates of cancer prevalence for 27 sites in the adult population in 2008. Int J Cancer 2013;132:1133-45
6. World Health Organization. Mortality database [Access date: 28-03-2013]. Available from: http://www.who.int/healthinfo/morttables/en/.
7. Bastos J, Barros H, Lunet N. [Breast cancer mortality trend in Portugal (1955-2002)]. Acta Med Port 2007;20:139-44
8. RORENO. Registo Oncológico Regional do Norte - 2008. Porto: Instituto Português de Oncologia do Porto, Março 2013.
9. Vogel VG. Epidemiology of breast cancer. 4th ed. Bland K, Copeland E, editors: Sauders; 2012.
10. Weiss JR, Moysich KB, Swede H. Epidemiology of Male Breast Cancer. Cancer Epidemiol Biomarkers Prev 2005;14:20-6
11. International Agency for Research on Cancer/World Health Organization (IARC/WHO). Breast cancer and screening. In: Vainio H, Bianchini F, eds. *Breast cancer screening*. Lyon: IARCPress, 2002
12. Colditz G, Baer H, Tamimi R. Breast cancer. In: Schottenfeld D, Fraumeni J, eds. *Cancer Epidemiology and Prevention*. 3rd ed. Oxford: Oxford University Press, 2006:995-1012
13. Nelson HD, Tyne K, Naik A, et al. Screening for Breast Cancer: Systematic Evidence Review Update for the U.S. Preventive Services Task Force. Evidence Review Update No. 74. Rockville, MD: Agency for Healthcare Research and Quality, 2009
14. AJCC Cancer Staging Handbook. Greene FL, Page DL, Fleming ID, editors. New York, NY: Springer; 2002.
15. Rakha EA, Reis-Filho JS, Baehner F, et al. Breast cancer prognostic classification in the molecular era: the role of histological grade. Breast Cancer Res 2010;12:207
16. Dijck J, Schouten L. Use of cancer registry data: prerequesites, limitations and solutions. In: Sankila R, Démaret E, Hakama M, Lynge E, Schouten L, Parkin DM, eds. *Evaluation*

and monitoring of screening programmes. Brussels, Luxembourg: Europe Against Cancer Programme, 2000:43-57

17. Day NE, Williams DR, Khaw KT. Breast cancer screening programmes: the development of a monitoring and evaluation system. Br J Cancer 1989;59:954-8
18. Carter CL, Allen C, Henson DE. Relation of tumor size, lymph node status, and survival in 24,740 breast cancer cases. Cancer 1989;63:181-7
19. Cianfrocca M, Goldstein LJ. Prognostic and predictive factors in early-stage breast cancer. Oncologist 2004;9:606-16
20. Fisher B, Bauer M, Wickerham DL, et al. Relation of number of positive axillary nodes to the prognosis of patients with primary breast cancer. An NSABP update. Cancer 1983;52:1551-7
21. Pereira H, Pinder SE, Sibbering DM, et al. Pathological prognostic factors in breast cancer. IV: Should you be a typer or a grader? A comparative study of two histological prognostic features in operable breast carcinoma. Histopathology 1995;27:219-26
22. Grann VR, Troxel AB, Zojwalla NJ, Jacobson JS, Hershman D, Neugut AI. Hormone receptor status and survival in a population-based cohort of patients with breast carcinoma. Cancer 2005;103:2241-51
23. Oldenhuis CN, Oosting SF, Gietema JA, de Vries EG. Prognostic versus predictive value of biomarkers in oncology. Eur J Cancer 2008;44:946-53
24. Perry N, Broeders M, de Wolf C, Tornberg S, Holland R, von Karsa L, editors. European guidelines for quality assurance in breast cancer screening and diagnosis. 4th ed. Office for Official Publications of the European Communities, Luxembourg; 2006.
25. Gondos A, Bray F, Hakulinen T, Brenner H, Group ESW. Trends in cancer survival in 11 European populations from 1990 to 2009: a model-based analysis. Ann Oncol 2009;20:564-73
26. Soerjomataram I, Louwman MW, Ribot JG, Roukema JA, Coebergh JW. An overview of prognostic factors for long-term survivors of breast cancer. Breast Cancer Res Treat 2008;107:309-30
27. Wilson JMG, Jungner G. Principles and practices of screening for disease. Public Health Papers, [Internet]. 1968 34 Access date: 2014/03/13;[Available from: http://whqlibdoc.who.int/php/WHO_PHP_34.pdf.
28. Bretthauer M, Kalager M. Principles, effectiveness and caveats in screening for cancer. Br J Surg 2013;100:55-65
29. Marmot MG, Altman DG, Cameron DA, Dewar JA, Thompson SG, Wilcox M. The benefits and harms of breast cancer screening: an independent review. Br J Cancer 2013;108:2205-40

30. Tabar L, Duffy SW, Vitak B, Chen HH, Prevost TC. The natural history of breast carcinoma: what have we learned from screening? Cancer 1999;86:449-62
31. Duffy SW, Smith RA, Gabe R, Tabar L, Yen AM, Chen TH. Screening for breast cancer. Surg Oncol Clin N Am 2005;14:671-97
32. Nelson HD, Tyne K, Naik A, Bougatsos C, Chan BK, Humphrey L. Screening for breast cancer: an update for the U.S. Preventive Services Task Force. Ann Intern Med 2009;151:727
33. Swedish Organised Service Screening Evaluation. Reduction in breast cancer mortality from organized service screening with mammography: 1. Further confirmation with extended data. Cancer Epidemiol Biomarkers Prev 2006;15:45-51
34. World Health Organization. Breast cancer: prevention and control [Access date: 1-5-2014]. Available from: http://www.who.int/cancer/detection/breastcancer/en/index1.html.
35. Berry DA, Cronin KA, Plevritis SK, et al. Effect of screening and adjuvant therapy on mortality from breast cancer. N Engl J Med 2005;353:1784-92
36. Brawley OW, Kramer BS. Cancer screening in theory and in practice. J Clin Oncol 2005;23:293-300
37. Drukteinis JS, Mooney BP, Flowers CI, Gatenby RA. Beyond mammography: new frontiers in breast cancer screening. Am J Med 2013;126:472-9
38. Obenauer S, Hermann KP, Grabbe E. Applications and literature review of the BI-RADS classification. Eur Radiol 2005;15:1027-36
39. International Agency for Research on Cancer/World Health Organization (IARC/WHO). Use of breast cancer screening. In: Vainio H, Bianchini F, eds. *Breast cancer screening*. Lyon: IARC Press, 2002:47-86
40. Klabunde CN, Sancho-Garnier H, Broeders M, Thoresen S, Rodrigues VJ, Ballard-Barbash R. Quality assurance for screening mammography data collection systems in 22 countries. Int J Technol Assess Health Care 2001;17:528-41
41. Hakama M, Coleman MP, Alexe DM, Auvinen A. Cancer screening: evidence and practice in Europe 2008. Eur J Cancer 2008;44:1404-13
42. Eurocourse: Europe Against Cancer. Recommendations on screening related items in European data set 2011 [Access date: 2014/04/01]. Available from: http://screening.iarc.fr/doc/RecommendationsWG.pdf.
43. von Karsa L, Arrossi S. Development and implementation of guidelines for quality assurance in breast cancer screening: the European experience. Salud Publica Mex 2013;55:318-28
44. Hakama M. Planning and designing of screening programmes. In: Sankila R, Démaret E, Hakama M, Lynge E, Schouten L, Parkin DM, eds. *Evaluation and monitoring of*

screening programmes. Brussels, Luxembourg: Europe Against Cancer Programme, 2000:13-28

45. Warner E. Breast-Cancer Screening. NEJM 2011;365:1025-32
46. Mandelblatt JS, Cronin KA, Bailey S, et al. Effects of mammography screening under different screening schedules: model estimates of potential benefits and harms. Ann Intern Med 2009;151:738-47
47. Hellquist BN, Duffy SW, Abdsaleh S, et al. Effectiveness of population-based service screening with mammography for women ages 40 to 49 years: evaluation of the Swedish Mammography Screening in Young Women (SCRY) cohort. Cancer 2011;117:714-22
48. Croswell JM, Ransohoff DF, Kramer BS. Principles of cancer screening: lessons from history and study design issues. Semin Oncol 2010;37:202-15
49. International Agency for Research on Cancer/World Health Organization (IARC/WHO). Efficacy of screening. In: Vainio H, Bianchini F, eds. *Breast cancer screening*. Lyon: IARC Press, 2002:87-117
50. Duffy SW, Nagtegaal ID, Wallis M, et al. Correcting for lead time and length bias in estimating the effect of screen detection on cancer survival. Am J Epidemiol 2008;168:98-104
51. Council of the European Union (2003). Council Recommendation of 2 December 2003 on cancer screening (2003/878/EC). Off J Euro Union; 2003. p. 34-8.
52. Kirkpatrick A, Törnberg S, Thijssen M, editors. European guidelines for quality assurance in mammography screening. 1st ed. Office for Official Publications of the European Communities, Luxembourg; 1993.
53. Zappa M, Federici A. The potential for systematic early detection and treatment of breast cancer to reduce the burden of disease is widely recognized in the European Union (EU). Introduction. J Med Screen 2012;19 Suppl 1:3-4
54. De Wolf C, Perry N, editors. European guidelines for quality assurance in mammography screening. 2nd ed. Office for Official Publications of the European Communities, Luxembourg; 1996.
55. Perry N, Broeders M, De Wolf C, Törnberg S, editors. European guidelines for quality assurance in mammography screening. 3rd ed. Office for Official Publications of the European Communities, Luxembourg; 2001.
56. Freedman DA, Petitti DB, Robins JM. On the efficacy of screening for breast cancer. Int J Epidemiol 2004;33:43-55
57. Health Council of the Netherlands. Population screening for breast cancer: expectations and developments. The Hague: Health Council of the Netherlands, 2014
58. Giordano L, von Karsa L, Tomatis M, et al. Mammographic screening programmes in Europe: organization, coverage and participation. J Med Screen 2012;19 (Suppl 1):72-82

59. 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
60. Gotzsche PC, Nielsen M. Screening for breast cancer with mammography. Cochrane Database Syst Rev 2011:CD001877
61. Jorgensen KJ, Keen JD, Gotzsche PC. Is mammographic screening justifiable considering its substantial overdiagnosis rate and minor effect on mortality? Radiology 2011;260:621-7
62. Gotzsche PC. Mammography screening: truth, lies, and controversy. Lancet 2012;380:218
63. Duffy S, Chen T, Smith R, Yen AM, Tabar L. Real and artificial controversies in breast cancer screening. Breast Cancer Manage 2013;2:519-28
64. Duffy SW, Tabar L, Olsen AH, et al. Absolute numbers of lives saved and overdiagnosis in breast cancer screening, from a randomized trial and from the Breast Screening Programme in England. J Med Screen 2010;17:25-30
65. Ursin G. Mammographic screening debate on study design: a need to move the field forward. BMC Med 2012;10:164
66. Broeders M, Moss S, Nystrom L, et al. 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
67. Harris R, Yeatts J, Kinsinger L. Breast cancer screening for women ages 50 to 69 years a systematic review of observational evidence. Prev Med 2011;53:108-14
68. Puliti D, Miccinesi G, Zappa M, Manneschi G, Crocetti E, Paci E. Balancing harms and benefits of service mammography screening programs: a cohort study. Breast Cancer Res 2012;14:R9
69. Autier P, Boniol M. Breast cancer screening: evidence of benefit depends on the method used. BMC Med 2012;10:163
70. Njor S, Nystrom L, Moss S, et al. Breast cancer mortality in mammographic screening in Europe: a review of incidence-based mortality studies. J Med Screen 2012;19 Suppl 1:33-41
71. Puliti D, Zappa M. Breast cancer screening: are we seeing the benefit? BMC Med 2012;10
72. Paap E, Verbeek AL, Puliti D, Paci E, Broeders MJ. Breast cancer screening case-control study design: impact on breast cancer mortality. Ann Oncol 2011;22:863-9
73. Moss SM, Nystrom L, Jonsson H, et al. 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

74. Berry DA. Breast cancer screening: controversy of impact. Breast 2013;22 Suppl 2:S73-6
75. Barratt A, Mannes P, Irwig L, Trevena L, Craig J, Rychetnik L. Cancer screening. J Epidemiol Community Health 2002;56:899-902
76. Puliti D, Duffy SW, Miccinesi G, et al. Overdiagnosis in mammographic screening for breast cancer in Europe: a literature review. J Med Screen 2012;19 Suppl 1:42-56
77. The benefits and harms of breast cancer screening - Authors' reply. Lancet 2013;381:803-4
78. Biesheuvel C, Barratt A, Howard K, Houssami N, Irwig L. Effects of study methods and biases on estimates of invasive breast cancer overdetection with mammography screening: a systematic review. Lancet Oncol 2007;8:1129-38
79. Jorgensen KJ, Gotzsche PC. Overdiagnosis in publicly organised mammography screening programmes: systematic review of incidence trends. BMJ 2009;339:b2587
80. Duffy SW, Parmar D. Overdiagnosis in breast cancer screening: the importance of length of observation period and lead time. Breast Cancer Res 2013;15:R41
81. Wait S. Cancer registries in evaluation of breast cancer screening programmes. In: Sankila R, Démaret E, Hakama M, Lynge E, Schouten L, Parkin DM, eds. *Evaluation and monitoring of screening programmes*. Brussels, Luxembourg: Europe Against Cancer Programme, 2000:133-50
82. Zappa M, Falini P, Bonardi R, et al. Monitoring interval cancers in mammographic screening: the Florence District programme experience. Breast 2002;11:301-5
83. Houssami N, Irwig L, Ciatto S. Radiological surveillance of interval breast cancers in screening programmes. Lancet Oncol 2006;7:259-65
84. Britton PD, McCann J, O'Driscoll D, Hunnam G, Warren RM. Interval cancer peer review in East Anglia: implications for monitoring doctors as well as the NHS breast screening programme. Clin Radiol 2001;56:44-9
85. Gower-Thomas K, Fielder HM, Branston L, Greening S, Beer H, Rogers C. Reviewing interval cancers: time well spent? Clin Radiol 2002;57:384-8
86. 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;237:437-43
87. Gilliland FD, Joste N, Stauber PM, et al. Biologic characteristics of interval and screen-detected breast cancers. J Natl Cancer Inst 2000;92:743-9
88. Zackrisson S, Janzon L, Manjer J, Andersson I. Improved survival rate for women with interval breast cancer - results from the breast cancer screening programme in Malmo, Sweden 1976-1999. J Med Screen 2007;14:138-43
89. Hoff SR, Abrahamsen A-L, Samset JH, Vigeland E, Klepp O, Hofvind S. Breast Cancer: Missed Interval and Screening-detected Cancer at Full-Field Digital Mammography and

Screen-Film Mammography— Results from a Retrospective Review. Radiology 2012;264:378-86

90. Hoff SR, Samset JH, Abrahamsen AL, Vigeland E, Klepp O, Hofvind S. Missed and true interval and screen-detected breast cancers in a population based screening program. Acad Radiol 2011;18:454-60
91. Hofvind S, Geller B, Skaane P. Mammographic features and histopathological findings of interval breast cancers. Acta Radiol 2008;49:975-81
92. Stockton D, Jenny M. Cancer registries in monitoring, evaluating and planning breast cancer screening programmes. In: Sankila R, Démaret E, Hakama M, Lynge E, Schouten L, Parkin DM, eds. *Evaluation and monitoring of screening programmes*. Brussels, Luxembourg: Europe Against Cancer Programme, 2000:181-94
93. Hofvind S, Bjurstam N, Sorum R, Bjorndal H, Thoresen S, Skaane P. Number and characteristics of breast cancer cases diagnosed in four periods in the screening interval of a biennial population-based screening programme. J Med Screen 2006;13:192-6
94. Raja MA, Hubbard A, Salman AR. Interval breast cancer: is it a different type of breast cancer? Breast 2001;10:100-8
95. Tornberg 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
96. Rosenberg RD, Hunt WC, Williamson MR, et al. Effects of age, breast density, ethnicity, and estrogen replacement therapy on screening mammographic sensitivity and cancer stage at diagnosis: review of 183,134 screening mammograms in Albuquerque, New Mexico. Radiology 1998;209:511-8
97. 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;276:33-8
98. Porter GJ, Evans AJ, Cornford EJ, et al. Influence of mammographic parenchymal pattern in screening-detected and interval invasive breast cancers on pathologic features, mammographic features, and patient survival. AJR 2007;188:676-83
99. Sala E, Warren R, McCann J, Duffy S, Day N, Luben R. Mammographic parenchymal patterns and mode of detection: implications for the breast screening programme. J Med Screen 1998;5:207-12
100. Porter PL, El-Bastawissi AY, Mandelson MT, et al. Breast tumor characteristics as predictors of mammographic detection: comparison of interval- and screen-detected cancers. J Natl Cancer Inst 1999;91:2020-8
101. Porter GJ, Evans AJ, Burrell HC, Lee AH, Ellis IO, Chakrabarti J. Interval breast cancers: prognostic features and survival by subtype and time since screening. J Med Screen 2006;13:115-22

102. Bordas P, Jonsson H, Nystrom L, Lenner P. Survival from invasive breast cancer among interval cases in the mammography screening programmes of northern Sweden. Breast 2007;16:47-54

103. Rayson D, Payne JI, Abdolell M, et al. Comparison of clinical-pathologic characteristics and outcomes of true interval and screen-detected invasive breast cancer among participants of a Canadian breast screening program: a nested case-control study. Clin Breast Cancer 2011;11:27-32

104. McCann J, Stockton D, Day N. Breast cancer in East Anglia: the impact of the breast screening programme on stage at diagnosis. J Med Screen 1998;5:42-8

105. Baré M, Bonfill X, Andreu X, on behalf of the Breast Cancer Screening Programme of Sabadell-Cerdanyola Research Group. Relationship between the method of detection and prognostic factors for breast cancer in a community with a screening programme. J Med Screen 2006;13:183-91

106. Kalager M, Tamimi RM, Bretthauer M, Adami HO. Prognosis in women with interval breast cancer: population based observational cohort study. BMJ 2012;345:e7536

107. Koivunen D, Zhang X, Blackwell C, Adelstein E, Humphrey L. Interval breast cancers are not biologically distinct--just more difficult to diagnose. Am J Surg 1994;168:538-42

108. Shen Y, Yang Y, Inoue LY, Munsell MF, Miller AB, Berry DA. Role of detection method in predicting breast cancer survival: analysis of randomized screening trials. J Natl Cancer Inst 2005;97:1195-203

109. Schwartz LM, Woloshin S, Sox HC, Fischhoff B, Welch HG. US women's attitudes to false positive mammography results and detection of ductal carcinoma in situ: cross sectional survey. BMJ 2000;320:1635-40

110. Hofvind S, Ponti A, Patnick J, et al. False-positive results in mammographic screening for breast cancer in Europe: a literature review and survey of service screening programmes. J Med Screen 2012;19 (Suppl 1):57-66

111. Castells X, Roman M, Romero A, et al. Breast cancer detection risk in screening mammography after a false-positive result. Cancer Epidemiol 2013;37:85-90

112. OECD. Screening, survival and mortality for breast cancer. 2013 In: Health at a Glance 2013: OECD Indicators [Internet]. OECD Publishing, [Access date 2014/05/07]. Available from: http://dx.doi.org/10.1787/health_glance-2013-52-en.

113. Alves JG, Cruz DB, Rodrigues VL, Goncalves ML, Fernandes E. Breast cancer screening in the central region of Portugal. Eur J Cancer Prev 1994;3 (Suppl 1):49-53

114. Administração Regional de Saúde do Algarve. Rastreio do cancro da mama no Algarve [Access date: 2014/04/13]. Available from: http://www.arsalgarve.min-saude.pt/portal/?q=node/3658.

115. Dourado F, Carreira H, Lunet N. Mammography use for breast cancer screening in Portugal: results from the 2005/2006 National Health Survey. Eur J Public Health 2013;23:386-92

116. Departamento de Estudos e Planeamento. Manual Executivo do Programa de Rastreio do Cancro da Mama na Região Norte. Porto: Administração Regional de Saúde do Norte, 2009

117. Departamento de Estudos e Planeamento. Relatório de atividades 2012. Porto: Administração Regional de Saúde do Norte, 2013

118. Departamento de Estudos e Planeamento. Relatório de Actvidades 2009. Porto: Administração Regional de Saúde do Norte, 2010

119. Administração Regional de Saúde do Norte, Prescrição de Mamografia de rastreio em mulheres do grupo etário 45 a 69 anos, nas Unidades de Cuidados de Saúde Primários, Circular normativa nº 1 (2011).

120. International Agency for Research on Cancer/World Health Organization (IARC/WHO). Effectiveness of screening. In: Vainio H, Bianchini F, eds. *Breast cancer screening*. Lyon: IARC Press, 2002:119-56

121. Buiatti E, Barchielli A, Bartolacci S, et al. The impact of organised screening programmes on the stage-specific incidence of breast cancer in some Italian areas. Eur J Cancer 2003;39:1776-82

122. Schouten LJ, de Rijke JM, Schlangen JT, Verbeek AL. Evaluation of the effect of breast cancer screening by record linkage with the cancer registry, The Netherlands. J Med Screen 1998;5:37-41

123. Brownson C, Baker E, Leet T, Gillespie K. The need for evidence-based public health. *Evidence-based public health*. 1st ed. Oxford: Oxford University Press, 2003:3-23

124. Yankaskas BC, Klabunde CN, Ancelle-Park R, et al. International comparison of performance measures for screening mammography: can it be done? J Med Screen 2004;11:187-93

125. Klabunde CN, Ballard-Barbash R. Evaluating Population-Based Screening Mammography Programs Internationally. Semin Breast Dis 2007;10:102-7

126. Tornberg S, Codd M, Rodrigues V, Segnan N, Ponti A. Ascertainment and evaluation of interval cancers in population-based mammography screening programmes: a collaborative study in four European centres. J Med Screen 2005;12:43-9

127. Kirsh VA, Chiarelli AM, Edwards SA, et al. Tumor characteristics associated with mammographic detection of breast cancer in the Ontario breast screening program. J Natl Cancer Inst 2011;103:942-50

128. Musolino A, Michiara M, Conti GM, et al. Human epidermal growth factor receptor 2 status and interval breast cancer in a population-based cancer registry study. J Clin Oncol 2012;30:2362-8
129. Olsson A, Borgquist S, Butt S, Zackrisson S, Landberg G, Manjer J. Tumour-related factors and prognosis in breast cancer detected by screening. Br J Surg 2012;99:78-87
130. Nagtegaal ID, Allgood PC, Duffy SW, et al. Prognosis and pathology of screen-detected carcinomas: how different are they? Cancer 2011;117:1360-8
131. Caldarella A, Puliti D, Crocetti E, et al. Biological characteristics of interval cancers: a role for biomarkers in the breast cancer screening. J Cancer Res Clin Oncol 2013;139:181-5
132. Domingo L, Jacobsen KK, von Euler-Chelpin M, et al. Seventeen-years overview of breast cancer inside and outside screening in Denmark. Acta Oncol 2013;52:48-56
133. Hofvind S, Skaane P. Stage distribution of breast cancer diagnosed before and after implementation of population-based mammographic screening. Rofo 2012;184:437-42
134. Bulliard JL, Ducros C, Jemelin C, Arzel B, Fioretta G, Levi F. Effectiveness of organised versus opportunistic mammography screening. Ann Oncol 2009;20:1199-202
135. 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;244:708-17
136. Alberdi RZ, Llanes AB, Ortega RA, et al. Effect of radiologist experience on the risk of false-positive results in breast cancer screening programs. Eur Radiol 2011;21:2083-90
137. Hofvind S, Lee CI, Elmore JG. Stage-specific breast cancer incidence rates among participants and non-participants of a population-based mammographic screening program. Breast Cancer Res Treat 2012;135:291-9
138. Ouedraogo S, Dabakuyo TS, Gentil J, Poillot ML, Dancourt V, Arveux P. Attending breast cancer screening alone does not explain the detection of tumours at an early stage. Eur J Cancer Prev 2013;22:103-11
139. Smith RA. International programs for the detection of breast cancer. Salud Publica Mex 2011;53:394-404
140. Schapira DV, Levine RB. Breast cancer screening and compliance and evaluation of lesions. Med Clin North Am 1996;80:15-26
141. Paci E. 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
142. Paci E, Giorgi D, Bianchi S, et al. Assessment of the early impact of the population-based breast cancer screening programme in Florence (Italy) using mortality and surrogate measures. Eur J Cancer 2002;38:568-73

143. Verbeek AL, Broeders MJ. Evaluation of cancer service screening: case referent studies recommended. Stat Methods Med Res 2010;19:487-505
144. Paap E, Verbeek AL, Botterweck AA, et al. Breast cancer screening halves the risk of breast cancer death: A case-referent study. Breast 2014
145. Mook S, Van 't Veer LJ, Rutgers EJ, et al. Independent prognostic value of screen detection in invasive breast cancer. J Natl Cancer Inst 2011;103:585-97
146. Pacelli B, Carretta E, Spadea T, et al. Does breast cancer screening level health inequalities out? A population-based study in an Italian region. Eur J Public Health 2014;24:280-5
147. Pace LE, Keating NL. A systematic assessment of benefits and risks to guide breast cancer screening decisions. JAMA 2014;311:1327-35
148. Miller AB, Wall C, Baines CJ, Sun P, To T, Narod SA. Twenty five year follow-up for breast cancer incidence and mortality of the Canadian National Breast Screening Study: randomised screening trial. BMJ 2014;348:g366
149. Gotzsche PC, Jorgensen KJ. Screening for breast cancer with mammography. Cochrane Database Syst Rev 2013;6:CD001877
150. Biller-Andorno N, Juni P. Abolishing mammography screening programs? A view from the Swiss Medical Board. N Engl J Med 2014;370:1965-7
151. Schmidt C. Assessing mammography's benefits and harms. J Natl Cancer Inst 2014;106:dju103

Acknowledgements

Acknowledgements

This thesis was possible owing to the efforts of many persons and I would like to express my sincere gratitude to the people who have encouraged and supported me on this task.

Professor Vitor Rodrigues, my supervisor for accepting to guide me through this work and share with me his expertise on breast cancer screening.

Professor Guilherme Gonçalves, my co-supervisor, for all his encouragement, and epidemiological skills.

Dr. Vitor Veloso, president of the Portuguese Cancer League – North Branch and Dr. Laranja Pontes, director of the Portuguese Cancer Institute of Porto for their confidence in my epidemiological work.

Dr. João Carvalho (*in memorium*) and Dra. Ana Teresa Aguiar, respectively former and actual director of the Breast Cancer Screening Programme of the Northern Region, who from the beginning of this work expressed their support and put at my disposal all the necessary information.

My Collegues Dr. Eduardo Ferreira, Professor Vicente Gonçalves, Professor Carlos Lopes, Dr. Silva Caspurro, Dr. Isabel Carvalho, Dr. Maria Jose del Rio and Dr. Fernando Pardal, without their collaboration some studies would not have been accomplished.

D. Leopoldina Amaral from Breast Cancer Screening Programme from the Northern Region for sharing her deep practical knowledge of the screening organization.

Dr. Manuel António Silva and Doutora Joana Bastos from the ROR-Centro for their valuable and indispensable help with the Cancer Registry database.

Many thanks to the persons from Breast Cancer Screening Programme from the Centre Region, namely, Professor Carlos Oliveira, Dr. Miguel Pina, Eng. Pedro Praça and Dra. Margarida Veiga for providing quality data for this work.

I want to express my sincere thanks to Beatriz Serrão (always in my heart), Clara Castro, Filipa Gonçalves and Luís Antunes for making a great job, with lots of well-humoured moments, smoothing the harshness of the working days.

To Conceição Leal, for our long friendship, for sharing “bolas de Berlim” with me during coffee break and for her clever advise on breast cancer pathology.

Finally, my warmest thanks belong to my family, Dirk, Dinis and Matilde, their infinite comprehension, encouragement and help on all levels have kept me going on.