

## 6. SUMMARY

---

### 6.1 Breast cancer

Breast cancer is the most commonly diagnosed cancer in women and the most common cause of cancer death in women worldwide. Globally, it is estimated that in 2012 there were 1.68 million new diagnoses (25% of all new cancer diagnoses in women) and 0.52 million deaths (15% of all cancer deaths in women) from breast cancer, corresponding to age-standardized incidence and mortality rates of 43.3 and 12.9 per 100 000, respectively. Thus, in 2012 there were an estimated 6.3 million women alive who had had a diagnosis of breast cancer in the previous 5 years (more than one third of all 5-year prevalent cancer cases in women). The largest contributor to the global burden was East and Central Asia (including China and India), where more than one third of the cases and more than 40% of the deaths occurred. In 2012, more than a 3-fold variation in the age-standardized breast cancer incidence rates was recorded between countries in North America and western Europe (rates > 90 per 100 000) and countries in Central Africa and East and South-Central Asia (rates < 30 per 100 000). In many high-income countries, 5-year survival rates now reach 80–90% (with 10-year survival rates of 60–70%), whereas in low- and middle-income countries (LMICs), 5-year survival rates may be less than 60% and as low as 12%. Globally, about one third of breast cancer cases are diagnosed in women younger than 50 years, and about one half in women aged 50–74 years; however, the mean age of diagnosis

is lower in LMICs. In most countries, an increase in incidence rates and a decrease in mortality rates were evident over recent decades, beginning in many countries before the implementation of mammography screening programmes. In those countries where screening was introduced in the 1980s and 1990s, the increase in incidence rates has been most evident in the age group of invited women.

Invasive adenocarcinoma of the breast is a malignant tumour that penetrates the basement membrane and spreads via both the blood and lymphatic systems, progressing to regional lymph node and systemic metastasis. Invasive breast cancers vary in morphological and molecular genetic characteristics, clinical features, and prognosis. The main non-invasive form of breast carcinoma in situ, ductal carcinoma in situ, has at least a 40% likelihood of progression to invasive cancer when untreated. Most benign breast lesions have no known relationship to the development of invasive breast cancer. However, some forms of breast epithelial proliferation, such as usual epithelial hyperplasia and atypical hyperplasia, are associated with an increased risk of subsequent breast cancer (by 1.5–2.0-fold and 2.5–4-fold, respectively).

Established breast cancer risk factors include early menarche, late menopause, later age at first pregnancy, nulliparity and low parity, little or no breastfeeding, higher body mass index at postmenopausal ages, and tall stature. Other factors associated with an increased risk include

low physical activity levels, alcohol consumption, certain exogenous hormone therapies, mammographic density, a history of proliferative benign breast conditions, and a family history of breast cancer. Exposure to ionizing radiation is linearly associated with an increased breast cancer risk. The risk shows an inverse relationship with age at exposure, with very low or no risk for women exposed after age 50 years and an increase in risk for women exposed before age 40 years. In addition to the above-mentioned breast cancer risk factors, genetic factors are of particular importance. The risk increases with the number of affected first-degree relatives and is most pronounced in young adults. Mutations in the high-penetrance genes *BRCA1* and *BRCA2*, together with mutations in additional cancer susceptibility genes, account for approximately 27% of all hereditary breast cancer cases and 5% of all breast cancer cases. The majority of cancer susceptibility genes code for tumour suppressor proteins involved in critical DNA repair pathways, which may increase the radiosensitivity of women in this population.

In LMICs, breast cancer cases are frequently diagnosed at more advanced stages than those in high-income countries, mostly due to the lack of effective diagnostic services. The mortality and morbidity associated with advanced disease may be reduced through early diagnosis of symptomatic breast cancer or early detection of breast cancer by screening in asymptomatic women. Promotion of breast cancer awareness may be a feasible option for early detection in settings with limited resources where screening is not feasible.

Comprehensive quality assurance, via evaluation and monitoring of performance indicators, is essential to maintain an appropriate balance between the benefits (mainly reduced mortality from breast cancer) and harms of screening. Quality assurance of breast cancer screening requires appropriate, sustainable resources for planning, coordination, and training.

## 6.2 Implementation of breast cancer screening worldwide

There is a social gradient to participation in breast cancer screening. Income, education level, place of residence, age, health, access to general health services (including screening), and cultural factors are among the factors that influence participation. Knowledge about breast cancer and screening is associated with higher participation. Worry about breast cancer and perceived risk of breast cancer are also associated with higher participation, but fatalism about cancer is associated with lower participation. Acculturation among minority women and immigrant women in settings with access to screening is usually associated with higher rates of screening.

Informed decision-making is a principle that underpins participation in screening; however, laypeople may conceptualize informed choice differently from policy-makers. Professionals debate about what constitutes appropriate information to provide to women, especially about overdiagnosis and false-positive test results (see Section 6.3.3a, b).

Participation in breast cancer screening can have psychological or psychosocial consequences for women, either from the invitation to screening or from the outcome, which may in turn affect further participation in screening (see also Section 6.3.3d).

### 6.2.1 Europe

Breast cancer screening is well established in western Europe and is delivered according to a common pattern, which has been guided by the activity of projects funded by the European Union. Some countries, particularly those in central and eastern Europe, have less well developed programmes or have not yet implemented screening. Breast cancer screening is delivered mainly by organized programmes, as

encouraged by the European Commission, which has published quality assurance guidelines, now in their fourth edition.

Participation rates vary across Europe, from less than 20% in Poland to more than 85% in Finland, with an estimated average of just less than 50%. Commonly, participation rates are higher among more affluent and more educated women and lower among women of lower socio-economic status or from a minority or immigrant background.

### 6.2.2 North America

Breast cancer screening has been widely available in parts of Canada and the USA since the late 1980s or early 1990s and typically achieves population attendance rates of about 50%, varying from 30% to 60%. In Canada, breast cancer screening is delivered primarily through organized programmes, whereas in the USA, screening is opportunistic. Both countries have well-developed quality assurance programmes. In the USA, management of quality assurance is mandated by federal regulations. Both Canada and the USA have programmes to raise awareness. Women in Canada and the USA face similar barriers to breast cancer screening, including living in a rural area, low income, low education level, and minority status.

### 6.2.3 Latin America

In Latin America, there has been increasing activity in breast cancer screening during the past decade. Currently, almost all Latin American countries in which breast cancer is the leading cause of cancer mortality among women have national recommendations or guidelines; however, no country in the region has a screening system that meets all the criteria of organized screening programmes. Most countries use mammography screening combined with clinical breast examination (CBE) and breast

self-examination (BSE); half of the countries recommend mammography for women younger than 50 years. Screening participation rates vary enormously across and within countries, with large differences between urban and rural areas and by income level. There is intensive advocacy activity, and information is provided by governments, NGOs, and the media, which appear to have induced a good level of breast awareness, although in a non-coordinated manner.

### 6.2.4 Sub-Saharan Africa

With the exception of South Africa, no country in sub-Saharan Africa has developed national recommendations or guidelines for breast cancer screening; however, relevant activity by NGOs is present throughout the region, and a few governments have carried out periodic campaigns to promote breast awareness. No population-based data on screening participation were available for most countries, and the only available national survey from South Africa found that 15.5% of women reported having had a mammogram during their lifetime. Accordingly, diagnosis occurs at a late stage of the disease. Despite several initiatives to increase breast awareness and provide health education, poverty, the lack of governmental support, and sociocultural influences represent relevant barriers to breast cancer awareness and screening.

### 6.2.5 Central and West Asia and North Africa

Countries located in Central and West Asia and North Africa are heterogeneous, and this is reflected in terms of access to breast cancer screening. While high-income countries such as Israel, Kuwait, and Qatar have well-developed health services, most countries in this area are classified as LMICs, with limited resources allocated to health care. Recent and current emergencies in several countries in this area have

exacerbated previous problems, and screening is not available to women in these circumstances and is not a priority.

Some screening is available in the more affluent countries, and there is NGO activity in some areas. A few pilot and exploratory projects have taken place. Both awareness and participation rates are low. Israel has a well-developed breast cancer screening system, and participation is high.

### 6.2.6 South-East Asia

Among the four countries or areas that have national programmes based on cancer screening guidelines, organized screening is present in the Republic of Korea, Singapore, and Taiwan, China, but not in Japan. The age group younger than 50 years has been included in the target population for breast cancer screening, except in Singapore. Some countries, such as China and Indonesia, have local community-based screening programmes. In several countries, such as India, screening is performed only within research studies. For the efficient use of limited resources, Thailand is developing risk-prediction models to target only women at an increased risk. National programmes for cancer control and prevention of noncommunicable diseases have promoted breast cancer awareness in Asian countries.

### 6.2.7 Oceania

Australia and New Zealand provide organized screening programmes for breast cancer. The target age groups were expanded in the past decade, in Australia to women in their early seventies and in New Zealand to women in their late forties. In the past decade, the participation rate has remained about 50% in Australia and has increased from 50% to more than 70% in New Zealand. Australia, Fiji, and New Zealand have national programmes for breast care awareness.

Because minority groups have low participation rates in breast cancer screening, they have been the major target of programmes to promote breast cancer awareness.

## 6.3 Mammography screening

The technology, technique, and interpretation skills of mammography have advanced enormously since its early development, leading, chiefly, to improved sensitivity and specificity, and reduced radiation doses. Digital mammography provides improved sensitivity in moderately dense breasts. Digital breast tomosynthesis produces three-dimensional mammographic images, allowing better visualization and localization of potential lesions. The radiation dose of digital mammography with tomosynthesis is approximately twice that of mammography alone, but is significantly reduced by reconstruction of two-dimensional images from the three-dimensional images. Many countries have developed detailed guidelines for quality control in mammography screening.

### 6.3.1 Efficacy of mammography screening from randomized controlled trials

Efficacy of a specific intervention generally refers to its beneficial effect under ideal circumstances. In practice, it is rarely possible to assess true efficacy. Randomized controlled trials (RCTs), which have been designed initially to assess whether mammography screening may reduce breast cancer mortality, increase life expectancy, and reduce the number of women undergoing aggressive treatments, may suffer from a low compliance rate, contamination in the control arm, long screening intervals, or suboptimal quality.

The Working Group considered all 10 randomized trials of breast cancer screening that have been conducted to be eligible for evaluation. These trials, initiated from 1963 until

1991, are: the Health Insurance Plan trial (USA); the Malmö I and Malmö II trials (Sweden); the Two-County trial (Sweden); the Stockholm trial (Sweden); the Gothenburg trial (Sweden); the Canadian National Breast Screening Study trials, CNBSS 1 and CNBSS 2 (Canada); the Edinburgh trial (United Kingdom); and the United Kingdom Age trial (United Kingdom). Individual randomization was performed in the Health Insurance Plan, Malmö, CNBSS, and Gothenburg trials (the latter only in women aged 39–49 years), and cluster randomization in the other trials. The mean duration of follow-up for breast cancer mortality ranged from 9 years for the Malmö II trial to 25–29 years for the Two-County trial. All but two RCTs, which had screening of the control group by design, showed breast cancer mortality reductions of between 10% and 35% for women invited to screening (across the ages 39–74 years at entry); however, the reduction was statistically significant in only two trials (the Two-County and Edinburgh trials). Meta-analyses of the RCTs showed a statistically significant reduction in breast cancer mortality of about 23% in women invited to screening aged 50–69 years at entry. Concerns have been raised that cluster randomization may not achieve balance in critical risk factors for breast cancer. This effect was demonstrated as a bias in the Edinburgh trial; in the Two-County trial, substantial bias was found to be unlikely. For only the Health Insurance Plan and CNBSS trials were data obtained to confirm balance in the distribution of conventional risk factors for breast cancer in women in the compared arms.

Evidence from the RCTs for the efficacy of mammography screening of women starting at age 40 years and continuing to age 74 years in reducing breast cancer mortality is less extensive. The United Kingdom Age trial, which included women aged 39–41 years at entry, was the only trial aimed at answering the question of whether mammography screening at age 40–41 years is effective in reducing breast cancer mortality in

women diagnosed during their forties; a 17% statistically non-significant reduction in breast cancer mortality was found in the trial. For women aged 70–74 years, only in the Two-County trial was screening offered to this age group, and a 24% non-significant reduction in breast cancer mortality was reported.

The CNBSS trials incorporated screening by CBE and the teaching and reinforcement of BSE in both the mammography and the control arms. The CNBSS 2 trial for women aged 50–59 years specifically addressed the question of whether adding mammography screening to CBE leads to additional benefits, and found no difference in breast cancer mortality. By modelling of the individual data, it was estimated that a reduction of more than 20% in breast cancer mortality could have been derived from the CBE if compared with a no-screening arm.

Eight of the RCTs had reported cumulative incidence of advanced breast cancers (the Health Insurance Plan, Malmö, Two-County, CNBSS 1 and CNBSS 2, Stockholm, Gothenburg, and United Kingdom Age trials), with reductions varying from 3% to 31% in the individual trials.

It was not possible to estimate the average overdiagnosis in women screened from age 50 years to age 69 years (or 74 years), because many trials had provided screening also for the control group or had not reported data specifically for the screening of women in the age range 50–69 years.

Screening intervals were 12 months in the Health Insurance Plan, CNBSS, and United Kingdom Age trials, 18 months in the Gothenburg trial, 18–24 months in the Malmö trials, 28 months in the Stockholm trial, and 24 months for women aged 40–49 years and 33 months for those aged 50–69 years in the Two-County trial. However, the different designs of the trials preclude an assessment of the comparative efficacy of screening by different intervals. One additional trial in the United Kingdom was specifically designed to evaluate the effects of

varying screening frequency; reduction in breast cancer mortality was modelled based on results of tumour size, nodal status, and histological grade. No statistically significant difference was found between a 3-year and a 1-year interval for women aged 50–64 years.

### 6.3.2 Effectiveness of mammography screening

Evaluation of the effectiveness of screening on breast cancer mortality can use various design and analytical approaches. Incidence-based mortality (IBM) cohort studies and nested case-control studies are the most robust designs for evaluating the effectiveness of service mammography screening, when they achieve a sufficient follow-up time. All of the studies currently available for evaluation were performed in high-income countries.

#### (a) Incidence-based mortality cohort studies

##### (i) Women aged 50–69 years

Nineteen separate IBM cohort study analyses have estimated the overall effects on breast cancer mortality of invitation to mammography screening, with or without CBE, in women aged 50–69 years or in a wider age group including this range (beginning at < 50 years in eight analyses and ending at > 69 years in five analyses).

Given substantial overlaps in space and time among reports based on population-based mammography screening programmes in Sweden, Finland, and Norway, the Working Group considered only the more extensive studies for each country (two of the six analyses based on the Swedish mammography screening programme, one of the five analyses based on the Finnish breast cancer screening programme, and two of the three analyses based on the Norwegian breast cancer screening programme).

The relative risks from IBM studies for invitation to screening ranged from 0.58 (95% confidence interval [CI], 0.44–0.75), for year 8

to year 12 of screening in Navarre, Spain, to 0.94 (95% CI, 0.68–1.29), for the first 15 years of screening in Nijmegen, the Netherlands. The median relative risk of all studies considered was 0.77, between the values of 0.76 (95% CI, 0.53–1.09), based on the first 6 years of screening in Finland, and 0.78 (95% CI, 0.70–0.87), based on 12 years of screening in Finland (1992–2003). If all Norwegian studies are removed from the analysis, because of the introduction of multi-disciplinary breast cancer care centres in parallel with the roll-out of the organized screening programme, the median relative risk is 0.76. The remaining analyses included one each from Denmark, Italy, and the United Kingdom. The study in the United Kingdom, which included CBE (annual) with mammography (biennial), reported a relative risk of 0.73 (95% CI, 0.63–0.84). Lead-time bias was the most common residual bias and would be expected to be conservative.

Eleven independent informative IBM cohort study analyses of effects of participation in mammography screening on breast cancer mortality were considered, after exclusion of studies reporting a relative risk based on an analysis of invitation to screening multiplied by an estimate of the participation rate.

Two of the four analyses based on the Swedish mammography screening programme substantially overlapped in space and time, and thus the Working Group considered only the more extensive study. Two further studies, one in Sweden and one in Italy, were not, or were probably only partially, adjusted for self-selection bias. A third study, in Canada, although it was not adjusted for self-selection bias, provided an analysis of a small component of the data that suggested that self-selection bias (in populations in which only one third to one half of women had been screened) was small and conservative. The remaining five analyses included one each in Denmark, Finland, and Norway, and two in the USA. The relative risks for attendance to screening from these studies ranged from 0.57

(95% CI, 0.53–0.62), based on 11–22 years of screening in Sweden, to 0.80 (95% CI, 0.34–1.85), based on 3.5 years of organized screening in a health maintenance organization in the USA. The median relative risk was 0.60, from both the Danish programme (95% CI, 0.49–0.74) and the Canadian programme (95% CI, 0.52–0.67). The Norwegian study, on attendance in the breast cancer screening programme, is methodologically probably the best of the studies considered, since it was based on individually linked data for all women studied; the relative risk was 0.57 (95% CI, 0.51–0.64) for screening in the period 1996–2009. The two studies in the USA included CBE as part of the screening offered.

Overall, IBM cohort studies indicate reductions in breast cancer mortality of about 20% for women invited to screening and of about 40% for women who attend screening, in this age range.

(ii) *Women younger than 50 years or older than 69 years*

IBM analyses can provide evidence on the effectiveness of screening in women younger than 50 years if they are based on women who were only offered screening while they were younger than 50 years or are limited to women whose breast cancer was diagnosed while they were younger than 50 years. Similarly, to provide evidence on the effectiveness of screening in women older than 69 years, IBM analyses must be based on women first offered screening after age 69 years and limited to breast cancer deaths that followed a diagnosis of breast cancer when the women were older than 69 years.

Based on screening experience for most of Sweden in the period 1986–2005, the relative risk for invitation to screening at age 40–49 years was estimated to be 0.74 (95% CI, 0.66–0.83); it was 0.83 (95% CI, 0.70–1.00) for invitation at age 40–44 years, and 0.68 (95% CI, 0.59–0.78) for invitation at age 45–49 years.

Three studies provided evidence on the effectiveness of screening in women older than 69 years.

One study in the Netherlands reported an odds ratio of 0.89 (95% CI, 0.56–1.40) for women first invited to screening at age 68–83 years. One study in Sweden reported an odds ratio of 0.96 (95% CI, 0.73–1.25) for women first invited to screening at age 65–74 years. One study in Canada reported an odds ratio of 0.65 (95% CI, 0.56–0.74) for women first attending organized screening at age 70–79 years. An alternative analysis of the Swedish data, using estimated excess mortality from breast cancer instead of the number of breast cancers that were registered as the underlying cause of death, gave an estimated relative risk of 0.84 (95% CI, 0.59–1.19); this alternative could be justifiable if there was material error in assignment of underlying cause of death in older women in this study. The Canadian study was limited by lack of adjustment for self-selection bias and lack of consideration of probable opportunistic screening before acceptance of an invitation to organized screening at age 70 years or older.

(b) *Case-control studies*

(i) *Women aged 50–69 years*

Eleven separate case-control studies conducted in Europe and Australia provided relevant data on the effectiveness of mammography in service screening programmes. Most of these studies enrolled women invited for screening at ages 50–69 years; two included women younger than 50 years at invitation, and four included women older than 69 years at invitation. Although some studies were conducted in the same geographical area, the studies were judged to have no effective overlap and hence to be independent. In these case-control studies, odds ratios for all ages ranged from 0.24 (with correction for self-selection bias) to 0.75.

Eight additional case-control studies conducted in Europe and the USA provided relevant data on the effectiveness of mammography screening conducted in other settings. Three

of the studies included women younger than 50 years at invitation, and none included women older than 70 years. Odds ratios for the largest range of ages included in these studies ranged from 0.30 to 0.91.

Case-control studies typically provide estimates of the effect of screening for women who participated in screening compared with women who had been invited or to whom screening was otherwise offered but who did not participate. Non-participating women may have a different risk of dying from breast cancer, so this may result in selection bias in the absence of appropriate adjustment. Information bias can be considered minimal if the case-control study is based on systematic historical databases on screening, but may be larger in other types of case-control studies. Self-selection bias can be assessed by comparing breast cancer mortality rates in unscreened women with those in screened women just before service screening started; in practice, self-selection bias has been shown to be limited in service screening programmes with high attendance rates. The results of case-control studies indicate that breast cancer mortality is reduced by about 48% in screened women.

(ii) *Women younger than 50 years or older than 69 years*

Case-control study analyses can provide evidence on the effectiveness of screening in women younger than 50 years if they are based only on deaths from breast cancer of women whose cancer was diagnosed when they were younger than 50 years or whose last screening or invitation to screening before diagnosis of breast cancer was while they were younger than 50 years. Similarly, to provide evidence on the effectiveness of screening in women older than 69 years, analyses must be based on women first offered screening after age 69 years and limited to breast cancer deaths that followed a diagnosis of breast cancer when the women were older than 69 years.

Six case-control study analyses estimated the effectiveness of invitation to or attendance of screening at ages 40–49 years (five studies) or below age 50 years (one study) in reducing breast cancer mortality. Odds ratios for invitation or attendance ranged from 0.50 to 1.18, with only one greater than 1.0. The two studies in women invited to attend the screening programme in Nijmegen, the Netherlands, analysed some of the same breast cancer deaths.

One case-control study provided a potentially valid estimate of the effectiveness of first attendance of screening at age 65–74 years, with an odds ratio of 0.54 (95% CI, 0.31–0.95) in women ever screened in that age range. (The breast cancer deaths included as cases in this study probably include most of those in the Dutch cohort study of women first invited to screening at age 68–83 years referred to above.)

(c) *Ecological studies*

Despite their lower value, ecological studies may be appropriate for evaluating population-level interventions, such as screening, when geographical areas or population groups are expected to be similar in cancer risk except for the introduction of screening. The Working Group considered that accurate information on standards of breast cancer treatment in different regions analysed and careful matching of regions by treatment standards or adjustment for differences between regions in treatment standards are minimum criteria for validity of ecological studies. Of the 87 studies considered, 5 studies were included in the review. Of those, three found benefits from mammography screening and two did not. Thus, evidence from the small number of informative studies was consistent with that from cohort studies and case-control studies.

(d) *Stage-specific incidence*

Overall, studies that compared incidence rates of advanced breast cancer in screened versus unscreened populations showed significantly



lower rates of advanced cancers in screened women. Ecological studies, which are based on cancer registries, without distinction between breast cancer cases detected by screening or otherwise (intention to screen), reported smaller differences.

(e) *Effect of adjuvant therapy on effectiveness of screening*

Adjuvant systemic therapy has been increasingly used since the late 1980s, and has thus probably affected the effects of screening. Two important studies have recently reported on this issue. A study using micro-simulation modelling reported that in 2008, adjuvant treatment was estimated to have reduced the breast cancer mortality rate in the simulated population by 13.9%, compared with a situation without treatment; biennial screening between age 50 years and age 74 years further reduced the mortality rate by 15.7%. Another modelling study, which included six natural history models for the population in the USA and used very similar techniques, reported that in 2000, screening and adjuvant treatment were estimated to have reduced breast cancer mortality by 34.8%, compared with a situation with no screening or adjuvant treatment; a reduction by 15.9% was estimated to have been a result of screening, and 23.4% as a result of treatment.

### 6.3.3 Adverse effects of mammography screening

Early detection of breast cancer by mammography screening is associated with harms, of which the most important are false-positive results of the screening test, overdiagnosis, and possibly risk of radiation-induced breast cancer.

(a) *Cumulative risk of false-positive recall*

The cumulative risk of a false-positive recall, an important harm of screening, is defined as the cumulative risk of recall for further assessment

at least once during the screening period (usually 10 biennial screening episodes in organized programmes) minus the cancer detection rate over the same period. There is a similar definition for the cumulative risk of recall with a subsequent invasive procedure (needle biopsy or surgical biopsy) and a benign outcome. There are large differences in estimates of the cumulative risk between organized breast cancer screening programmes and opportunistic screening. The modelled estimate of cumulative risk of false-positive recall in organized screening programmes in Europe is about 20% for a woman who had 10 screenings between the ages of 50 years and 70 years; less than 5% have an invasive procedure. In opportunistic screening, such as in the USA, rates of recall are higher, and the protocols for assessment are different; the cumulative risk of having at least one false-positive recall after 10 years of screening has been estimated to be about 40% with biennial screening and about 60% with annual screening, and these rates are similar for women starting screening at age 40 years and at age 50 years.

(b) *Overdiagnosis*

Overdiagnosis refers to the detection by screening of breast cancers (ductal carcinoma in situ and invasive) that would never have been diagnosed clinically if the women had not been screened. Overdiagnosed breast cancers are treated because they cannot be distinguished from cancers that would progress if not treated; therefore, treatment is the main component of the harm of overdiagnosis. The epidemiological quantification of overdiagnosis in observational studies is important because estimates may be influenced by local screening practice and technological innovations.

The Working Group noted and endorsed the classification of measures of overdiagnosis suggested by the Independent United Kingdom Panel (measures A to D). Use of this classification when reporting overdiagnosis estimates

will enhance the prospects of valid comparison between overdiagnosis estimates made in different studies and in different screening programmes.

RCTs have shown that after the drop in incidence that follows the end of regular screening has occurred, there is a persistent excess of diagnosed cases, which can give an estimate of the number of overdiagnosed cases. Based on a start of screening at age 40–69 years and a follow-up time of at least 10 years after the end of the screening period, two RCTs with long follow-up periods estimated overdiagnosis to be 4–12% of all cancers detected in control (unscreened) women over the same follow-up period (measure A). As a proportion of screen-detected cancers only, the estimate was 22–29% (measure D). To obtain a truly valid estimate of overdiagnosis in RCTs, there should be no screening after the trial has ended in either the study or the control arm. It is doubtful whether any RCT has met this requirement. Moreover, the RCT estimates relate to screening performed in the 1980s, and there are no pooled age-specific estimates (e.g. for women aged 40–49 years or 50–69 years).

The methodology for evaluating overdiagnosis in observational studies, based mainly on organized programmes, has varied widely across studies. Two main approaches, aided by modelling, are currently proposed. The cumulative incidence approach follows a population (cohort or dynamic) over time, including over the period of the compensatory drop in incidence after the end of screening. Models have estimated that breast cancers may be screen-detectable up to 10 years before they would present clinically (i.e. screening has a lead time of up to 10 years), although the issue is controversial and others have argued for shorter lead times. Assuming a lead time of up to 10 years, a follow-up period of at least 5–10 years after the end of screening attendance is needed to include the compensatory drop. The second approach involves statistical adjustment for the lead time that has produced the excess of cases

initially. A further challenge in estimating overdiagnosis is proper allowance for any underlying trend in incidence with time or adjustment for exposure to factors confounded with screening (e.g. hormone replacement therapy) that may cause such a trend. Studies evaluate incidence rates in populations invited and not invited to screening, or screened and not screened, and in the latter case bias from self-selection for screening should be taken into account.

The Working Group considered 30 observational studies that reported estimates of overdiagnosis. Their results varied widely; estimates of the overdiagnosis risk, principally the Independent United Kingdom Panel's measure A, ranged from –0.7% to 76% for invasive cancer only and from 1% to 57% for in situ and invasive cancers together. For 13 of these studies that were considered to be adequately adjusted for underlying trend in breast cancer incidence and for lead time, the measure A estimates ranged from 2% to 25% for invasive cancer only and from 2% to 22% for in situ and invasive cancers together.

### (c) *Risk of radiation-induced breast cancer*

The low dose of X-ray photon radiation received during mammography is a potential adverse effect of breast cancer screening, since exposure of the breast to ionizing radiation may induce breast cancer. The number of breast cancers induced by mammography is estimated through risk assessment approaches, which use a range of hypotheses about risk model, latency time, correction factor for low dose and dose rate, mean glandular dose to the breast during mammography, targeted population, and screening modalities. For biennial screening from age 50 years to age 74 or 80 years (with follow-up until age 85 years or older), the estimated number of breast cancer deaths induced by mammography screening ranges from 1 to 7 per 100 000 women screened. These estimates are smaller than estimates of breast cancer deaths prevented by mammography screening

by a factor of at least 100. For 10 years of annual screening from age 40 years to age 49 years (with follow-up until age 85 years or older), the estimated number of breast cancer deaths induced by mammography screening ranges from 8 to 20 per 100 000 women screened.

#### (d) *Psychosocial consequences*

Studies of the psychological impact of false-positive mammography, which were summarized in seven reviews, showed varied results. Some studies reported that women who have further investigations after a routine mammogram experience anxiety in the short term, and possibly in the long term. Also, some studies reported that some women with false-positive results conducted more frequent BSE and had higher levels of distress and anxiety, although not apparently pathologically so, and thought more about breast cancer than did those with normal results; in other studies, the effects were limited to breast cancer-specific outcomes. Two of the reviews concluded that the process decreased women's quality of life for weeks and even months.

#### 6.3.4 *Cost-effectiveness of mammography screening*

Decisions about implementation of health-care interventions are based primarily on health benefits and a favourable harm-benefit ratio, but – to use limited resources efficiently – are also often based on cost-effectiveness analyses. A cost-effectiveness analysis compares different policies, including the current one, with no intervention (average cost-effectiveness) or compares a more-intensive programme with a less-intensive programme (incremental cost-effectiveness). Effects are often defined as disease-specific deaths prevented and life years gained but are ideally adjusted for quality of life, resulting in quality-adjusted life years.

Ideally, all possible screening policies that are of relevance are compared in a cost-effectiveness

analysis. However, it is not feasible to compare all scenarios of interest in an RCT or observational study. By the use of mathematical models, findings from screening trials and observational studies are extrapolated to simulated populations. Numerous cost-effectiveness analyses showed that organized mammography screening, often biennially, is cost-effective. Despite their greater effectiveness, screening strategies that consist of annual screening are often found to be less efficient and less cost-effective, due to a disproportionate increase in costs or due to diminishing returns; about 80% of the effect of annual screening is retained when screening is performed every 2 years.

Several studies have assessed the cost-effectiveness of CBE, mass media awareness-raising campaigns, limited mammography screening, and increasing the coverage level of treatment in LMICs. However, evidence on the effectiveness of these approaches in these countries is still absent.

## 6.4 Other imaging techniques

### 6.4.1 *Techniques*

Ultrasonography is performed using handheld ultrasonography (also called two-dimensional [2D] ultrasonography) or automated breast ultrasonography (also called three-dimensional [3D] ultrasonography). Since with handheld ultrasonography only a very small selection of images seen during acquisition is recorded for interpretation, image acquisition requires high diagnostic skills to minimize selection error. This problem may be eliminated by using automated breast ultrasonography, in which all images are recorded. Screening with ultrasonography has been used mostly as an adjunct to mammography in women with dense breasts. In addition, use of ultrasonography as a primary tool for breast cancer screening has been reported recently in China. Knowledge about quality assurance of

image acquisition or reading of breast ultrasonography is still limited.

Digital breast tomosynthesis, a three-dimensional approach to digital mammography, is described in Section 6.3.

Magnetic resonance imaging (MRI) without contrast agent and MRI spectroscopy have not been applied or validated for screening use, and their application is being tested for diagnostic use. Contrast-enhanced MRI has been evaluated as an adjunct to mammography in studies of women at an increased risk (see Section 6.5). Potential side-effects of the magnetic field (in women with metallic devices) must be considered. Contrast-enhanced MRI screening also leads to risk of severe kidney disease and severe allergy. Costly equipment, false-positive test results, and the expensive assessment of MRI-only detected lesions result in high costs for this technique. No quality assurance programme has yet been established for MRI screening.

Positron emission tomography (PET) and positron emission mammography (PEM) involve intravenous application of radioactively marked [<sup>18</sup>F]-fluorodeoxyglucose to measure glucose metabolism, which is assumed to be higher in tumours. Other metabolites could be measured but have not been validated for clinical use. PET has a lower resolution and signal-to-noise ratio than PEM. No study has evaluated screening by PET or PEM. In the diagnostic situation, PEM has sensitivity and specificity comparable to those of MRI. Due to the slow clearance time of the radioactive marker from the body, PEM (like PET) is associated with a radiation dose to the whole body 16 times that for mammography.

Scintimammography measures the uptake of radioactively marked <sup>99</sup>Tc-sestamibi, which binds to mitochondria. The density of mitochondria is assumed to be increased in tumours. A single study assessed the validity of scintimammography for screening, but it included a high percentage of women at an increased risk of breast cancer. In that study, sensitivity and

specificity were comparable to those of MRI. The radiation dose received for scintimammography and similar technologies is 9–20 times that for mammography.

Infrared spectroscopy measures spectral differences in the examined tissue, and the proportions of haemoglobin and deoxyhaemoglobin have been suggested to differ between benign and malignant tissue. Thermography measures temperature distribution in the examined tissue, assuming that malignant tissue has a higher temperature. Electrical impedance imaging measures conductivity and impedance, relying on the assumption that cancer cells have increased conductivity and thus decreased impedance. Initial clinical experience and/or attempts to use these methods for screening have generally yielded lower sensitivity and specificity than those of standard imaging technologies. None of these methods has been validated for screening.

Molecular imaging uses vectors that emit a fluoroscopic or scintigraphic signal attached to targeting agents, which might identify molecules within the cell membrane or cellular matrix of tumours. Development of such agents is in the preclinical stage.

#### 6.4.2 Effectiveness in screening

##### (a) Ultrasonography

Nine observational studies (the majority retrospective) conducted in Austria, Italy, and the USA assessed ultrasonography as an adjunct to mammography for breast cancer screening in women with dense breast tissue and negative mammography. The incremental breast cancer detection rate ranged from 1.9 per 1000 screens to 4.0 per 1000 screens. In one additional prospective study in China in women screened with mammography and ultrasonography without restriction to those with dense breasts, adjunct ultrasonography detected additional cancers in 1 per 1000 screened women. However, none of

the studies had a comparison or control group, and some included women at an increased risk of developing breast cancer. Ultrasonography-only detected cancers were frequently early-stage cancers, generally at a comparable or earlier stage than cancers detected by mammography. Two of these studies reported estimates of interval cancer rates of 1.1 per 1000 screens and 1.7 per 1000 screens at 12 months of follow-up, but interpretation of these estimates is limited due to the lack of a comparison group and to substantial heterogeneity in the underlying breast cancer incidence rates in study populations.

All available studies consistently showed that adjunct ultrasonography substantially increases rates of false-positive recall or testing. Five studies reported incremental rates of false-positive biopsy (mostly surgical biopsy) of between 1.2% and 2.8%, and seven studies reported additional false-positive testing or follow-up in 1.7% to 7.5% of screens.

There were no observational studies assessing screening efficacy in terms of mortality reduction or assessing screening impact using surrogate end-points for screening efficacy.

#### (b) *Digital breast tomosynthesis*

In five non-randomized studies of digital mammography with tomosynthesis (also referred to as integrated 2D/3D mammography), two of which were prospective trials within population-based programmes, the incremental breast cancer detection rate relative to digital mammography ranged from 0.5 per 1000 screens to 2.7 per 1000 screens. Two of four observational studies reporting cancer stage distribution showed that the incremental detection was of invasive tumours, whereas the other two studies showed incremental detection of in situ and invasive tumours. One observational study reported an estimated interval cancer rate of 0.8 per 1000 screens at 12 months of follow-up, but interpretation of this estimate is limited due to the lack of a comparison group.

Digital mammography with tomosynthesis reduced rates of false-positive recalls in four informative observational studies, with absolute decreases in false-positive recalls ranging from 0.8% to 3.6% of screened women, representing reductions of 15% to 36% in false-positive recalls.

Given the dual acquisition of images, digital mammography with tomosynthesis increases the radiation dose received by approximately doubling the mean glandular dose; however, this will depend on the exact technology used and the number of acquisitions. Based on one observational study, reconstruction of the 2D images from the tomosynthesis acquisition decreases the radiation dose by 45% compared with the dual acquisition and yields similar incremental cancer detection to that from the dual acquisition.

## 6.5 Screening of women at an increased risk

### 6.5.1 *Women with a BRCA1/2 mutation*

Fourteen prospective cohort studies of women with a *BRCA1* or *BRCA2* mutation assessed the screening performance of MRI plus mammography performed in the same screening round, with a review of the diagnostic test performed. The sensitivity and specificity of mammography in this population of women were about 40% and 95%, respectively; corresponding values for MRI plus mammography were about 95% and 80%, respectively, showing a clear increase in sensitivity and decrease in specificity compared with mammography alone.

Four prospective cohort studies assessed reduction in breast cancer mortality in women with a *BRCA1* or *BRCA2* mutation screened with mammography. The studies reported varying results, from a 5-year all-cause survival of 63% in *BRCA1* mutation carriers to a 6-year all-cause survival of 93% in *BRCA1/2* mutation carriers. In the only study in which the breast cancer-specific survival of women with a *BRCA1* mutation

screened annually with MRI plus mammography was compared with that in unscreened women with a *BRCA1* mutation, a significant difference in 10-year breast cancer-specific survival was found (95.3% in the screened group vs 73.7% in the unscreened group).

### 6.5.2 Women with a high familial risk without a *BRCA1/2* mutation

Two prospective cohort studies of women with a high familial risk without a *BRCA1* or *BRCA2* mutation assessed the screening performance of MRI plus mammography performed in the same screening round, with a review of the diagnostic test performed. The reported estimates for the sensitivity and specificity of mammography were 25–46% and 95–97%, respectively; corresponding values for adjunct MRI were 73–100% and 89–98%, respectively.

### 6.5.3 Women with a high familial risk with or without a *BRCA1/2* mutation

One observational study with long-term follow-up reported a shift to a lower stage of the tumours detected in women with annual MRI and mammography screening compared with women without intensified screening.

In the 10 studies that evaluated the sensitivity of ultrasonography in women with a high familial risk with or without a *BRCA1* or *BRCA2* mutation, the sensitivity was comparable to or lower than that of mammography and was always lower than that of MRI. No study assessed the specificity of ultrasonography.

Seven prospective cohort studies assessed the incremental cancer detection rate of CBE in women with an increased familial risk screened with MRI plus mammography, with or without ultrasonography. None of the studies addressed the effect of CBE alone. Five of the studies did not detect any additional cancers; in the remaining two studies, which reported a lower screen

detection rate, a total of 4 out of 243 cancers (1.6%) were found by CBE only.

### 6.5.4 Women with a personal history of breast cancer (*invasive or in situ*)

One large multicentre study assessed mammography screening in women with a personal history of breast cancer compared with those without such a history (58 870 screens in each group). The sensitivity and the specificity of mammography were significantly lower in women with a personal history of breast cancer compared with those without such a history.

One comparative study assessed the value of adding ultrasonography to annual mammography in women with a personal history of breast cancer versus women with various types of risk factors for breast cancer. The incremental cancer detection rate was comparable between the two groups; when ultrasonography was added to mammography, the recall rate increased significantly, from 11.5% to 26.6%.

In a small substudy that assessed the value of adding MRI to annual mammography plus ultrasonography in women with a personal history of breast cancer versus those without such a history, the recall rate increased significantly, from 16.3% to 36.3%.

### 6.5.5 Women with lobular neoplasia or atypical proliferations

One large multicentre comparative study assessed mammography screening in women with lobular carcinoma in situ (LCIS) or atypical proliferations compared with women without such lesions (2505 and 12 525 screens, respectively). The sensitivity of mammography in women with LCIS or atypical proliferations was not statistically significantly lower than that in matched controls; however, the specificity was lower. Four studies (two comparative and two non-comparative) evaluated a series of patients

to examine the sensitivity of MRI in screening women with LCIS or atypical hyperplasia. In the non-comparative studies, high sensitivities were reported for the MRI screening in women with LCIS. In the comparative studies, women with such lesions selected to undergo MRI screening were younger and had stronger family histories of breast cancer. In addition, MRI screening generated more recall biopsies compared with mammography.

## 6.6 Clinical breast examination

CBE is a simple technique involving visual inspection and systematic palpation of both breasts and nipples by a trained health-care provider. This technique has a moderate sensitivity (range, 50–60%) and a specificity of more than 85%.

Three RCTs, two conducted in India and one in the Philippines, assessed the efficacy of CBE alone versus no screening. All three studies reported a significant shift to a lower stage of the tumours detected (early detection). Although the study in the Philippines was stopped after one round of screening, the two studies in India are currently under way and the effect of CBE on breast cancer mortality in these studies is awaited.

Two RCTs showed that CBE in combination with mammography reduced breast cancer mortality compared with no intervention in women older than 50 years. In the earlier study, conducted in 1963 in the USA, 67% of the tumours were detected by CBE and mammography, and 45% were detected by CBE alone. In the other study, conducted in 1979 in the United Kingdom, 74% of the tumours were detected by CBE and mammography, and 3% by CBE alone. In an RCT conducted in Canada, CBE plus mammography screening did not show a significant mortality benefit compared with CBE alone. In addition, five observational studies, conducted mostly in the 1970s, reported that CBE contributed 5–10%

in incremental detection rate over and above mammography.

CBE is a low-cost intervention and thus a feasible screening modality in LMICs.

## 6.7 Breast self-examination

Several techniques for BSE have been described, with the number of steps ranging from 21 to 34. Women are unlikely to perform such elaborate techniques, and hence simpler techniques have been recommended. Structured training and individual instruction have been shown to improve compliance with BSE practice. Sensitivity, specificity, and positive predictive value of 58.3%, 87.4%, and 29.2%, respectively, have been reported for BSE. Breast cancer awareness, socioeconomic status, level of education, and availability of privacy are the principal determinants of BSE practice.

Two RCTs of BSE have been conducted. A study in St Petersburg, Russian Federation, compared women who received intensive instruction in BSE and annual reinforcement sessions, plus annual CBE, with women who received only annual CBE. A study in Shanghai, China, compared women who received intensive BSE instruction, periodic reminders, two reinforcement sessions 2 years and 4 years after initial instruction, and periodic practice sessions under the supervision of a medical worker, with women who received no BSE instruction or any other type of breast cancer screening. In both studies, after about 10 years of follow-up, there were no differences between the instruction and control arms in breast cancer mortality rates, in breast cancer incidence rates, in the size or stage of the breast cancers, or in survival rates in the cancer cases. In both RCTs, more benign lesions were detected in the instruction arms than in the control arms. In the St Petersburg trial, the frequency of BSE practice declined with time after initial instruction and after a re-education programme; in the Shanghai trial, no

information on compliance was collected. One possible explanation for the trial results is poor compliance. Both trials were conducted in populations with easy access to diagnostic and treatment facilities, and the women in the control groups of both studies presented with relatively small tumours.

Two of three observational cohort studies showed reduced mortality from breast cancer in women who received BSE instruction, but the results are likely to be due to factors unrelated to BSE practice. Results of four case–control studies provided inconsistent results with regard to the relationship between the frequency of BSE practice and the risk of fatal or advanced breast cancer (as a surrogate for breast cancer death). However, two studies showed weak decreasing trends in the risk of fatal or advanced disease with increasing level of proficiency of BSE. In a study at Duke University, USA, women at moderate to high risk of breast cancer who received annual screening by mammography and MRI were given detailed BSE instruction in conjunction with CBE two or three times a year. All 12 interval cancers were detected in women who reported practising BSE competently and regularly, and 6 of the cancers were initially detected by BSE.

Surveys of BSE practice in the general population in LMICs as well as surveys in women before and after receiving BSE instruction have generally shown that the percentages of women who report practising BSE are too low to be likely to have a meaningful impact on mortality from breast cancer.