REVIEW ARTICLE

Benefit and Risk of Mammography Screening

Considerations From an Epidemiological Viewpoint

Nikolaus Becker, Hans Junkermann

SUMMARY

Introduction: An organized mammography screening program is being implemented in Germany since the end of 2005. Even before its inception, a heated debate surrounded the question of how to communicate its rationale appropriately.

Methods: Selective literature review identifying facts relevant to informing decision making about mammography attendance.

Results: Decision making about individual attendance for mammography screening is crucially dependent on the reduction of risk of death from breast cancer for a diseased woman. This can be communicated in the following way: out of 100 women with a diagnosis of breast cancer 31 will die within the following 10 years without screening and 20 with screening (35% fewer).

Discussion: The actuarial concept of risk is defined by the probability of the undesired event multiplied by the expected magnitude of loss. This approach appears more appropriate for the appraisal of risk in the context of mammography screening than a pure event probability. It could help understand why women attend screening even if they perceive the risk of getting breast cancer as low – namely, that they rank the consequences of getting the disease as highly significant for themselves and others in their social environment.

Dtsch Arztebl Int 2008; 105(8): 131–6

DOI: 10.3238/arztebl.2008.0131

Key words: epidemiology, breast cancer, early detection, screening, quality assurance, randomized studies, survival time

For some time there have been calls for a change in the quality of the presentation of the benefits and disadvantages of screening. These demands are accompanied by the claim that the conventional epidemiological presentation of the rationale for screening is unclear or even misleading and therefore precludes the provision of accurate information (1). This position is accompanied by the aspiration to develop scientifically based methods of communicating risks (2). However, it misinterprets fundamental relationships and cannot remain unchallenged in the light of increasing public interest in early diagnosis and prevention.

Material and methods

A selective analysis of the literature was performed to determine the main considerations that motivate the decision to attend mammography screening, and to rate the evidential value of different modes of presentation. Based on this literature, the benefits and risks are quantified by our own calculations.

Results

Purpose of early diagnosis

The purpose of screening examinations is to detect a disease in an early, still preclinical stage in order to treat it more effectively and thereby lower morbidity and mortality (3). Mammography screening is the method used to identify still unidentified breast cancer in the asymptomatic population. The goal is to prolong the survival of subjects with already identifiable breast cancer by providing early treatment.

Survival time is defined as the period elapsing between diagnosis and death. In breast cancers discovered by early detection programs, survival times are also prolonged by bringing forward the time of diagnosis in cases in which the disease course is not influenced by providing treatment earlier. The effectiveness of screening therefore cannot be gauged from the survival times. Mortality is the only variable quantifiable without bias for studies on the effectiveness of early detection activities (3–5).

Results of randomized studies

The question addressed in the randomized studies (6) on mammography screening was whether periodic systematic mammography screening is a suitable

Deutsches Krebsforschungszentrum, Abteilung Krebssepidemiologie, Heidelberg: Prof. Dr. rer. nat. Becker

Sektion Senologische Diagnostik, Universitätsfrauenklinik Heidelberg: Prof. Dr. med. Junkermann

means of significantly reducing mortality associated with breast cancer in the general population. Since the target population in the studies was first randomized and only women assigned to the "active" study group were invited for screening, it was possible to simultaneously quantify the effectiveness of the screening modality "mammography screening" and the effectiveness of an invitation to periodic mammography screening as a health care measure (6). The former is expressed as a reduction in mortality among the women who actually followed the invitation, the latter as a mortality reduction among all invited women.

The re-evaluation of all studies on this topic which was performed by the WHO's cancer research institute (International Agency for Research on Cancer, IARC) in response to the re-emergence of a heated debate concerning the effectiveness of mammography screening in 2002 revealed a mean reduction in breast cancer mortality of 25% for subjects in the 50 to 69 years age range who were invited to screening and of about 35% among those who actually attended (6). The rate of participation in the studies was about 70%.

In consideration of the goal of early detection programs, the results may be interpreted as follows: early diagnosis of still asymptomatic breast cancer by systematic mammography screening during the study follow-up period prevented a fatal disease course in approximately 25% and 35% of cases, respectively. In other words: mortality is the product of the number of diseases multiplied by the probability of survival and death in the event of breast cancer occurring. Based on a mean 10-year survival probability of 69% (7), the result of the studies can be summarized as follows (*table 1*): of 100 women with a diagnosis of breast cancer, 31 will die during the following 10 years without offered screening and 23 with offered screening (25% fewer). This is the message for health policy for which the effectiveness of a screening program at population level is decisive. The second of the two figures is relevant for decision making about individual attendance for mammography screening: of 100 women with a diagnosis of breast cancer, 31 will die during the following 10 years without screening attendance and 20 with screening attendance (35% fewer) (*box*). The probability of developing the disease is also an important factor. It may vary between countries and therefore cannot be meaningfully derived from the Swedish studies. For Germany, it is about 5% for the age range 50 to 69 years.

In terms of health policy it is interesting to define how many cases of death from breast cancer can be prevented by a screening program: the figure is in the range of about 2500 to 3000 cases annually. This estimate is based on an assumed rate of approximately 55 000 new cases per year, of which about 40% to 50% occur between the ages of 50 to 69 years (8), a 70% screening attendance rate, and the mentioned 10-year survival rate of 69%.

A 10-year period, however, is an arbitrary figure on which to base communication of the effect of

mammography screening and results in considerable underestimation especially when considering the absolute effects. Follow-up periods of 16 years and more are now available which show that the effect is also maintained after such long periods (9). The actual number of deaths from breast cancer that can be prevented by screening is therefore higher than estimated above. Moreover, even these follow-up periods of 16 years and more do not provide a complete picture of the screening effect, because in all studies screening was also commenced in the control groups after a few screening cycles. A complete presentation of the study outcomes in absolute figures would require the publication of lifelong breast cancer specific survival times, but such data are not available. All the observed absolute figures therefore underestimate the achievable effects.

Disadvantageous effects

The benefits of screening are balanced by unavoidable disadvantageous effects:

- False-positive findings
- Tissue biopsies of benign lesions
- Radiation exposure
- Overdiagnosis.

The quantitative extent of false-positive findings, associated biopsies, and the extent of overdiagnosis is shown in *table 2*. The figures are based on the outcome variables for mammography screening among 1000 attendees as specified in the European guidelines.

Accordingly, within a two-year screening interval an abnormal mammography gives rise to 30 to 50 further investigations, such as magnification mammography, ultrasound or clinical examination, and, to resolve further existing abnormalities, 10 biopsies. A mean 0.5 cases of breast cancer resulting from overdiagnosis may be expected per 1000 attendees in the two-year screening interval.

Cumulated over 10 screening cycles – equivalent to 20 calendar years – these figures represent the anticipated disadvantageous effects from the viewpoint of 50-year-old women: with reference to 1000 women, 50 women develop breast cancer during these 20 years. About 260 to 400 women are invited to further evaluations following discovery of mammographic abnormalities, and 222 to 362 prove to be false-positive. In 100 women in whom the finding still remains abnormal after the follow-up examinations, tissue biopsies, e.g., in the form of core or vacuum biopsies, are recommended. These procedures show another 62 abnormal findings to be false-positive. Up to 10 women are advised to have surgical tissue excision due to suspected breast cancer, with benign postoperative findings. 38 women are diagnosed with breast cancer (true positive finding). In five women breast cancer is diagnosed which would never have appeared without screening (so-called overdiagnosis).

Combining these results and the finding that about five deaths from breast cancer are prevented among

TABLE 1

Estimated breast cancer mortality1**

Event	Number	Explanation
Without screening		
Number of women followed up	100 000	Age range 50 to 69 years
New breast cancer cases per year	250	0.25% of 50 to 69-year-olds
New breast cancer cases in 2 years	500	2 x 0.25% of 50 to 69-year-olds
Of these, death within 10 years after diagnosis	155	About 31% (see 10-year survival rates; [7])
Number of breast cancer patients who survive for at least 10 years – without screening –	500 – 155 = 345	About 69% of affected persons
With screening		
Number of women attending screening	100 000	
New breast cancer cases in 2 years (one two-year screening cycle)	500	Of these, about ¾ detected in screening (375 cases) and ¼ in the interval (125 cases, see [12])
"Overdiagnoses" which would not have appeared as disease without screening	50	10% overdiagnosis with reference to incidence without screening (see [13]) ^{*2}
Of these, death within 10 years from breast cancer	$0.65 \times 0.31 \times 500 = 101$	65% of mortality without screening due to 35% mortality reduction by screening; 31% mortality with 69% 10-year survival rate (see above); overdiagnosed cases do not die of the disease
Mortality balance		
Breast cancer mortalities prevented by one screening cycle	155 – 101 = 54	Per 100 000 screening attendees and 500 persons with disease, 54 breast cancer mortalities are prevented within a 10-year period Or: per 100 persons with disease, 31 without screening and 20 with screening die within 10 years of breast cancer
Breast cancer mortalities prevented by 10 screening cycles (total duration of screening program)	540	540 breast cancer mortalities are prevented (about 5 per 1000) per 100 000 screening attendees on attendance at all 10 screening cycles offered

*1 with reference to new cases of breast cancer diagnosed in 50 to 69-year-old women taking into account a follow-up period of 10 years without screening and in screening attendees (own calculations based on the cited literature)

*2 This is the only randomized study to date from which overdiagnosis could be directly quantified

1000 screening attendees (table 1), this means that, for one prevented death from breast cancer, up to 10 breast operations with benign findings are performed and one case of overdiagnosis occurs. By far the commonest disadvantageous effect, however, is a false-positive mammography result.

Radiosensitivity of mammary tissue is age dependent. It is highest in the age group between 10 and 20 years and decreases with age. The German Radiation Protection Commission (Strahlenschutzkommission, SSK) states that two-yearly mammography screening in the 50 to 69-year-old age group cumulatively leads to an additional risk of 0.01% to 0.24% of a 50-year-old woman developing breast cancer during her lifetime. The benefit (number of prevented deaths from breast cancer) therefore exceeds the risk (number of breast cancer deaths caused by radiation exposure) at least 100 fold.

Presentation based on absolute risks

The method of presentation used by Mühlhauser and Höldke (1) is based on the view that statements of relative risk – such as the statement that mammography screening reduces the risk of dying of breast cancer by 30% – are unclear or misleading.

The authors therefore avoid stating relative probabilities and report the results of the randomized studies to which the relative risk data refer, in the form of absolute percentages or frequency data. For example, it is stated: "Without mammography screening, 4 of 1000 women die of breast cancer within a 10-year period. With mammography screening, 3 of 1000 women die of breast cancer within a 10-year period". Or: "Without mammography screening, 996 of 1000 women do not die of breast cancer within a 10-year period. With mammography screening, 997 of 1000 women do not die of breast cancer within a 10-year period". Or finally, described by the authors as an

BOX

Key statements

- The effectiveness of an invitation to periodic mammography screening is the result of the reduction of breast cancer mortality among all invited women (not all of whom follow the invitation) and is 25%. The effectiveness of the screening procedure is the result of the reduction of breast cancer mortality among the women who did follow the invitation and is 35% (5).
- For individual decision making, the risk reduction in the event of illness is decisive and can be expressed as a frequency: of 100 women with a diagnosis of breast cancer, 31 will die within the following 10 years without screening and 20 with screening (35% fewer).
- In the actuarial definition of risk the probability of the undesired event is multiplied by the expected magnitude of loss if the undesired event occurs. This takes into account the fact that it is also advisable to insure against very rare events if they could result in significant loss. In mammography screening this explains why women decide in favor of screening even if they rate the probability of developing breast cancer as relatively low.

"unconventional, but accurate" presentation: "Of 1000 women with mammography screening over 10 years, 999 have no benefit because they would also not have died of breast cancer without mammography screening" (1).

The presentation of the results in the form of relative probabilities (breast cancer mortality decreases by 35%) is also open to doubt (1, 2) because studies have shown that both physicians and potential screening attendees were unable to correctly convert the usual relative percentages into absolute figures.

Discussion

In view of the possible modes of presentation, the question arises as to which is most suitable and is properly understood by physicians and the lay public who are relatively unfamiliar with epidemiological topics.

Conditional probabilities, relative risk, and the actuarial definition of risk

Various concepts are available for the assessment of risk. The epidemiological definition of risk comprises only the likelihood of an undesired event occurring. The actuarial definition of risk also includes the magnitude of loss: the probability of the undesired event is multiplied by the presumed magnitude of loss if the event occurs. This takes into account the fact that it is also advisable to insure against very rare events which, if they occur, could spell economic ruin for the person concerned.

The assessment of risks is a task which has to be continuously mastered in everyday life. When assessing for what risks provisions should be made, attention is focused mainly on the magnitude of the potential harm and only secondarily on the probability of its occurrence. In these cases one usually thinks in terms of conditional probabilities (e.g. "if the house burns down, then I'll be ruined"). Although the event is extremely unlikely in individual cases, building insurance is virtually universal and in Baden-Württemberg, Germany, is even compulsory. Other examples could be given which prove that thinking in terms of conditional probabilities and a concept of risk that includes the extent of possible harm is not unusual and, on the contrary, permeates social life on many levels.

There is thus good evidence that for many women the reason for attending mammography screening is not the actual probability of the event, i.e. the likelihood of becoming ill or dying, but the thought of the consequences of the illness. The high participation rates in the Scandinavian countries and in the German model projects, as well as in the first screening units of the nationwide screening program in the Federal Republic of Germany support this interpretation.

This means: a 50-year old woman who perceives the approximately 5% probability of developing breast cancer over the next 20 years as comparatively low, may possibly nevertheless decide in favor of screening because she rates the consequences of a disease as significant both for herself and her social environment. From this perspective, the benefit of screening emerges as the conviction that – in the event of becoming ill – one has done everything possible according to existing medical knowledge to limit the serious consequences. All attendees in screening programs benefit from this aspect, not only the women who actually become ill, and among those who do become ill, not only those whose death could be prevented.

The method of presentation propagated by Mühlhauser and Höldke (1) completely disregards this aspect. The authors combine the two variables most important for deciding whether to attend – the risk of becoming ill and the increased probability of survival in the event of disease onset during screening attendance – in a manner which is misleading for individual decision making. In responding to a woman's question about the benefits of screening attendance, it is unsuitable to quote a figure that relates the affected person's improved chances of cure to all women entitled to attend, i.e., to the 95% who do not develop breast cancer. This is the basis on which Mühlhauser and Höldke ([1], table 1) arrive at the stated 0.07% absolute risk reduction by screening with reference to breast cancer mortality (see also [10] and the response in [11]). The question remains as to whether a presentation of the numerical data with conditional probabilities need necessarily be difficult to understand.

Early diagnosis is concerned with prolonging the survival of the affected persons, but this cannot be

TABLE 2

Disadvantageous effects in a mammography screening program*1

Event	Number	Explanation
Number of women attending screening	1000	
Screening cycle		
Conspicuous findings per screening cycle	30–50	3% to 5% of attendees within 2 years (12)
Biopsy based diagnoses	10	1% of attendees within 2 years (Swedish guidelines, see also [14])
Overdiagnoses	0.5	10% overdiagnosis with reference to the incidence without screening*2
Number of new disease cases and events of undesirable adverse reactions within 10 screening cycles in 50-year-old women (with reference to 1000 women)		
Number of women who develop breast cancer within the next 20 years (=10 screening cycles)	50	Cumulative probability of 5% of developing disease ($20 \times 0.25\%^{*2}$)
Number of women with breast cancer diagnosis due to overdiagnosis	5	10% of breast cancer incidence without screening*2
Number of women with positive mammography findings	260–400	Cumulative probability of 26% to 40%; consequence is repeat invitation for further diagnostic evaluations
Number of women with false-positive mammography findings	223–363	Cumulative probability of 22.25% to 36.25% Positive findings minus correct positive findings ($10 \times 1.5 \times \text{incidence} = 3.75\%^{*2}$)
Number of women with biopsy	100	Cumulative probability of 10% ($10 \times 1\%^{*2}$)
Number of women with a false-positive indication for a biopsy	63	Cumulative probability of 6.25% Proportion of biopsies minus the biopsies with correct positive findings ($10 \times 1.5 \times \text{incidence}^*2$)
Number of women with breast cancer operation resulting in <i>benign</i> findings	<10	Cumulatively < 25% of the detection rate ($10 \times 1.5 \times \text{incidence}$, see [12])

*1 Own calculations based on the cited literature.

*2 see table 1

directly quantified empirically for the reasons explained. The consideration of mortality is only the unavoidable detour required in order to nevertheless obtain an answer to the question whether survival is prolonged by early detection.

A presentation in the form of conditional probabilities (reduction of lethality in the event of disease onset) is the most suitable way of approaching the problem. The fact that an incorrect understanding of such variables has been observed in surveys cannot be used as a justification for arguing against its use if it is appropriate for the purpose. Since conditional probabilities frequently play a role in decision making in daily life, the main concern is to achieve understanding of the facts about screening by establishing a relationship to such everyday life experiences.

It should be emphasized once again at this point that the figures presented in *table 1* are fictional values synthesized from partial results as a means of

presenting and quantifying the basic approaches in a simplified form. Data that are repeatedly requested in order to illustrate the procedures followed in screening and which are to be seen in *table 1* cannot be obtained empirically. Survival times in screening for example are necessarily systematically biased. Constructive procedures are therefore required.

The problems addressed are not peculiar to an organized screening program but reflect the difficulties generally involved in screening. They occur in the same manner in colorectal and cervical screening, and become quantitatively visible in an organized, quality assured program in contrast to opportunistic screening. In opportunistic screening, which does not allow systematic quality assurance, much more unfavorable figures must be assumed. The unavoidable side effects described underline the necessity for strict quality assurance (in all kinds of screening) in order to minimize their extent as far as possible.

Conflict of interest statement

The authors declare that no conflict of interest exists according to the guidelines of the International Committee of Medical Journal Editors.

Manuscript received on 8 May 2007; revised version accepted on 22 November 2007.

Translated from the original German by mt-g.

REFERENCES

1. Mühlhauser I, Höldke B: Information zum Mammographiescreening – vom Trugschluss zur Ent-Täuschung. *Radiologe* 2002; 42: 299–304.
2. Gigerenzer G, Edwards A: Simple tools for understanding risks: from innumeracy to insight. *BMJ* 2003; 327: 741–4.
3. Morrison AS: Screening in Chronic Disease. Second edition. *Monographs in Epidemiology and Biostatistics Volume 19*. New York, Oxford: Oxford University Press 1992.
4. Becker N: Screening aus epidemiologischer Sicht. *Radiologe* 2002; 42: 592–600.
5. Becker N: Die Rolle der epidemiologischen Qualitätsparameter im Mammographie-Screening. *Radiologie* 2006; 46: 984–92.
6. IARC: Breast Cancer Screening. IARC Handbooks of Cancer Prevention, 2002; Volume 7. International Agency for Research on Cancer. World Health Organization. Lyon: IARC Press 200; 179.
7. Brenner H, Stegmaier C, Ziegler H: Long-term survival of cancer patients in Germany achieved by the beginning of the third millenium. *Ann Oncol* 2005; 16: 981–6.
8. Gesellschaft der epidemiologischen Krebsregister in Deutschland e.V. und das RKI: Krebs in Deutschland. 5. Ausgabe. Saarbrücken 2006.
9. Nyström L, Andersson I, Bjurstam N, Frisell J, Nordenskjöld B, Rutqvist LE: Long-term effects of mammography screening: updated overview of the Swedish randomised trials. *Lancet* 2002; 359: 909–19.
10. Mühlhauser I: Ist Vorbeugen besser als Heilen? *Dtsch Arztebl* 2007; 104(25): A 1804–7.
11. Willich SN: Gesundheitliche Wertschöpfung. *Dtsch Arztebl* 2007; 104(26): A 1893–5.
12. Perry N, Broeders M, de Wolf C, Törnberg S, Holland R, von Karsa L, Puthaar E (eds.): European guidelines for quality assurance in breast cancer screening and diagnosis – Fourth Edition. Luxembourg 2006: Office for Official Publications of the European Communities.
13. Zackrisson S, Andersson I, Janson L, Manjer J, Garne JP: Rate of over-diagnosis of breast cancer 15 years after end of Malmö mammographic screening trial: follow-up study. *BMJ* 2006; 332: 689–92.
14. Junkermann H, Becker N, Peitgen HO: Konzept und Durchführung der Modellprojekte für Mammographiescreening in Deutschland. *Radiologe* 2001; 41: 328–36.

Corresponding author

Prof. Dr. rer. nat. Nikolaus Becker
Deutsches Krebsforschungszentrum
Abteilung Krebs Epidemiologie
Im Neuenheimer Feld 280
69120 Heidelberg
Germany
n.becker@dkfz.de