Chapter 4: How is statistical information typically represented in medical risk communication? The case of mammography pamphlets

In the remainder of this dissertation, I will focus on the question how natural frequencies can be used to educate medical lay people about the uncertainties and risks associated with diagnostic tests. To explore this question, I chose one specific example of a diagnostic test: the screening mammography.

In a screening mammography, women who do not show any symptoms of breast cancer[11] get an x-ray picture of their breasts (a mammogram), with the goal to detect breast cancer in its early stages to reduce mortality. As the screening is performed on apparently healthy individuals, the obligation to thoroughly inform potential participants about benefits and risks is seen to be even stronger than for tests and treatments that are performed on symptomatic persons (Marshall, 1996; McQueen, 2002). The reason is that in a screening, the number of participants who benefit from the test (those who have an early stage of the disease and would profit from early treatment) is rather small, whereas the side effects of the test (e.g., exposure to x-rays during mammography) affect all participants. Consequently, women should be explicitly informed about the benefits, risks, and efficiency of mammography screening before they decide to participate in it (Gigerenzer, 2002; Karsa, 1995; Marshall, 1996; Mühlhauser & Höldke, 1999). Informing women about these issues means giving them quite a number of facts about the test and the disease, and many of these facts include statistical information, such as the reduction of mortality through screening, the predictive values of positive and negative mammograms, the likelihood of false positive results and their consequences. There has been extensive research on what women know about breast cancer and mammography, and the conclusions are not optimistic: Women's knowledge about breast cancer and mammography screening is generally not very accurate, even though the topic is particularly popular in the media. For example, women have been repeatedly shown to overestimate the benefits of mammography screening on the one hand, while underestimating the risks on the other hand (e.g., Black, Nease, & Tosteson, 1995; Schwartz, Woloshin, Sox, Fischhoff, & Welch, 2000).

---

[11] If there were symptoms such as a palpable lump in the breast, mammography would also be used, but this would be *clinical* mammography, not *screening* mammography. I will only consider screening mammography here.

Can natural frequencies be used to improve women's knowledge and understanding of information about mammography screening? To answer this question, I focus on one specific source of information about mammography screening, namely, information pamphlets. I start by summarizing the most important facts about mammography screening and then illustrate the most frequent problems women have in interpreting these facts.

Facts about mammography screening

Benefits. To assess the benefit of mammography screening, one has to compare the breast cancer mortality in women who participate in screening to that of women who do not participate. The raw result from four Swedish randomized trials for women between 40 and 74 years of age was the following (Nyström et al., 1996, in Mühlhauser & Höldke, 1999): Out of 1,000 women who *did not* participate in mammography screening, 4 died of breast cancer. Out of 1,000 women who *did* participate in mammography screening, 3 died of breast cancer. The difference between the two groups, 4 breast cancer deaths versus 3, can be quantified in different ways: One way is the *absolute risk reduction*, that is, screening saved the life of 1 out of 1,000 women who participated in screening, a reduction of 0.1%. Another quantification is the *relative risk reduction*: Screening saved the life of 1 out of 4 women who would otherwise have died from breast cancer, which is a reduction of 25%. The relative risk reduction is the most frequent way of presenting the benefit of mammography screening in the media. A typical sentence is: "Mammography screening reduces breast cancer mortality by 25%." But although this sentence is frequently reported, the debate about the question if mammography screening programs really reduce breast cancer mortality and for what groups of women is ongoing (for reviews, see Kerlikowske, 2000; Olsen & Gøtzsche, 2001).

Risks. The cost of participating in mammography screening includes side effects during mammography such as exposure to x-rays or pain due to the pressure applied to the breast during the test, and other risks that can manifest after the screening. These risks include (see also Gigerenzer, 2002; Karsa, 1995; Mühlhauser & Höldke, 1999):

- *Psychological and physiological strain due to false positive results.* Women who receive a positive mammogram are called back for further investigation. Many of these women have a false-positive result. For women who undergo mammography screening for the first time, 9 out of 10 positive results prove to be false positives; and 1 in 4 women who have biannual mammograms over 10 years will receive at least one false-positive result (Mühlhauser & Höldke, 1999). Almost all women with false positives have to undergo an additional mammogram or an ultrasound scan. About 1 in 5 women with a false-positive result undergoes a biopsy (Mühlhauser & Höldke, 1999) that bears, as an invasive treatment, its own risks such as wound infections and scarring. False positives can also have psychological costs. Women experience a considerable amount of stress and anxiety in the weeks between the false-positive mammogram and the negative result of the biopsy, and while some are simply relieved afterwards and go back to normal life (Scaf-Klomp, Sandermann, van de Weil, Otter, & van den Heuvel, 1997), others experience anxiety about breast cancer and mood impairment that can persist for up to 3 months (Lerman et al., 1991). Women with false positives have to undergo these examinations although they do not benefit from them and even bear new risks and stress. Of course, it can only be determined post hoc if the examinations were justified, and women might be willing to accept this "just to make sure," but nevertheless the potential consequences of receiving a positive mammogram should be made clear before women decide to participate in mammography screening.

- *Radiation-induced breast cancer.* This risk depends strongly on the dose of radiation and the age of the woman at exposure. It is estimated that out of 10,000 women, between 2 and 4 women who started to have annual mammograms at the age of 40 will develop radiation-induced breast cancer, and 1 to 2 of them will die (Mühlhauser & Höldke, 1999).
- *Unwanted early detection of precancerous lesions.* Because improved mammograms show lesions in ever more early stages of development, there is a danger of overtreatment (Napoli, 1997; Olsen & Gøtzsche, 2001). There is the possibility that a woman will receive massive treatment, even a mastectomy, because of a lesion that would have never developed into cancer in her remaining lifetime (e.g., slow-growing tumors, or the "carcinoma in situ," a lesion that does not progress into cancer in 5–9 out of 10 cases; Gigerenzer, 2002).
- *Early detection of breast cancer does not equal longer life-expectancy*. An early diagnosis does not automatically prolong the life-expectancy of women who have breast cancer. But these women have to live longer with the diagnosis (Gigerenzer, 2002; Karsa, 1995).

Test efficiency. This aspect includes information about the sensitivity and specificity of the test, as well as the predictive values of positive and negative results (see Chapter 1). In a large American study with over 26,000 women between the ages of 30 and over 70 years who participated in a first mammography screening, the sensitivity was 90% and the specificity was 93.5% (Kerlikowske, Grady, Barclay, Sickles, & Ernster, 1996). A meta-analysis over several systematic screening programs found over all age groups and for a one-year interval sensitivities between 83% and 95% and specificities between 93% and 99% (Mushlin et al., 1998, cited from Mühlhauser & Höldke, 1999). The error rates, especially the sensitivity, are influenced by the age of the women, due to changes in breast tissue (better sensitivity in older women), but also by the radiological criteria being used and the training and experience of the radiologist (Mühlhauser & Höldke, 1999). As outlined earlier, the positive and negative predictive values depend on the error rates and the prevalence of breast cancer. In the American study mentioned above, the positive predictive value was 10% and the negative predictive value was 99.9%. The test efficiency is relevant not only for the evaluation of the quality of the diagnostic test itself (and the decision to participate in it), but also for interpreting the test results (General Medical Council, 1998; Gigerenzer, 2002; Mühlhauser & Höldke, 1999; Slaytor & Ward, 1998). Although a positive mammogram is a stressful event for any woman, it is plausible to assume that the interpretation of the meaning of a positive mammogram can influence its perceived threat: Women who know that 9 out of 10 positive results later prove to be false positives might be less shaken by a positive mammogram than women who believe that a positive result indicates breast cancer with very high certainty (Gigerenzer, 2002; see also Marteau, 1995).

Misperceptions about mammography screening

Many studies have assessed how much women know about mammography screening and how accurate this knowledge is. These studies have identified several misperceptions. First, most women tend to overestimate the risk of developing and dying from breast cancer and are not aware of the fact that these risks are age dependent (Black et al., 1995; Dolan, Lee, & McDermott, 1997; Paepke, Schubert, Hüttner, Blohmer, & Lichtenegger, 2000). Second, most women overestimate the benefit of screening (Black et al., 1995; Schwartz et al., 1997; Woloshin et al., 2000). In one study, the majority of the 302 participating women clearly overestimated the absolute and the relative risk reduction through mammography screening even though they had read one of the two risk reduction rates just before estimation (Schwartz et al., 1997). Third, a majority of women is not informed about the risks of mammography screening. 60% of the participating women in an Australian study (Cockburn, Pit, & Redman, 1999) and 92% in an American study (Schwartz et al., 2000) said that a mammography has *no* potential negative effects for a woman without breast cancer.

These are the most frequently reported misperceptions about mammography screening. How accurate is the knowledge of women concerning the test efficiency of screening mammograms? Only few studies have addressed this question. Most women seem to know that false-negatives and false-positives can occur (Schwartz et al., 2000). One study found that a third of the participating women thought that the false-negative rate was 20% or higher, while it is actually around 10% (Cockburn, Redman, Hill, & Henry, 1995). But to the best of my knowledge, there are no studies that asked women to estimate the error rates and predictive values of screening mammograms. There is, however, one study that assessed these estimates from patients for five different diagnostic tests (Hamm & Smith, 1998). This study found that medical lay-people assumed similar error rates and positive predictive values for all diagnostic tests, independent of actual test efficiency. They expected rather low error rates (false-negatives were perceived to be more likely than false-positives) and very high positive predictive values. If women applied this rationale to the test efficiency of screening mammograms, then one could expect that they would also overestimate the positive predictive value of screening mammograms.

Why do these misperceptions occur?

Several reasons for the above-mentioned misperceptions about mammography screening have been proposed in the literature.

The overestimation of breast cancer risk and mortality is typically attributed to the widespread use of misleading risk figures in publications about breast cancer and mammography screening (Baines, 1992; Black et al., 1995; Dolan et al., 1997; Gigerenzer, 2002; Napoli, 1997; Phillips et al., 1999). The estimates of women for breast cancer incidence *and* mortality within the next one or two decades were often very close to the highly publicized “1-in-10” figure (“One of every 10 women will be diagnosed with breast cancer during her life”). This figure actually refers to a woman’s cumulative chances of developing breast cancer by the age of 85, and it can be found in the large majority of publications about breast cancer by public health institutions as well as by the popular media (it was first used by the American Cancer Society in the early 1990s to improve compliance with mammography screening programs; Phillips et al., 1999). This is the most dramatic way of describing the risk, that is, the risk figure with the largest number possible. That the 1-in-10 figure refers to a cumulative risk is typically not made clear in the media, but without such information, most women are not able to see that although the cumulative lifetime risk is 10%, the risk of developing breast cancer in any given decade of life never exceeds 3.4% (Gigerenzer, 2002; Phillips et al., 1999). Put differently, the reference class of the 1-in-10 figure is often not specified clearly enough. And as was described in Chapter 3, ambiguous reference classes often lead to misunderstandings. The results of the above-mentioned studies indicate that women misinterpret the lifetime risk as a short-term risk of developing or even dying from breast cancer (Black et al., 1995; Gigerenzer, 2002; Phillips et al., 1999).

In another study, the overestimation of the benefit of mammography screening was explained with innumeracy, since most of the participants in the study had low scores in some very basic mathematical tasks (Schwartz et al., 1997). But the overestimation of the benefit was also found in numerate women (Black et al., 1995). One explanation could be that the benefit of mammography screening is usually expressed as the relative risk reduction, but almost never as the absolute risk reduction (Gigerenzer, 2002; Slaytor & Ward, 1998; Steckelberg, Balgenorth, & Mühlhauser, 2001). As can be seen from the example above, the relative risk reduction contains a larger number than the absolute risk reduction (25% vs. 0.1%). It therefore suggests a larger amount of benefit, and people indeed prefer interventions that are advertised with their relative risk reduction over those that are advertised with absolute risk reduction (Hux & Naylor, 1995; Malenka, Baron, Johansen, Wahrenberger, & Ross, 1993; Sarfati, Howden-Chapman, Woodward, & Salmond, 1998). Moreover, the absolute risk reduction is seen as the more relevant information for women who consider participation in mammography screening (Atkins, 1997; Gigerenzer, 2002; Malenka et al., 1993), because it refers to the group of women who participate in screening (rather than the group of women who would die without screening, as the relative risk reduction), which is exactly the group that the women belong to at this point.

The lack of knowledge about risks of mammography screening is usually attributed to the fact that there is too little coverage of the limitations and disadvantages of mammography screenings in campaigns by cancer organizations and the popular media, compared to the coverage of benefits (Napoli, 1997). An analysis of 58 Australian mammography pamphlets showed that information about the accuracy of mammography screening was only provided occasionally (Slaytor & Ward, 1998). While the sensitivity was mentioned in 26% of the pamphlets, none of them gave information about the specificity or the positive predictive value. Another finding of the Australian pamphlet analysis was an emphasis on incidence rather than mortality to communicate the risk of breast cancer to women. The lifetime risk of developing breast cancer was stated in 60% of the pamphlets, whereas only 2% mentioned the lifetime risk of dying from breast cancer. The authors call this partiality “worrying”, because the goal of mammography screening is to reduce mortality, not incidence (Slaytor & Ward, 1998). The lack of information about test efficiency, mortality, and risks was also found in another sample of pamphlets on cancer screening, here screening for colorectal cancer (Steckelberg et al., 2001). The reason for the omission of certain information about cancer screening tests seems to be the concern that information about disadvantages would lower participation rates in screening.

To sum up, most of the explanations for the misperceptions focus on the question *what* information about mammography screening is provided to women in health information materials and the popular media. The answer has been mostly that not enough relevant information has been given to the women. The question *how* information about mammography screening is presented has not received the same amount of attention, but some authors suggested that misleading reference classes could be responsible for misunderstandings of the 1-in-10 figure and the benefit of screening. The use of different statistical formats in information about mammography screening has so far not been investigated.

As mentioned above, I would like to explore how natural frequencies could be used to improve understanding of mammography information. I will approach this question in two steps. First, I will analyze one source of information about mammography screening, namely mammography pamphlets. The descriptive analysis of the currently available mammography pamphlets has two goals, (a) to examine how often and in what form the most important topics about mammography screening are covered in the pamphlets, and (b) to identify potential sources of misperceptions that should be eliminated in the design of future pamphlets (Secker & Pollard, 1995). The second step will be to design an improved version of a mammography pamphlet and to test how different statistical formats influence understanding and evaluation of the pamphlet (Chapter 5).

Study 4: Analysis of German mammography pamphlets

Health information pamphlets are, after the physician and the popular media, the most important source of information on the early detection of breast cancer and mammography screening for women of all age groups, in Germany as well as in the US (Metsch et al., 1998; Paepke et al., 2001). Because they are relatively inexpensive and easy to distribute, they are particularly suitable for information about mass screenings such as mammography screening (Drossaert, Boer, & Seydel, 1996).

In the present study, I analyzed how often a sample of German mammography pamphlets mentioned specific topics about mammography screening, and which statistical formats were used to represent statistical information.

I also added a third level of analysis. To what extent is statistical information about mammography screening actually expressed numerically? Although the 1-in-10 figure and the 25-30% mortality reduction are widely publicized, health information materials often present general verbal rather than precise numerical information, for instance “Mammography is not perfectly accurate ” rather than “Mammography detects than 90% of breast tumors” (Slaytor & Ward, 1998; a similar observation was made for German pamphlets on HIV testing, Gigerenzer et al., 1998). While medical lay-people prefer to receive statistical information in a numerical mode, medical experts often prefer verbal expressions (Hallowell, Statham, Murton, Green, & Richards, 1997; Fox & Irwin, 1998; Erev & Cohen, 1990). One explanation for this finding is that numerical expressions are perceived by the speaker to be unnaturally precise, and the vagueness inherent in verbal expressions helps capture their uncertainty about the accuracy of the estimate (Fox & Irwin, 1998; Hamm, 1991; Heilbrun, Philipson, Berman, & Warren, 1999; Merz, Druzdzel, & Mazur, 1991). However, verbal expressions are less informative than numerical expressions. The translation of statistical information into simplified verbal expressions can be adequate in contexts where only a vague, summary representation of the information is sufficient. But when extensive and at the same time precise information is necessary, as in the case of mammography screening or other medical procedures, numerical expressions are better suited (Hamm, 1991; Hamm & Smith, 1998; Merz et al., 1991; Nakao & Axelrod, 1983; Gigerenzer et al., 1998). If physicians use verbal expressions, they should at least supplement them with precise numerical specifications from the literature (or, if not available, a numerical estimate that is clearly identified as an estimate) to reduce fuzziness in communication (Merz et al., 1991; Fischer & Jungermann, 1996; Nakao & Axelrod, 1983).

Method

In spring 2001, I contacted 34 institutions from the German health care system by phone and fax and asked them to send me all their pamphlets and written information on the topic of mammography and the early detection of breast cancer[12]. In October 2001, the

[12] The following institutions were contacted: Deutsche Krebshilfe, Deutsche Krebsgesellschaft mit regionalen Verbänden, Deutsches Krebsforschungszentrum, Krebsinformationsdienst, Bundesgesundheitsministerium, Bundeszentrale für gesundheitliche Aufklärung, Bundesamt für Strahlenschutz, Berufsverband der Frauenärzte, Berufsverband der Radiologen, Deutsche Gesellschaft für Senologie, Deutsche Gesellschaft für Gynäkologie

information providers were contacted again to update the sample of pamphlets obtained to date, i.e. to include the latest editions of the pamphlets previously obtained. Information from the Internet pages of the institutions (or from the Internet pages of third parties, if the institutions were linked to them; links from these pages to still others were not included) was entered into the sample of pamphlets as long as the content was not identical to printed pamphlets.

For the analysis of the pamphlets, a list of 30 possible aspects that could be covered in the pamphlets was constructed, based on the criteria of two previous analyses of health pamphlets (Slaytor & Ward, 1998; Steckelberg et al., 2001). A new item, "unspecified mortality reduction," was added to the list to categorize how often the ambiguous phrasing of the mortality reduction was used (see above); the final list with its 31 entries is displayed in Table 4.1. Out of these 31 aspects, 19 can be backed up by statistical data from the literature (subsequently called “quantifiable items”). Whenever a pamphlet mentioned a quantifiable item, it was noted if the statement was presented purely verbally or numerically. In the latter case, the statistical format of the given numerical information was also noted.

To be included in the analysis, a pamphlet had to fulfill two criteria: (a) It had to mention mammography in the context of the early detection of breast cancer, and (b) it had to communicate at least one of the items listed in Table 4.1. This double criterion was fulfilled by 21 of the pamphlets received by mail and 6 of the "on-line pamphlets" obtained from the Internet. Thus, altogether 27 pamphlets from 20 different information providers were analyzed. Thirteen of the 27 pamphlets were from 2000 or 2001; the oldest pamphlet in the sample was published in 1989 (but still in use); eight pamphlets did not mention the publication year. The pamphlets were analyzed independently by a co-worker[13] and myself. We agreed in 93% of the categorizations of the pamphlet contents; differing decisions were discussed until agreement was reached.

Results

Frequency of topics covered in the pamphlets

Table 4.1 shows how often the 31 topics were mentioned by the 27 pamphlets. Mentioned in 19 pamphlets, the yearly incidence of breast cancer and the improved survival rate through screening were the most frequent topics. The mortality reduction by

und Geburtshilfe, Mammographie-Screening-Planungsstelle, Fachverband Elektromedizinische Technik, gesetzliche Krankenkassen.

[13] I am grateful to Julie Pflaum for the co-analysis.

mammography screening was included in 11 pamphlets. It was notably mentioned in the newer pamphlets: of these 11 pamphlets, 8 were from 2000 or 2001. However, there was a problem in the formulation of the mortality reduction: In 7 of the 11 cases, it was formulated ambiguously. It was therefore not clear if the expression related to the absolute or the relative risk reduction, although the number mentioned (mostly "30%") points to the latter; these expressions are coded in Table 4.1 as "unspecified mortality reduction".

Table 4.1
Frequency of 31 topics in 27 German mammography pamphlets

Topics		N	%
Frequency of developing breast cancer	Breast cancer incidence per year	19	70%
	Relationship breast cancer risk and age	18	67%
	Cumulative lifetime risk of breast cancer	10	37%
Breast cancer mortality	Breast cancer mortality per year	3	11%
	Relationship breast cancer mortality and age	1	4%
	Cumulative lifetime risk of breast cancer death	1	4%
Benefit of the screening	Improved survival rate with screening	19	70%
	Unspecified mortality reduction	7	26%
	Absolute mortality reduction	3	11%
	Relative mortality reduction	2	7%
	Prospect of gentler treatment	2	7%
	Number needed to screen	1	4%
Risks of mammography screening	Pain during mammography	14	52%
	Health risks associated with radiation	12	44%
	Debate about the utility of mammography screening	6	22%
	Physical and/or psychological strain due to false-positive results	3	11%
	Unwanted detection of pre-cancerous lesions	3	11%
	Early detection of breast cancer does not equal longer life-expectancy	3	11%
Test efficiency	Sensitivity or false-negative rate	11	41%
	Specificity or false-positive rate	6	22%
	Criteria to judge the quality of the screening institution	6	22%
	Positive-predictive value	4	15%
	Proportion of women who get a positive test result in the screening	3	11%
	Negative-predictive value	1	4%
Additional information	Other methods to detect breast cancer early (breast self-examination, clinical examination)	24	89%
	Screening interval	23	85%
	Contact addresses for more information	21	78%
	Diagnostic tests after positive mammogram	20	74%
	Publication year of pamphlet	19	70%
	Description of mammography procedure	16	60%
	References to literature used	7	26%

Note. Percentages refer to N = 27 pamphlets.

Only few pamphlets mentioned more than one or two risks and side-effects of the screening. The most frequently mentioned disadvantages of mammography were that the test could be painful (14 pamphlets) and that the radiation poses additional health risks (12 pamphlets). Regarding the test efficiency of the mammography screening, the specificity and the positive and negative predictive values of mammography were only rarely communicated (although general expressions such as "Mammography is not 100% certain" could be found frequently). The positive predictive value was communicated in four pamphlets (in the Australian mammography pamphlet sample from Slaytor & Ward, 1998, it was not mentioned even once).

Table 4.2 Frequency of verbal and different types of numerical expressions of statistical information in the mammography pamphlets

			Numerical		
Topic	N	Verbal	Freqs	Percs	Freqs & percs
Frequency of developing breast cancer	*47*	*20*	*26*	*1*	*0*
Breast cancer incidence per year	19	5	14	0	0
Relationship breast cancer risk and age	18	15	3	0	0
Cumulative lifetime risk of breast cancer	10	0	9	1	0
Breast cancer mortality	*5*	*0*	*5*	*0*	*0*
Breast cancer mortality per year	3	0	3	0	0
Relationship breast cancer mortality and age	1	0	1	0	0
Cumulative lifetime risk of breast cancer death	1	0	1	0	0
Benefit of mammography screening	*32*	*15*	*4*	*13*	*0*
Improved survival rate with screening	19	11	1	7	0
Unspecified mortality reduction	7	2	0	5	0
Absolute mortality reduction	3	2	1	0	0
Relative mortality reduction	2	0	1	1	0
Number needed to screen	1	0	1	0	0
Risks of mammography screening	*18*	*15*	*1*	*1*	*1*
Health risks associated with radiation	12	10	0	1	1
Physical and/or psychological strain due to false-positive results	3	2	1	0	0
Early detection of breast cancer does not equal longer life-expectancy	3	3	0	0	0
Test efficiency	*25*	*11*	*6*	*7*	*1*
Sensitivity or false-negative rate	11	6	0	5	0
Specificity or false-positive rate	6	5	1	0	0
Positive-predictive value	4	0	3	0	1
Proportion of women who get a positive test result in the screening	3	0	1	2	0
Negative-predictive value	1	0	1	0	0
Overall	*127*	*61*	*42*	*22*	*2*

Note. Freqs = Frequencies, Percs = Percentages

Frequency and representation of numerical information

Nineteen of the 31 topics could be backed up by statistical data from the literature (Table 4.2). Overall, these 19 points were mentioned 127 times. Sixty-six of these statements included numerical, the remaining 61 statements included only verbal information. However, Table 4.2 shows that there were differences between the content domains. Basic information about breast cancer occurrence and mortality was mostly backed up by numerical data; only the short-term risk of developing breast cancer related to age was the exception here (typical verbal expression: "The risk of developing breast cancer increases with age."). For the benefit and the test efficiency of the mammography screening, numerical and verbal statements were about equally frequent. The only content domain in which verbal expressions were clearly more prevalent than numerical expressions were the risks of mammography screening.

Looking at those expressions that were presented numerically, which formats were used? In 42 of 66 numerical expressions, almost two-thirds of the numerical information was presented in absolute frequencies (Table 4.2). Twenty-two expressions were presented as percentages (relative frequencies), and 2 expressions used both representations. Probabilities were not used at all. The majority of percentage expressions could be found in the benefit domain. I also checked whether any other formats were in use such as chances or odds, but this was not the case.

Summary and discussion: Guidelines for pamphlet design

Let me start with an illustration of the information that was *typically* presented in the analyzed mammography pamphlets. If one collects from Table 4.1 the items that were presented most frequently, e.g. those that were mentioned in at least 40% of the pamphlets (excluding items of the section "Additional information"), the following text results:

Every year, about 43,000 women are diagnosed with breast cancer. The risk of getting breast cancer increases with age. Detected early, a breast tumor is today curable in most of the cases. Studies have shown that mammography screening can decrease breast cancer mortality by about 30% for women between the ages of 50 and 69. Mammography can detect most of the tumors, but not all of them. The pressure performed on the breast during mammography can be unpleasant, sometimes even painful. The radiation from mammography is very small. The benefit of mammography screening is much higher than the risk of getting breast cancer from the radiation.

When interpreting the results of this study, the heterogeneity of the pamphlet sample has to be kept in mind. As there is as yet no official mammography screening in Germany, information about mammography is often contained in other pamphlets (e.g. on breast cancer or cancer prevention in general). Hence, not all of the pamphlets analyzed here had the goal to treat the topic of mammography extensively. Nevertheless, it is plausible to assume that the present material will serve as a basis for the design of future mammography pamphlets and that, therefore, it is useful to specify recommendations on how these pamphlets can be improved (especially with regard to the potential introduction of mammography screening in Germany).

When designing a pamphlet that allows its readers to make informed decisions, the pamphlet designer's goal has to be twofold. The pamphlet has to provide all the information necessary to the reader, and at the same time, the information should be presented in a way that is as comprehensive as possible: "comprehension is as essential as disclosure" (Marshall, 1996, p. 379). The importance of these two aspects is stressed in the guidelines of both the British General Medical Council and the German Medical Chamber (Bundesärztekammer, 1990; General Medical Council, 1998).

The design of a health pamphlet that reaches the two goals involves careful consideration of many factors, from the size of the font used to the design of tables and graphics to the overall discourse structure of the pamphlet (Reschke, 1990; Wright, 1999a, 1999b). As mentioned above, I focus on two more general aspects, the choice of topics for the pamphlets and the representation of statistical information. In the following, will propose a set of guidelines concerning the two aspects, based on the results of the pamphlet analysis on the one hand and the literature review on the other hand.

Relevance. An ideal mammography pamphlet should present all the facts that are relevant for women considering participation. However, the analysis of German mammography pamphlets showed, just like in the Australian study (Slaytor & Ward, 1998), that the presentation of information in the pamphlets was not balanced. On the one hand, a large majority of the pamphlets did provide information about the incidence of breast cancer (70% of pamphlets), the benefit of higher survival rates through mammography screening (70%), and the recommended screening interval (85%). On the other hand, only a minority of the pamphlets informed women about the frequency of false-positive results (22%), the risk of psychological and physical strain due to such results (11%), and the predictive value of positive and negative (15% and 4%) mammograms. All these aspects of mammography screening should of course be included in mammography pamphlets to allow informed decision making, but simply adding a few phrases not sufficient.

Moreover, since pamphlets should include all the information the audience considers relevant for their decisions, it is important to involve the audience in the process of determining the contents for the pamphlets (Schriver, 1997; Wright, 1999a, 1999b). Readers of health information materials often have specific questions and directly skip straight to places in the document where they believe the relevant information is to be found; if they do not find relevant information (either because it is not in the pamphlet, or because they do not find it due to the structure), they will ignore the pamphlet (Wright, 1999a, 1999b). This implies that women could not profit even from an improved mammography pamphlet that includes all the aspects listed above (that is, topics that *experts* consider to be relevant for the screening decisions of women), as long as it still fails to answer those questions that women consider to be most relevant. Such a pamphlet would be inefficient because it would not reach the goal of enabling women to make an informed decision about mammography screening. Assessment of the patients' information need is therefore one of the very first steps in guidelines about the development of evidence-based patient information (Coulter, 1997b; Secker & Pollard, 1995). The information need of women concerning pamphlets about mammography screening will be assessed in the next chapter.

Up-to-dateness. Obviously, the information in the mammography pamphlets should be correct. The pamphlets have to be checked regularly to ensure correctness and updated if necessary. This was not the case for some of the analyzed pamphlets. For example, two pamphlets mentioned the "prospect of gentler treatment" as an advantage of mammography screening. A recent meta-analysis has shown, however, that the opposite seems to be true, namely, that the detection of a tumor in a screening leads to more aggressive treatment (Olsen & Gøtzsche, 2001).

Transparency. To make the information and recommendations given in a pamphlet more transparent to the reader, recency of the information and the literature used should be mentioned in any case (Ollenschläger, 2000). If the current state-of-the-art is unclear or disputed, this should also be made transparent in the pamphlet. This is especially important for the continuing debate about the utility of mammography screening (the debate has been mentioned in 6 of the 27 pamphlets), to allow women to obtain a clear understanding of the situation as a basis for their decisions, in spite of the numerous and partly contradictory media reports on the topic (on women's understanding of the mammography screening debate, see Woloshin et al., 2000).

Unambiguous reference classes. In the pamphlet analysis, the specification of the reference class was checked for two pieces of information, namely the cumulative lifetime risk of breast cancer (the 1-in-10 figure) and the mortality reduction of mammography screening. First, 10 pamphlets mentioned the "1 in 10" figure; only 1 of them added in parentheses "computed up to the 80th year of life" to indicate that the figure is the cumulative lifetime risk of developing breast cancer, rather than a short-term risk. Second, most statements about mortality reduction through mammography screening (7 out of 12) were ambiguous, e.g. "Mammography screening reduces breast cancer mortality by 25%". This statement leaves open to which group of women the reduction of 25% refers and can thus easily be misinterpreted. The size of the figure "25%" points to the relative risk reduction, but although experts might recognize this immediately, most lay people will not. Both cumulative lifetime risk and relative risk reduction should be clarified to avoid misunderstandings. A frequency table can be added to the 1-in-10 figure in which the lifetime risk of a population of concrete cases is broken down into risk estimated for several subgroups, for example, for women between 50 and 54 years, 55 and 59 years, and so on (Gigerenzer, 2002). The relative risk reduction should be explained together with, if not replaced by, the absolute risk reduction (Slaytor & Ward, 1998; Gigerenzer, 2002).

Precise and comprehensive numerical information. The pamphlet analysis showed that, out of the statements that could be backed up by statistical data, about 50% were expressed numerically. About two thirds of the statements that were expressed numerically were stated as absolute frequencies, the remaining one third as percentages. The use of frequencies and percentages depended clearly on the content domain: Breast cancer incidence and mortality were almost exclusively expressed as frequencies, whereas the benefit of screening was in most cases presented in terms of percentages. Based on the literature reviewed above, this differential use of frequencies seems suboptimal[14], because if natural frequencies are the best way to facilitate comprehension of statistical information in health information pamphlets, then central information such as the benefit of screening should not be excluded from this way of representing information. The theoretical prediction that a pamphlet that represents statistical information in terms of natural frequencies should be easier to understand than one that uses only percentages will be tested empirically in the following chapter.

Chapter 5 will also address a related question: What would be the recommended proportion of numerical versus verbal expressions of quantifiable information in health

[14] See also Chapter 3 on the effects of mixing statistical formats

information pamphlets? On the one hand, the arguments made above – numerical information is more precise than verbal statements; patients prefer to get numerical information – suggest that the proportion of numerical information should be as high as possible, and thus the 50% found in the mammography pamphlets would be insufficient. On the other hand, in the case of mammography screening this would mean to put a considerable amount of numerical information into mammography pamphlets to reduce fuzziness in communication. The question is whether the readers would be willing to read through all this information. A pamphlet that included a maximum of precise numerical information but would not be accepted by the intended audience would be inefficient. Thus, to develop a guideline on the amount of numerical information in mammography pamphlets, it would be useful to know more about the information demand of readers of mammography pamphlets.