Chapter 5: Can natural frequencies improve understanding of mammography pamphlets? The role of statistical formats in learning from health information pamphlets

"Educated individuals do not always make wise decisions. Still, without an understanding of the risk, appropriate decisions about personal action can come only from luck or from following someone else's advice or example, neither of which is a dependable guide."
(Weinstein, 1999, p. 15)

Study 5: Evaluation of a model mammography pamphlet

In the previous chapter, I discussed several suggestions on how the design of mammography pamphlets should be improved. In the present chapter, I asked women between 40 and 69 years of age to evaluate an improved mammography pamphlet. This "model" pamphlet was developed as a preliminary version of a text that tries to implement the suggestions made earlier, for instance provides information about a broad range of relevant topics, specifies this information with many pieces of statistical information, and clarifies the respective reference classes. The model pamphlet should be seen as an imperfect prototype that will be improved further on the basis of this study.

Study 5 had two main goals. The first goal was to explore whether *understanding* of the model pamphlet depends on the statistical format used in the pamphlet. The second goal was to assess the *information demand* of the main readership of mammography pamphlets.

Understanding vs. knowing risk information

The central task for readers of health information pamphlets, after having found the information they were looking for, is "to understand it in the way the message-giver intended" (Wright, 1999a, p. 715). Two versions of the model pamphlet were constructed to explore the impact of statistical formats on understanding. One gave statistical information in terms of natural frequencies, and the other gave the same information in terms of percentages (a pamphlet with probabilities was not evaluated since probabilities were not used in the German mammography pamphlet sample). Based on the literature reviewed earlier in this dissertation

(see Chapter 3), natural frequencies have two features that should facilitate the correct understanding of statistical information in the model pamphlet: natural frequencies always specify the reference class, and they are cardinal numbers that are generally easier to process than fractions which constitute percentages.

But how can the understanding of information in a pamphlet be measured? Previous studies mainly used knowledge questions to assess if people understood the information given to them. Typically, knowledge about risks was tested with open-ended questions (Cockburn et al., 1999), forced-choice questions (Dolan et al., 1997; Schwartz et al., 2000), and questions that ask for likelihood estimates for the risks (Black et al., 1995; Cockburn et al., 1995; Hamm & Smith, 1998).

In the present study, I used three different types of knowledge measures. Women were asked (a) to name the most important risks and side effects of mammography screening in an open-ended question, (b) to identify the correct interpretation of one specific piece of information from a set of three possible interpretations (they were given paraphrases of the ambiguous mortality reduction phrase found in the German mammography pamphlets: "Screening reduces breast cancer mortality by 30%"), and (c) to make nine quantitative estimates about benefit, risk and test efficiency of mammography screening. Women received these questions at the beginning of the study and a second time after reading the model pamphlet (which informed about all of these points).

According to Weinstein (1999), measure (c) has to be interpreted cautiously. A person who is able to state the correct likelihood of a risk has not necessarily understood this information: she could simply reproduce a number she had previously heard without realizing the magnitude of the risk. I agree that performance in these three tasks might provide only a modest picture of *real* understanding of risk information. But a satisfactory performance in these three tasks is a necessary first step to the goal of understanding the information Thus, when I speak of "understanding" of risk information in this chapter, I mean *basic* understanding as captured in the three knowledge measure mentioned above.

It is also relevant for the effectiveness of a pamphlet whether the readers themselves have the impression that the pamphlet was comprehensive and efficient (here: in enabling them to make an informed decision about participation in mammography screening). I therefore added a number of evaluation questions. Women were asked (a) whether they felt that they could make an informed decision based on the pamphlet, (b) to evaluate the model pamphlet on eight dimensions (understandability, clarity, amount of information, information

value, relevance for decision, interestingness, objectivity, and relevance for mood), and (c) to give an overall grade and positive and negative feedback on the pamphlet.

Information demand

As mentioned in the previous chapter, if a pamphlet fails to address the information demand of its audience, it is likely to be inefficient. Therefore, the second goal of the present study was to assess the *information demand* of the main readership of mammography pamphlets. Women were asked which topics should, in their view, be included in a good mammography pamphlet, and how the information should be presented.

Information demand is relative to the target group. The model pamphlet evaluated here is intended to inform women about the pros and cons of participating in mammography screening. The official German mammography screening trials target women between 50 and 69 years of age, because screening has the biggest benefit for this age group (Karsa, 1995, 1998; Kerlikowske, 2000; Olsen & Gøtzsche, 2001). These official trials are only conducted in three regions of Germany because they are still in the test phase. In the remaining regions, the unofficial screening practice of the last years is still in use, which is rather permissive and also includes a substantial number of women in their forties (it is estimated that 25% of the German women between 40 and 69 years of age regularly undergo mammography screening; Gibis et al., 1998). Thus, the question of participating in mammography screening is currently most relevant for women between 40 and 69 years of age (see also Banks et al., 1995) so that I invited women from this age group to participate in the study.

I used three types of questions to assess the information demand of the women. First, they were asked to indicate their preferences concerning four transparency issues in pamphlet design (whether disadvantages of screening, differing expert opinions about screening, expert recommendations and precise numerical information should be included or not). Second, in an open question at the beginning of the study, women were asked to list the information they want to receive in a mammography pamphlet to be able to make an informed decision. The question was meant to elicit those topics that women are primarily interested in when they pick up a pamphlet about mammography screening, that is, before they start to think more deeply about the topic. Third, at the end of the study, women could choose from a list of 35 potential topics those that they considered relevant for a subjectively "ideal" mammography pamphlet. This was done to assess the information demand of women who had already thought more deeply about mammography screening (due to the knowledge questions and the model pamphlet). The idea here was to supplement the spontaneous information demand with

a more informed relevance judgment. For 12 of the 35 items, women could also choose between different formulations[15].

Intention to participate in mammography screening

Some proponents of mammography screening seem to discourage a transparent discussion of the risks of mammography screening, because they fear that doing so would keep women from participating in the screening (Napoli, 1997). If this assumption was true, then it should be possible to detect at least a small effect of reading the model pamphlet on women's intention to participate in mammography screening, because the model pamphlet used in this study included more – and more precise – information about the risks of mammography screening than almost any other of the currently available mammography pamphlets. Yet, cognitive models on the determinants of health behavior suggest that such an effect of merely reading a pamphlet is rather unlikely because "changing one's health behavior is considered to be a difficult self-regulation process" (Schwarzer & Gutiérrez-Doña, 2000, p. 460), and not one but several factors determine intention formation (e.g., Schwarzer, 1992). Determinants of the intention to participate in mammography screening were not the main focus of this study. But to get a first impression of potential effects of the model pamphlet on participation intention, I included a simple question about women's current thoughts about future participation in mammography screening at the beginning and at the end of the study.

Method

Design

Participants were randomly assigned to one of three groups[16]. In the *frequency pamphlet* (FP) group, women received a mammography pamphlet that used natural frequencies to represent statistical information. In the *percentage pamphlet* (PP) group, women received a mammography pamphlet that used percentages to represent statistical information. Women in the *no pamphlet* (NP) group did not receive a pamphlet. The NP

[15] In the present study, I decided to let women compare single original statements, rather than whole original pamphlets, because the pamphlets differed in too many respects to be directly comparable (e.g. length, use of graphics, etc.).

[16] In the original design, there was a fourth group of 25 women who received a pamphlet that contained both natural frequencies and percentages. Unfortunately, due to some technical errors in the pamphlet that were only discovered after the study, the data of this group cannot be interpreted and are thus omitted from this presentation.

group was added because women were asked in the last part of the study to construct an ideal pamphlet, and their choices could be influenced by reading one of the model pamphlets.

Participants

Seventy-five women between 40 and 69 years of age participated in the study, 25 in the FP group, 24 in the PP group, and 26 in the NP group. Three women, one in each group, had a prior diagnosis of breast cancer and were excluded from the analysis (because the target group for the mammography pamphlet consists of women who typically participate in mammography screening, that is, women without symptoms or a prior diagnosis of breast cancer). The following analyses are based on data from the remaining 72 women.

The women were recruited by a newspaper advertisement and a flyer that was distributed in gynecologists' practices. The text of the ad and flyer specified the purpose of the study, namely to learn from women between the ages of 40 and 69 years how pamphlets about mammography screening should be designed according to their view. The text stated that the survey was conducted at the Max Planck Institute for Human Development, that the session would take about two hours, and that participation would be reimbursed with 20 Euro.

Table 5.1
Description of the 72 participating women

Sample characteristics		N
Age groups [a]	Women between 40 and 49 years	23
	Women between 50 and 59 years	27
	Women between 60 and 69 years	21
Education [a]	Secondary school	33
	General qualification for university entrance (Abitur)	18
	University degree	20
Location before re-unification [a]	East Germany	5
	West Germany	64
	Foreign country	1
Prior experience with mammography	Women who have previously had a mammogram	65
Description of subjective information status about mammography screening [a]	"I feel sufficiently informed"	15
	"I want to have more information"	51
	"I do not know"	5
Description of prior interest in the topic of mammography screening [a]	"Never thought about the topic before"	9
	"Sometimes think about it (e.g., when it was just in the news)"	28
	"Often think about it"	30

Note. [a] Numbers in this row do not add to the total number of 72 participants due to missing values.

Table 5.1 summarizes the characteristics of the whole sample (there were no significant differences between the three groups). The mean age of the women was 53.9

years. Most women were located in the former West Germany (88%). Twenty-eight percent of the women had a university degree. A large majority of the women had already had at least one mammogram (90%); the mean age at first mammogram was 39.6 years. The mean number of previous mammograms was 4.7 (*SD* = 6.4). Due to one woman with an extremely high number of prior mammograms, the mean number of prior mammograms was with 6.4 higher in the PP group than in the other two groups (FP: 3.6; NP: 4.0; $F(1,44)=1.55$, $p > .05$). Without this outlier, the mean in the PP group was 4.3, and the overall average was 4.0 (*SD* = 3.2).

The women in this sample were rather interested in the topic of mammography screening: 42% stated that they often thought about the topic, and 71% said they wanted to have more information about it (Table 5.1). Consequently, most women had experience with mammography pamphlets as a source of information about mammography screening. They had read on average three mammography pamphlets before (*SD* = 4.2, *Min* = 0, *Max* = 30). As Table 5.2 shows, pamphlets rank third in a list of the most frequently used sources of information about mammography screening, after the gynecologist and TV/radio. This finding is consistent with a previous study with a representative sample of German women (Paepke et al., 2001; similar result for American women in Metsch et al., 1998). Those women who read a mammography pamphlet before had received it in most cases (64%) from their gynecologist, followed by health insurance companies (26%) and cancer societies (11%; Table 5.3).

Table 5.2
Information sources about mammography screening that women used in the past

Information source	N	%
Gynecologist	61	85%
TV/Radio	46	64%
Pamphlets	45	63%
Magazines	39	54%
Relatives, friends, acquaintances	26	36%
Medical literature	17	24%
Radiologist	10	14%
General practitioner	4	6%
Internet	3	4%
Other physician	2	3%
Other information source	7	10%

Note. Numbers are absolute frequencies of women who used the source. Percentages refer to the total number of N = 72 women.

Table 5.3
Providers of mammography pamphlets that women used in the past

Provider	N	%
Gynecologist	46	64%
Health insurance company	19	26%
Cancer society	8	11%
Relatives, friends, acquaintances	3	4%
General practitioner	1	1%
Other physician	1	1%
Radiologist	0	0%
Other	11	15%

Note. Numbers are absolute frequencies of women who used the source. Percentages refer to the total number of N = 72 women.

Procedure

Women were invited to the Max Planck Institute in groups of 2–7, yet they worked individually through the questionnaires. Upon arrival at the Institute, they were welcomed by a female experimenter. Women were seated at separate tables; on each table they found a folder that contained all materials (instructions, questionnaires, pamphlet) in the order of the procedure. The order in which the materials were presented in the three groups is displayed in Table 5.4. In a short introductory speech to all participants, the experimenter first introduced the goal of the study: to learn from the women how, in their view, a pamphlet on mammography screening should be designed to enable them to make a subjectively good decision about participation in screening. Women received a written copy of this so-called "leading question" of the study that remained on their tables and could be reread anytime during the session. The experimenter then explained the terms "screening mammography" (as opposed to clinical mammography) and "informed decision making" and finally the procedure of the study.

Table 5.4
Order of materials presented in the three experimental conditions

Material	FP group	PP group	NP group
Questionnaire 1 *(information demand, knowledge, IP)*	✓	✓	✓
Pamphlet	✓	✓	–
Questionnaire 2 *(evaluation of pamphlet, IP)*	✓	✓	–
Questionnaire 3 *(knowledge)*	✓	✓	–
Questionnaire 4 *(information demand)*	✓	✓	✓
Questionnaire 5 *(personal information, IP)*	✓	✓	✓

Note. IP = Intention to participate in mammography screening.

After that, women started to work on the questionnaires individually. The experimenter was available for individual questions at all times. A time limit of two hours

was mentioned to the women, but they could work through the folder at their own pace. In the FP and PP groups, most women were finished after about 110 minutes; women in the NP group were finished after about 75 minutes. When finished with the last questionnaire, the women went to the experimenter, received their reimbursement, and left.

Materials: Model pamphlets

The content of the frequency pamphlet and the percentage pamphlet were identical; the only difference between the pamphlets was the numerical format. The pamphlets did not include graphics or tables, only text. As a template, the pamphlet of the official German mammography screening trials was used that had also been analyzed in Study 4 (Mammographie-Screening Planungsstelle Köln, 2001). The template was edited and supplement with other information, partly taken from other original pamphlets, partly newly constructed. The section on the test efficiency was entirely new; many new aspects were also added to the sections about the benefits and risks of mammography screening. Table 5.5 lists the topics that were presented in the pamphlet and indicates which of the items were specified with numerical information (altogether, 28 different pieces of numerical information were provided). Before reading the pamphlet, women received the following instructions:

The goal of this study is to develop, with your support, a pamphlet about mammography screening that is comprehensive and useful. On the basis of currently available mammography pamphlets, we developed several drafts of such a pamphlet that are supposed to be evaluated in this study. You will find one of these drafts on the following pages. Please, read the pamphlet draft first and then work through the two questionnaires. If you do not like the draft, please do not hesitate to say so. And if you cannot answer some of the questions about information given in the pamphlet, please do not hesitate to make this clear – we will certainly not draw any negative conclusions about you, but we will get valuable information about what has to be improved in the draft. But please let us also know what you like about the draft.

These instructions were meant to clarify that the pamphlet was not finished, but in a stage where the women's feedback would be most valuable. We also wanted to encourage them to give us negative feedback and to avoid the impression that we only wanted to have their approval for our draft

Table 5.5
Topics covered in the model pamphlet

Heading	Topics
Introduction	– Goal and intended audience of the pamphlet
Breast cancer	– Breast cancer incidence per year*
	– Short-term risk of developing breast cancer for six age groups*
	– Breast cancer mortality per year*
	– Short-term risk of dying from breast cancer for six age groups*
	– Risk factors for breast cancer
What is mammography screening?	– Definition of screening mammography and diagnostic mammography
	– Goal of mammography screening
Who can participate in the screening?	– Debate about the utility of mammography screening
	– Intended audience for mammography screening and recommended interval
What is the benefit of the screening?	– Relative mortality reduction*
	– Explanation of relative mortality reduction with absolute mortality reduction*
What are disadvantages and limitations of the screening?	– Amount of radiation per mammogram*
	– Comparison with natural radiation*
	– Risk of radiation-induced breast cancer*
	– False-positive rate in first screening*
	– Number of women with at least one false-positive in 10 years*
	– Risk of physical and/or psychological strain after false-positive result
	– Number of women who receive a biopsy due to a false-positive result*
	– Unwanted detection of precancerous lesions
	– Early detection of breast cancer does not equal longer life expectancy
	– Breast cancer between screenings
How efficient is mammography screening as a diagnostic method?	– Sensitivity and false-negative rate*
	– Specificity and false-positive rate*
	– Positive predictive value of mammogram*
	– Numerical example that illustrates the relationship between predictive values, error rates and the prevalence of breast cancer *
	– Negative predictive value of mammogram*
Having a mammogram	– Description of mammography procedure
	– Pain associated with having a mammogram
The test result – what happens afterwards?	– Recommended screening interval
	– Symptoms of breast cancer
	– Self- and clinical examination of breast as other early detection methods
	– Diagnostic tests that are typically performed after a positive mammogram
Screening and quality control	– Relevance of quality control
	– Criteria for quality-controlled mammography
Contact addresses	– Contact addresses for more information

Note. Pamphlet topics are presented in the order in which they were given in the pamphlet. * This item was specified with numerical information.

Results

The results will be presented separately for four groups of dependent variables: knowledge (R1), evaluation of model pamphlet (R2), information demand (R3), and participation intention (R4).

R1. Results: Knowledge about benefits and test efficiency of mammography screening

Prior knowledge

The first questionnaire in the study was the same for all women and contained a number of knowledge questions about mammography screening.

First, women were asked if they thought that false positives, or false negatives could occur at all in mammography screening, and if they thought that there were any risks associated with mammography screening. For the latter question, women were also asked to name the risks they had heard about. Fifty-nine of the 72 women (82%) had heard about false-positive results, and 62 had heard about false-negative results (86%) before. Fewer women had heard about mammography risks, only 32 out of 72 (44%). When asked to list the risks they knew, the 32 women mentioned radiation (24), incorrect diagnoses (6), bruises (4), and pain (3) due to the pressure during mammography. None of the women mentioned more than two risks ($M = 1.2$).

Second, women were asked to indicate their understanding of the ambiguous phrasing of the relative risk reduction that was frequently found in the German mammography pamphlets: "Regular attendance of mammography screening can reduce breast cancer mortality by 30%". The question was if women knew what the intended meaning of this sentence was, namely the *relative* mortality reduction. Women had to identify the correct interpretation of the sentence from three alternatives, one correct and two incorrect.[17] Of the 72 women, 28 (39%) chose the correct interpretation, 42 women (58%) chose one of the incorrect interpretations or more than one interpretation, which was also scored as incorrect. Both incorrect interpretations were chosen about equally often (absolute mortality reduction: 20 women; reduction of incidence: 19 women).

Third, women were asked to give numerical estimates for nine pieces of statistical information about mammography screening (see Table 5.6). The numerical format of the questions was the same for all three groups, that is, all women were asked to give their estimates in the form of absolute frequencies.

Table 5.7 shows the results for the whole group of 72 women. Because there were a number of outliers for most questions, both the mean and the median estimates, together with the standard deviation, are presented. The table also displays the information that was later

[17] Correct: "Of every 30 women who would die without regular screening, 10 can be saved by regular screening". Incorrect 1 (absolute mortality reduction): "Of every 1000 women who participate in screening, the screening saves 300 women from dying from breast cancer." Incorrect 2 (reduction of breast cancer incidence): "Women who regularly attend mammography screening develop breast cancer less often."

given in the pamphlet to give an indication of the range of the correct values (values adapted from Mühlhauser & Höldke, 1999).

Table 5.6
Nine questions asking for quantitative estimates of information concerning mammography screening

Original question	Topic
In a mammography-screening program, the participating women can receive a mammogram every two years, that is about 5 mammograms in 10 years. Imagine now 1000 women your age, who have not participated in such a program so far and who do not have any symptoms of breast cancer. How many of these 1000 women…	
– … are expected to develop breast cancer within the next 10 years, if they do not participate in such a mammography-screening program?_____of 1000	Breast cancer incidence in 10 years WITHOUT screening
– … are expected to develop breast cancer within the next 10 years, if they do participate in such a mammography-screening program?_____of 1000	Breast cancer incidence in 10 years WITH screening
– … are expected to die of breast cancer within the next 20 years, if they do not participate in such a mammography-screening program?_____of 1000	Breast cancer deaths in 10 years WITHOUT screening
– … are expected to die of breast cancer within the next 20 years, if they do participate in such a mammography-screening program?_____of 1000	Breast cancer deaths in 10 years WITH screening
– … will receive at least one false-positive mammography result within the next 10 years, if they do participate in such a mammography-screening program?_____of 1000	At least one false-positive in 10 years of screening
Is it possible that women receive a positive (= abnormal) screening mammogram, although they actually do not have breast cancer? ❑ No ❑ Yes; if yes: How often does this (mistake) occur? – Of 1000 women who do not have breast cancer, _____women receive a false-positive screening mammogram result.	False-positive rate
Is it possible that women receive a negative (= normal) screening mammogram, although they actually do have breast cancer? ❑ No ❑ Yes; if yes: How often does this (mistake) occur? – Of 1000 women who do have breast cancer, _____women receive a false-negative screening mammogram result.	False-negative rate
– Imagine a group of women your age who receive a positive (= abnormal) screening mammogram. How many of these women do actually have breast cancer?_____of 1000 women	Positive predictive value
– Imagine a group of women your age who receive a negative (= normal) screening mammogram. How many of these women do actually not have breast cancer?_____ of 1000 women	Negative predictive value

With one exception (for the positive predictive value, see below), there were no significant differences between the three groups. A comparison of the estimates with the correct values yielded the following results:

Breast cancer incidence: Women overestimated breast cancer incidence for women who do not participate in mammography screening by on average 9 percentage points (four- to ninefold). The mean estimate for breast cancer incidence for women who participate in

screening was slightly lower, indicating that some women incorrectly believed that participation in screening would lower breast cancer incidence.

Table 5.7
Mean estimates for nine quantitative aspects of mammography screening *prior* to reading the model pamphlet

Aspect	Pamphlet information	Mean	Median	SD
Breast cancer incidence in 10 years WITHOUT screening	1.2%-2.5% [a]	11%	8%	13%
Breast cancer incidence in 10 years WITH screening	1.2%-2.5% [a]	8%	5%	11%
Breast cancer deaths in 10 years WITHOUT screening	0.4%-0.8% [a]	11%	5%	15%
Breast cancer deaths in 10 years WITH screening	0.4%-0.8% [a]	6%	2%	9%
False-positive rate	9%	9%	10%	18%
False-negative rate	10%	6%	10%	15%
Positive predictive value	10%	42%	40%	35%
Negative predictive value	99.9%	73%	90%	35%
At least one false-positive in 10 years of screening	25%	4%	10%	13%

Note. [a] Values differed for specific age-groups.

Mortality and mortality reduction of screening. Women overestimated breast cancer mortality for women who do not participate in screening by about 10 percentage points (14- to 27-fold). Moreover, mortality rates were comparable to incidence rates. It is not clear if women really believed that almost all women who develop breast cancer will die from it, or if they interpreted the question differently. Women also overestimated the benefit of screening. That is, although the absolute level of estimated mortality for women who participate in screening was too high, the mortality reduction that women assumed for this group, compared to a group of women who do not participate in screening, was larger than the highest reduction reported in the literature (around 25 - 30%). The median computed relative risk reduction for all women was 50% (*SD* = 83%, *Min* = -5%, *Max* = 100%).

Test efficiency. Women had rather adequate prior knowledge about the false-positive and the false-negative rate, but they overestimated the positive predictive value by on average 32 percentage points (fourfold). Women also underestimated the negative predictive value. The median estimate was in an acceptable range, but due to some extreme answers and a large variance, the women on average underestimated this value by 26 percentage points. Finally, women underestimated how many women receive at least one false-positive result after 10 years of mammography screening by on average 21 percentage points.

Knowledge after the pamphlet

After they read through the model pamphlet and worked through Questionnaire 2 (see below), the 47 women in the FP and PP group received the same knowledge questions a

second time. The goal was to detect if the pamphlets were successful in improving the knowledge of women about mammography screening. The numerical format for the questions was now consistent with the pamphlet format, that is, women in the FP group were again asked to give their estimates in absolute frequencies, women in the PP group in percentages.

First, the women were asked again to list the risks and side-effects of mammography screening that they knew (they were not asked again *if* there were errors or risks, since they had received this information in the pamphlets). The model pamphlets mentioned eight different risks. As Table 5.8 shows, only five of them were mentioned in recall (not mentioned were physiological strain due to false-positive results [e.g., biopsy], unwanted detection of precancerous lesions, pain during mammography). The two most frequently mentioned risks and side-effects were incorrect diagnosis (mentioned by 24 of the 47 women) and radiation (22 of 47). Radiation risks had been the most frequently mentioned risk before the pamphlet (PP: 8, FP: 10 women)[18]. But the risk of incorrect diagnosis was mentioned by significantly more women after the pamphlet (PP: 12, FP: 12) than before (PP: 1, FP: 3), $\chi^2(1,47) = 20.35$, $p < .001$, $\phi = .47$.

Table 5.8
Women's statements about risks of mammography screening *after* reading the model pamphlets

Statement	PP	FP	Overall [a]
Risks according to model pamphlet			
Incorrect diagnosis (false positive and/or negative results)	12	12	24 (51%)
Radiation	11	11	22 (47%)
Psychological distress due to false-positive results	3	2	5 (11%)
Early detection of breast cancer does not equal longer life-expectancy	2	0	2 (4%)
Other statements			
Breast cancer between screenings	3	2	5 (11%)
Mammography has no risks	0	2	2 (4%)
Psychological distress when waiting for results	0	1	1 (2%)
Mammography cannot prevent breast cancer	1	0	1 (2%)
Mammography cannot say if lump is benign or malign	0	1	1 (2%)
Costs for health insurance	0	1	1 (2%)
Unclear statement	1	2	3 (6%)
Number of statements	33	34	67
Number of women who made a statement	19	18	37

Note. [a] The percentages in this column refer to the total number of women in the two groups, N = 47.

Compared to the answers before reading the pamphlet, the maximum number of statements per woman increased from 2 to 4, and the number of different statements made by

[18] Please note that all references in this section to results concerning knowledge *before* reading the pamphlet refer to $N = 47$ (PP and FP group). The results might thus differ from the results reported in the previous section, which referred to $N = 72$ (all three groups).

the whole group increased from 4 to 10. However, not all of the statements referred to risks of mammography screening (see "Other statements" in Table 5.6). For instance, some statements dealt with effects that mammography screening *does not* have such as the prevention of breast cancer. To conclude, although reading the model pamphlets made the answers more diverse, there was only one notable difference compared to the answers before reading the pamphlet: Significantly more women said that the possibility of getting an incorrect mammogram is a risk of mammography screening. This effect was independent of pamphlet format.

Second, women were again asked to identify the correct interpretation of the ambiguous phrasing of the relative risk reduction from one correct and two incorrect alternatives. The meaning of the sentence had been explicitly explained in the pamphlet. After the pamphlet, 20 of the 47 women (PP: 9, FP: 11) were able to identify the correct paraphrase (the two incorrect paraphrases were chosen equally often in both groups), compared to the 16 before the pamphlet. Hence, a McNemar-test for dependent measurements shows that reading any of the two model pamphlets had overall only a small effect on the ability to identify the correct paraphrase, $\chi^2(1, 47) = .55$, $p = .50$, $\phi = .11$.

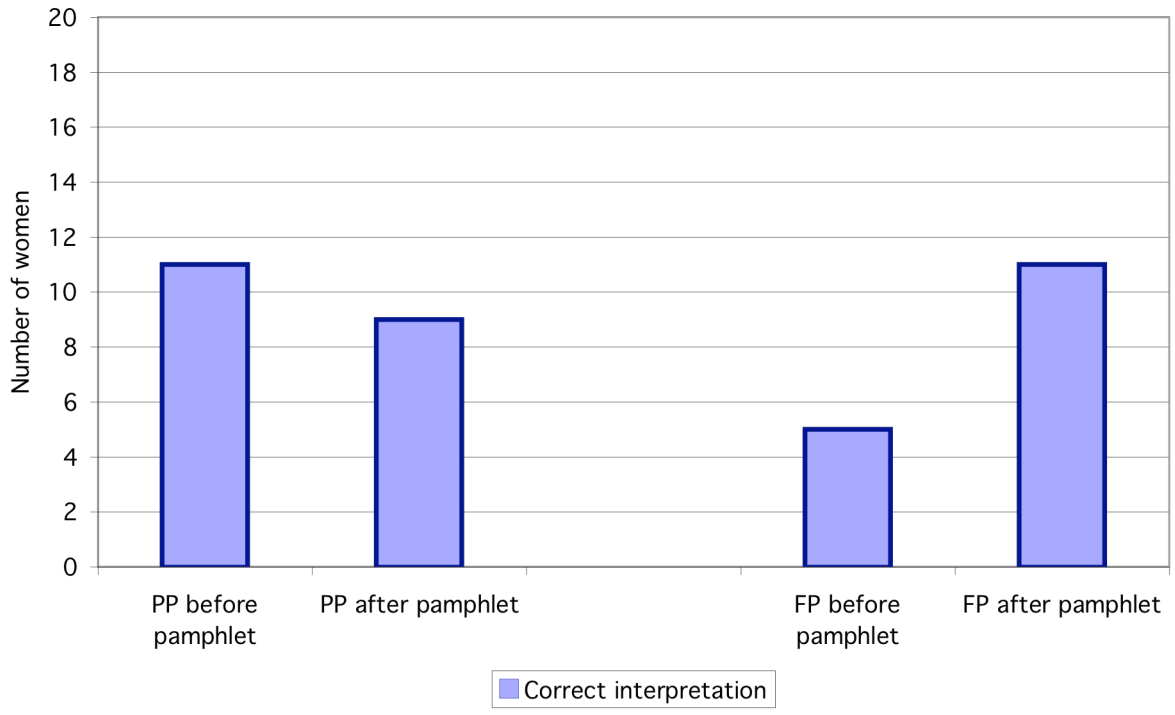

Figure 5.1 Correct interpretation of an ambiguous risk reduction statement in the FP and PP group before and after reading the pamphlet

However, there were differential effects within the two groups (Figure 5.1). There had been a prior difference between the two groups: Only 5 women in the FP group chose the correct interpretation, whereas 11 in the PP group (χ^2 (1, 47) = 3.81, p < .06, ϕ = .28) did so. After reading the pamphlet, the number of women who identified the correct answer in the PP group was even slightly reduced from 11 to 9. In the FP group on the other hand, the number of correct interpretations doubled from 5 to 11. Still, the learning effect resulting from the frequency pamphlet was only small (χ^2(1, 24) = .51, *McNemar p* = .11, ϕ = .15).

Third, women were again asked to give numerical estimates for the same nine pieces of statistical information about mammography screening on which they were questioned before reading the pamphlet. Seven of the nine estimates were directly given in the pamphlet, the other two could be inferred from the text: (a) the breast cancer mortality for women who do participate in screening can be inferred from the general mortality rates and the explanation of the absolute and relative risk reduction, and (b) the incidence of breast cancer for women who participate in screening can be inferred from the explanation that mammography screening affects only mortality, not incidence, and the general incidence rates were given.

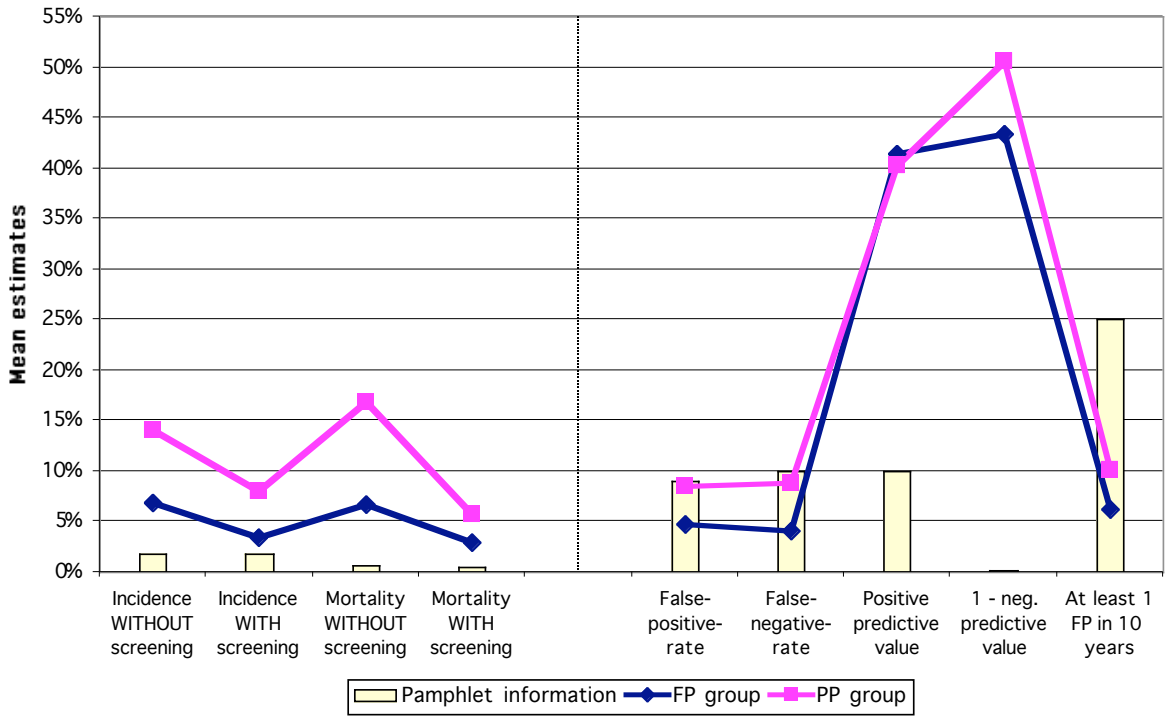

Figure 5.2 Mean estimates for nine quantitative aspects of mammography screening in the FP and PP group *after* reading the model pamphlet

Figure 5.2 shows the mean estimates for the 47 women in the FP and PP groups after the pamphlet. Figures 5.3 and 5.4 illustrate the mean estimates in the two groups before and after the pamphlet, respectively. In all three figures, the bars represent the information given in the pamphlet (which is taken here as "correct"), and the lines represent the mean estimates given by the two pamphlet groups. Prior to reading the pamphlet, the had been only one difference between the two groups: Women in the FP group overestimated the positive predictive value of mammography screening much stronger than the women in the PP group (mean estimates FP: 49.8%, PP: 29.5%; $F(1,43) = 4.27$, $p < .05$). How did the estimates of the two groups differ after reading the pamphlet?

As Figure 5.2 shows, four out of the nine estimates (mortality and incidence) were more accurate in the FP than in the PP group, three estimates were more or less comparable, and two were more accurate in the PP than in the FP group (error rates). Let us have a more detailed look at the results (tables with medians, means, standard deviations are provided in the Appendix).

Breast cancer incidence, mortality and mortality reduction through screening. Figure 5.2 (left-hand side) shows that the mean estimates of breast cancer incidence and mortality were higher than the respective values given in the pamphlet, thus women in both groups still overestimated these variables. However, there was a notable difference between the FP and the PP group, because the mean estimates in the FP group were closer to the correct values than those in the PP group (mortality without screening: $F(1,35) = 7.00$, $p < .05$; mortality with screening: $F(1,35) = 3.51$, $p < .07$; incidence with screening *and* incidence without screening: $F(1,34) = 3.15$, $p < .09$, respectively). As Figures 5.3 and 5.4 show, not only did the mean estimates in the FP group get closer to the correct values compared to the prior estimates, but they also slightly departed from the correct values in the PP group (see Appendix for values of t and p for paired sample t-tests; none of the differences was significant on the $p = .05$ level). In both groups, the mean estimates for the two incidence rates differed, indicating that a number of women still believed that mammography screening could lower breast cancer incidence (even though the pamphlet had explained that this was not the case). The median computed relative risk reduction remained 50% for the FP group (SD = 30%, Min = 0%, Max = 90%), but increased from 50% to 68% for the PP group (SD = 205%, Min = -7%, Max = 90%).

Test efficiency. The mean estimates for the two error rates in the PP group matched the actual values precisely, whereas they were too low in the FP group (false-negative rate: $F(1,40) = 6.24$, $p < .05$; false-positive rate: $F(1,40) = 3.01$, $p < .10$). Figure 5.3 shows that this

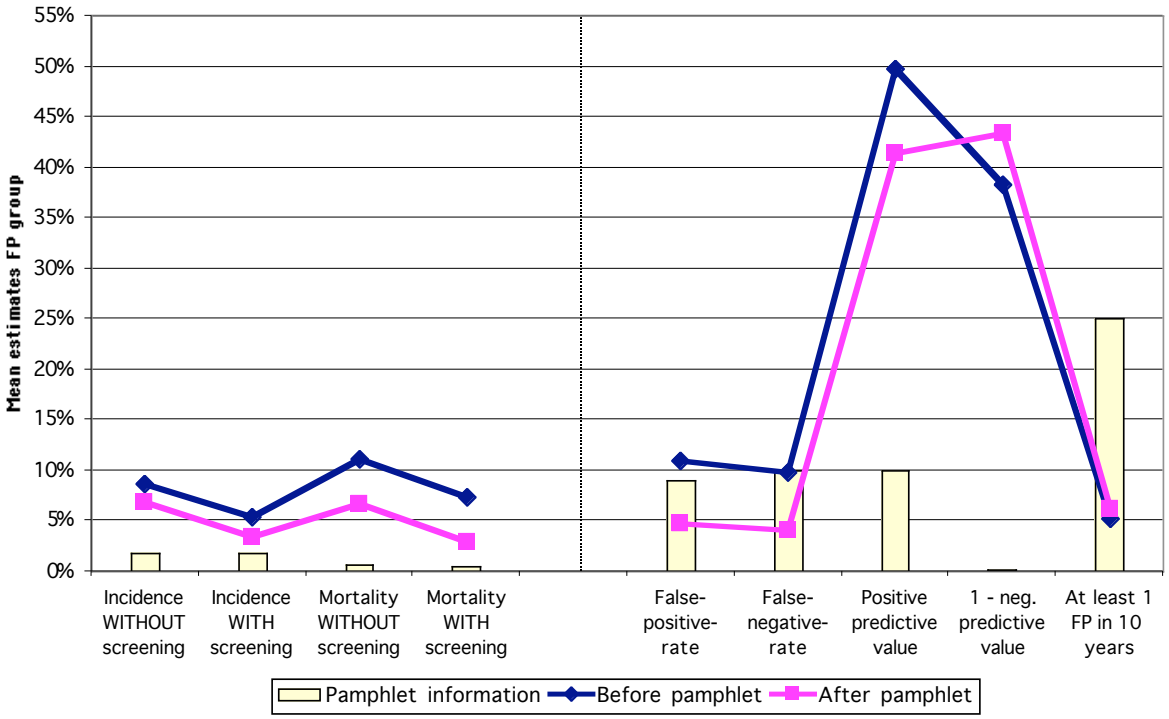

Figure 5.3 Learning in the FP group: Comparison of the mean estimates before and after reading the frequency pamphlet

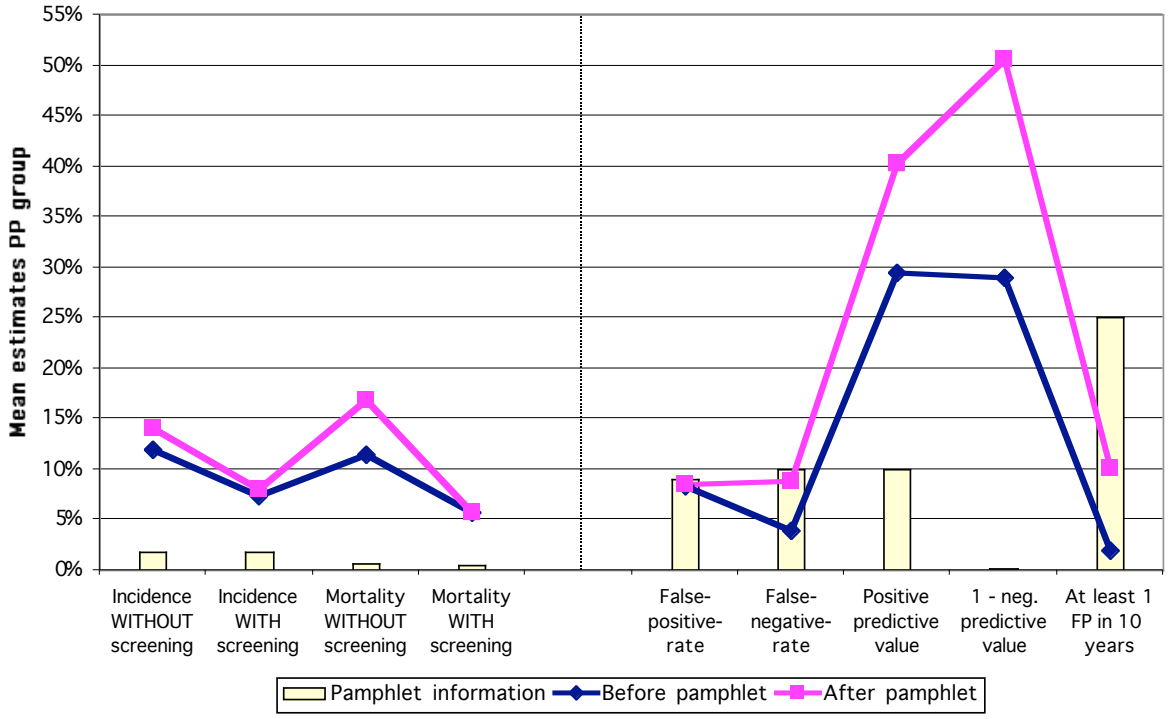

Figure 5.4 Learning in the PP group: Comparison of the mean estimates before and after reading the percentage pamphlet

was mainly due to lower, that is, less accurate estimates in the FP group after reading the pamphlet. For the remaining three estimates, there was no effect of pamphlet format on

accuracy. The positive predictive value was still overestimated in both groups by about 30 percentage points; the prior difference for this estimate disappeared as estimates in the two groups converged. Also, the negative predictive values and the number of women who received a false positive result in 10 years were, as before, strongly underestimated.

The above analyses have been performed on the group level. When interpreting the group means and medians of the estimates, it should be noted that the standard deviations were often very large (see Appendix). Thus, means and medians might not provide an adequate picture of how close the estimates of the women were to the values given in the pamphlet. How many women could recall the information given in the pamphlet in a way that their estimates fell within an acceptable range of the correct answer? The size of an "acceptable" range is of course not predetermined and depends, among other things, on the goal of the communicator of how precise the knowledge of women should be after reading the pamphlet. One might also think of different criteria ranges for different pieces of information within a pamphlet. For the present analysis, I scored how many of the women in the two pamphlet groups gave answers within a range of ± 50% of the correct value. A 50% deviation from the correct value might seem rather large at first. But in my view, estimates within this range still represent the actual values in a satisfying way, at least those that are not larger than 10% (see Table 5.9). This is the case for seven of the nine items. For the remaining two, the negative predictive value (99.9%) and the number of women who receive at least one false-positive result in 10 years of screening (25%), I also present results with a stricter 10% range.

Table 5.9
Boundaries of the 50% range for nine numerical values given in the model pamphlet

Aspect	Pamphlet	Lower bound	Upper bound
Breast cancer incidence in 10 years WITHOUT screening	1.9%	0.9%	2.8%
Breast cancer incidence in 10 years WITH screening	1.9%	0.9%	2.8%
Breast cancer deaths in 10 years WITHOUT screening	0.6%	0.3%	0.9%
Breast cancer deaths in 10 years WITH screening	0.6%	0.3%	0.9%
False-positive-rate	9.0%	4.5%	13.5%
False-negative-rate	10.0%	5.0%	15.0%
Positive predictive value	10.0%	5.0%	15.0%
Negative predictive value	99.9%	50.0%	149.9%
At least one false-positive in 10 years of screening	25.0%	12.5%	37.5%

Figure 5.5 shows the number of answers in the two groups that fell into the 50% range for each of the nine items. The dotted lines refer to the answers given prior to reading the pamphlet, the bold lines to the answers given afterward. Let us first compare the two bold lines: Did the proportion of answers in the acceptable range (subsequently called "acceptable answers") differ depending on the format of the pamphlet? Overall, there was only a very

small effect of the format of the pamphlet. Across all nine items, 30% of the answers in the FP group (65 out of 9 x 24) and 23% of the answers in the PP group (47 out of 9 x 23) fell within the acceptable range, $\chi^2(1, 423) = 2.96$, $p > .05$, $\phi = .08$. On the left-hand side, the pattern is relatively similar to that of Figure 5.2, that is, there are more acceptable answers for the FP group, especially for the mortality rates. On the right-hand side, there is an interesting difference to Figure 5.2: Although the mean estimates for the error rates were worse in the FP group than in the PP group, the number of acceptable answers was almost the same in both groups (false-negative rate: PP 8, FP 8; false-positive rate: PP 10, FP 8 acceptable answers).

Next, compare each bold line with the dotted line of the same color. The comparison of the number of acceptable answers before and after the pamphlet shows again only a small difference between the two groups. Across all nine items, the number of acceptable answers did hardly increase in the PP group (46 acceptable answers before vs. 47 after the pamphlet), because for some items, the number of acceptable answers after reading the pamphlet was in fact lower than before. In the FP group, a small increase in the number of acceptable answers took place (46 acceptable answers before vs. 65 after the pamphlet).

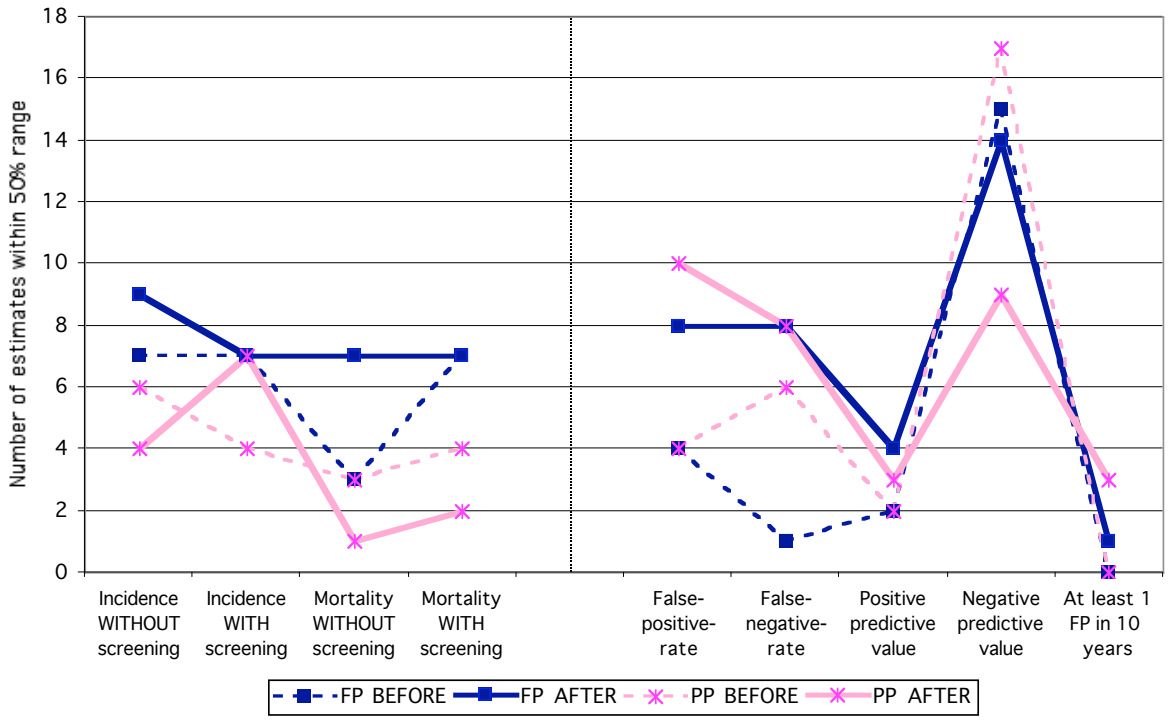

Figure 5.5 Number of estimates within 50% range of correct value for the FP and PP group before and after the pamphlet

As mentioned above, I also applied a stricter 10% range to the two items with the biggest numerical values. The results were not different from those with the 50% criterion

because the number of acceptable answers decreased only minimally (negative predictive value: 21 rather than 23 acceptable answers; number of women who receive at least one false-positive result in ten years of screening: 2 rather than 4 acceptable answers).

Summary of results for knowledge

Before reading the model pamphlets, those women who knew that mammography screening also involves risks could typically name only one, namely radiation. Reading the model pamphlets which contained information about several risks led to only one substantial change: significantly more women said that the possibility of getting an incorrect mammogram is a risk of mammography screening. This effect was independent of pamphlet format.

Before reading the model pamphlets, about 80% of the women had heard about false-positive and false-negative results. Their estimates for the likelihood of the two error rates were rather adequate, but they underestimated how many women receive at least one false-positive result after 10 years of mammography screening, and they misjudged the predictive value of a mammogram. Positive mammograms were believed to be more predictive for breast cancer than they actually are, and negative mammograms were believed to be less predictive for the absence of breast cancer than they actually are. Reading the model pamphlet, no matter what format, did not improve these estimates. The mean estimates for the two error rates were even less accurate after reading the frequency pamphlet, although the number of acceptable estimates (i.e., within a range of ± 50% of the correct value) for the two error rates was comparable in both pamphlet groups.

Initially, most women overestimated breast cancer mortality and the amount by which mammography screening could reduce it. Reading the model pamphlet with natural frequencies made the estimates on average more accurate than before and also more accurate than in those in the PP group.

The frequency pamphlet also lead to a small increase of the number of women who could identify the correct meaning of an ambiguous sentence about the risk reduction of mammography screening. The percentage pamphlet had no effect on the number of correct interpretations.

R2. Evaluation of model pamphlets

First, the women were asked if they felt that they could make an informed decision about screening participation based on the model pamphlet (in other words, if the pamphlet

reached its goal to enable informed decision-making). Eighteen of 23 women in the PP group (78%), and 21 of 24 women in the FP group (88%) said "yes". Thus, women from both groups seemed to be satisfied with the model pamphlets. This was also reflected in the grades given to the pamphlets: On a scale of 1 to 5 in which 1 is best, the percentage pamphlet received a mean grade of 2.3 (SD = 0.9), and the frequency pamphlet received a mean grade of 2.2 (SD = 0.7).

In addition to these global measures, women were asked to evaluate the model pamphlet on eight dimensions with 7-point Likert scales (Table 5.10). The two pamphlets were evaluated as equally understandable. For most of the dimensions, the two pamphlets were evaluated comparably. Yet, there were some differences: The frequency pamphlet was found to be significantly more interesting and more objective than the percentage pamphlet ($p < .05$ for $F(1,41) = 6.83$ and $F(1,40) = 4.76$, respectively), but women also thought that the frequency pamphlet contained rather too much than too little information ($p < .05$ for $F(1,38) = 7.16$).

Another question asked whether women felt emotionally influenced by the pamphlet. Every third woman reported some mood change (PP: 6, FP), most of these women were worried, but some felt also calmed. Women who were worried (FP: 5, PP: 5) mostly attributed this to the error rates (e.g., "I am worried because I could be the one with a false-positive or false-negative diagnosis. I feel very insecure about the quality of the method."); women who felt calmed (FP: 5, PP: 1) said this was because they felt better informed now.

Table 5.10
Women' s mean evaluations of the model pamphlets

	Mean		*SD*	
Dimensions	PP	FP	PP	FP
Was the pamphlet understandable (1) or not (7)?	2.3	2.3	1.1	1.4
Was the pamphlet clear (1) or not (7)?	2.3	2.7	1.3	2.0
Did the pamphlet give too much (1) or too little information (7)?	3.6	2.6	0.9	1.4
Was the pamphlet relevant for the decision (1) or not (7)?	2.8	2.0	1.4	1.3
Was the pamphlet informative (1) or not (7)?	2.1	1.5	1.4	1.2
Was the pamphlet calming (1) or alarming (7)?	4.0	4.5	1.3	1.8
Was the pamphlet interesting (1) or boring (7)?	3.0	1.9	1.7	1.1
Was the pamphlet objective (1) or biased (7)?	3.1	2.1	1.5	1.2

Note. Scales went from 1 to 7; location of the attributes is indicated in the first column.

Finally, to get more specific suggestions on how the model pamphlets could be improved, women were asked to give their positive and negative feedback concerning the pamphlets in some open-ended questions. The open feedback did not differ for the two pamphlets, and also the number of women who made positive and negative statements were

comparable (positive: PP 21, FP 23; negative: PP 16, FP 14 statements). The most frequent positive comments were that the pamphlets mentioned specific topics with which women had previously been unfamiliar (23 of 44 women who made a positive comment), that the pamphlets were comprehensive, objective, and answered all questions (12 of 44, respectively). The most frequent negative comments were that the pamphlets were confusing, contained too many numbers (10 of 30 women who made a negative comment, respectively), and too much information (8 of 30). Some women also remarked that the layout should be improved and that graphics should be added (7 and 4 of 30, respectively).

R3. Information demand

The second goal of the study was to assess what information (topics and formulations) women want to receive to make a subjectively well-informed decision about participating in mammography screening.

Preferences concerning four transparency issues in pamphlet design

The preferences elicited here concerned four topics related to transparency and objectiveness of the pamphlet. For each topic, women could decide between two opposing options. All options and the frequency with which they were chosen are listed in Table 5.11. Since there were no substantial differences between the treatment groups, the results are again presented for the whole group.

A clear majority of the 72 women (92%) said that mammography pamphlets should inform about all disadvantages of screening, even if this could lower the number of participating women. Most women (81%) also said that they want to learn about the different points of view that different experts might have about certain aspects of mammography screening. There was no clear preference for the topic of recommendations. About half of the women (46%) agreed with the position that no participation recommendation should be given to women in mammography pamphlets, whereas the other half (47%) would welcome such a recommendation.

Finally, 40 of the 72 (56%) women said they would prefer to get precise numerical information in mammography pamphlets whenever possible. Among the reasons for this preference were that numbers would help them to make the decision because they could better compare pros and cons of the screening (22), be more precise (11), and also more objective (6). Ten women did not answer the question. The main reasons for the remaining 22 women (31%) who wanted to receive as few numbers as possible were that they found numbers to be

confusing (20), not relevant for the single-case (11), and often manipulative (6). Since the preference for numerical versus verbal statements is of high relevance for the present study, I also checked whether this preference was related to other variables such as age. I splitted the group along the median age (54 years) and checked whether older women would be less in favor of numerical information, however, there was no relationship at all between the two variables, $\chi^2(1, 61) = .00$, $p > .05$, $\phi = .00$. It was also not the case that especially those women who were more interested in the topic of mammography screening (see above) preferred to receive numerical information, $\chi^2(1, 61) = .86$, $p > .05$, $\phi = .12$. However, more women with university degrees preferred numerical information (13 out of 15), compared to women with qualification for university entrance (10 out of 18) and women with secondary school degree (17 of 29), $\chi^2(2, 62) = 4.29$, $p < .12$, $\phi = .26$).

Table 5.11
Preferences concerning four transparency issues in pamphlet design

Topic	N	%
Disadvantages: Pamphlets should clearly mention all disadvantages (risks, costs) of the screening mammography,		
1 even if this could lower the number of participating women.	66	92%
2 only if this does not lower the number of participating women.	4	6%
Differing opinions: If experts disagree in their evaluation of some aspect of mammography screening (e.g. risks, utility),		
1 pamphlets should depict all differing points of view.	58	81%
2 pamphlets should depict only one point of view, because the presentation of all points of view would make the reader uncertain.	10	14%
Statistics: Pamphlets should		
1 support their statements as often as possible with statistical information.	40	56%
2 contain as few numbers as possible.	22	31%
Recommendations: Pamphlets should		
1 not contain a recommendation (pros or cons) for the reader concerning the participation in the screening. It is better when the reader can make her decision as uninfluenced as possible.	33	46%
2 contain a recommendation (pros or cons) for the reader concerning the participation in the screening. It is better when the reader can use such a recommendation as an orientation for her own decision.	34	47%

Note. Numbers are absolute frequencies. Percentages refer to the total number of N = 72 women. The percentages within each of the four topic do not add up to 100% due to missing values.

Open question for topics of interest at the beginning of the study

The very first question of the study asked women for the information they would want to receive in a mammography pamphlet to be able to make an informed decision about participation in mammography screening (open response format)[19].

[19] Wording of the question: "If you had to decide now if you want to receive a screening mammogram: What information do you need to judge whether a participation makes sense for you or not? Think about open questions you have concerning the screening mammography as well as facts you already know and would

The question was meant to elicit those topics that women are primarily interested in when they pick up a pamphlet about mammography screening, that is, before they start to think more deeply about the topic.

Overall, the 72 women mentioned 31 different aspects of mammography screening that they would like to be informed about, on average 4 per woman (*SD* = 2.5). Because of the open-ended question, the level of specificity of the answers varied; for example, while some women mentioned the sensitivity as relevant information, others more generally wrote that they wanted to be informed about errors the test could make.

Table 5.12
The Top 5 topics that women considered relevant for screening participation at the beginning of the study

Rank	Topic	N	%
1.	General: risks and side effects of screening	40	56%
2.	General: What is the benefit of screening	23	32%
3.	What should be the interval between mammograms	22	31%
4.	Radiation	21	29%
5.	Who should and who should not participate in screening (contra-indications, age)	18	25%

Note. Percentages refer to the total number of participants (N = 72).

Table 5.12 with the five most frequently mentioned topics gives a first impression of what was at the top of women's heads when thinking about screening participation: questions about risks, benefits and procedural aspects of mammography screening. Most frequent was the general question about "risks and side effects of mammography screening", it was mentioned by 56% of the women. Table 5.13 includes the complete list of all topics that the women mentioned. I categorized them according to the content domains that were used in the analysis of the German mammography pamphlets, and added new domains if necessary. It can be seen that women found topics from a broad range of domains to be relevant. They were most interested in descriptions of the screening procedure itself (30%, or 87 of all 291 statements), e.g. who pays for the screening, where and when mammograms should be done, and so forth. The second most important content domain was risks and side effects of mammography screening (23% of all statements), other domains were the test efficiency (14%), the benefit of screening (11%), information about breast cancer (8%), and about the frequency of developing breast cancer (2%).

consider for your decision. Please write down the information you need for your decision and which therefore should be included in every good information pamphlet about mammography screening."

Table 5.13
Open question: What information do women consider relevant for making an informed decision about participation in mammography screening?

Topics [a]		N	%
Risks and side effects of screening (67)	General: risks and side effects of screening	40	56%
	Radiation	21	29%
	Pain	6	8%
Procedure of screening (58)	What should be the interval between mammograms	22	31%
	Description of mammography procedure	13	18%
	Who pays for screening	10	14%
	Where are mammograms done	10	14%
	How does the mammography machine work	2	3%
	In which phase of the menstrual cycle should the mammogram be done	1	1%
Test efficiency (40)	Sensitivity	13	18%
	Minimal size of tumor that can be detected	10	14%
	General: How often do errors occur	8	11%
	Positive predictive value	4	6%
	How to identify doctors that are qualified to interpret mammograms	3	4%
	False-positive rate	2	3%
Benefit of screening (32)	General: What is the benefit of screening	23	32%
	To what extent does screening reduce breast cancer incidence	6	8%
	To what extent does screening reduce breast cancer mortality	3	4%
What is screening and who can participate (29)	Who should and who should not participate in screening (contra-indications, age)	18	25%
	Definition of screening mammography	11	15%
Information about breast cancer (24)	Risk factors for developing breast cancer	10	14%
	What are the implications of having a malign/benign tumor	5	7%
	Breast cancer therapy	4	6%
	Symptoms of breast cancer	3	4%
	Survival rate of breast cancer	2	3%
Frequency of developing breast cancer (7)	How frequent is breast cancer in particular age groups	4	6%
	General: How frequent is breast cancer	3	4%
Further questions	Alternatives to screening	6	8%
	Anatomy of breast	2	3%
	Contact addresses for more information	1	1%
Unrelated statements		19	26%
Sum of statements		291	

Note. Six women did not answer the question. Percentages refer to the total number of participants (N = 72). [a] Numbers in parentheses are the sum of the statements per domain.

Choice of topics for a subjectively ideal pamphlet at the end of the study

At the end of the study, women could choose from a list of 35 potential topics those that they considered relevant for a subjectively ideal mammography pamphlet. The choice of topics at about mammography screening could at this point be made from a more informed point of view (however, comparisons between the relevance of topics at the beginning and at the end of the study can only be made for those topics which had been mentioned at the

beginning *and* were included in the list of 35 potential topics). Women were asked to choose as many topics as they considered relevant, but they were also told that their ideal pamphlet should only be as extensive as necessary for reaching its goal. Women were reminded at the end of the questionnaire that they should only choose as many topics as necessary and that they should, if they now felt that they did not follow this rule, go back and delete those items that were not as relevant as they first thought.

Overall, a participating woman chose on average 27 topics out of the 35 (*SD* = 5.1, *Min* = 14, *Max* = 35). Put differently, the average frequency with which a topic from the list was chosen was 55.1 (*SD* = 12.8, *Min* = 25, *Max* = 71); this corresponds to being chosen on average by 76.6% of the participants. Thus, women chose most of the items from the list and excluded topics rather selectively.

Table 5.14
The Top 10 topics that women considered relevant for screening participation at the end of the study

Rank	Topic	N	%
1.	Mortality reduction through screening	71	99%
1.	False-negative rate	71	99%
1.	Amount of radiation per mammogram	71	99%
4.	Breast cancer between screening tests	69	96%
5.	Relationship of breast cancer risk and age	68	94%
6.	Cumulative lifetime risk of breast cancer	67	93%
6.	Symptoms of breast cancer	67	93%
6.	Survival rate of breast cancer	67	93%
6.	Clinical breast exam as other early detection method	67	93%
10.	Description of mammography procedure	66	92%
10.	Breast self-examination as other early detection method	66	92%

Note. Percentages refer to the total number of participants (N = 72).

Table 5.14 summarizes the ten most frequently chosen topics (which were actually eleven, due to an equal number of votes for two topics on rank 10). As before, information about the benefit of screening (mortality reduction) and the radiation involved in getting mammograms was seen as extremely relevant. A topic that had not been mentioned explicitly before, but was now considered relevant by 99% of the women was the false-negative rate. Table 5.15 lists all 35 topics and the frequency with which they were chosen by the women. Choices are presented for the whole group since there were, with three exceptions, no significant differences between the groups (exceptions will be discussed below). I would like to highlight the results for three content domains.

- *Benefit.* The mortality reduction through screening was one of the three most important topics (chosen by 71 of 72 women). The different ways to represent the mortality reduction were presented as formulation choices. The majority of the women who chose

the topic voted for the inclusion of the relative risk reduction in the pamphlet (41 of 71 women). Of the remaining women, still more preferred to have a vague verbal statement included in the pamphlet ("For the group of women who participate in mammography screening, the probability of dying from breast cancer is reduced"; 13 of 71 women), rather than the absolute mortality reduction or both absolute and relative mortality reduction (6 of 71 women, respectively; the latter option had been used in the model pamphlet).

- *Test efficiency.* This was, after risks and procedural aspects of mammography screening, the third most important domain at the beginning of the study, and women still thought it was relevant after reading more about it. Seventy-one of 72 (99%) women wanted to have information about the false-negative rate included in an ideal mammography pamphlet. The false-positive rate was chosen significantly less often (57 of 72 women, 79%; $\chi^2(1, 72) = 13.78$, $p < .01$, $\phi = .31$). The topic that is clearly least relevant to women is the relationship between the predictive values of the mammograms with the base rate of breast cancer and the error rates of the test. Only 41 of 72 women (57%) found this information to be relevant, and 30 of 41 preferred the shorter verbal explanation to the longer explanation with the numerical example.
- *Risks and side effects.* Women were most interested in information about the amount of radiation (chosen by 71 of 72 women) and breast cancer that could develop between screenings (69 of 72 women). The latter shows that women were very concerned with the possibility that breast cancer could be overlooked, as was also indicated in the high relevance attributed to the false-negative rate. More women were interested in the *amount* of radiation, rather than in the health risks associated with radiation. To represent the latter information, women chose about equally often a vague verbal statement and a detailed numerical specification of the risk of radiation-induced breast cancer (both statements were taken from original pamphlets).

The differences in the choices between the three treatment groups were checked with Chi-square-tests. Significant differences were found in only three cases. Women who had not read a model pamphlet were less likely to choose the following topics: "early detection of breast cancer does not equal longer life-expectancy" (NP = 5, vs. FP = 14 and PP = 15, $\chi^2(2, 72) = 11.61$, $p < .001$, $\phi = .40$), "unwanted detection of precancerous lesions" (NP = 7, vs. FP = 19 and PP = 20, $\chi^2(2, 72) = 21.69$, $p < .001$, $\phi = .55$), and "debate about utility of mammography screening" (NP = 11, vs. FP = 16 and PP = 18, $\chi^2(2, 70) = 7.12$, $p < .05$, $\phi = .32$). These three topics described risks and limitations of mammography screening, they had

Table 5.15
Choice of topics for a subjectively ideal mammography pamphlet

	Topic	N	%
Benefit of screening	Mortality reduction through screening	71 [a]	99%
	- Relative mortality reduction	41	
	- Absolute mortality reduction	6	
	- Absolute and relative mortality reduction	6	
	- General verbal statement	13	
Procedure of screening	Description of mammography procedure	66	92%
	Screening interval	62	86%
Frequency of developing breast cancer	Relationship of breast cancer risk and age	68	94%
	Cumulative lifetime risk of breast cancer	67	93%
	Breast cancer incidence (per year)	59	82%
	Frequency of benign lesions in breast	55	76%
	Increasing incidence in younger women	53	74%
	Frequency of genetically determined breast cancer	47	65%
Test efficiency	False-negative rate	71	99%
	Positive predictive value	63	88%
	False-positive rate	57	79%
	Negative predictive value	53	74%
	Relationship PPV with base rate and error rates	41 [a]	57%
	- With numerical example	11	
	- Without numerical example	30	
Risks and side effects	Amount of radiation per mammogram	71	99%
	Breast cancer between screening tests	69	96%
	Comparison of radiation risk with benefit of screening	57	79%
	Physical and/or psychological strain due to false-positives	54	75%
	Risk of radiation-induced breast cancer	51	71%
	- With numerical information	23	
	- Without numerical information	27	
	Unwanted detection of precancerous lesions	46*	64%
	Debate about utility of mammography screening	45*	63%
	Early detection of breast cancer does not equal longer life-expectancy	34*	47%
Information about breast cancer	Symptoms of breast cancer	67	93%
	Survival rate of breast cancer	67	93%
	Risk factors for breast cancer	60	83%
	Frequency of breast-saving operations	42	58%
	Breast cancer therapy	26	36%
Breast cancer mortality	Breast cancer mortality (per year)	59	82%
	Relationship of breast cancer mortality and age	59	82%
	Cumulative lifetime risk of breast cancer death	29	40%
Further topics	Clinical breast exam as other early detection method	67	93%
	Breast self-examination as other early detection method	66	92%
	Contact addresses for more information	60	83%
	Diagnostic tests after positive mammogram	57	79%
	Relevance of quality control for screening	57	79%

Note. * These topics were chosen significantly less often by women in the NP group. [a] The formulation choices do not add up to the total number of choices because some women did not decide between formulations.

been included in the model pamphlet, but not in the first knowledge questionnaire. Thus, it seems that some risk topics are evaluated as less relevant by women who are not familiar with these topics than by women who are more familiar with them.

Choice of formulations

For 14 of the potential pamphlet topics, women could decide among two or more formulations. Three have been presented in the previous section, the presentation of the mortality reduction, risk of radiation-induced breast cancer, and the method of explaining the predictive values of mammograms. The formulation choices for the 11 remaining topics are presented in this section. The number of formulation alternatives offered ranged from 2 to 5. The women were told to choose those formulations that were most comprehensive. The differences between the provided formulation alternatives refer to two factors relevant for pamphlet design that were also addressed above with the attitude questions: Use of numerical versus verbal statements, and the amount of transparency.

Let us first look at the use of numerical versus verbal statements. Women could choose for nine topics between numerical (frequency or percentage) and purely verbal statements. More specifically, as Table 5.16 shows, they could either choose between frequency and percentage statements (four topics), frequency and verbal statements (three topics), and frequency, percentage and verbal statements (three topics). In the cases where both a frequency and a percentage statement were offered, these statements had been newly constructed to ensure parallel wordings. The remaining cases, including all verbal statements, were adapted from the original German mammography pamphlets (Chapter 4).

Table 5.16 presents the results for all 72 women (there were no significant differences between the three conditions). The upper part of the table shows that there was no notable preference for one of the two numerical formats: Of all 204 choices made, in 108 (53%) a frequentistic phrasing was chosen, and in 96 (47%) a percentage statement. The lack of a consistent preference can be demonstrated, for example, if one compares the choices for the positive and negative predictive values: although the alternatives in both cases were constructed to be completely parallel, a majority wanted to have frequency information for the positive predictive value (60%), but percentage information for the negative predictive value (62%).

The middle part of the table shows that, when women could choose between a frequency and a verbal statement, there was no clear preference for one of the two. Of the 172 choices of this type, 82 (48%) were assigned to frequency and 90 (52%) were assigned to

verbal statements. When women could choose between verbal, percentage and frequency statements (lower part of Table 5.16), verbal and percentage statements were about equally attractive, while frequency statements were chosen less often (43%, 40% and 16%, respectively). However, although the formulation alternatives were constructed to be parallel for both topics (false-negative and false-positive rate), the distribution of choices differed considerably, indicating again that women had no consistent preference for one of the three ways of representing statistical information.

Table 5.16
Women's choices from frequency, percentage and verbal statements

Topic	Frequency statement	Percentage statement	Verbal statement	Total [a]
Frequency vs. percentage				
Cumulative lifetime risk of breast cancer	37 (1)	26 (1)	-	63
Cumulative lifetime risk of breast cancer death	15 (1)	14 (1)	-	29
Positive predictive value	37 (2)	23 (2)	-	60
Negative predictive value	19 (2)	33 (2)	-	52
Total	*108 (6)*	*96 (6)*		*204*
Frequency vs. verbal				
Risk of having to undergo a biopsy due to a false-positive result	27 (1)	-	26 (1)	53
Relationship breast cancer risk and age	33 (1)	-	31 (2)	64
Relationship breast cancer mortality and age	22 (1)	-	33 (2)	55
Total	*82 (3)*		*90 (5)*	*172*
Frequency vs. percentage vs. verbal				
False-negative rate	5 (2)	41 (2)	20 (1)	66
False-positive rate	15 (2)	7 (2)	32 (1)	54
Total	*20 (4)*	*48 (4)*	*52 (2)*	*120*

Note. Numbers in parentheses refer to the number of statements that were offered in each cell. [a] The total in this column might differ from the total number of choices of each topic mentioned in Table 5.15 due to missing values.

Secondly, there were two topics for which the formulation alternatives were either more or less transparent. Earlier in the study, 92% of the women had said that mammography pamphlets should clearly mention all disadvantages of mammography screening (see above). When given a choice between a formulation that mentioned pain during mammography and one that did not mention pain (but was otherwise equivalent), only 62% women picked the formulation that mentioned pain (i.e., 41 of the 66 women who chose the topic "procedure of mammography"). Similarly, 81% of the women had said that the pamphlets should depict all differing points of view if experts disagree in their evaluation of some aspect of mammography screening. In the formulation choices, however, only 52% of the women preferred a statement that mentioned differing recommendations concerning the screening interval over one that only depicted one recommendation (i.e., 32 of the 62 women who chose

the topic "screening interval"). Thus, the reported preferences concerning transparency in mammography pamphlets made earlier were only moderately predictive for the formulation choices made here.

Summary of results for information demand

A majority of the women said that mammography pamphlets should inform about all disadvantages and about all different points of view that experts might have about certain aspects of mammography screening. However, when they later had to choose statements for a subjectively good mammography pamphlet, only half of the women included a statement that featured different viewpoints. Women were undecided on the question whether mammography pamphlets should contain recommendations about screening participation or not.

Overall, women considered many pieces of information about mammography screening relevant for inclusion in a pamphlet. At the beginning of the study, women were most interested in learning about risks and side-effects, procedure, test efficiency, and benefit of mammography screening. Out of a list of 35 potential pamphlet topics, they later chose on average 27 to be included in a subjectively ideal mammography pamphlet. The amount of radiation and the likelihood of breast cancer between screenings were the most relevant risk information aspects. Women were also highly interested in the false-negative rate of mammography screening, more so than in the false-positive rate.

The acceptance of detailed information seemed to depend on the topic. For instance, many women wanted to be informed about the abstract measure of radiation, but most did not want an extensive explanation of predictive values of test results. However, the reason for the latter could also be that this explanation contained more pieces of statistical information. Women were divided on the question of relevance of statistical information in mammography pamphlets. About half thought that precise numerical information was helpful and should be included often in pamphlets, whereas the other half thought numbers were confusing and preferred verbal statements. This was also reflected in the choices of formulations for the ideal mammography pamphlet that differed in the way statistical information was represented, verbal and numerical statements were chosen equally often. Among numerical statements, there was also no consistent preference for either frequency or percentage statements.

R4. Participation intention

All women were asked at the beginning (T1, Questionnaire 1) and at the end of the study (T3, Questionnaire 5) how they currently thought about participating in mammography screening; they could indicate whether they currently intended to participate, did not intend to participate or were not sure about participation. Women in the FP and PP groups received the participation question also immediately after reading the pamphlet (T2, Questionnaire 2).

As Figure 5.6 shows, at all three times, the majority of women said that they planned to participate in mammography screening (T1: 53%, T2: 50%, T3: 58%). Only few women said they would not participate (T1: 8%, T2: 6%, T3: 4%). The remaining women were insecure about the decisions (T1: 32%, T2: 40%, T3: 38%). There were no substantial differences between the time points for the three conditions, and thus overall no effect of the model pamphlet on participation intention.

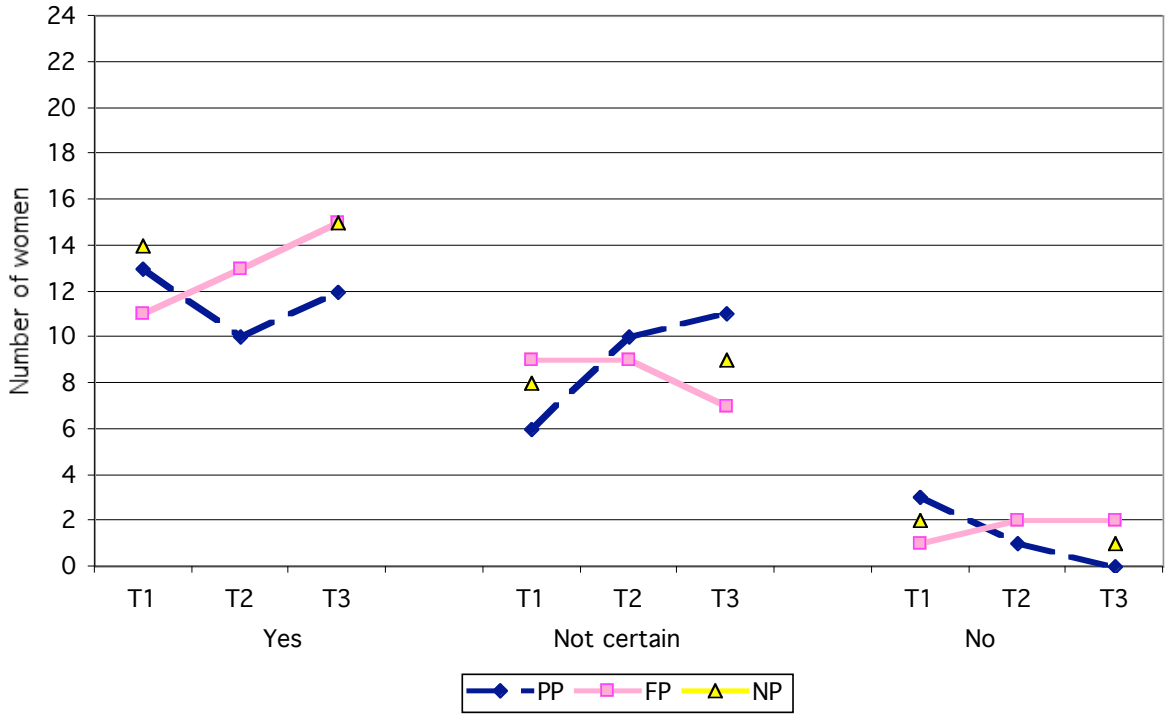

Figure 5.6 Women's intention to participate in mammography screening at three time points during the study

To check whether eventual changes in participation intention were hidden on the group level, an analysis on the individual level was made for the decisions at T1 and T3. But Figure 5.7 confirms the impression that little change took place and most women were willing to participate. 47 of 72 women did not change their intention (65%), and two thirds of these

women intended to participate. 20 of 72 women reported some change (27%), 11 to uncertain and 8 to yes. There was no effect of the three conditions on participation intention.

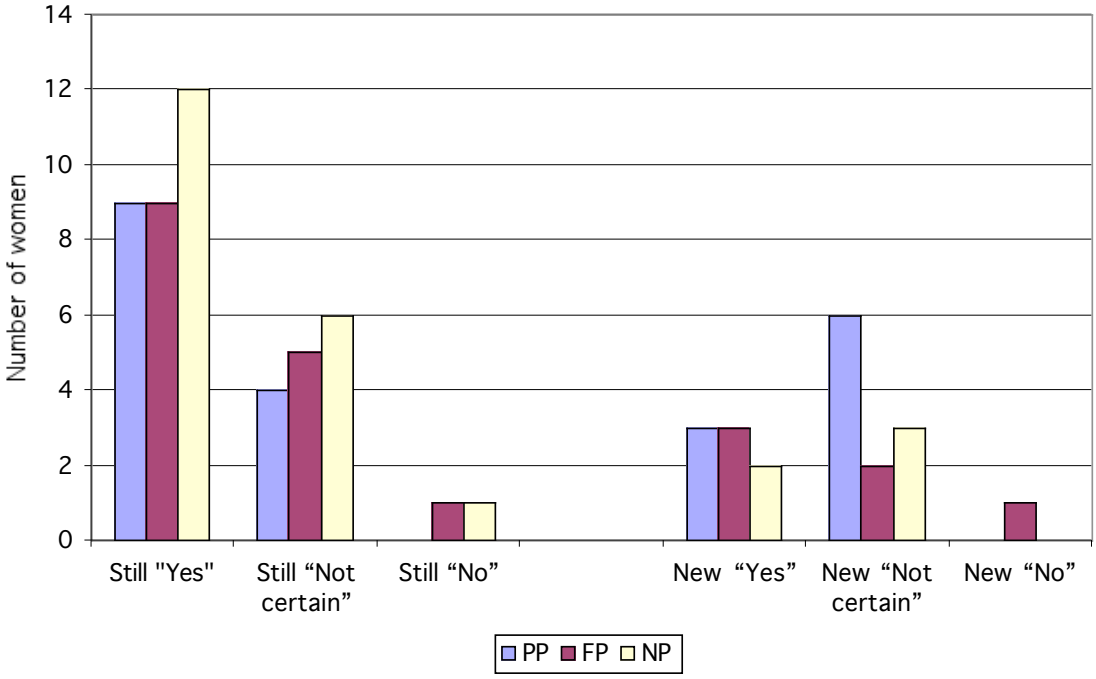

Figure 5.7 Individual changes in the intention to participate in mammography screening from the beginning (T1) to the end (T3) of the study

Discussion

„It seems to be a widespread belief that anyone can write a leaflet or specify the content and style of a video. This myth persists in spite of considerable evidence to the contrary."

(Wright, 1999a, p.718)

The main goal of this study was to find out whether natural frequencies could help to improve understanding of mammography pamphlets. The results showed overall only a small effect of the factor statistical format on learning from the pamphlets. Compared to the percentage pamphlet, the natural frequency pamphlet

- *did not* increase the number of risks recalled from the pamphlet;
- *did not* improve the over- and underestimations of the predictive values and the number of false positives in 10 years;
- *impaired* the estimations of the error rates of mammography screening; but

- *improved* estimations of breast cancer incidence, mortality, and mortality reduction through screening;
- *increased* the number of women who could identify the correct meaning of an ambiguous sentence about the risk reduction of mammography screening (small effect).

Thus, natural frequencies had some advantage over percentages, but the advantage was rather small and could be found for some, but not all measures. Simply presenting statistical information in a mammography pamphlet in terms of natural frequencies is not a silver bullet to eliminate, or even reduce, all misperceptions of women about mammography screening. Still, almost every time there was a difference between the percentage and the frequency group, it was the frequency group that had more accurate knowledge. Especially the understanding of information about breast cancer mortality with and without screening (and thus the benefit of screening) was facilitated by a representation in terms of natural frequencies. However, something must have been confusing about the presentation of the error rates in the frequency pamphlet, because the estimates here got worse after reading. This is somewhat surprising, because the presentation of the test efficiency in the model pamphlet was very similar to the typical diagnostic inference problems where natural frequencies facilitate statistical thinking. Some possibilities to improve especially this part of the model pamphlet will be discussed below. Nevertheless, natural frequencies were clearly more helpful than percentages in explaining the mortality reduction of mammography screening and should thus be the preferred format for this topic. Note that the women did not consistently prefer one format over the other, be it in the evaluation of the pamphlets or in the choice of formulations for their subjectively ideal pamphlet, so that we can indeed focus on the knowledge variables to decide whether one specific format should be recommended in pamphlet design.

Why was the difference between the two formats so small, given all the evidence for an advantage of natural frequencies that were presented earlier? One might suspect that the specific problem considered here, namely recalling statistical information from a long text, is simply a problem too different from those studied before, that is, a problem where an ecological representation of information is *not* crucial for good performance (other variables such as attention and short-term memory might be much more important).

But it would be premature, in my view, to draw this conclusion now. Women in both pamphlet groups had problems in recalling the values given in the pamphlets; in fact, the pamphlets seemed to yield only a very limited learning effect. This is especially worrying since the women who participated in this study were probably more interested and had better

prior knowledge about mammography screening than the average reader of a mammography pamphlets. However, it is not an uncommon finding that patients have substantial problems in recalling health information, provided either in written or oral form (e.g., Hopwood, 2000; Lloyd et al., 1999; Morrell, Park, & Poon, 1989). I suspect two factors that might have contributed to the disappointing learning effects of both pamphlet versions and that therefore made it difficult to detect potential format effects in the present study.

First, the amount of information in the pamphlet might have simply been too great, so that recalling nine out of 28 pieces of statistical information was too difficult for the participants (who were not the usual psychology students, but women between 40 and 69 years of age). An easy way to test this hypothesis would be to present a text with much less information, say, only the four pieces of information on test efficiency, or only the information on the benefits of screening. But even if an advantage of natural frequencies over percentages could be found in this simplified text, the problem remains: how could such a format effect be made beneficial for pamphlets with more information? One way could be to focus the attention of the reader with an improved structure and layout of the text and especially the addition of graphics (Reschke, 1990; Schriver, 1997; Wright, 1999a, 1999b). Graphics were also requested by a number of women, and I agree that some should be included in the next version of the model pamphlet. The reasons why I did not use graphics in the present study were that (a) I did not want to introduce another potential source of variance at this point, (b) the usefulness of adding illustrations to written materials for patients appears to be mixed (Wright, 1999), and (c) almost none of the available mammography pamphlets used graphics to illustrate statistical information given in the text.

The second factor that might have affected recall is that many of the participating women might not have been motivated to read the statistical information carefully. Remember that every third woman said that she wanted to see as few numbers as possible in a good mammography pamphlet, and the proportion of women who preferred verbal over numerical formulations was even higher. Shaw and Dear (1990, cited from Hallowell et al., 1997) reported that, in genetic counseling, women distinguished between information used for decision-making and information “for the record”. The women argued that in the former case, quantitative estimates should be given, as it is essential that risk information be understood precisely, whereas less precise qualitative descriptions of risk will suffice for the latter. Previous research showed that readers of health information materials direct their attention selectively to the points that are most relevant to them (Wright, 1999a). Therefore, it could be that the women in this study found the statistical information not relevant for their

participation decision – maybe because they are not used to this level of precision from the available pamphlets – and thus ignored the respective details in the pamphlet (some remarks of the women after the study support this hypothesis). Future research has to show to what extent this hypothesis is factual and how the problem can be overcome. This is especially important because the average reader of a mammography pamphlet might be even less interested in the topic than the women who participated in this study. One idea could be to include explicit motivational appeals in the pamphlets that explain why it is important to read the entire pamphlet carefully (see also below).

The second goal of the study was to assess the *information demand* of women between 40 and 70 years concerning mammography screening. Assessment of the information demand of the main audience is generally important in patient-information design (Coulter, 1997b; Secker & Pollard, 1995), because a pamphlet that ignores it is likely to be inefficient: Either because the reader still feels that she cannot make an informed decision after reading it, or because she does not even read the (whole) pamphlet because she thinks it is irrelevant for her personal questions. I also wanted to answer a more specific question of information demand in this study: Would the readers accept a pamphlet that, unlike the available pamphlets, quantifies risks and uncertainties as often as possible?

The results do not answer this question unequivocally. On the one hand, both pamphlet versions received rather good grades from the women and were evaluated positively. Only 10 women said that there were too many numbers included. Prior to reading the model pamphlet, 56% of the women said that they wanted to have as much precise numerical information as possible in a mammography pamphlet (and two women explicitly welcomed in the open evaluation questions that the model pamphlet realized this). On the other hand, the formulation choices for their subjectively ideal pamphlet showed a preference for verbal over numerical statements. Some women remarked in the general feedback that they did not read through all the numbers because there were too many, and/or because they found them irrelevant, and the bad recall for some of these numbers supports the impression that this was a problem. Informally, the women mentioned that they especially disliked the numerical example of the test efficiency, and this example was indeed only rarely chosen for the subjectively good pamphlet.

Taken together, these results suggest that there is no single way of presenting statistical information that would satisfy all women. It is encouraging that a majority of the women seem to be positive about the inclusion of statistical information in the pamphlets. However, the 31% women who explicitly dislike numbers have to be addressed in pamphlet

design. One way could be the motivational appeals mentioned above that explain why this unfamiliar way of presenting information is relevant. Another way could be to avoid clustering of numbers within the pamphlet and to distribute them more evenly over the text to prevent complete paragraphs from being skipped (as was the case for the detailed numerical example about the test efficiency). Finally, it would be helpful to find out more about why these women dislike numerical information to address their concerns more specifically in the pamphlets. We saw that age and level of interest were not associated with the preference for numerical information, but level of education was. However, we saw above that even highly educated professionals show similar reservations against numerical information in risk communication. To elucidate the role of education in this context, a more comprehensive study of the group of people who dislike numerical information is necessary.

Another important result of this study was that the information demand of the participating women was quiet high, they wanted to be informed about a number of different aspects to be able to make an informed decision about participation in mammography screening. Women were most interested in learning about risks and side effects, the procedure, benefit, and test efficiency of mammography screening. It seems that the women were especially concerned with the error rates of mammography (visible in the reported mood influence, risk knowledge after the pamphlet, choices, etc.), and more in the false-negative rate than in the false-positive rate (interestingly, also many physicians seem to worry more about false negatives than about false positives; Gigerenzer, 2002). This is an example for one result of the study that could lead directly to a recommendation on how to improve the currently available mammography pamphlets: Information about error rates should be featured much more prominently in the pamphlets (e.g., only 11 of the 27 pamphlets analyzed in Chapter 4 mentioned the false-negative rate or the sensitivity; in 2 out of 11, this information was not mentioned in the main text, but in the Appendix under the heading "future developments").

However, some of the choices that women made for their subjectively ideal pamphlet are less easy to convert into recommendations for pamphlet design. Some of the topics or representations that women preferred were, according to the discussion in Chapter 4, not the most relevant ones: for instance, clearly more women chose the relative over the absolute mortality reduction to represent the benefit of screening, and more women were interested in the absolute amount of the radiation in mammography screening, rather than in the actual health risks associated with this radiation. There is no general solution to the question on how to bring together the opinions of the audience on the one hand and those of experts on the

other hand (see Wright, 1994), especially opinions about what information is relevant to ensure truly informed decision-making. It is straightforward to supplement and structure the topics proposed by experts according to the priorities of the audience, but obviously pamphlet designers should not choose a representation of a topic that has been shown to produce misunderstandings (e.g., the relative risk reduction) just because a majority of the audience prefers this representation over a more comprehensive one. The literature on the design of patient information has, to the best of my knowledge, not yet addressed this specific problem. My tentative answer would be that in all cases the most comprehensive representation (based on empirical evidence) should be used in patient information, even if it is not the most preferred representation of the audience, *given* that it has at least some basic acceptance in the audience (because if there is no acceptance at all, the information will probably be skipped).

Related to that, there is no clear consensus on how much information exactly should be given. Some say the patient should get "as much information as the patient freely chooses to have" (O'Hagan, 1991, cited in Marshall, 1996). The British General Medical Council (1999) gives a list of points that should be mentioned, while the German Medical Council says that the amount of disclosure depends on "factual and temporal necessities", such as acuteness and severity of the health threat (Bundesärztekammer, 1990). There should be clear guidelines to help doctors communicate correctly, guidelines concerning both content and representation. The guidelines should be developed together with the users, patients and doctors, to ensure usability (Coulter, 1997b; Schriver, 1997; Secker & Pollard, 1995; Wright, 1999a, 1999b).

Finally, I would like to conclude this chapter with another interesting observation. More than 90% of the women said that mammography pamphlets should inform about all disadvantages and about all different points of view that experts might have about certain aspects of mammography screening. Note that the participating women were relatively "pro" mammography screening, given their prior mammography experience and the high proportion of women who wanted to participate in screening at any time during the study. Still, they voted for maximum transparency (a similar finding has been reported in Cockburn et al., 1999). This is an important counter-argument against those proponents of mammography screening who discourage a transparent discussion of the risks of mammography screening, because they fear that doing so would keep women from participating in the screening (Napoli, 1997). From an ethical point of view, this position is problematic. The results reported here suggest that it could also be incorrect.