LUND UNIVERSITY

Psychosocial consequences of false-positive mammography among women attending breast cancer screening. Assessment, prediction, and coping.

Bolejko, Anetta

2014

Psychosocial consequences of false-positive mammography among women attending breast cancer screening

Assessment, prediction, and coping

Anetta Bolejko

LUND
UNIVERSITY
Faculty of Medicine

DOCTORAL DISSERTATION
by due permission of the Faculty of Medicine, Lund University, Sweden.
To be defended in Lilla Aulan, Jan Waldenströms gata 5, MFC SUS Malmö
the 7th of February 2014 at 9.00 am.

Faculty opponent
Docent Claudia Lampic
Karolinska Institutet

<table>
<tr><td>Organization
LUND UNIVERSITY
Department of Health Sciences
Author(s) Anetta Bolejko</td><td colspan="2">Document name
DOCTORAL DISSERTATION
Date of issue: February 7, 2014
Sponsoring organization</td></tr>
<tr><td colspan="3">Title and subtitle: Psychosocial consequences of false-positive mammography among women attending breast cancer screening: Assessment, prediction, and coping.</td></tr>
<tr><td colspan="3">Abstract: One side-effect of breast cancer (BC) screening is a false-positive mammogram among healthy women. That is, findings on a screening mammogram that lead to additional breast examination(s) but where the woman is eventually considered free from BC. There is evidence of short-term psychosocial consequences of false-positive BC screening. Regarding long-term consequences, research findings are inconsistent. Lack of validated condition-specific questionnaires targeting such consequences has been postulated as a potential reason for the inconsistencies. Therefore, the Consequences of Screening – Breast Cancer (COS-BC) questionnaire was developed in Denmark. However, before the COS-BC can be used for studying psychosocial consequences of false-positive BC screening across countries, it needs to be adapted and psychometrically (validity and reliability) tested therein. Furthermore, studies of prediction of long-term psychosocial consequences of false-positive BC screening and coping with such consequences might identify women at risk as well as interventions to prevent consequences of screening. Thus, the aim of this thesis was to validate measures of and study the psychosocial consequences of false-positive mammography among women in a Swedish breast cancer screening programme, and to explore how women cope with such a situation. Interviews with 26 women experiencing false-positive screening mammography (Paper I) provided support for the content validity of a Swedish version of the COS-BC; questionnaire items were generally found relevant, understandable, and covering the psychosocial consequences of false-positive BC screening. Psychometric tests (Paper II) of the COS-BC among 1442 women with false-positive or negative mammography demonstrated support for five COS-BC scales (Sense of dejection, Anxiety, Behavioural, Sleep, and Existential values) for cross-sectional and longitudinal group assessments. The remaining seven COS-BC scales should be used more cautiously. One year follow-up study (Paper III, framework) of 399 recalled women and 449 controls showed that women experience psychosocial consequences targeted by the COS-BC scales, except for breast self-examination consequences. Early recall for subsequent mammography demonstrated the strongest prediction of long-term consequences. Dissatisfaction with information at recall, worry about BC, lack of social support, and being foreign-born were also identified as potential predictors. Interviews with 13 women (Paper IV) experiencing psychosocial consequences of false-positive screening mammography revealed that coping with the situation implied *a roller coaster of emotions and sense*. Social support, sisterhood, and being professionally taken care of were identified as important aspects of coping with the perceived psychosocial consequences of false-positive BC screening (Paper IV). In conclusion, findings of this thesis confirm the occurrence of short-term psychosocial consequences and demonstrated long-term consequences of false-positive screening mammography among women. Early recall should be avoided and personalized information and communication could be of value in order to diminish the risk of long-term psychosocial consequences of false-positive BC screening. Further research is needed to investigate adequate communication styles, especially in order to face multicultural populations in the context of BC screening.</td></tr>
<tr><td colspan="3">Key words: women's health, false-positive screening mammography, breast cancer screening, predictors, coping, Consequences of Screening – Breast Cancer questionnaire, Rasch model, psychometric evaluation</td></tr>
<tr><td colspan="3">Classification system and/or index terms (if any)</td></tr>
<tr><td colspan="2">Supplementary bibliographical information</td><td>Language: English</td></tr>
<tr><td colspan="2">ISSN and key title: 1652-8220
Lund University Faculty of Medicine Dissertation Series 2014:13</td><td>ISBN 978-91-87651-37-3</td></tr>
<tr><td rowspan="2">Recipient's notes</td><td>Number of pages
198</td><td>Price</td></tr>
<tr><td colspan="2">Security classification</td></tr>
</table>

Signature  Date February 7, 2014

Psychosocial consequences of false-positive mammography among women attending breast cancer screening

Assessment, prediction, and coping

Anetta Bolejko

LUND
UNIVERSITY
Faculty of Medicine

English language revision: Christopher Kennard

Faculty of Medicine, Department of Health Sciences
ISBN 1652-8220
ISSN 978-91-87651-37-3
Printed in Sweden by Media-Tryck, Lund University
Lund 2014

"No matter how confused, self-doubting, or ambivalent we are about what's happening in our interactions with other people, we can never entirely silence the inner voice that always tells us the truth. We may not like the sound of the truth, and we often let it murmur just outside our consciousness, not stopping long enough to listen. But when we pay attention to it, it leads us towards wisdom, health, and clarity. That voice is the guardian of our integrity."

Susan Forward

Contents

Abstract

One side-effect of breast cancer (BC) screening is a false-positive mammogram among healthy women. That is, finding(s) on a screening mammogram that lead to additional breast examinations but where the woman is eventually considered free from BC. There is evidence of short-term psychosocial consequences of false-positive BC screening. Regarding long-term consequences, research findings are inconsistent. Lack of validated condition-specific questionnaires targeting such consequences has been postulated as a potential reason for the inconsistencies. Therefore, the Consequences of Screening – Breast Cancer (COS-BC) questionnaire was developed in Denmark. However, before the COS-BC can be used for studying psychosocial consequences of false-positive BC screening across countries, it needs to be adapted and psychometrically (validity and reliability) tested therein. Furthermore, studies of prediction of long-term psychosocial consequences of false-positive BC screening and coping with such consequences might identify women at risk as well as interventions to prevent consequences of screening. Thus, the aim of this thesis was to validate measures of and study the psychosocial consequences of false-positive mammography among women in a Swedish breast cancer screening programme, and to explore how women cope with such a situation. Interviews with 26 women experiencing false-positive screening mammography (Paper I) provided support for the content validity of a Swedish version of the COS-BC; questionnaire items were generally found relevant, understandable, and covering the psychosocial consequences of false-positive BC screening. Psychometric tests (Paper II) of the COS-BC among 1442 women with false-positive or negative mammography demonstrated support for five COS-BC scales (Sense of dejection, Anxiety, Behavioural, Sleep, and Existential values) for cross-sectional and longitudinal group assessments. The remaining seven COS-BC scales should be used more cautiously. One year follow-up study (Paper III, framework) of 399 recalled women and 449 controls showed that women experience psychosocial consequences targeted by the COS-BC scales, except for breast self-examination consequences. Early recall for subsequent mammography demonstrated the strongest prediction of long-term consequences. Dissatisfaction with information at recall, worry about BC, lack of social support, and being foreign-born were also identified as potential predictors. Interviews with 13 women (Paper IV) experiencing psychosocial consequences of false-positive screening mammography revealed that coping with the situation im-

plied *a roller coaster of emotions and sense*. Social support, sisterhood, and being professionally taken care of were identified as important aspects of coping with the perceived psychosocial consequences of false-positive BC screening (Paper IV). In conclusion, findings of this thesis confirm the occurrence of short-term psychosocial consequences and demonstrated long-term consequences of false-positive screening mammography among women. Early recall should be avoided and personalized information and communication could be of value in order to diminish the risk of long-term psychosocial consequences of false-positive BC screening. Further research is needed to investigate adequate communication styles, especially in order to face multicultural populations in the context of BC screening.

Abbreviations

ANOVA	Analysis of variance
BC	Breast cancer
COS-BC	Consequences of Screening – Breast Cancer questionnaire
COS-BC 1	Consequences of Screening – Breast Cancer questionnaire part 1
COS-BC 2	Consequences of Screening – Breast Cancer questionnaire part 2
COS-LC	Consequences of Screening – Lung Cancer questionnaire
CTT	Classical test theory
CVI	Content validity index
DIF	Differential item functioning
DP	Dual-panel translation
ICC	Item characteristic curve
I-CVI	Content validity index of an item in a rating scale
k*	Modified kappa
LC	Lung cancer
NHP	Nottingham Health Profile
PCQ	Psychosocial Consequences Questionnaire
OR	Odds ratio
RM	Rasch measurement model
S-CVI	Content validity index of a rating scale
SD	Standard deviation
T1, T2	Administration of the study questionnaires

Glossary of terms

Abnormal screening mammogram	A mammogram with radiological findings leading to recall for further diagnostic work-up.
Condition-specific scale	A rating scale developed for a specific target population and/or context.
Construct	A latent variable which is not directly observable, but only by its manifestations (typically operationalized by items in a rating scale).
Content validity	The extent to which items in a rating scale are relevant to and cover the construct that the scale intends to measure.
Coping	"Constantly changing cognitive and behavioural efforts to manage specific external and/or internal demands that are appraised as taxing or exceeding the resources of the person" (1, page 141)
Early recall	Subsequent mammography performed with a shorter notice than routine screening.
False-positive screening mammography	Course of events which follows an abnormal screening mammogram that, after additional breast examination(s) is considered free from breast cancer.
Item	In a rating scale, a question or statement to respond to.
Mammographic service screening	Population-based screening programme where each eligible woman who is registered in the region served by the programme is individually invited to attend screening.

Negative screening mammography	An examination of a woman's breasts using X-ray (a screening mammogram) which after evaluation is considered free from breast cancer.
Prevalence	The number of events; a disease or condition, in a given population at a specific time point.
Predictor	A test result or other condition that is considered to forecast an event; for example estimating a risk of developing a condition.
Psychometric properties of rating scales	The extent to which a rating scale is successful and reliable in measuring the construct that the scale intends to measure.
Rasch measurement model	A psychometric measurement model that mathematically defines data requirements for objective measurement. Whether rating scales yield valid measurement depends on the extent to which data fit the Rasch model. The model postulates that the probability of a certain item response is a logistic function of the difference between the level of the measured construct represented by the item and that possessed by the person.
Screening	A test applied to "a defined group of persons in order to identify an early stage, a preliminary stage, or a risk factor of a disease. The object of screening as a service is to identify a certain disease or risk factor for a disease before the affected person spontaneously seeks treatment, in order to cure the disease or prevent or delay its progression" (2, page 2)
Unidimensionality	Items in a rating scale representing a single common construct.

List of papers

This thesis is based on the following papers, which will be referred to in the text by their Roman numerals.

I Bolejko A, Wann-Hansson C, Zackrisson S, Brodersen J, Hagell P. Adaptation to Swedish and further development of the Consequences of Screening – Breast Cancer questionnaire: a multimethod study. Scandinavian Journal of Caring Sciences, 2013;27:475-486.

II Bolejko A, Brodersen J, Zackrisson S, Wann-Hansson C, Hagell P. Psychometric properties of a Swedish version of the Consequences of Screening – Breast Cancer questionnaire. Submitted for publication.

III Bolejko A, Hagell P, Wann-Hansson C, Zackrisson S. Women experience long-term psychosocial consequences of false-positive mammography – can we predict them? A cohort study in a population based screening programme. Submitted for publication.

IV Bolejko A, Zackrisson S, Hagell P, Wann-Hansson C. A roller coaster of emotions and sense – coping with the perceived psychosocial consequences of a false-positive screening mammography. Journal of Clinical Nursing, doi:10.1111/jocn.12426.

Introduction

Breast cancer (BC) is the most frequent cancer among women worldwide and the leading cause of women's death from cancer (3). Although some factors have been associated with the risk of getting BC (4), no prevention strategy is recognised at present. Thus, early detection of the disease is recommended in order to promote successful treatment (5) and thereby reduce disease mortality rates (6, 7). Therefore, mammographic screening has been established in many countries during the last few decades (6, 8). However, screening also has disadvantages (9), such as radiological findings on a screening mammogram leading to additional breast examination(s) but where the woman is eventually considered free from BC (referred to as false-positive screening mammography) (6). It has been estimated that for every 1000 women participating in ten biennial BC screening tests in Europe, 200 women will have false-positive screening mammography (7, 10).

While false-positive results are a recognised side-effect of the mammographic screening process (7), less is known about the psychosocial consequences among women who have experienced such circumstances. Assessment of the psychosocial consequences of screening is challenging (11, 12) and studies addressing such consequences in a context of false-positive mammographic screening have shown inconsistent results (13-16). That is, there is evidence of short-term consequences of emotional dysfunction and anxiety following a recall letter for additional breast examination and at the diagnostic work-up (13, 16). Regarding long-term consequences, some studies have shown BC-related worry and distress, whereas others have either found no psychosocial consequences or have reported mixed findings (13, 14, 16). Lack of validated condition-specific questionnaires targeting the consequences of false-positive mammographic screening has been postulated as a potential reason for these inconsistencies (13). Therefore, the Consequences of Screening – Breast Cancer (COS-BC) questionnaire was developed in Denmark to study such consequences (17). This is the first condition-specific tool of its kind. Yet, validation studies in non-Danish settings are needed to investigate the usefulness of the COS-BC before its wider international use (18).

A recently published study based on the COS-BC confirmed short-term psychosocial consequences and also found long-term effects following false-positive screening

mammography among women in Denmark (19). The prevalence, magnitude, and longitudinal development of such consequences among participants in BC screening services in other countries are, however, still to be further examined. In addition, identifying women at potential risk would provide valuable knowledge to facilitate directed interventions to prevent psychosocial consequences of false-positive mammographic screening and provide support for those with a potentially compromised ability to overcome such consequences. For that purpose, socio-demographic and psychological predictors of long-term consequences have been investigated (16). For instance, younger age, living alone, low level of education (20, 21), and dissatisfaction with client-provider communication (22) and distress at screening (20) have been hypothesised as potential predictors of long-term consequences of false-positive screening mammography. These hypotheses warrant further investigation, in particular when the consequences are assessed by the use of the COS-BC.

In addition to population-based studies, a nuanced understanding of women's experiences of false-positive screening mammography is needed in order to reveal the complexity of the matter (15, 16). Therefore, qualitative studies have also been conducted (23-26). However, even though women's experiences of recall examinations following BC screening appear to have been revealed, they have not been fully elucidated in the context of false-positive screening, because most of the previous qualitative studies also included women diagnosed with BC (23, 25, 26). Women with a false-positive result do not have BC and screening asymptomatic people for disease implies that they do not perceive themselves as ill (9). Hence, it might be anticipated that women experiencing false-positive screening mammography react rather differently than women diagnosed with BC (16). Knowledge about their explicit experiences and how they cope in such a situation might reveal further aspects of screening, and also provide clues for interventions. Coping with the perceived psychosocial consequences of false-positive screening mammography as described by women themselves, does not seem to have been explored yet.

Background

Mammographic screening

The concept of screening

As early as in 1861 a physician at London's Royal Hospital for Chest Diseases argued for periodical examinations among asymptomatic adults and children in order to give recommendations for the prevention of ill health and disease (27). During the twentieth century, the concept of promoting public health and preventing disease and illness became increasingly widespread (9). Screening for syphilis, tuberculosis, and diabetes were examples of early screening practices for infectious or chronic diseases (28). Screening for cancer has been implemented progressively; cervical and breast cancer to mention a few types (28), and new screening programmes for additional cancer types are being considered (29). As a result of the increasing interest in medical screening, the World Health Organisation commissioned guidelines on the principles and practice of screening for disease, which were published in 1968 (30). Since then, many national guidelines for screening have been proposed, as the development of diagnostic screening technologies has continued (31, 32). In 1994, the Committee of ministers of the European member states presented recommendations for medical screening for chronic diseases (2). According to these recommendations, screening is defined as *applying a test to a defined group of persons in order to identify an early stage, a preliminary stage, a risk factor or a combination of risk factors of a disease. The object of screening as a service is to identify a certain disease or risk factor for a disease before the affected person spontaneously seeks treatment, in order to cure the disease or prevent or delay its progression or onset by (early) intervention* (2, page 2). From the report it follows that although screening has the potential of improving public health, it also has adverse effects; psychological consequences among those with limited health gain and side-effects of invasive follow-up of false-positive screening have been mentioned. It has been postulated that although the advantages of screening are usually well described, it is also important to be aware of the disadvantages; they should be evaluated in relation to the target population and the individual person, and be reported to the public (2). In 2003, the Council of the

European Union expanded on the previous recommendations, proposing guidelines for cancer screening as well (33).

Mammographic screening

In 1963, Sam Shapiro and co-workers initiated the first randomised controlled trial exploring the effects of frequent mammography and clinical breast examination on BC mortality reduction (34). Since then, several randomised controlled trials have been conducted, including four in Sweden (35-39), in order to investigate the efficacy of BC screening with or without clinical breast examination (6). Swedish trials were similar in that they were population-based, comparing mammographic screening alone with no screening (6). A meta-analysis from 1993 of Swedish studies showed a significant 29% relative risk reduction in BC mortality among women aged 50-69 at entry to the studies (40). Subsequent Swedish studies have supported the efficacy of mammographic screening among women 40-49 years of age (41-43). An additional long-term follow-up study from 2002 showed a significant 21% relative risk reduction in BC mortality in women 40 years of age or older at entry to mammographic screening (44). The results from the Swedish trials allow an estimation of the absolute risk reduction in BC mortality over a period of ten years, indicating the prevention of death from BC for five out of 10 000 women aged 40-49 at entry to screening, and for nine out of 10 000 women aged 50-69 (6).

Based on the results from the randomised controlled trials many European countries, several states in the U.S., and other countries such as Canada and Australia introduced screening programmes during the 1990s (6). Later on, following the recommendation of the Council of the European Union (33), screening programmes have been implemented in many member states (8). The programme of population-based service screening means that each eligible woman registered in the region, served by screening, is individually invited to attend BC screening (8).

Criticism has been raised regarding the randomised trials and BC screening has been questioned (45-47). In the Cochrane meta-analyses of the randomised trials the assumption of a 15% relative risk reduction in BC mortality would mean prevention of death from BC for five out of 10 000 women participating in screening in ten years (46, 47). It has also been argued that the BC mortality reduction that has been observed since the introduction of population-based BC programmes might be an outcome of improved treatment and management of health care systems rather than a result of screening (48). Critics also point out aspects of over-diagnosis and over-treatment as significant adverse consequences of screening (46, 47). This means that some women attending mammographic screening are diagnosed with BC that would most likely not have developed into a clinical disease in their lifetime. The

troublesome aspect is that there exists no possibility of verifying whether a woman diagnosed with BC has been over-diagnosed or not. Yet, when the abnormality is detected, it might also lead to over-treatment and therefore contribute to side-effects of mammographic screening. Another disadvantage of screening is a false-positive result (6). That is, radiological findings on a screening mammogram (an abnormal screening mammogram) that lead to recall for additional breast examination(s) and where the woman is eventually considered free from BC. According to the European guidelines for quality assurance in BC screening the recall rate for additional examinations should be below 5% at initial screening and below 3% at subsequent screening rounds, though not lower than 1% as it might be associated with a reduced cancer detection rate (49). The recall rate refers to the proportion of women recalled for additional diagnostic assessment owing to an abnormal screening mammogram among all women attending screening. Screening programmes across countries have reported that between 3.5% and 54% of recalled women are diagnosed with BC (6). Consequently, a considerable number of recalled women have false-positive screening mammography. It has been estimated that for every 1000 women aged 50-51 at entry to screening and participating in ten biennial BC screening tests in Europe, seven to nine women are prevented from dying from BC, four are over-diagnosed, and 200 have false-positive screening mammography (7).

Mammographic service screening in Sweden and in the study setting

Sweden was one of the first countries to implement a population-based mammographic screening programme that has been in operation nationwide since 1997 (6, 50). In 1986, the Swedish National Board of Health and Welfare recommended organized BC screening for women 40-74 years of age, but shortly thereafter revised the recommendations to invite women aged between 50 to 69 (50). On the basis of additional results for women under the age of 50, mammographic screening was yet again recommended in 1998 for women from the age of 40 (50), and current recommendations include women between the ages of 40 and 74 (51, 52).

Mammographic screening has a long tradition in Malmö. The Malmö Mammographic Screening Trial was one of the first randomised controlled trials conducted in Sweden (35, 41). After the termination of the trials a population-based screening programme was established in 1990. The programme included women 50-69 years of age, and after 1996 women aged up to 74 were also invited. Since 2009, women between 40 and 54 years of age have been invited via mail at 18-month intervals, while those aged 55 to 74 at 24-month intervals (53). The rationale behind the different screening intervals is that breast density in general is higher in pre-menopausal women, reducing the sensitivity of mammography within this group.

When invited to mammographic screening, women receive a letter that includes the time and place for the examination, the purpose of mammographic screening, and a brief description of the examination. Women can call and reschedule their appointment. No reminders are used. At screening, each woman receives oral and written information about the potential need for additional breast examinations. Approximately two weeks later each woman receives either written information that BC was not found (negative screening mammography) or a recall letter to have an additional examination within one week. At recall, usually a clinical mammogram is conducted along with ultrasound breast examination, which provides sufficient diagnostic work-up for most women free from BC. Women are informed at the visit that additional examinations did not show BC (having had a false-positive screening mammogram). Some women need further breast examinations, such as fine needle/core biopsy and follow-up by the surgical breast clinic. Depending on the type of assessment and the probability of cancer, women are either scheduled for a telephone call within approximately two weeks or for a visit to the surgical breast clinic. The appointment at the surgical breast clinic is provided at recall. Women who eventually are considered free from BC following these additional procedures are invited to subsequent screening mammography according to the standard practice. In some cases, early recall for subsequent mammography is recommended; that is, recall for further mammographic testing with a shorter interval, usually 4-6 months, than at the 18 or 24-month interval. All women with negative and false-positive screening mammography are recommended to regularly self-examine their breasts and consult the health care service if they feel a lump. The abovementioned clinical practice in Malmö relates to the period between 2008 and 2011.

Psychosocial consequences of false-positive screening mammography

Quantitative studies on psychosocial consequences

Psychosocial consequences of false-positive screening mammography among women have mainly been investigated by means of quantitative study designs. Both general outcomes and those related to feelings and thoughts about BC have been in focus (14, 16, 54). For example, anxiety and depression following false-positive screening mammography have been examined in a general sense, as well as BC-related distress, worry, fear, and anxiety (54). However, the concept of psychosocial consequences of false-positive mammographic screening had not been defined until the Consequences of Screening – Breast Cancer questionnaire (COS-BC) for targeting such consequences was developed. The theoretical framework of the concept is presented in the next chapter, as all studies of the psychosocial consequences of false-

positive mammographic screening, except for one (19), were conducted before the COS-BC was available.

All studies addressing the short-term psychosocial consequences of false-positive screening mammography appear to have found adverse effects (13, 16). These effects were reported by women after receiving a recall letter for additional breast examination following an abnormal screening mammogram and also in relation to the diagnostic work-up. The consequences appeared for instance as general anxiety (55, 56), or emotional dysfunction (57, 58) and distress related to thoughts about BC (21, 59). General anxiety has for example been reported among 46% of women waiting for the recall examination (55) and moderate to severe BC-related distress was found in 45% of women within one week after recall investigation (21).

Studies of long-term consequences of false-positive screening mammography have, on the other hand, shown inconsistent results (13, 15, 16). For example, distress related to BC was found in some studies (21, 22, 60), and intrusive thinking (61), worry about BC (61, 62) or BC anxiety (20) have also been reported. Conversely, other studies have found no effects on distress (57, 63) or BC worry and fear (63, 64). Depression and general anxiety were investigated in some studies, but generally no evidence of long-term effects was found (55, 56, 63, 65). Altogether, a meta-analysis of 17 studies published up to 2007 found that long-term psychosocial consequences of false-positive mammographic screening were limited to BC-related outcomes, such as anxiety, worry and distress about BC, perceived likelihood of BC, and frequency of breast self-examination (54). Outcomes in these studies have been investigated at various time intervals; from approximately one month after false-positive screening mammography, up to three years later (14). These studies were mostly conducted during the 1990s in different countries in Europe, including two in Sweden (21, 55, 66), as well as in the U.S., Canada, and Australia. A recently published study confirmed short-term psychosocial consequences and found BC-related effects up to three years following false-positive screening mammography among women in Denmark (19).

Assessment of psychosocial consequences

Previous study results should be evaluated in the light of the assessment tools that have been applied (13, 54). For example, the Hospital Anxiety and Depression Scale, the General Health Questionnaire, and the State Trait Anxiety Inventory have commonly been used (13). Although these scales have been psychometrically tested in the general sense, evidence regarding their appropriateness and psychometric properties among women who had experienced false-positive screening mammography is lacking (13). For example, the lack of documented content validity of

these tools among women having experienced false-positive screening mammography should be considered a particular disadvantage. Some studies have also used the Psychological Consequences Questionnaire (PCQ) (67); a condition-specific questionnaire. However, this assessment tool was developed to investigate the psychosocial consequences of mammographic screening, and was not initially intended for the assessment of psychosocial consequences of false-positive mammographic screening. In the development process interviews with women attending screening were conducted, including women with false-positive screening results. However women undergoing surgical biopsy or who were scheduled for early recall for subsequent mammography were not included (67). In addition, evidence of the adequacy of the PCQ to assess long-term consequences is lacking (13). Many studies among which long-term effects have been found, have also applied single questions, which were related to feelings and thoughts about BC (54). These were generally investigator-devised questions and have not been tested regarding their validity and reliability. On the whole, results from studies using such data can therefore be viewed as incomplete and partly inconclusive (13), which points to the need for further investigations of the psychosocial consequences of false-positive mammographic screening using different approaches. That is, prior to further assessment studies, the concept of the psychosocial consequences of false-positive screening mammography should be explored, followed by development of a condition-specific questionnaire targeting such consequences. The recently published Danish study (19) used such a questionnaire; the COS-BC. This questionnaire was developed and validated specifically for the assessment of the consequences mentioned above (17, 68, 69). The COS-BC appears to be the first of its kind and offers a new approach to study psychosocial consequences of false-positive cancer screening. However, validation studies in non-Danish settings are needed in order to investigate the usefulness of the COS-BC before use across countries (18).

Qualitative studies on psychosocial consequences

While the vast majority of studies addressing the psychosocial consequences of false-positive mammographic screening have been conducted by means of various assessment tools, only a few have employed qualitative methods to elucidate women's experiences following recall breast examinations (23-26). Consequently, a call for qualitative studies to further explore psychosocial consequences has been postulated (15, 16). Previous studies have reported that women tend to experience mixed and intense emotions related to the follow-up examinations following an abnormal screening mammogram. For example, a rapid diagnostic work-up has been perceived as reassuring, but might as well be seen as an indication of malignancy (25). Women felt like they were losing control, not knowing whether to imagine the worst or to accept the limitation of the facts at hand (25). Information about recall

and BC rates was hopeful for some, but for others it did not make much sense since numbers were perceived as irrelevant when a fatal disease became a possible reality (25). Studies have also revealed perceived dissatisfaction with client-provider communication and respect (23, 24, 26). Conversely, women having timely follow-ups found efforts from health care staff to be supportive and those who were confident in their ability to advocate for themselves experienced a feeling of being in control of their situation (23). Anxiety over the effects of repeated breast examinations was mixed with doubts about the diagnostic work-up (24). Consequently, receiving the diagnosis of not having BC not only led to a sense of relief, but also evoked further questions of whether additional diagnostic procedures were worth going through (24).

Available qualitative studies seem to have revealed women's experiences of recall examinations following BC screening, but less is known in the context of false-positive mammographic screening, since most of the studies also included women with a high suspicion of, or diagnosed with BC (23, 25, 26). It might be anticipated that women experiencing false-positive screening mammography react rather differently than women diagnosed with BC (16). Another aspect to be aware of is that in some studies only ethnically diverse women were included (23, 24). These studies were conducted in the U.S. where the health care system differs from that in several European countries, including in Sweden (27). Altogether, it follows that explicit experiences of false-positive screening mammography among women remain to be explored, and this should be done without any particular socio-demographic focus.

Women experiencing false-positive screening mammography do not have BC and if they were to experience psychosocial consequences following recall examinations it may be presumed that they have been harmed by screening. Exploring how women cope with the situation might identify support procedures for preventing consequences of false-positive mammographic screening or for providing support for those with a potentially compromised ability to overcome such consequences. Even though qualitative studies seem to have elucidated women's experiences of recall examinations following BC screening, studies addressing coping with psychosocial consequences of false-positive screening mammography as described by women themselves appear to be lacking. Coping might be defined as *constantly changing cognitive and behavioural efforts to manage specific external and/or internal demands that are appraised as taxing or exceeding the resources of the person* (1, page 141). In the context of false-positive mammographic screening, the *demand* is referred to as the experience of false-positive screening mammography. The *demand*, when *appraised as taxing or exceeding the resources of the person* can be perceived of as decreasing the person's psychosocial well-being. Coping, as *constantly changing cognitive and behavioural efforts to manage the demands* entails the effort the individual makes to manage the demands, without necessarily mastering the situation (1).

Predictors of long-term psychosocial consequences of false-positive screening mammography

Identifying women at risk of experiencing psychosocial consequences of false-positive screening mammography would provide valuable knowledge about how to allocate support resources. However, in contrast to numerous studies of long-term psychosocial consequences of false-positive mammographic screening, predictors of such consequences have been scarcely studied. Emphasis has been put on socio-demographic and psychological variables, for example, younger age, living alone, or perceived vulnerability. It has been found that a low level of education (20, 21), living in a high density urban area (21), dissatisfaction with client-provider communication (22), and distress, anxiety and worry at screening and diagnostic work-up (20-22, 60) were potential predictors of long-term consequences. Having invasive breast examinations (fine needle/core biopsy) following an abnormal screening mammogram or being scheduled for early recall for subsequent mammography have also been identified as potentially influencing the psychosocial consequences of false-positive screening mammography (60). However, these results warrant further investigation, particularly when the problem is addressed using the COS-BC.

Development of the Consequences of Screening – Breast Cancer (COS-BC) questionnaire

As noted above, the PCQ was one of the questionnaires used in studies of short- and long-term consequences of false-positive mammographic screening (13). The questionnaire assesses negative and positive physical, emotional, and social consequences of attending BC screening (67). However, its usefulness for the assessment of psychosocial consequences of false-positive mammographic screening has been questioned, in particular regarding long-term consequences (13); hence it was validated in the context of both abnormal and false-positive mammographic screening (17). First, focus group interviews with women facing such circumstances were conducted in Danish screening (68). Interviews revealed the ambiguity and irrelevance of several questionnaire items along with a compromised ability of the questionnaire to cover short- and long-term psychosocial consequences of false-positive screening mammography. Consequently, several items were reworded or excluded from the questionnaire, and a range of new items were developed to cover the concept of psychosocial consequences of false-positive mammographic screening. One additional item (*sick leave*) was included to address psycho-economic consequences. The interviews also revealed that prior to the final diagnosis only negative consequences were experienced, but following the diagnosis both positive and negative consequences might occur. Because of this, response categories to items addressing consequences

after the diagnosis were changed to response options representing both directions of potential consequences (17, 68). Altogether, the adaptation of the PCQ resulted in a new questionnaire, the Consequences of Screening – Breast Cancer (COS-BC), comprising two parts (17, 68). Part 1 consisted of items intended to represent the psychosocial consequences of abnormal mammographic screening prior to the final diagnosis. As such, they represented short-term consequences. However, these consequences may persist and some may not occur until after the final diagnosis. Part 2 consisted of items intended to represent only long-term (after the diagnosis) psychosocial consequences of false-positive mammographic screening.

Subsequent validity and reliability tests of the COS-BC provided initial support for its psychometric properties as an assessment tool of psychosocial consequences of false-positive mammographic screening in a Danish context (17, 69). The concept of short- and long-term consequences that had emerged in the interviews was supported by the psychometric tests (17, 68, 69). Consequently, six scales were developed in part 1 of the questionnaire; *Sense of dejection, Anxiety, Behavioural, Breast examination*, and *Sexuality*. Six items were proposed to be used as single items or in need of further evaluation. Responses to questionnaire items, except for the *sick leave* item, represented either intensity or frequency of experienced consequences, where higher scores denoted more negative consequences. Tests of part 2 items resulted in four scales; *Existential values, Social relations, Relaxed/calm*, and *Anxiety/reassurance about BC*. Response categories represented both directions of potential consequences; much less, less, the same as before, more, and much more of the consequences. These are recommended to be re-coded into three categories; the same as before, less/more, and much less/much more, where the categories represent changes in experienced consequences (regardless of direction) compared to before screening (17, 70). The final original Danish COS-BC consists of 30 (part 1) and 13 (part 2) items, respectively.

Some COS-BC items are not considered BC-specific and have also been applied in other cancer screening studies since they were hypothesised to be relevant regardless of the cancer type screened for (70). Focus group interviews have supported this hypothesis among people participating in a Danish lung cancer (LC) screening study, where the Consequences of Screening – Lung Cancer (COS-LC) questionnaire was developed (70). In addition, new potential items representing long-term consequences of LC screening were identified. However, some of these items were not LC-specific. They represented experiences of impulsivity, empathy, and relief. In total, 10 items representing long-term consequences of cancer screening in general were proposed (70). However, these items remain untested in other contexts, such as BC screening.

Theoretical framework of the concept of psychosocial consequences

The theoretical framework of the concept of psychosocial consequences operationalized in the COS-BC was the biopsychosocial medical model (71). The model postulates a basis for understanding various factors concerning illness and its treatment, taking into account the individual and social context one lives in. According to the model, nothing in nature can be described without defining the system it coexists with. Analogically, the person must be viewed in the social system he/she lives in, which includes the person's experiences and relationships as well as destabilizing events in the environment (72). Taking such a perspective, the developers of the questionnaire argue that as an abnormal screening mammogram raises a suspicion of BC in a healthy woman, it triggers a response from the woman that affects her life (17). The psychosocial consequences of false-positive screening mammography thus include psychological, emotional, sexual, social, behavioural, physical, and cultural characteristics that might influence the woman's well-being (17).

Latent variable measurement

The COS-BC is an example of a multidimensional questionnaire to target latent variables. Each dimension in the questionnaire is supposed to represent a *single* latent variable. Latent variables are considered phenomena not directly observable, but apparent only by their manifestations (18). For example, observable common behaviours and emotions when feeling threatened by BC might represent manifestations of the latent variable of BC anxiety. A model of observable manifestations of the latent variable is usually referred to as a construct. Manifestations of the latent variable are represented by questions or items, which are combined into rating scales. Other terms are also common, such as scales, measurement scales, instruments, and measurement tools. Rating scales of latent variables are supposed to represent a unidimensional construct (18, 73). That is, items in a scale should represent observable manifestations of a *single* construct. To obtain an assessment of the construct, scores for each item within a rating scale are usually added to provide a single total score of the latent variable (18). A questionnaire might consist of either one or multiple rating scales, while the latter provides several dimensions of different aspects of the same concept, for instance psychosocial consequences of false-positive screening mammography.

The quality of a rating scale is evaluated in terms of psychometric properties, such as validity and reliability, which is essential to determine the usefulness of the scale (18). Validity refers to whether items in a scale sufficiently cover the construct, and the extent to which the scale is successful in measuring the construct intended to be

measured. Reliability reflects the amount of error in measurement. Tests of validity and reliability provide evidence for the legitimacy in computing a total score from items in the scale, and the interpretability and precision (lack of measurement error) of that score.

Objective measurement can be defined as *the repetition of a unit amount that maintains its size, within an allowable range of error, no matter which instrument, intended to measure the variable of interest, is used and no matter who or what relevant person or thing is measured* (74). This means that rating scales of latent variables should generate measurements on an interval level, and the performance of the scale should be independent of the items in the scale and the sample the scale is applied to (73). It follows that, when these requirements are not satisfied, the numbers that a scale yields cannot be considered measures. This does not mean that the scale is not useful, but it should rather be considered an assessment represented by numbers that have no more than ordinal properties.

Psychometric properties of rating scales

Whereas content validity is a prerequisite for valid measurement, other psychometric properties such as for example construct validity and reliability, are also needed to be evaluated in order to determine the scale's usefulness for studying the construct it is supposed to measure. Validity and reliability are evaluated by a range of a priori hypotheses tests, of which no single test is either necessary or sufficient. Psychometric properties can be assessed by a variety of approaches; for example classical test theory (CTT) or the Rasch model (RM) (73, 75, 76). Whereas CTT is the most commonly used approach, the RM is considered preferable (73, 77).

Content validity

When developing as well as adapting and evaluating a translated rating scale it is important to ensure that its items are relevant with respect to the target construct and, conversely, that it does not contain items of poor relevance. Similarly, items should provide sufficient coverage of the construct that is intended to be measured (18). This is referred to as content validity. To assess the content validity of a scale, interviews and/or ratings by experts are recommended (78). For rating scales intended to investigate experiences and highly personal outcomes, representatives of the scale's target population are considered to be the experts. Content validity can be assessed by, for example, open-ended questions used in an interview and calculation of the content validity index (CVI) (79). The CVI is based on item relevance ratings and can be computed for each item (I-CVI) as well as for the overall scale (S-CVI) (80). It is recommended that the understandability of the items, response categories, and

format of the questionnaire is included in the evaluation process in order to minimize the risk of misunderstanding and maximize user-friendliness (81).

Psychometric tests according to the classical test theory

The CTT postulates that an observable score consists of a true unobservable score and measurement error (73). The errors are not correlated with the true or observable scores. However, these assumptions cannot be verified. Another troublesome aspect of the theory is that the distribution of the scale items on the continuum of the construct is dependent on the distribution of the sample the scale is applied to, and vice versa, as the observed score includes both the item and sample parameters. To investigate the quality of a rating scale, correlation statistics are usually used. For example, the relationship between scores on each item in a scale and the total score of remaining items (corrected item-total correlations) might be calculated in order to assess whether the scale items define a single construct (73). Another approach is to articulate hypotheses regarding the relationships between scores of the target scale and other variables, followed by assessment of the empirical correlations. Known-groups validity is also commonly used, where differences in scores between groups of people that are hypothesized to differ regarding the construct are assessed. Empirical observations that accord with a priori hypotheses are interpreted as support for the scale construct validity (18, 73). Reliability is expressed as a coefficient that ranges between 0 (no reliability) and 1 (perfect reliability) and can be assessed in different ways within the CTT, of which internal consistency and test-retest reliability are among the most common. Internal consistency is typically assessed by Cronbach's α (18) and test-retest reliability by the intraclass correlation coefficient between scores derived from the same individuals on two different occasions (typically 1-2 weeks apart) under the assumption that the measured construct has not changed (18, 82).

Psychometric tests according to the Rasch model

The RM articulates a mathematic definition of an objective measurement, which provides a means to test whether rating scales satisfy the requirements of objective measurements. According to the RM, a person's response to a scale item is a logistic function of the difference between the person's level of, for example BC anxiety, and the level of BC anxiety represented by the item. RM locates independently each person and scale item on a shared continuum (a logit metric; interval level) according to how much of the measured construct each person possesses in relation to each other, and, conversely, how much item represents in relation to one another (73, 75, 76). As such, RM states the fundamental requirement of measurement. From this it follows that observed data from rating scales can be tested against the model, which is generally referred to as test of fit, and potential problems of the scale functioning can be identified. Whether rating scales yield valid measurement depends on the extent to which data fit the RM (73, 75, 76). Consequently, sufficient data model fit

implies a justification to compute total scores across items, and that the scale satisfies the requirements of objective measurement.

The RM offers a range of analyses based on various hypotheses. For example, responses to an item in a rating scale are not supposed to differ for subgroups of people, for example for women who had experienced false-positive and negative screening mammography, given they have the same level (location) on the continuum of the construct. This requirement, referred to as the absence of differential item functioning (DIF), can be empirically tested by the use of the RM (73). Furthermore, ordered scale response categories are supposed to correspond to increasing or decreasing levels of the measured construct. The empirical functioning of response categories can be assessed as an additional aspect of model fit (73). However, it should be kept in mind that inconsistencies between the data and the RM are signs of differences against a perfect measurement model (77).

Cultural adaptation of rating scales

Translation of ratings scales is common before they are used in research studies or clinical practice, since they are usually developed in a single language. However, it is recognized that translation alone is not sufficient to adapt a rating scale to another language and/or culture (18, 83). One troublesome aspect is that the translated version might not be expressed in comprehensible lay language, as the scales are usually translated by highly educated bilinguals. Another issue might be cultural differences in defining and expressing the construct intended to be measured, which the translation process cannot capture since only the scale items that operationalize the construct in the source culture and language are available in the translation process. Furthermore, the psychometric properties of the translated version are unknown. The adaptation process should thus aim to develop a scale version that is expressed in lay language and that is conceptually, semantically, operationally, and psychometrically equivalent with the original scales (18). Conceptual equivalence means that the construct covered by the scale and its items exists and is relevant in the target population. Semantic equivalence is supported when items have the same meaning in the source and target populations. Operational equivalence refers to the appropriateness of instructions, format, and intended administration of the scale. Measurement equivalence is determined when the psychometric properties of the translated scale are tested within the target population, and are equivalent to those in the source version (18).

Rationale

Population versus individual perspective on mammographic screening

The key feature of mammographic screening is its detection of BC in an asymptomatic stage of the disease. Initiated treatment is thus expected to be more effective, thereby reducing BC mortality in the female population. To date, irrespective of whether the data was analysed in research studies, reviewed by either the proponents or critics of screening, or reported from population-based programmes, the results indicate that BC screening plays a role in preventing death from the disease, in particular for women 50-69 years of age at screening (6, 7, 46). An individual woman attending screening hopes, given she has BC, that she will benefit from early detection of the disease (84, 85). As long as the individual woman believes there is a chance to benefit from attendance, she will probably value screening and accept the risks (84). An individualised approach to a woman diagnosed with BC considers the psychosocial consequences of false-positive screening mammography as less important in relation to the life that might be saved (86). On the other hand, the critics of BC screening stress the considerable number of false-positive results and the significance of the psychosocial consequences of false-positive mammographic screening (46). The challenging aspect is that current knowledge about the psychosocial consequences of false-positive results in a context of population-based mammographic screening is still unclear, particularly with regard to long-term effects. The majority of previous studies in the field were conducted in the 1990s at the initiation of screening programmes (16). Relatively little is known about the psychosocial consequences of false-positive screening mammography from a woman's point of view. Furthermore, it is important to bear in mind that mammographic screening cannot be performed without risk, and the advantages and disadvantages of screening must be taken into account from the perspective of the population as well as of the individual person (2). These two perspectives cannot be separated (9). That is, knowledge about the prevalence, magnitude, and longitudinal development of the psychosocial consequences of false-positive mammographic screening among women provides clues to address *the extent* of the side-effects of screening. The extent of such effects might be further discussed in relation to BC mortality reduction

in the population. To discuss *the relevance* of psychosocial consequences of false-positive screening mammography, the perspective of the individual woman also needs to be explored. Even though mammographic screening is a population-based programme, it is provided to individuals with their own cultural values, beliefs, and perceptions of screening (49).

Furthermore, it has been postulated that there is a difference from an ethical point of view between providing diagnostic examinations to an individual invited to screening and to someone who consults the health care service due to symptoms (87). Delivery of screening raises the issue of responsibility for the side-effects of the programme (32). In addition, the individual should be informed that the screening service involves risks (2, 27, 33, 52). From this perspective, it appears crucial to monitor the potential psychosocial consequences of false-positive mammographic screening in an ongoing population-based screening programme. Gained knowledge might provide a basis for informed decision-making (88). In addition, providers of screening might develop interventions minimizing the risks among women who do not necessarily benefit from the programme.

Despite the numerous studies conducted in the field, there are still many questions that remain unanswered. Do women experience psychosocial consequences of false-positive screening mammography in the long-term? What are the characteristics of these consequences? How do women cope with their experiences? What are the predictors of the psychosocial consequences of false-positive screening mammography? Are we able to identify women at risk of developing psychosocial consequences following false-positive BC screening, and support those with compromised ability to overcome such consequences? What intervention strategies are likely to have the ability to reduce the consequences and their impact? Do we understand the woman's perception of false-positive screening mammography?

As these questions still need to be addressed, the current thesis is intended to deliver some answers.

Aim

The overall aim of this thesis was to validate measures of and study the psychosocial consequences of false-positive mammography among women in a Swedish breast cancer screening programme, and to explore how women cope with such a situation.

Specific aims

- to evaluate the content validity and other psychometric properties of the Swedish version of the Consequences of Screening – Breast Cancer questionnaire (Papers I and II)
- to investigate the prevalence, longitudinal development, and predictors of the psychosocial consequences of false-positive mammographic screening (Paper III)
- to elucidate women's perceived psychosocial consequences of experiencing false-positive screening mammography and to explore how they cope with the situation (Paper IV)

"There are no shortcuts to any place worth going"

Beverly Sills

Methods

Design

The study design applied in this thesis was chosen based upon the pragmatic world-view assumption of the nature of knowledge (89). In this perspective, when a research problem is identified, multiple methods are equally chosen to address the problem in different ways with the purpose of gaining an enhanced understanding of the problem. Thus, within the pragmatic paradigm the researcher does not need to be loyal to one research method, but can incorporate both quantitative and qualitative methods with regard to data collection, analysis, and interpretation of the results.

A flowchart of the studies is provided in Table 1. A mixed qualitative and quantitative approach was chosen to explore the construct of psychosocial consequences of false-positive screening mammography in a Swedish context, and to investigate whether the COS-BC items, together with ten items from the COS-LC that were not considered LC-specific, comprehensibly and sufficiently operationalized the construct (Paper I). A cross-sectional and test-retest study (Paper II) was conducted to test the psychometric properties of the questionnaire proposed in Paper I. The COS-BC scales with good validity and reliability were applied in a cohort study (Paper III) to investigate the prevalence and longitudinal development of the psychosocial consequences of false-positive mammographic screening and to examine predictors for such consequences. In the interest of obtaining the individual's viewpoint on the research question, women's perceived psychosocial consequences of experiencing false-positive screening mammography and their ways of coping with the situation were explored by means of a qualitative method design (Paper IV).

Table 1: The study design

	Paper I	Paper II	Paper III	Paper IV
Objectives	Content validity of the COS-BC and ten items from the COS-LC Understandability of items and response categories	Internal and external construct validity, and reliability of the COS-BC	Prevalence, longitudinal development, and predictors of the psychosocial consequences of false-positive mammographic screening	Perceived psychosocial consequences of false-positive screening mammography and coping with the situation
Design	Mixed methods study	Psychometric cross-sectional and test-retest study	Cohort study	Qualitative study
Sample	26 women with false-positive screening mammography	640 women with false-positive screening mammography 802 matched women with negative screening mammography	399 women with false-positive screening mammography (subgroup enrolled in Study II) 449 matched women with negative screening mammography (subgroup enrolled in Study II)	13 women with false-positive screening mammography who experienced psychosocial consequences as indicated by the COS-BC (subsample from Study II)
Data	Individual interviews Items relevance ratings	Swedish COS-BC 1 and COS-BC 2 NHP	Swedish COS-BC 1 and COS-BC 2 NHP Additional questions Register data	Individual interviews
Analysis	Content analysis CVI	Rasch and classical test theory analyses: internal validity, external validity, reliability	Bivariate hypotheses testing Logistic regression	Inductive content analysis

CVI, Content Validity Index; NHP, Nottingham Health Profile; COS-BC 1 and COS-BC 2, Consequences of Screening – Breast Cancer part 1 and 2, respectively; COS-LC, Consequences of Screening – Lung Cancer

Context and sample

The study sample is presented in Figure 1. The sample consisted of Swedish-speaking women with false-positive or negative mammography attending BC screening at a facility in Malmö, Sweden. The facility provides mammographic screening and diagnostic work-up as well as clinical breast imaging. Women invited to screening were registered in the municipality of Malmö, Trelleborg and Vellinge In 2008-2011, 86193 women (50-74 years of age and since 2009 from the age of 40) were screened at the facility, and the recall rate was on average 3.3%.

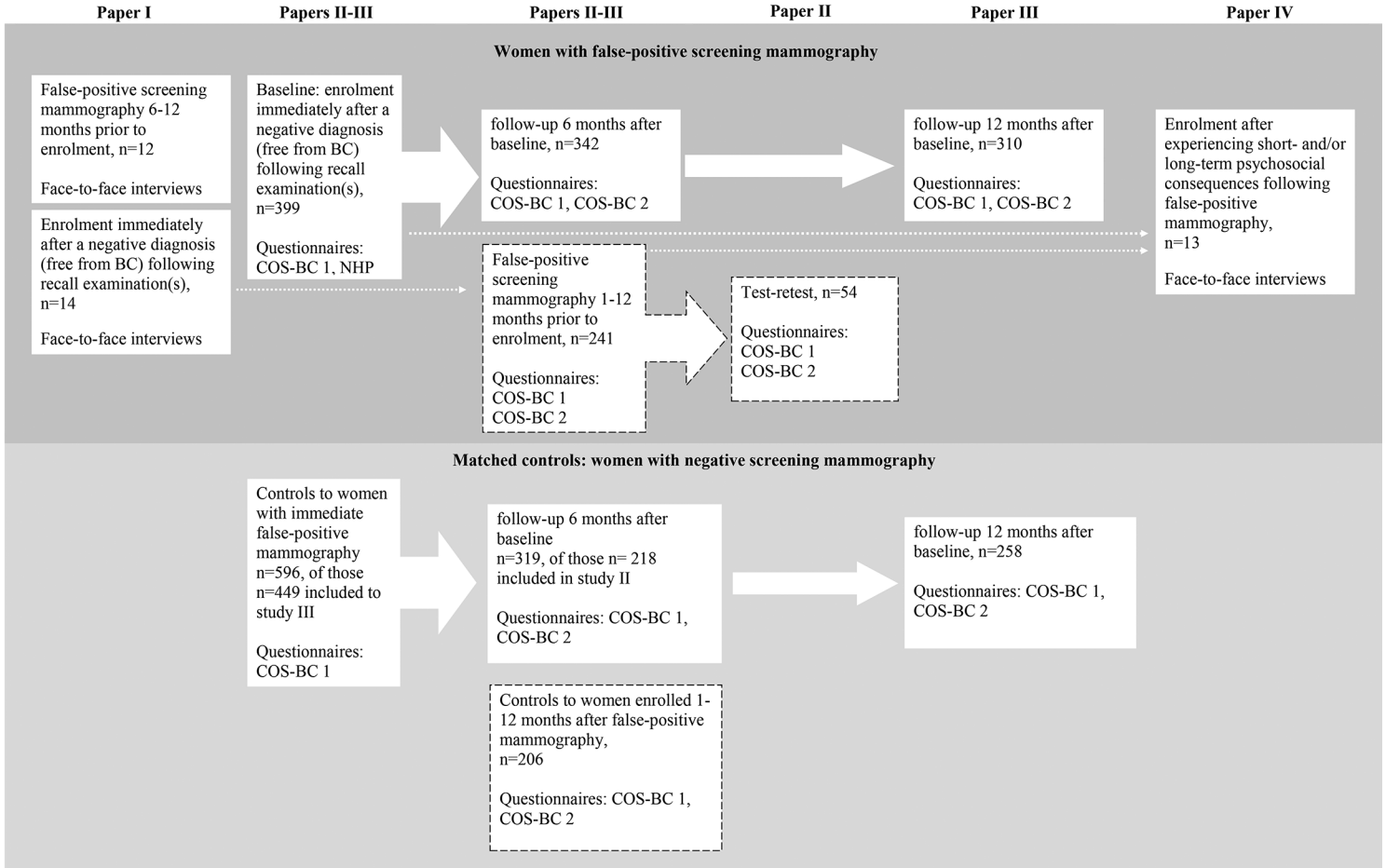
Women with false-positive screening mammography

This group consisted of women with an abnormal screening mammogram who after recall mammography and ultrasound and other supplementary breast examination(s) (fine needle/core biopsy, follow-up by the surgical breast clinic, and scheduled for early recall to subsequent mammography) were told that BC had not been found (referred to as false-positive screening mammography). Women who had discovered a potential abnormality by means of breast self-examination prior to mammographic screening (routinely recalled) were excluded. Altogether, during the study period from September 2008 to June 2011, 987 women fulfilled the study criteria and were asked to participate, of which 658 (40-75 years of age) agreed to an interview and/or responded to the study questionnaires.

Women with negative screening mammography

This group consisted of women with negative (no BC found) screening mammography who were matched with women with false-positive screening mammography according to age (+/- 5 years) and time of screening mammography. A total of 802 women (40-76 years of age) out of 1428 were enrolled.

Figure 1: Sample of the study



COS-BC 1, Consequences of Screening – Breast Cancer, part 1; COS-BC 2, Consequences of Screening – Breast Cancer, part 2; NHP, Nottingham Health Profile; BC, Breast cancer; Sample presented in the boxes with broken lines was not included in Paper III

Questionnaires and other quantitative data

The COS-BC (Paper I-III)

The Danish COS-BC

Following the development process of the Danish COS-BC (68) the questionnaire was psychometrically tested by means of the RM and CTT among women with abnormal, false-positive, and negative screening mammography (17, 69).

RM analysis provided evidence for six scales in part 1 of the COS-BC; *Sense of dejection, Anxiety, Behavioural, Sleep, Breast examination,* and *Sexuality*. Seven items in the questionnaire demonstrated model misfit; however, two of them were retained as single items due to content validity (68). Another item showed signs of DIF, thus was excluded from the *Sense of dejection* scale but retained as a single item. Scores from the scales correlated with domains of a generic health status questionnaire, the Nottingham health profile (NHP), in a predicted pattern. Two items from the sleep section of the NHP were proposed for inclusion in the COS-BC in order to provide better coverage of this construct. Scores from groups of people hypothesized to differ, women with abnormal and negative screening mammography, did so for all scales. Cronbach's α reliability for all scales ranged between 0.71 and 0.92.

Analyses of part 2 of the COS-BC were performed in a similar manner. RM analysis generated four scales; *Existential values, Social relations, Relaxed/calm*, and *Anxiety/ reassurance about BC*. For all scales there were observed differences between groups of women hypothesised to differ. Cronbach's α reliability ranged between 0.81 and 0.92.

The final Danish COS-BC comprises 30 (part 1) and 13 (part 2) items representing the psychosocial consequences of abnormal and false-positive mammographic screening in the short and long term, respectively (Figure 2). Responses (not at all, a bit, quite a bit, a lot; scored 0-3) to part 1 items (except for *sick leave*) are summed for each scale, where higher scores denote more negative consequences. In part 2, responses (much less, less, the same as before, more, much more) are first re-coded into three categories (the same as before remains unchanged, less and more become less/more, and much less and much more become much less/much more) scored 0-2, and then summed for each scale; higher scores represent a higher degree of change (regardless of direction) of psychosocial consequences (17, 70).

The Danish COS-LC

In the development process of the questionnaire targeting the consequences of LC screening, 10 items not considered LC-specific were proposed as potential items representing long-term consequences of cancer screening in general (Figure 2) (70). RM analysis of the items demonstrated model fit for nine items of two scales; *Impulsivity* and *Empathy*. The Cronbach's α for the scales was 0.88 and 0.69, respectively. The remaining item, representing experiences of being relieved, did not fit the hypothesised construct of feeling relaxed and calm. Therefore, it was proposed the item be considered as a single item (70). The response categories to the items were the same as for the items in the COS-BC part 2 and thus it was recommended that they should be rescored accordingly.

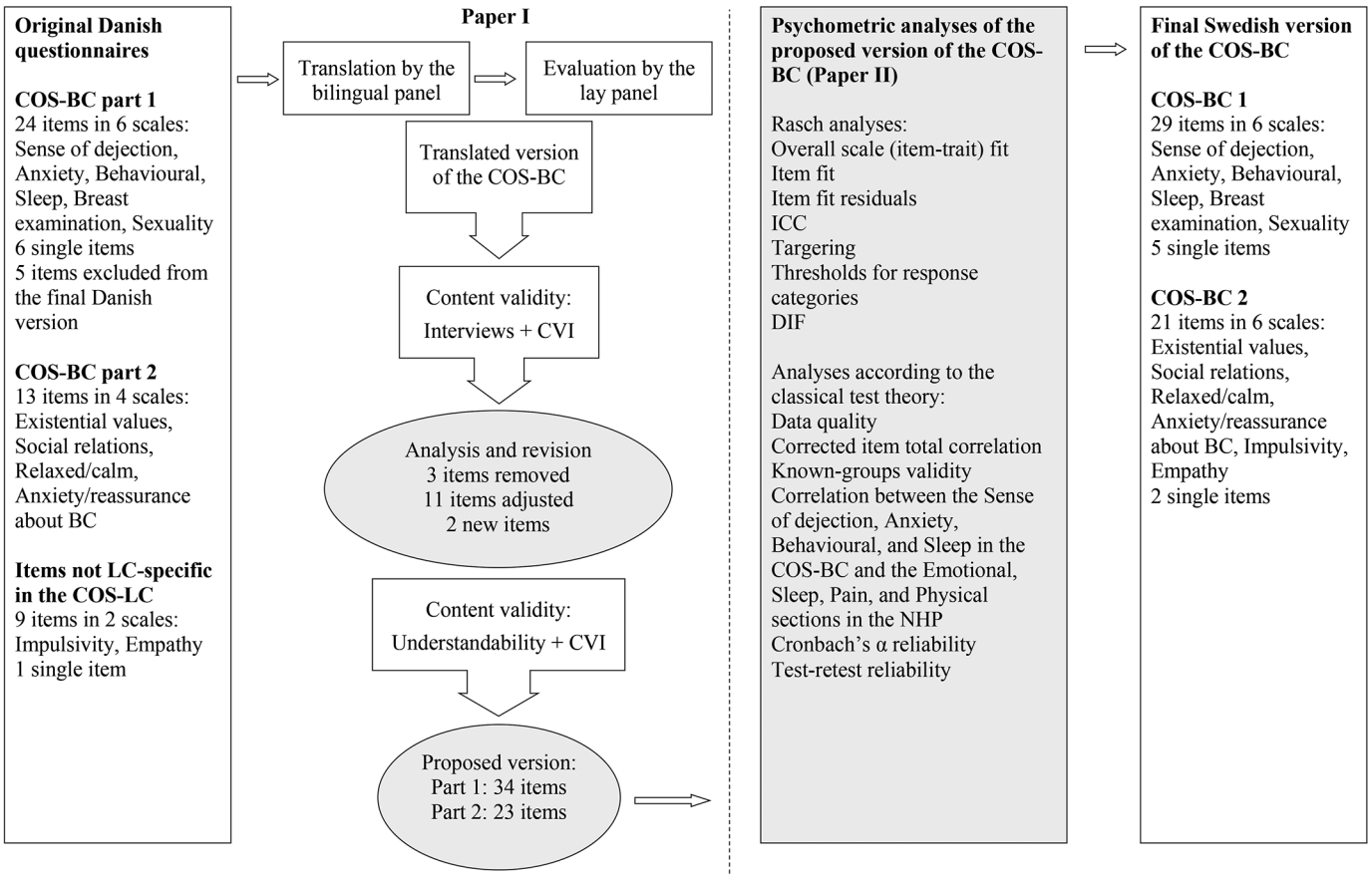
Translation

In addition to the final Danish COS-BC, items excluded from the original version due to suboptimal content validity and lack of fit to the Rasch model (17) were translated into Swedish and tested in Paper I (Figure 2). This was done in order to investigate whether they were also causes of concern among women in Sweden. As such, they served as an additional test of content validity in a Swedish setting. Furthermore, the ten potential general cancer screening items developed in an LC screening setting (70) and related to long-term consequences of cancer screening were included. In total, the COS-BC put forward for translation comprised 35 (part 1) and 23 (part 2) items related to short-and long-term psychosocial consequences, respectively.

Translation was conducted according to the dual-panel (DP) methodology (90). In this method, one bilingual and one lay panel is used with the aim of developing a translation that is expressed in lay language and which is conceptually, semantically, and operationally equivalent to the original questionnaire.

Using the DP approach, a panel of three Swedish bilingual women (41-58 years of age) produced a first draft of the Swedish COS-BC version (Figure 2). This translation was forwarded to a lay panel consisting of four Swedish women (46-69 years of age) for evaluation regarding wording, interpretation, and acceptability. Only the first panel had access to the source version of the questionnaire. A representative from the developers of the original questionnaires was present with both panels.

Figure 2: Adaptation and testing of the Swedish COS-BC



CVI, Content validity index; ICC, Item characteristic curve; DIF, Differential item functioning; COS-BC 1 and COS-BC 2, Consequences of Screening – Breast Cancer part 1 and 2, respectively; COS-LC, Consequences of Screening – Lung Cancer; NHP, Nottingham Health Profile

The Swedish COS-BC

The translated version of the COS-BC, including ten items from the COS-LC, was tested regarding content validity and understandability in Paper I (Figure 2). The validation process resulted in a questionnaire version consisting of 34 (part 1, also referred to as the COS-BC 1) and 23 items (part 2, also referred to as the COS-BC 2) that was tested psychometrically in Paper II. The final Swedish COS-BC 1 and COS-BC 2 questionnaires (Table 2) are presented in Appendix 1 and 2, respectively. Details of content validity, understandability, and other psychometric analyses of the Swedish questionnaire version are presented in the results section in the framework of this thesis. The COS-BC 1 was used in Paper III for the assessment of short- and long-term psychosocial consequences of false-positive screening mammography, whereas the COS-BC 2 was applied for targeting long-term consequences not represented by the COS-BC 1.

The Nottingham Health Profile (Paper II-III)

The Nottingham Health Profile (NHP) (Appendix 3) is a generic health status questionnaire which encompasses 38 items representing Sleep, Energy, Emotional reactions, Social isolation, Physical mobility and Pain (91) Each section is scored from 0 to 100 (100=worse health). The NHP has been adapted and validated for use in Sweden (91). The questionnaire was applied in Paper II to assess the external construct validity of the COS-BC 1, as was previously done with the original Danish questionnaire (69). In Paper III, the questionnaire was used to assess the health status of women who had experienced false-positive screening mammography.

Socio-demographic variables and other potential predictors (Paper III)

Variables considered potential predictors were identified based on previous studies (20-22, 60), and collected for women who had experienced false-positive screening mammography using registries and self-reported questionnaire items (Table 3). Socio-demographic register data from the end of years 2008 and 2009 was linked from Statistics Sweden for women enrolled in 2009 and 2010, respectively. Data from the Swedish Cancer Registry on malignant cancer diagnoses prior to study entry was also included. In addition, screening history and type of current supplementary breast examination(s) were collected from clinical records. At baseline, women answered questions addressing perceived vulnerability to BC, attitudes towards the health care system, information about BC and screening, and communication at the screening and diagnostic work-up (Appendix 4). These questions were developed in the assessment of attendance patterns in Swedish mammographic screening (92).

Socio-demographic and cancer register data were also collected for women with negative screening mammography in order to control for potential group differences at study entry. To perform a non-participation analysis, aggregated data from Statistics Sweden was used for women with false-positive screening mammography who had declined participation.

Table 2: The final Swedish COS-BC

COS-BC 1	COS-BC 2
Scale **Item**	**Scale** **Item**
Sense of dejection ➢ worried ➢ uneasy ➢ time passed slowly ➢ sad ➢ depressed ➢ unable to cope	Existential values ➢ value life ➢ broader aspects of life ➢ enjoyment of life ➢ well-being ➢ awareness of life ➢ thoughts about future
Anxiety ➢ worried about the future ➢ scared ➢ nervous ➢ upset ➢ restless ➢ terrified	Social relations ➢ family ➢ other people ➢ friends Relaxed/calm ➢ relaxed ➢ calm
Behavioural ➢ busy to take mind off things ➢ hard to concentrate ➢ quieter that normal ➢ irritable ➢ withdrawn into myself ➢ difficulty dealing spare time ➢ change in appetite ➢ difficulty dealing things ➢ difficulty dealing work	Anxiety/reassurance about BC[b] ➢ not BC ➢ anxiety about BC Impulsivity ➢ desire to venture into sth new ➢ desire to venture into sth risky ➢ lived life to the full ➢ being impulsive ➢ overstep one's bounds
Sleep ➢ slept badly ➢ taken long time to fall asleep ➢ wake up early in the morning ➢ lie awake most of night	Empathy ➢ understanding people's problems ➢ ability to listen ➢ responsibility for family
Breast examination ➢ examined my breasts ➢ examined my breasts in the mirror	Single items ➢ relieved ➢ energy
Sexuality ➢ less interest in sex ➢ not breast caressed	
Single items ➢ tired ➢ symptoms from the breasts ➢ thoughts about bodily changes ➢ less attractive ➢ sick leave[a]	

[a] included in the Danish COS-BC to address psycho-economic consequences; [b] referred to as Breast cancer scale in Paper I

Response categories to the COS-BC 1: not at all, a bit, quite a bit, a lot; scored 0-3; response categories to the COS-BC 2 after re-coding: the same as before=0, less/more=1, much less/much more=2

Table 3: Socio-demographic variables and other potential predictors

Variables	Data level	Applied in the analysis	Variables	Data level	Applied in the analysis
Demographic variables			**Previous cancer diagnosis**		
age	interval	age in years	breast cancer	nominal	not included due to low occurrence
living area	nominal	rural=0, urban=1	other malignant cancer	nominal	not included due to low occurrence
country of origin	nominal	Sweden=0, other=1	**Screening history**	nominal	first screening mammography yes=1, no=0
				nominal	experienced previous recall, yes=1, no=0
Socio-economic variables			**Diagnostic follow-up**	nominal	fine/core needle biopsy, yes=1, no=0
level of education	nominal	compulsory school[a]		nominal	breast surgery clinic, yes=1, no=0
		secondary school[a]		nominal	early recall, yes=1, no=0
		higher education=reference	**Perceived vulnerability**		
employment	nominal	employed=0, not employed=1	worry; three items	ordinal, score 0-3	summed score between 0 and 9[c]
main source of income	nominal	gainfully employed=reference	susceptibility; two items	ordinal, score 0-3	total score between 0 and 6[c]
		retired/other source of income[a]	lack of social support, three items	ordinal, score 0-3	total score between 0 and 9[c]
		no income[a]	a close one having BC, either now or previously	nominal	yes=1, no=0
family income (hundred SEK)	interval	categorised according to quartiles[b]			
accommodation	nominal	owning a house=reference	**Attitudes, information and communication**		
		owning an apartment[a]	not trusting the healthcare system	ordinal, score 0-3	not included due to low occurrence
		renting[a]	usually do not read a news article about BC	ordinal, score 0-3	dichotomised, dissatisfied =1 versus not =0[d]
marital status	nominal	married/cohabitant=0	dissatisfied with information at BC screening	ordinal, score 0-3	dichotomised, dissatisfied =1 versus not =0[d]
		divorced/separated/widowed/single=1	dissatisfied with information at recall	ordinal, score 0-3	dichotomised, dissatisfied =1 versus not =0[d]
type of family	nominal	children < 18 years old living at home[a]	dissatisfied with own knowledge about BC	ordinal, score 0-3	dichotomised, dissatisfied =1 versus not =0[d]
		children ≥ 18 years old living at home[a]	dissatisfied with own knowledge about BC screening	ordinal, score 0-3	dichotomised, dissatisfied =1 versus not =0[d]
		no children living at home=reference[a]			

[a] Dummy coded with either category coded as 1 and the reference category coded as 0; [b] Quartiles (hundred SEK): ≤2433, 2434-3933, 3934-5733, ≥5734; [c] Higher score represents higher degree of BC worry, BC susceptibility, lack of social support, respectively; [d] Satisfied =0 (score 0-1), dissatisfied =1 (score 2-3)

Data collection

In Paper I, 21 women who had experienced false-positive screening mammography were interviewed individually to assess the content validity of the translation of the COS-BC, and to investigate whether items expressing long-term consequences of false-positive LC screening results are relevant in a BC screening context. The women were allowed to speak freely, yet were encouraged to maintain focus on their short- and long-term experiences of false-positive screening mammography. They also completed the draft version of the COS-BC together with ten items from the COS-LC. This was followed by an interview to investigate the women's understanding of the items and their opinion about the instructions, response categories, and lay-out of the questionnaire (Figure 2). The two abovementioned parts of the interview were conducted for the purpose of investigating whether items provided coverage of the intended construct and whether the scales, their items, and response categories appeared understandable. Finally, the women were asked to assess the relevance of each item in relation to the intended construct. Each item was rated as not relevant, somewhat relevant, quite relevant, or highly relevant (79). To evaluate the COS-BC revision, which resulted from the data collected in the interviews, five women with false-positive screening mammography were interviewed to obtain their views on the understandability and relevance of the revised or new items (Figure 2).

In Papers II-III, women with false-positive screening mammography (n=399) were approached immediately after they had been informed that no malignancy had been found (referred to as baseline, Figure 1). They were then asked to complete the COS-BC 1 (administration T1) according to their experiences prior to the final diagnosis (free from BC). At the same time, they also responded to the NHP. These women were followed-up by using both parts of the COS-BC regarding their experiences six (administration T2) and 12 months later. Women scheduled for early recall (subsequent mammography) responded to the questionnaires immediately after their appointments; that was at approximately six and 12 months after baseline. Women who had experienced false-positive screening mammography 1-12 months prior to enrolment (n=241) responded to the COS-BC 1 and COS-BC 2 according to their experiences at the time they received the questionnaires (administration T2). In addition, 54 of these women completed both parts of the COS-BC two weeks later (Test-retest, Figure 1). These women were also asked to respond to a transition question which had five ordered response categories (much better, better, the same as before, worse, much worse) regarding their well-being compared to when they completed the COS-BC the first time (administration T2). Women with negative screening mammography (n=802) received the COS-BC questionnaire(s) on the same occasions as women with false-positive screening mammography (Figure 1) and were instructed to respond accordingly to their experiences at the time they received the questionnaire(s).

In Paper IV, face-to-face interviews with 13 recalled women were conducted 3-11 months after their final diagnosis of no evidence of BC (Figure 1). The recruited women had participated in study II and had shown some degree of short- and/or long-term psychosocial consequences following false-positive screening mammography, as indicated by a COS-BC score >0 (0 = no reported psychosocial consequences). Purposeful selection was used to obtain variation in sampling regarding age (40-74 years), residential area (urban, rural), ethnic background (native, immigrant), and experience of any kind of additional breast examination (mammography, ultrasound, fine needle/core biopsy, follow-up by the surgical breast clinic and/or early additional mammography). A question concerning the first thoughts that came to mind when recalling the preceding false-positive screening mammography opened the interview. Each woman was then allowed to speak freely about her experiences and emotions of false-positive screening mammography. Probing questions concerning the woman's thoughts in dealing with her experiences and emotions, as well as what she was doing to manage these experiences and emotions, were asked. The interview was closed by asking about their attitudes towards BC screening. All interviews were audio recorded (ranging from 23 to 74 minutes) and transcribed verbatim, including notes indicating laughter, crying, sighs, etc.

Ethical considerations

The study was conducted according to the Declaration of Helsinki (93). All women received verbal and/or written information about the study, and signed informed consent was a prerequisite for entering the study. The women might have experienced the questionnaires and interviews as an intrusion upon their privacy. However, they might also have felt that someone was interested in their situation, and it might have provided an opportunity to reflect upon their experiences. Because of this, along with the confidentiality, the informed consent and the right to terminate participation at any time, the advantages of the study appeared to outweigh the disadvantages. In addition, the women might have felt that their contribution would benefit both themselves as well as other women in the future.

Furthermore, owing to ethical considerations all the women were invited to participate in the study after they had been informed that BC had not been found. Micro data from Statistics in Sweden and the Swedish Cancer Registry was received in an anonymous mode. In addition, careful consideration was given to the decision about linking register data for women who declined participation. That is, as the research method applied in Paper III has its limitation in sample selection, a non-participation analysis is recommended to control for potential selection bias. However, such a procedure challenges women's choice of non-participation. On the other hand, evidence provision (Paper III) of no selection bias straightens the

study results and their implications for women experiencing false-positive screening mammography. Consequently, it was decided to use register data regarding socio-demographic variables but only in an aggregated mode. Screening history and type of current additional breast examinations were not collected being aware of that such a decision would limit the study findings, but respect the privacy of women's medical journals.

The study was approved by the Regional Ethical Review Board (no. 373/2008), Statistics Sweden (no. 218299/871210-1 and 220843/8721813), and the Swedish National Board of Health and Welfare (no. 59062/2012).

Analyses

Content validity of the Swedish COS-BC (Paper I)

Content validity evaluation started with the analysis of the interviews. A qualitative content analysis was applied to analyse the transcripts (94). The aim of the analysis was to assess the extent to which women's experiences were congruent with the consequences operationalized in both parts of the COS-BC, and to explore whether any additional experiences related to the false-positive mammographic screening, but not included in the questionnaires became apparent. Thus, statements in the text referring to the same content and corresponding to the experiences of having false-positive screening mammography were identified. The emerging meaning units were labelled by a short phrase or a word; a code. Later on, it was verified whether the items in the COS-BC were represented by the codes and, conversely, whether there were any codes not covered by the scales. This analysis was conducted by the first author of Paper I, and was followed by a discussion between all authors. The transcripts were also re-checked to confirm that no experiences related to the intended construct remained unrevealed. All comments on questionnaire instructions, items, response categories, and lay-out were considered in order to identify aspects that were misunderstood or otherwise in need of revision (18, 78).

The CVI, a quantitative approach to assessing the content validity of individual items (I-CVI), was calculated based on the respondents' relevance ratings (79, 80). I-CVI is expressed as the proportion of respondents considering each item as quite or highly relevant. A minimum I-CVI value of 0.78 or higher was considered acceptable (80). To adjust each I-CVI for chance agreement on relevance, a modified kappa statistic (k*) was computed; values between 0.40-0.59 were considered fair, 0.60-0.74 good, and >0.74 excellent (80). The content validity index of the COS-BC 1 and COS-BC 2 was also calculated for each questionnaire (S-CVI). The

S-CVI represented the average I-CVI across items. A minimum S-CVI value of 0.90 was considered acceptable (80).

Lastly, the content validity of the COS-BC was evaluated by merging the results from the qualitative and quantitative analyses, and also taking the results from the RM of the original Danish questionnaires into account (69, 70), before a decision on adjustment, inclusion, or rejection of an item was taken (Figure 2). For example, (i) if an item was poorly covered in the interviews but showed acceptable I-CVI or good to excellent k*, and fitted the RM in Denmark, the item was retained; (ii) in cases of problematic understandability of an item covered in the interviews, the item was modified to improve its conceptual and semantic equivalence; (iii) ambiguousness or insufficient content validity among women in Sweden and Denmark, and the RM misfit in a Danish setting resulted in the item being omitted. In addition, records from the translation process were re-checked as an aid to the evaluation of the results. For example, if an item appeared problematic and alternative translations had been suggested by the panels, the alternative was considered instead.

Psychometric testing of the Swedish COS-BC (Paper II)

Rasch analyses

The internal construct validity of the Swedish COS-BC scales was tested with the RM for ordered response categories (73, 95). The following approach was applied; (i) all scales were analysed according to the original Danish questionnaire structure (17), (ii) single items, including the COS-LC items, were introduced to scales, based on conceptual considerations (Paper I) (17, 70) followed by repeated analyses of the scales, (iii) any misfit to the RM was explored before the final Swedish version of the scales was suggested.

The study of RM fit is a process where no individual fit statistic is either necessary or sufficient (73, 75, 95). Therefore, a range of analyses were conducted and interpreted interactively. In this perspective, no analysis is to be considered a more important aspect of RM fit than another. The final interpretation is a collective one of a variety of indices, particularly since no data ever fit any model, given sufficient observations (73, 75, 76).

Overall item-trait and individual item chi-square statistics were derived for all scales. These tests are based on differences between the observed item responses and model expectations. Non-significant statistics support model fit (73, 95). Standardized item fit residuals (reflecting discrepancies between observed item responses and model expectations) were inspected. Values between +/- 2.5 are generally considered acceptable (95); values >2.5 suggest multidimensionality, and values <-2.5 indicate

item redundancy (referred to as local dependency). To further examine the presence of local dependency, the correlations between item residuals were examined within scales with fit residual values <-2.5; inter-item residual correlations >0.3 are considered indicative of local dependency (73). Model fit was also assessed graphically by inspecting the accordance between data and the item characteristic curve (ICC). The ICC represents model expected item responses at various locations on the measured continuum. Thus, a reasonable accordance between model expectations (the ICC) and responses from groups of people at similar locations on the measured continuum (class intervals) supports model fit (73, 95).

Targeting, that is the relative distribution of estimated person and item locations on the measured continuum, was examined to explore whether the scales represented the levels of the measured construct experienced by the sample and vice versa. In a well-targeted scale the mean sample location should approximate the mean item location (zero); poor targeting compromises measurement precision (73, 95).

The empirical functioning of response categories can be assessed as an additional aspect of model fit (73, 95). That is, ordered response categories are supposed to correspond to increasing or decreasing levels of the measured construct. This assumption was tested by examining the thresholds between adjacent response categories (the locations on the continuum with a 50% probability of responding in either of two adjacent response categories). Disordered thresholds suggest that response categories are not working as intended.

An additional aspect of fit is differential item functioning (DIF), which means that responses to an item are not supposed to differ for subgroups of people in the same class interval (73, 95). A significant systematic response discrepancy between groups across class intervals indicates uniform DIF, whereas interacting discrepancies (groups x class intervals) indicate non-uniform DIF (95). DIF was tested by two-way ANOVA of the residuals across class intervals for diagnosis (false-positive versus negative mammography) and age (younger versus older as defined by the median, <55 versus ≥55 years old), followed by Bonferroni correction (the critical P-value, 0.05 divided by the number of hypothesis tests). In the presence of uniform DIF, this was adjusted for by splitting the affected item into two subgroup specific items (77). Potential DIF-induced group-level bias was explored by estimating the effect sizes of the differences between person locations estimated from adjusted and non-adjusted total scores (differences in mean person locations/pooled SD) (96). Effect sizes of 0.2, 0.5, and 0.8 were regarded as small, moderate, and large, respectively (97).

DIF by administration time was also tested and no evidence for DIF was found. Data was therefore pooled and a sample size of 200 responders per scale was used. That is, all available responses were included in the analysis of each scale, but the

sample size was algebraically adjusted to n=200 in the calculation of P-values (95). Algebraic adjustment maintains unaltered the full sample and all other derived statistics, such as locations, residuals, but the sample size is adjusted in the calculation of P-values. This was done to avoid statistical misfit as a result of a too large sample size and to avoid interpretation problems due to unequal statistical power across scales. This is of relevance because the effective n is not only related to the number of respondents but also to targeting, since people scoring at the minimum/maximum is not included in the RM estimations (75, 95). The number of 200 was set according to a recommendation to use 10 to 30 responders for each item and response category threshold in a scale (98), the observation that samples >200 often flag minor misfit as significant (99), and general estimations suggesting that stable calibration within 0.5 logits requires ≥150 subjects (100). However, analyses based on unaltered sample sizes were also conducted, and potential DIF-induced group-level bias was explored as described above.

Classical test theory analyses

Data quality and internal construct validity of the final Swedish COS-BC scales were also analysed in terms of CTT as follows (73, 101); (i) data completeness (<5% missing item responses were considered acceptable), (ii) percentage of the sample with computable total scores for each scale, (iii) floor/ceiling effects (i.e. proportions of the sample at the minimum/maximum score), and (iv) corrected item-total correlations (should be at least 0.30) as a test of internal construct validity. As CTT assumes complete data, women with missing item responses within a scale were not included in the analysis of that scale.

External construct validity of the final COS-BC scales was assessed by Spearman correlations and known-groups approaches (Mann-Whitney test) (18, 102). Based on observations from the original Danish questionnaires it was hypothesized that T1 scores from the *Sense of dejection, Anxiety, Behavioural,* and *Sleep* scales would show stronger correlations with the NHP Emotional reactions and Sleep sections than with the Pain and Physical mobility sections (17, 69). It was further anticipated that women who had experienced false-positive screening mammography would score higher on COS-BC 1 and COS-BC 2 scales than those with negative mammography at T1 and T2, respectively.

Internal consistency reliability of the final versions of the COS-BC scales was evaluated by Cronbach's α (18). Test-retest reliability was assessed by the one-way ANOVA intraclass correlation coefficient (82) between the T2 and Test-retest questionnaire administrations. Responses to the transition question were used to control for potential change in well-being between questionnaire administrations (103). Reliability should preferably be around 0.7 or above (73).

Data was analysed using SPSS version 17.0 and RUMM2030. Two-tailed P-values ≤0.05 following Bonferroni correction (due to the large number of hypothesis tests conducted in the analyses) were considered significant.

Statistical analysis (Paper III and the framework)

Paper III was based on COS-BC scales with supported psychometric properties according to analyses in Paper II. In addition, results from all scales in the COS-BC 1 and COS-BC 2 are presented in the framework of this thesis. Total scores were dichotomised for each scale; women scoring >0 were considered experiencing psychosocial consequences and those scoring 0 were not considered experiencing any consequences. No imputation of the scores was applied. Between group differences addressing the prevalence of psychosocial consequences were analysed by chi-squared tests. Changes of the prevalence over time for women with false-positive screening mammography were investigated using the McNemar's test. Logistic regression analysis was conducted to investigate the likelihood of each outcome at all assessments (baseline, six and 12 months later) depending on the diagnosis (women with false-positive versus negative screening mammography) controlled for age.

To identify potential socio-demographic and other predictors of long-term psychosocial consequences of false-positive screening mammography the following approach was applied in Paper III; (i) a univariate logistic regression of each variable was conducted to predict each outcome (COS-BC scales with supported psychometric properties according to analyses in Paper II) at the six-month follow-up among women with false-positive screening mammography, (ii) variables demonstrating P-values <0.25 were included in a multivariate forward stepwise likelihood ratio logistic regression in order to define a final prediction model of each outcome. The threshold for the P-value was chosen in order to avoid that a variable alone might associate weakly with the outcome, but become significant when analysed together with other potentially important predictors (104). The multivariate regression models were controlled for age and experienced consequences at baseline. For example, the presence or absence of *Sleep* consequences at baseline was used as a covariate (together with age) when modelling potential predictors for experiencing *Sleep* consequences six months later. The same approach was used to investigate predictors of consequences experienced at the 12-month assessment and presented in the framework of this thesis.

The student's t-test or chi-squared test was applied to test for group differences at study entry and potential differences between women with false-positive screening mammography and those who declined participation. Control for bias due to potential systematic drop-out among women with false-positive screening mammo-

graphy was performed as well. That is, the prevalence of psychosocial consequences of false-positive screening mammography captured by each scale at baseline was compared (chi-squared test) between responders and non-responders at the six and 12-month follow-ups. Data were analysed using IBM SPSS statistics, version 20.0. Two-tailed P-values ≤0.05 were considered significant.

Qualitative analysis (Paper IV)

The interview transcripts were analysed according to the inductive approach to content analysis methodology (105) using the analysis procedure described by Graneheim and Lundman (94). The analysis started with a reading of all transcripts, and this was followed by a discussion between two authors of Paper IV. The process resulted in a range of tentative codes (a word or a short word sequence) mirroring the emerging content. Following this, two interviews were analysed independently by the two authors. Statements (meaning units) related to the same manifest content were identified and labelled by the tentative codes. New codes were applied if needed. Subsequently, underlying meanings of the units began to emerge, and tentative categories as well as a theme were discussed. The first author of the paper then analysed the remaining interviews in the same manner. The tentative categories evolved to represent groups of codes sharing the same meaning as the latent content of the text. The theme demonstrating the highest abstraction of the meaning related to coping with the perceived consequences of false-positive screening mammography remained throughout the analysis. After a third discussion between the two authors, two interviews, along with all identified meaning units and the code-category tree, were passed over to the remaining authors of the paper, who agreed with the analysis results.

"All this he knows but will not tell
To those who cannot question well"

Percy Bysshe Shelley

Results

Psychometric properties of the Swedish COS-BC (Papers I-II)

The interviews (Paper I) supported the presence of all domains in the COS-BC 1 and COS-BC 2 except the *Sexuality* scale in part 1. However, the two *Sexuality* items showed excellent I-CVI values (0.86). The contents of three other items in the COS-BC 1 and ten items in COS-BC 2, including six general items developed in a LC screening setting, were not clearly reflected in the interviews. However, most of these items showed good to excellent I-CVIs (≥0.67) (Paper I).

The interviews also revealed experiences that were not covered by the original Danish questionnaire. These included difficulties participating in spare time activities, having thoughts about possible bodily changes, and a need for debriefing. Thus, two items representing the two former aspects were added to the COS-BC 1, followed by an additional evaluation that showed adequate understandability of the items and excellent I-CVIs (1.0). The need for debriefing was not added to the questionnaire since it was considered a coping strategy rather than a consequence, and therefore not a manifestation of the construct. A further revision resulted in the rejection of three items (*keeping things from those who are close to you, headache, taking sleeping tablets*), which had also been excluded from the Danish COS-BC 1 due to misfit to the RM (69). The items were considered either ambiguous to interpret or of suboptimal content validity. Five other misfitting items, of which two were removed from the Danish version (69), exhibited content validity and were therefore retained in COS-BC 1 (*less attractive, tired, symptoms from the breasts, busy to take mind off things*) and COS-BC 2 (*relieved*). Four items in COS-BC 1 were misunderstood and two were considered inappropriately located in the questionnaire. However, content validity was supported for these items and they were therefore retained but reworded or relocated. The additional evaluation of the reworded items showed adequate understandability and excellent I-CVIs (0.8-1.0). The COS-BC 2 remained unchanged except for five items expressed in the past tense. This was seen as a potential cause of erroneous responses by the interviewees and the wording was thus changed to the present tense. Altogether, the Swedish COS-BC 1 and COS-BC 2

proposed in Paper I consisted of 34 and 23 items, respectively. The instructions were found easy to understand and the questionnaire lay-out was considered acceptable. The response categories were considered comprehensive and unambiguous. S-CVI was 0.87 and 0.81 for the COS-BC 1 and COS-BC 2, respectively.

Iterative RM analyses (Paper II) of the Swedish questionnaire version resulted in nine COS-BC scales demonstrating scale (item-trait) and item level RM fit (Table 4) but all twelve scales exhibited poor targeting with relatively large floor effects (8.9-93.3%). Items demonstrating large negative fit residuals (in the *Sense of dejection, Anxiety, Behavioural* scales) were further examined; residuals exhibited varying degrees of local dependency (residual correlations >0.30) in complex patterns that involved most items in these scales. However, inspection of the respective ICCs showed reasonable concordance between observed and expected responses, except for the lowest class interval. Marginally reversed response category thresholds were observed for two items (*quieter than normal, difficulty dealing work*). There was no significant DIF with n set at 200, except for one item in the *Empathy* scale (effect size <0.1; DIF by age). According to full sample analyses, significant uniform DIF appeared by diagnosis (false-positive versus negative screening mammography) for five and two items in the COS-BC 1 and COS-BC 2, respectively, and by age for two items in the COS-BC 1. However, except for the *Sense of dejection* scale (effect size, 0.5; DIF by diagnosis) these did not appear to bias the respective total scores (effect sizes <0.1). Taken together for the final Swedish version (Paper II), five items (*sad, difficulty dealing spare time, wake up early, lie awake most of night, busy to take mind off things*) were added to the *Sense of dejection, Behavioural,* and *Sleep* scales of the Swedish COS-BC 1. One item (*energy*) was removed from the Impulsivity scale in the COS-BC 2, but retained in the questionnaire as a single item owing to god I-CVI (Paper I). Other scales remained unchanged compared to the Danish original. Five items remained as single items (*tired, symptoms from the breast, thoughts about bodily changes, less attractive, relieved*) due to excellent I-CVIs (Paper I).

Missing item responses were randomly distributed (0-3.7%) among all women. Across all data administrations among women with false-positive screening mammography, the total scale scores of the final Swedish version could be computed for >95% of respondents, except for the *Sexuality* scale (83-89%). Corrected item-total correlations of the final Swedish version were ≥0.36. Differences in COS-BC scores between women with false-positive and negative screening mammography, and correlations with NHP scores followed an expected pattern (Paper II). Cronbach's α was ≥0.70 for ten final scales and test-retest reliability was ≥0.75 for six scales (Table 4). Although causes for concern were identified, the psychometric analysis provided support for cross-sectional and longitudinal use of four COS-BC 1 scales (*Sense of dejection, Anxiety, Behavioural,* and *Sleep*) and one COS-BC 2 scale (*Existential values*) for group level assessments of the psychosocial consequences of false-positive mammographic screening (Table 4).

Table 4. Summary of psychometric properties of the final Swedish COS-BC 1 and COS-BC 2

Final Swedish version	No. of items	Scale RM fit[a] P-value	Known-groups validity[b] P-value	Cronbach's α reliability[c]	Test-retest reliability[d]
COS-BC 1					
Sense of dejection	6	0.013	<0.001	0.92 / 0.88	0.75
Anxiety	6	0.083	<0.001	0.92 / 0.90	0.84
Behavioural	9	0.531	<0.001	0.90 / 0.86	0.85
Sleep	4	0.815	<0.001	0.93 / 0.91	0.77
Breast examination	2	**<0.001**	<0.001	0.66 / 0.75	0.68
Sexuality	2	0.841	<0.001	0.86 / 0.90	0.43
COS-BC 2					
Existential values	6	0.464	<0.001	0.87	0.80
Social relations	3	**<0.001**	<0.001	0.88	0.39
Relaxed/calm	2	0.689	0.001	0.70	0.43
Anxiety/reassurance about BC	2	0.514	<0.001	0.53	0.58
Impulsivity	5	0.076	<0.001	0.84	0.55
Empathy	3	**<0.001**	<0.001	0.76	0.90

[a] Chi square test of scale RM fit as assessed with n set at 200 across all scales. Critical P-value after Bonferroni correction was 0.004 and 0.006 for COS-BC 1 and COS-BC 2, respectively; values in bold are significant following Bonferroni correction
[b] Mann-Whitney test between women with false-positive and negative screening mammography at administration T1 for the COS-BC 1, and at administration T2 for the COS-BC 2.
[c] Cronbach's α at questionnaire administrations T1 and T2.
[d] Intraclass correlation coefficient between administration T2 and Test-retest.
COS-BC 1, Consequences of Screening – Breast Cancer, part 1; COS-BC 2, Consequences of Screening – Breast Cancer, part 2; RM, Rasch model; CTT, classical test theory; BC, breast cancer.

Psychosocial consequences of false-positive screening mammography (Paper III and the framework)

Socio-demographic characteristics and the prevalence of previous malignant cancer diagnosis did not differ between women with false-positive (mean [SD, range] age, 53 [9, 40-74]) and negative (mean [SD, range] age, 53 [9, 40-74]) screening mammography. Low total scores in the NHP sections (median, 0) indicated good health among women at study entry. Women with false-positive screening mammography who declined participation did not differ in socio-demographic variables except for being foreign-born and being unemployed, which indicated a slight underrepresentation of these groups (Paper III). There was no significant difference in the prevalence of consequences of false-positive screening mammography at baseline between responders and non-responders to the COS-BC 1 at the six and 12-month follow-ups (P≥0.087), except for *Behavioural* and *Sexuality* consequences (P≤0.038) indicating underrepresentation in the follow-up assessments of those who experienced consequences at baseline. The same drop-out analyses of the COS-BC 2 con-

sequences indicated overrepresentation of *Anxiety/reassurance about BC* and *Empathy* consequences at the 12-month follow-up (P≤0.044).

Sense of dejection, Anxiety, Behavioural, Sleep, and Existential values consequences

For five COS-BC scales with supported psychometric properties, the prevalence of consequences was significantly higher at all assessments (P<0.001) among women with false-positive screening mammography versus the controls (Paper III). The prevalence of the outcomes among women with false-positive screening mammography decreased significantly (P<0.001) between baseline and six months later, but levelled out between the six and 12-month assessments (P≥0.136) (Figure 3). Among those who experienced consequences at the 12-month follow-up, between 58% (n=53) and 77% (n=125) also had experienced such consequences at the six-month follow-up. Experiencing false-positive screening mammography resulted in significantly higher odds ratios (OR) for psychosocial consequences (P<0.001) than for women with a negative diagnosis for all assessments (Table 5). The odds ratios were at least five times higher at baseline (OR, 5.47-15.52) and at least twice as high (OR, 2.75-4.32) at later time points compared to women with negative screening mammography.

Breast examination, Sexuality, Social relations, Relaxed/calm, Anxiety/reassurance about BC, Impulsivity, and Empathy consequences

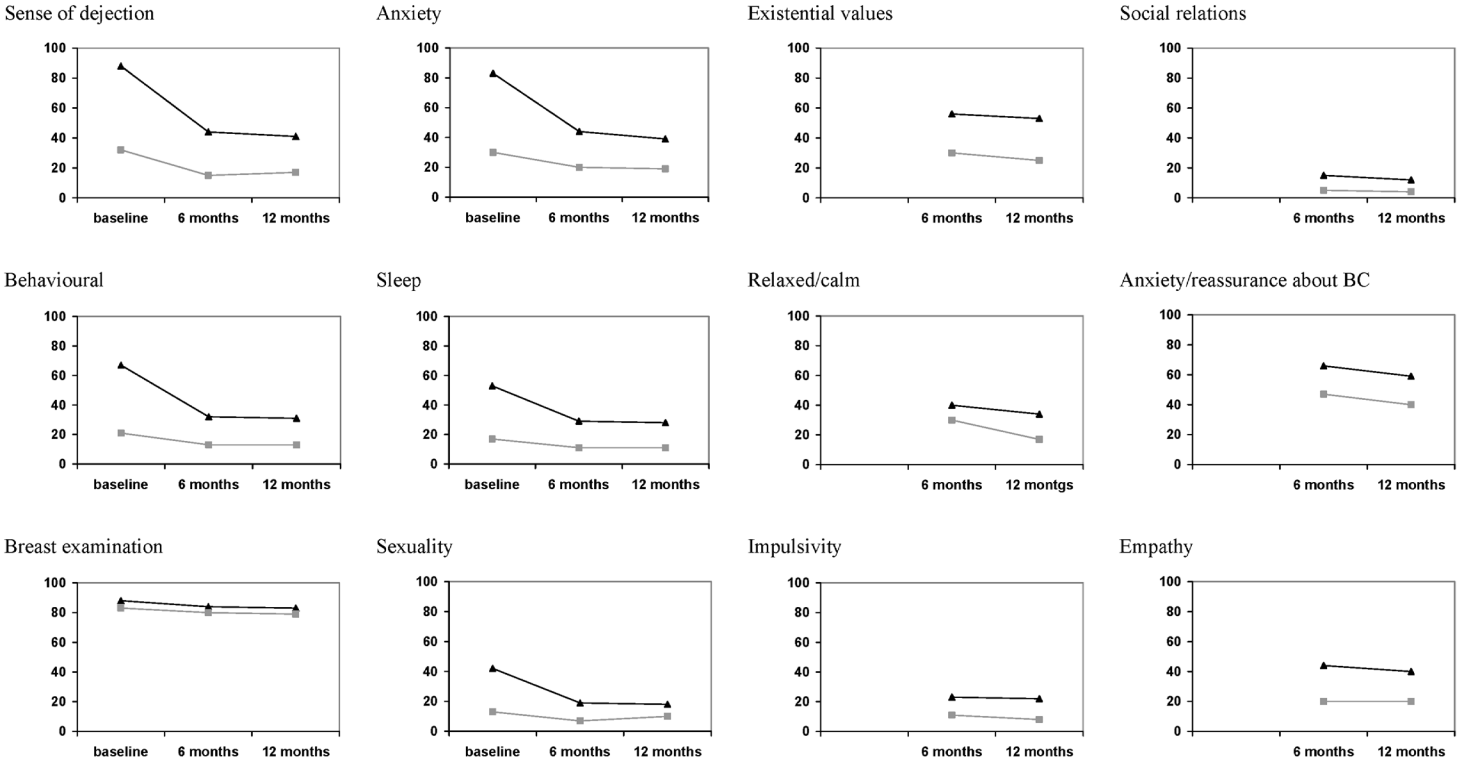
For the remaining COS-BC scales, the prevalence of psychosocial consequences was significantly higher at all assessments (P≤0.008) among women with false-positive screening mammography versus the controls, except for *Breast examination* consequences (P≥0.064) (Figure 3). The prevalence of *Sexuality* consequences among women with false-positive screening mammography decreased significantly (P<0.001) between baseline and at the six-month follow-up, but levelled out between six and 12 months (P=0.568). A similar pattern of the absence of changes over time for long-term consequences was observed for *Social relations* and *Impulsivity* consequences (P≥0.291). Experiencing false-positive screening mammography resulted in significantly higher odds ratios for psychosocial consequences (P≤0.010) than for women with a negative diagnosis for all assessments, except for *Breast examination* consequences (OR, 1.33-1.46) (Table 5).

Table 5: Psychosocial consequences due to false-positive (=1) versus negative (=0) screening mammography controlled for age, prior to the diagnosis of not having breast cancer, and six and 12 months later

	Odds ratio (95% CI) P-value (N)		
	baseline	**6-month follow-up**	**12-month follow-up**
Sense of dejection	15.52 (10.79-22.33) P<0.001 (834)	4.32 (2.96-6.29) P<0.001 (651)	3.55 (2.37-5.31) P<0.001 (557)
Anxiety	11.09 (7.94-15.49) P<0.001 (833)	3.10 (2.18-4.41) P<0.001 (648)	2.75 (1.86-4.06) P<0.001 (557)
Behavioural	7.58 (5.54-10.36) P<0.001 (824)	3.23 (2.16-4.83) P<0.001 (644)	2.93 (1.88-4.55) P<0.001 (550)
Sleep	5.47 (3.98-7.50) P<0.001 (842)	3.39 (2.22-5.18) P<0.001 (651)	3.07 (1.93-4.88) P<0.001 (559)
Breast examination	1.46 (0.99-2.17) 0.057 (839)	1.33 (0.89-2.00) 0.169 (656)	1.37 (0.90-2.09) 0.147 (560)
Sexuality	4.86 (3.79-7.00) P<0.001 (729)	2.99 (1.77-5.05) P<0.001 (599)	2.05 (1.20-3.48) 0.008 (516)
Existential values	N.A.	3.04 (2.19-4.21) P<0.001 (643)	3.47 (2.41-5.00) P<0.001 (561)
Social relations	N.A.	3.42 (1.88-6.23) P<0.001 (652)	3.79 (1.79-8.01) P<0.001 (565)
Relaxed/calm	N.A.	1.54 (1.11-2.13) 0.010 (650)	2.46 (1.64-3.69) P<0.001 (562)
Anxiety/reassurance about BC	N.A.	2.11 (1.54-2.90) P<0.001 (648)	2.09 (1.49-2.93) P<0.001 (563)
Impulsivity	N.A.	2.56 (1.65-3.99) P<0.001 (647)	3.35 (1.97-5.70) P<0.001 (565)
Empathy	N.A.	3.05 (2.15-4.32) P<0.001 (650)	2.48 (1.70-3.62) P<0.001 (564)

Hosmer-Lemeshow goodness-of-fit p-values ranged between 0.09 and 0.99, except for Impulsivity 0.031 at the six-month follow-up; N.A., not applicable

Figure 3: Percentages of women with false-positive screening mammography (black curves) experiencing psychosocial consequences at baseline (prior to the final negative (free from BC) diagnosis following diagnostic work-up), six and 12 months later versus their controls (grey curves).



Prediction of long-term psychosocial consequences of false-positive screening mammography (Paper III and the framework)

All studied predictor variables except age, marital status, and attending BC screening first time met the P<0.25 criterion for inclusion in multivariate analysis of one or more outcomes at the six-month follow-up (Paper III). Data regarding previous cancer diagnosis and not trusting the health care system were not included in the analyses due to low prevalence. The final multivariate models for the five COS-BC outcomes at six months are reported in Table 6. Early recall demonstrated the largest prediction (OR range; 5.24-10.31) for *Sense of dejection, Anxiety,* and *Sleep* consequences. Susceptibility (OR, 1.50) and worry about BC (OR, 1.46) were the most evident predictors for *Behavioural* and *Existential value* consequences, respectively. Dissatisfaction with information at recall (OR, 2.28-2.56) was also identified as a predictor for most outcomes. Among potential socio-demographic predictors two were identified; being foreign-born (OR, 2.40-3.71) which predicted most of the outcomes and level of education (OR, 1.84-2.89) which influenced *Sense of dejection*.

Table 6: Multivariate logistic regression models[a] of predictors of the long-term (six months) psychosocial consequences of false-positive screening mammography.

	Sense of dejection	**Anxiety**	**Behavioural**	**Sleep**	**Existential values**
Potential predictors			Odds radio (95% CI) P-value (Wald)		
Worry (total score 0-9)		1.20 (1.06-1.37) 0.005 (8.051)			1.46 (1.31-1.64) <0.001 (42.431)
Susceptibility (total score 0-6)	1.49 (1.20-1.84) <0.001 (13.558)	1.53 (1.21-1.95) <0.001 (12.229)	1.50 (1.21-1.87) <0.001 (13.535)		
Lack of social support (total score 0-9)			1.15 (1.03-1.29) 0.016 (5.780)	1.20 (1.06-1.35) 0.003 (8.553)	
Dissatisfied with own knowledge about BC	2.08 (1.02-4.26) 0.045 (4.017)				3.11 (1.45-6.67) 0.004 (8.514)
Dissatisfied with information at recall	2.28 (1.05-4.95) 0.037 (4.328)	2.56 (1.17-5.61) 0.019 (5.493)	2.42 (1.12-5.24) 0.025 (5.049)	2.38 (1.09-5.24) 0.031 (4.675)	
Early recall	10.31 (5.01-21.23) <0.001 (40.159)	6.25 (3.16-12.38) <0.001 (27.607)	3.21 (1.68-6.14) <0.001 (12.469)	5.24 (2.72-10.07) <0.001 (24.632)	
Country of origin (Sweden =0, other =1)		2.40 (1.11-5.19) 0.026 (4.976)	2.96 (1.36-6.45) 0.006 (7.478)	3.71 (1.62-8.51) 0.002 (9.547)	
Level of education					
➢ compulsory school	2.89 (1.15-7.27) 0.024 (5.079)				
➢ secondary school	1.84 (1.05-3.23) 0.034 (4.520)				
➢ college/higher education/university	Reference				
Sample size in the analysis (n)	323	322	313	327	329
Hosmer-Lemeshow test (P-value)	0.862	0.367	0.058	0.428	0.268
Nagelkerke's pseudo R-square	0.363	0.379	0.317	0.367	0.239

[a] Forward stepwise likelihood ratio multivariate logistic models; all models were controlled for age and prevalence of the psychosocial consequences at baseline.
BC, breast cancer

All studied predictor variables except for living area, marital status, attending BC screening first time, and a close one having BC, either now or previously, met the P<0.25 criterion for inclusion in multivariate analysis of one or more outcomes at the 12-month follow-up (Framework). The final multivariate models for the five COS-BC outcomes at 12 months are reported in Table 7. Early recall demonstrated the largest prediction (OR, 3.67) for *Sense of dejection* and *Sleep* consequences. Worry about BC was the most evident predictor for *Anxiety* (OR, 1.47) and B*ehavioural* (OR, 1.28) consequences, whereas susceptibility (OR, 1.50) predicted *Existential values* consequences. Lack of social support (OR, 1.13-1.25) was also identified as a predictor for most outcomes. Among potential socio-demographic predictors being foreign-born (OR, 2.35-3.02) demonstrated the largest prediction, which influenced *Anxiety* and *Sleep* consequences.

Table 7: Multivariate logistic regression models[a] of predictors of the long-term (12 months) psychosocial consequences of false-positive screening mammography.

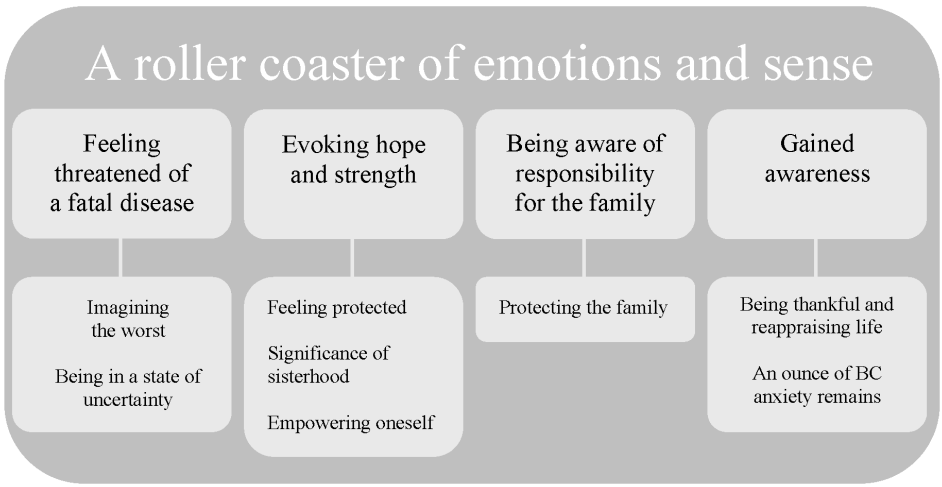
	Sense of dejection	Anxiety	Behavioural	Sleep	Existential values
Potential predictors			Odds radio (95% CI) P-value (Wald)		
Worry (total score 0-9)	1.30 (1.14-1.48) <0.001 (15.146)	1.47 (1.27-1.70) <0.001 (25.648)	1.28 (1.13-1.46) <0.001 (14.206)		
Susceptibility (total score 0-6)	1.28 (1.01-1.62) 0.039 (4.241)	1.30 (1.00-1.69) 0.048 (3.902)		1.46 (1.15-1.85) 0.002 (9.703)	1.50 (1.19-1.88) 0.001 (11.912)
Lack of social support (total score 0-9)	1.13 (1.01-1.26) 0.028 (4.824)	1.22 (1.08-1.38) 0.001 (10.179)	1.17 (1.04-1.31) 0.010 (6.608)	1.25 (1.10-1.41) <0.001 (12.217)	
Early recall	3.67 (1.95-6.91) <0.001 (16.318)	3.21 (1.62-6.38) 0.001 (11.068)	3.05 (1.55-6.04) 0.001 (10.318)	3.67 (1.82-7.40) <0.001 (13.274)	
Follow-up by the breast surgery clinic				2.71 (1.10-6.70) 0.031 (4.650)	
Country of origin (Sweden =0, other =1)		2.35 (1.04-5.34) 0.042 (4.187)		3.02 (1.32-6.92) 0.009 (6.852)	
Type of family					
➢ children <18 years old living at home				0.70 (0.30-1.63) 0.405 (0.694)	
➢ children ≥18 years old living at home				2.76 (1.17-6.53) 0.021 (5.333)	
➢ no children living at home				Reference	
Accommodation					
➢ owning a house		Reference	Reference		
➢ owning an apartment		1.90 (0.94-3.85) 0.074 (3.200)	2.24 (1.11-4.53) 0.025 (5.021)		
➢ renting		2.89 (1.33-6.32) 0.008 (7.103)	2.35 (1.11-4.96) 0.026 (4.979)		
Sample size in the analysis (n)	295	294	286	299	283
Hosmer-Lemeshow test (P-value)	0.239	0.874	0.142	0.062	0.469
Nagelkerke's pseudo R-square	0.289	0.428	0.317	0.372	0.375

[a] Forward stepwise likelihood ratio multivariate logistic models; all models were controlled for age and prevalence of the psychosocial consequences at baseline, except for Existential values, which was controlled for consequences at six months.

Perceived psychosocial consequences of false-positive screening mammography and coping with the situation (Paper IV)

Coping with the perceived psychosocial consequences of false positive screening mammography implied *a roller coaster of emotions and sense* (Figure 4). That is, women spoke of how they imagined the worst scenarios and how they felt insecure in face of *the threat of a fatal disease*. Thoughts about women who had died from BC and identifying with those with bad prognoses were revealed. Seeking information on the Internet was one way of dealing with the situation, although women realized that they eventually could not decrease their anxiety by reading internet blogs. Being selective about what to reveal and to whom, was another way of dealing with the situation. Keeping gnawing thoughts to oneself and avoiding the sympathy of others was also expressed. Being in a state of uncertainty, such as not knowing the diagnosis when waiting for examination results, was seemingly unbearable. Recall examinations were performed at short notice, but due to the shifting perception of time this sometimes felt like endless waiting, and sometimes seemed like an indication of urgency due to a high cancer risk.

Figure 4. Perceived psychosocial consequences of false-positive screening mammography and coping with the situation



On the other hand, women felt protected being surrounded by family and being professionally taken care of, which together with perceived sisterhood and self-empowerment *evoked strength and hope*. As the unpleasant situation progressed while waiting for the diagnosis, women became closer to their spouses by opening up to talk about matters of which they had not spoken before. In contrast,

although talking to relatives provided a sense of relief, conversations with family members were not always easy. Instead, strength and hope in the company of peers was evoked. Peers who had experienced false-positive screening mammography understood what the women were going through and the company of female friends was an opportunity to debrief. Being professionally taken care of was encouraging and the encounters with the female health care staff were described as comforting and significant. Reading between the lines of what the health care staff members were saying was also expressed. Both objective and personalised information was sought. Pep-talk, meditation, and convincing oneself that there was, so far, nothing to worry about were expressed. Women assumed that as the diagnostic work-up had taken a considerable amount of time, their case could not possibly be a fatal one. Checking that no family member had died from cancer was another example of seeking reassuring signs.

Being aware of responsibility for the family was revealed to be a crucial matter. Hence, women concealed their anxiety, particularly from their children. Behaving as usual and making an effort to be strong were other stances taken. Securing the children's future, and deciding who would take care of them in case of death from BC was one of the meaningful issues.

Experiencing false-positive screening raised thoughts of thankfulness and a reappraisal of life, although an ounce of BC anxiety remained. That is, women were grateful for the careful examinations and felt reassured. On the other hand, doubts regarding the diagnosis were apparent. Those who were scheduled for early subsequent mammography found the need for frequent examinations confusing, given that the recall examination showed that they did not have BC. The experience of false-positive mammography influenced the habits of breast self-examination and women felt motivated to do it more frequently and more thoroughly. However, some women felt that breast self-examination would not make any difference. Those, who already self-examined their breasts on a frequent basis, maintained their routine. Although women had unpleasant memories, they believed in the value of mammographic screening and intended to participate in subsequent screening. A negative screening examination was perceived as providing assurance of not having BC and security when feeling uncertain while self-examining one's breasts. In case of a positive diagnosis it was considered beneficial to get treatment early, which in light of women's experiences, was a major reason for attending mammographic screening. Women felt that in some way they had got their life back. They started to appreciate it more and valued things that they had not contemplated before. Consequently, *gained awareness* about the BC screening program and values in life emerged.

"What you see in the present is created by the past. If you define reality by what you see in the present you create a future that will be no different. If you want a different future, you have to change your perception of what you actually see now."

Lara Honos-Webb

Discussion

Usefulness of the Swedish COS-BC to study psychosocial consequences of false-positive screening mammography

The Swedish versions of the COS-BC 1 and COS-BC 2 proposed in this thesis consist of 34 and 23 items, respectively. Although causes for concern were identified, the findings support cross-sectional and longitudinal use of the *Sense of dejection, Anxiety, Behavioural*, and *Sleep* scales (COS-BC 1) and the *Existential values* scale (COS-BC 2) for group level assessments of the psychosocial consequences of false-positive mammographic screening. The remaining scale scores should be interpreted more cautiously. The purpose of the validation studies (Paper I, II) was to propose a conceptually, semantically, operationally, and psychometrically equivalent questionnaire version to the original COS-BC. Conceptual, semantic, and operational equivalence was supported for the Swedish COS-BC items (Paper I), and Paper II proposed a questionnaire version with supported psychometric properties for five scales. From this perspective, despite some differences between the original and adapted questionnaires, as long as the conceptual, semantic, operational, and psychometric properties of the scales have been supported, both versions are considered equivalent (18). For example, two items omitted from the final Danish COS-BC remained in the Swedish version and two new items were developed (Paper I). Furthermore, following iterative psychometric analyses, five items were added to the *Sense of dejection, Behavioural,* and *Sleep* scales in the final Swedish COS-BC (Paper II). This demonstrates that the construct covered by the original scales might, to some extent, manifest in different ways across cultures. However, the results also suggest that although there are some cultural differences between countries, equivalence is supported. For example, the vast majority of the COS-BC items were found relevant and provided coverage of the psychosocial consequences of false-positive screening mammography, which supports the content validity of the Swedish questionnaire version (Paper I). Furthermore, all scales supported validity criteria according to CTT and the majority of the scales generally met the RM requirements (Paper II).

The main cause of concern is the test-retest results for five scales (*Sexuality, Social relations, Relaxed/calm, Anxiety/reassurance about BC,* and *Impulsivity*), of which four

are from the original COS-BC (Paper II). However, it should be noted that the test-retest sample was limited and that a two week interval may have been inappropriate due to possible changes in psychosocial consequences of false-positive screening mammography. An attempt to control for this by applying a transition question was made but it is possible that this general well-being transition question was unable to reflect changes regarding specific aspects such as sexuality or social relations. It is also noteworthy that scales exhibiting reliability problems were those with only two or three items. Having fewer items and response category thresholds is known to compromise measurement precision (as well as validity) (18, 73), and may therefore also have contributed to reliability problems. On the other hand, lack of DIF over time (Paper II) supported test-retest reproducibility (73). Nevertheless, until further evidence has been provided, longitudinal data from these scales should be interpreted with caution.

Another source of concern is the poor targeting for all scales, manifested through large floor effects and negative person locations (Paper II). Well-known effects of this are that measurement precision and the ability to separate people is compromised, as is the ability to fully determine psychometric properties (73, 75, 77). However, these data must be interpreted in view of the intention and use of the questionnaire, and with an instrument such as the COS-BC, targeting problems are not surprising. The COS-BC questionnaire is intended for screening settings among generally healthy women, where a subset is assumed to experience negative consequences due to false-positive screening mammography. Although poor targeting is probably a general issue with this type of tools, its effect on the COS-BC measurement precision remains. It follows that COS-BC scores, despite reasonable RM fit, should not be regarded as precise measures but rather seen as assessments of coarser levels of psychosocial consequences. This implies that the Swedish COS-BC has limitations for measurement of the magnitude of the consequences.

The use of a mixed method approach to investigate the content validity of the questionnaire should be seen as an advantage (Paper I). That is, potential limitations of one method can be compensated for by another (89). For example, items in the *Sexuality* scale were poorly reflected in the interviews while exhibiting excellent I-CVI values and good RM fit in the original Danish version, arguing for their retention in the Swedish version (Paper I). Six other items failed to fit their hypothesized scales (Paper II). This might argue for exclusion of these items from the questionnaire. Furthermore, single items commonly demonstrate limitations such as compromised ability to discriminate between different levels of the construct and they are at greater risk of random error compared to rating scales (106). However, content validity supported their relevance (Paper I), suggesting that they represent important aspect of the psychosocial consequences of false-positive screening mammography. Until further evidence emerges it is therefore suggested that these items remain, but as single items, and they should be interpreted with caution.

The lack of meaningful DIF, except for the *Sense of dejection* scale, provided support for equal item functioning between diagnostic groups (women who had experienced false-positive and negative screening mammography) (Paper II). The *Sense of dejection* scale is therefore not recommended for comparing women with false-positive versus negative screening mammography unless DIF is accounted for (107). Another approach to eliminate the consequences of DIF when testing a hypothesis of between group differences is to dichotomise scale scores for each group prior to applying inferential statistics (Paper III). An additional argument for such an approach is that women who experience false-positive screening mammography are generally healthy people, and if they experience psychosocial consequences they have been affected by screening, and the magnitude of their experiences can be considered of less relevance. On the other hand, dichotomisation of the scores leads to loss of information.

It is important to bear in mind that the RM requirements are rigorous and difficult to satisfy (73, 77). It follows that even though these requirements were not fully supported the data provides useful information about the psychosocial consequences of false-positive mammographic screening. Another fact to keep in mind is that all Swedish COS-BC scales met the CTT assumptions. The main causes of concern were test-retest results for five scales and poor targeting for all scales, which challenges longitudinal stability and precision of scale scores. Consequently, the Swedish COS-BC scales should be considered assessment rather than measurement tools. Although single items were found to manifest important psychosocial consequences of false-positive screening mammography, their measurement properties remain limited.

Usefulness of general cancer screening items

The Council of the European Union recommends population-based screening for cervical, breast, and colorectal cancer (33), although the two latter are still being debated (46, 47, 108). Other screening programmes for additional cancer types are being considered (29). As well as the advantages of screening to decrease cancer mortality rates, it is also important to be aware of the disadvantages, and they should be accounted for (32, 33). From this perspective, a core questionnaire targeting the psychosocial consequences of false-positive results appears to be valuable in order to study adverse effects across cancer screening programmes (70). Such a questionnaire could be applied in research studies on the efficacy of new screening programmes as well as for monitoring ongoing programmes regardless of the cancer type screened for.

Two scales targeting long-term consequences of LC screening, *Impulsivity* and *Empathy*, including items not considered LC-specific (70), were evaluated in this thesis in a context of Swedish BC screening (Paper I, II). The aim was to investigate whether they would also be causes of concern among women experiencing false-

positive screening mammography, and thus to further explore the hypothesis that the scales and their items would be relevant and useful regardless of the cancer type screened for (70). As previously mentioned, the COS-BC items not considered BC-specific have been tested and found relevant in a Danish LC screening trial (70). This fact, together with the findings of this thesis, provide further evidence of the relevance of a set of core items that works across cancer screening programmes (70). Nevertheless, further research is needed to fully understand the psychosocial consequences of a false-positive result within cancer screening and how the consequences should be sufficiently assessed. Although the majority of the general items developed in LC screening were found to be relevant and to express long-term consequences (Paper I), causes of concern regarding their psychometric properties were also identified (Paper II). That is, the *Empathy* scale did not fit the RM, and the *Impulsivity* scale exhibited compromised test-retest reliability. From this perspective, as noted above for the Swedish COS-BC scales, validity and reliability problems of the general COS-LC items need further investigation to determine whether these occur in other cancer screening settings, and if they are true issues in BC screening or related to study design. The exact reason for the compromised test-retest reliability of the *Impulsivity* scale cannot be determined. However, as discussed above for some Swedish COS-BC scales, it could indicate that this variable fluctuates over time. The RM misfit of the *Empathy* scale could be due to multidimensionality, i.e. that the scale items (*understanding people's problems, ability to listen*, *and responsibility for family*) represent different constructs. Although the interviews with women in Paper IV did not aim to evaluate content validity of the items, the findings indicate that there were empathy issues related to both sisterhood (e.g. feelings of empathy for women in the same situation) and reappraisal of life (e.g. being humble about ill people). As for being aware of family responsibilities, this was an aspect of protecting the family rather than a matter of empathy. This supports the interpretation that the *Empathy* scale may have dimensionality problems.

Relevance of the psychosocial consequences of false-positive screening mammography

It seems reasonable to consider the short- and long-term psychosocial consequences of false-positive screening mammography as less important in relation to the life that might be saved. However, such consequences appear to be common and persistent over time (Paper III), which is in line with a recently published Danish study, the only other study that has used the COS-BC (19). Furthermore, the results presented here demonstrate at least twice as high odds ratios for the occurrence of psychosocial consequences (sense of dejection, anxiety, impact on behaviour, sleep, and existential values) for recalled women compared to their controls (Paper III). Other psychosocial

consequences, such as impact on sexuality or impulsivity, were also common among women who had experienced false-positive screening mammography, although these results should be interpreted more cautiously due to the compromised psychometric properties of these scales. Only the prevalence of breast self-examination did not differ between women with false-positive and negative screening mammography (Paper III). Presumably, this was owing to recommendations which all women received at screening and/or recall examination(s). The study findings demonstrate that women might perceive false-positive screening mammography as taxing, which activates profoundly excruciating experiences and *a roller coaster of emotions and sense* to handle the situation (Paper IV). Similar results have been found in other studies. For example, various and intense experiences have been described by recalled women in a Norwegian BC screening programme (25). A study on coping strategies that women apply in such a situation found significant associations between anxiety and a variety of behavioural and cognitive approaches (109). Taken together, the results of this thesis confirm the evidence for short-term psychosocial consequences of false-positive screening mammography and support the hypothesis of the occurrence of long-term adverse effects. Importantly, these consequences relate to generally healthy women who do not necessarily benefit from the programme.

In keeping with this line of reasoning it is essential to be aware that delivery of screening implies a responsibility for unnecessary consequences (32); thus there is a duty to introduce interventions in order to minimise the adverse consequences and provide support for women with a compromised ability to overcome them. This relates to ethical principles such as nonmaleficence and beneficence. That is, an obligation not to harm individuals but to contribute to their well-being (110). However, the principle of nonmaleficence does not need to be violated in BC screening if interventions are provided to minimise adverse consequences, and benefits are delivered and outweigh the risks (49). From such a perspective, it is vital to monitor screening programmes continually, with regard to both benefits and risks, in order to provide a basis for quality evaluation of the programme as well as development of adequate interventions.

Interventions on a basis of prediction and coping with the consequences

Several variables have been tested for prediction of the long-term psychosocial consequences of false-positive screening mammography in this thesis. Early recall to subsequent mammography demonstrated the strongest prediction for experiencing such consequences (Paper III), which differed from previous results. That is, although studies have found that women experienced long-term distress following early recall, their experiences generally did not differ from those having fine or core needle biopsy (22, 60). The findings presented here (Paper III) that fine/core needle biopsy generally did not predict long-term consequences seem reasonable since the gold standard in assessment of suspicious BC is invasive procedures, combined with

imaging and clinical examination (triple assessment) (49), which probably provide confidence in the diagnosis for most women. On the contrary, as suggested by the results in Paper IV, early recall seems to create confusion and doubt about the diagnosis. These observations (Paper III, IV) support the European guidelines to avoid early recall (target <1% of screened women) (49). The importance of finding cancers through early recall should of course be balanced against the drawbacks. However, evidence on the diagnostic yield of early recall is limited, and the general opinion is therefore that the frequency of early recall should be kept as low as possible (49, 111). Those who are recommended for early recall should be fully informed and preferably provided with an opportunity to consult health care professionals as support induces feelings of comfort and security (Paper IV).

Dissatisfaction with information at recall also appears to influence long-term psychosocial consequences (Paper III), which is in line with previous findings (22). This highlights the significance of providing information that can help women to cope with their situation. Interviews with women with false-positive screening mammography (Paper IV) have found that recalled women look for adequate, and whenever possible, individualised information. Personalised information and communication is therefore most likely both preferable and needed in order to prevent or lessen psychosocial consequences of false-positive mammographic screening. The information that health professionals provide and the way in which they communicate influence how women cope with their experiences (Paper IV). However, although women found strength and hope from being professionally taken care of they still sought information from a variety of other sources (Paper IV). Professional care includes a range of tasks where information provision and communication is emphasised (49). However, customising information to each individual might be challenging. People obviously vary in their needs, knowledge, values, beliefs, and cultural preferences. From this it follows that in meeting women's needs health care professionals should be offered support, such as discussion sessions, exchange of experiences, and training in communication skills. Furthermore, the most adequate and effective communication styles in a context of BC screening remain to be investigated. For example, the Internet probably offers a new approach to reach a younger generation of women in BC screening. Women sought objective information from websites and blogs (Paper IV), which seems to support the value of exploring the full potential of the Internet for improving client-provider communication in the screening service.

It is also noteworthy that among several socio-demographic variables investigated in this thesis, being foreign-born influenced the prevalence of long-term psychosocial consequences of false-positive mammographic screening (Paper III). This might further support the argument for the existence of different communication needs within this group, but the matter also warrants further research as such a prediction has not been found before (20). However, experiences of foreign-born women do not seem to differ from those of native women (Paper IV). On the other hand, ac-

cording to the narratives of the immigrants, they seemed to have lived in Sweden for a long time. It may therefore be assumed that they had adjusted to the Swedish health care system. Nevertheless, the importance of communicating information becomes especially apparent in multicultural populations with potentially different needs and attitudes towards screening. The application of relevant, understandable, and phase-specific information and communication have been emphasized (49). Previous qualitative studies of ethnically diverse women in the U.S. revealed perceived dissatisfaction with client-provider communication and disrespect (23, 24). Such results might reflect misunderstandings related to cultural differences and communication preferences, which appears to further support the significance of individually tailored information. The logic here is essentially the same as that underpinning adaptation and content validation of rating scales in the target population. Mere translation of the information is not enough, since it might overlook cultural factors, preferences for delivering information, and how information is understood and valued. To investigate adequate communication styles in order to deal with multicultural populations in the context of mammographic screening seems also to respect the ethical principle of justice (110). In this context, that means to prevent unnecessary risks among subgroups of women (49).

Another important finding was the significance of sisterhood when coping with the perceived psychosocial consequences of false-positive screening mammography (Paper IV). It was important to talk to peers as they understood the unpleasant situation and provided a means to forget the unpleasant circumstances of the diagnostic work-up, and also to share feelings of joy following the diagnosis of not having BC. These results, together with findings of feeling protected when being listened to by relatives (Paper IV) correspond to a study of coping strategies on well-being following false-positive screening mammography (112). In that study, social support was found to be an effective strategy to prevent adverse consequences on women's well-being in those who experienced a false-positive result as stressful. Such results seem to support the argument for the provision of social support to women who experience susceptibility and worry about BC (Paper III). Consequently, women should be encouraged to utilize social support. Such a practice is a common recommendation when women are followed-up by the surgical breast clinic and might also be useful in the early stage of the diagnostic work-up following abnormal screening mammography. However, this will need to be investigated empirically.

Promoting informed decision-making

Within research and health care we strongly defend the principle of not restricting the autonomy of any individual. Consequently, women invited to attend screening should receive clear and balanced information about the benefits and harm of screening in order to promote informed decision-making about attendance (27, 33,

88). Action should be taken to ensure that the information is understandable and to provide knowledge and awareness of advantages and shortcomings of the programme (49). On the other hand, information about the risks of BC screening, including psychosocial consequences of false-positive screening mammography, might challenge attendance at screening, which is one of the prerequisites for the effectiveness of the programme (6). However, interviews with recalled women showed that they are prone to continue their participation in screening regardless of their agonising experiences following false-positive screening mammography (Paper IV). This result corresponds to a meta-analysis of European studies showing no effects on re-attendance at BC screening following false-positive mammography (113). Women seem to value screening given that in the case of a positive diagnosis it is considered beneficial to get treatment early (Paper IV). They also trust the health care system (Paper III). On the other hand, it is important to be aware that mammographic screening has been in operation for over three decades (6) and the influence of the mass media is substantial, thus a potential social pressure to participate might restrict a person's autonomy (9). Consequently, it is of paramount importance to provide evidence-based information about screening in order to respect individual informed choice about attendance (49).

Recently, the Swedish National Board of Health and Welfare has preliminary updated the recommendations for BC screening in Sweden, which now also include recommendations to improve information about screening both in the invitation to attend and on the websites of mammographic facilities (52). Among the consequences of mammographic screening, aspects of over-diagnosis and risk of overtreatment are emphasised. Furthermore, it is stated that psychosocial consequences of false-positive screening mammography are moderate and transient for the majority of women but systematic evaluations within ongoing programmes have not yet been conducted (52). Scientific knowledge about BC screening is complex and somehow difficult to communicate. Mammographic screening is delivered to the population, but scientific knowledge needs to be customized to the individual woman to help her apply the information to her own situation. For example, comprehensive and understandable information booklets or websites about the advantages and disadvantages of screening might be of value to achieve a better informed decision regarding attendance. Two Australian randomised controlled trials on the effects of providing comprehensive internet and booklet-based information for women aged 40 and 70, respectively, who were invited to BC screening demonstrated that their knowledge about BC screening was increased without influencing anxiety (114, 115). Women were more likely to make an informed decision compared to their controls, and attendance at screening did not differ for women aged 70, whereas those aged 40 were less likely to start screening.

Beyond the current context

The most debated harmful consequence of cancer screening is the *extent* and *relevance* of over-diagnosis and over-treatment in relation to the benefits of screening (9). However, other adverse effects such as psychosocial consequences of false-positive screening should also be considered, although it seems reasonable to value such consequences as less relevant in relation to over-diagnosis and over-treatment. On the other hand, as demonstrated in this thesis, their prevalence might be high. There is also the aspect of multiple cancer screening programme attendance, such as participation in both BC and cervical cancer screening, which increases the risk of a false-positive result for an individual. The troublesome issue is that the consequences of false-positive results outside BC screening have been scarcely studied. For example, it was concluded in a Cochrane review on randomised controlled trials for colorectal cancer screening using a faecal occult blood test that studies on the psychosocial consequences of a false-positive test are limited (116). On the basis of the trials, it was found that only between 5.2% and 18.7% of individuals receiving abnormal screening results actually had the disease, which means that a considerable number of people screened for colorectal cancer experience a false-positive result (108). Regarding the benefits of colorectal cancer screening, randomised controlled studies found a significant 16% reduction in the relative risk of colorectal cancer mortality following ten years of screening, which corresponds to avoiding approximately one of six deaths from the disease (116). Furthermore, estimates for cervical cancer screening on the basis of an ongoing programme demonstrated that for each 10 000 screened women in 35 years, 10 out of 25 would be prevented from dying from the disease (117). For each life saved, 1955 women had an abnormal cytology. A study among women with mild atypical cytological results found that women were poorly informed about the programme and it has been suggested that adequate information prior to screening and at follow-up would most probably decrease women's anxiety at the diagnostic work-up (118).

As noted previously, the Council of the European Union recommends population-based screening for cervical, breast, and colorectal cancer (33). Many countries have implemented the programmes (8). In Sweden, cervical cancer screening has been provided since the 1960s (119) and recently the Swedish National Board of Health and Welfare has given a preliminary recommendation for screening for colorectal cancer (120). The council of the European Union states that screening should be offered to fully informed people where the benefits and risks are well known (33). Furthermore, people invited to screening should be informed about the benefits and harm of screening in order to ensure they can make an informed decision about attendance (27, 33). In a survey among 317 cancer-screened Americans aged 50 to 69, 80% reported that they wanted to be told about the possible harmful consequences of screening, and 59% said they would continue attendance even if they

were informed about the risks of over-diagnosis and over-treatment (121). Studies on the psychosocial consequences of false positive cancer screening are needed, preferably using a core questionnaire targeting such consequences across countries and screening programmes (70). General items from the original Danish COS-BC and COS-LC (17, 68-70) and their Swedish adaptations (Paper I, II) provide initial step towards international harmonization in the monitoring of the psychosocial consequences of false-positive screening. As such, they might deliver additional knowledge for the evaluation of the screening programmes. Evidence-based, honest, and unbiased information provision about the advantages and disadvantages of cancer screening respects the autonomy of individuals. Delivery of cancer screening should include interventions to minimise the adverse effects of the programmes.

Policy-makers are usually under strong pressure to introduce screening programmes. In general it is believed that medical screening has the potential to contribute to the prevention of ill health and save lives (27). However, the downside of screening might be considerable, as for example in the context of LC screening. It has been estimated on the basis of the National LC Screening Trial among former or current heavy smokers (122) that for 1000 screened persons, three would be prevented from dying from LC, but also another 231 would have a false-positive result (123). In addition, there would be 40 additional invasive procedures. However, LC is the leading cause of death from cancer worldwide (3). Even so, the best preventive strategy might not be screening for LC but to take organised action across countries to prevent or at least profoundly reduce smoking (32).

Conclusions

The findings of this thesis confirm previous evidence of short-term psychosocial consequences among women with false-positive screening mammography, and add further knowledge about the prevalence and prediction of long-term effects. Women experience consequences such as a sense of dejection, sleep disturbance, and thoughts about existential values up to one year following false-positive screening mammography. Women most at risk of experiencing such consequences are those who are scheduled for early recall to subsequent mammography, who are worried about BC, dissatisfied with information received at recall, foreign-born, or lacking social support. Importantly, study assessments were conducted by means of a condition-specific questionnaire validated among women to target the psychosocial consequences of false-positive screening mammography. The construct of such consequences in the context of Swedish BC screening has been studied here. It can therefore be assumed that the outcomes are to be considered relevant from women's perspective, but also for the providers of BC screening. In addition, the general absence of signs of selec-

tion bias indicates sufficient representativeness of the study sample in relation to the background population attending BC screening from the municipalities of Malmö, Trelleborg, and Vellinge. The thesis also provides clues about women's perception of the consequences of false-positive screening mammography and how they cope with the situation. Their experiences might be excruciating and trigger a range of various coping strategies, such as information-seeking and evoking strength and hope following social and professional support. Coping with the perceived psychosocial consequences of false-positive screening mammography implied *a roller coaster of emotions and sense*. However, it is important to bear in mind that these results refer to women who have experienced some degree of the consequences. Although they experienced false-positive screening mammography, the women seem to value BC screening and believe they will benefit from early detection of the disease. They are grateful for being carefully examined and intend to continue their attendance. Altogether, the findings presented here suggest that the occurrence of long-term psychosocial consequences of false-positive BC screening should be acknowledged and reported to the consumers and providers of BC screening. Delivery of screening implies a duty to introduce interventions to minimize adverse consequences and to provide support for women with a compromised ability to overcome them. Intervention provision might not only be of value from the women's perspective but might also contribute to the cost-effectiveness of screening. For example, reducing the number of those subjected to early recall not only has the potential to decrease the prevalence of psychosocial consequences of false-positive mammographic screening but also to reduce the costs of diagnostic follow-up, particularly since the positive predictive value of early recall is considered to be low. However, it is suggested that this should be investigated further. Furthermore, information about the benefits and risks of BC screening increases the awareness of the programme, which is of value from the perspective of the population and the individual.

Implications for research and practice

A considerable proportion of women with false-positive screening mammography experience psychosocial consequences for up to one year and the consequences might profoundly influence the woman's life evoking a variety of coping strategies. Further research is needed to investigate the magnitude of the consequences and whether they generalise nationwide and outside of Scandinavia. The Swedish National Board of Health and Welfare has recommended a nationwide BC screening register in order to provide a basis for evaluation of the programme. A Swedish version of the COS-BC validated in this thesis provides a means for monitoring the consequences of false-positive mammographic screening. Interventions of avoiding early recall to subsequent mammography and providing individualised, clear, objective in-person

information at the diagnostic work-up are advised. Research studies to investigate adequate communication styles are needed, especially in order to face multicultural populations in the context of mammographic screening. Provision of a supportive environment evokes strength and hope among recalled women. Women should be encouraged to use social support.

Women invited to mammographic screening should be informed about the potential benefits and harm of the programme. In doing so, the risk of long-term psychosocial consequences of false-positive screening mammography should be acknowledged. The information should be evidence-based, unbiased, and understandable in order to promote informed decision-making about attendance. This respects the autonomy of each individual. Studies are needed to investigate the best format and effectiveness of decision aids, for example booklets or websites. It is suggested that the value of the Internet is explored as a potential means to support health care professionals in communicating with women invited to screening and those at recall.

Five Swedish COS-BC scales; *Sense of dejection*, *Anxiety*, *Behavioural*, *Sleep*, and *Existential values* are proposed as primary tools for monitoring the psychosocial consequences of false-positive mammographic screening. The other questionnaire scales should be interpreted more cautiously. Although the COS-BC fills an important gap, it should be considered for group level assessments. It is advised that the Swedish version of the questionnaire is subjected to further research to determine whether its validity and reliability problems occur in other settings; i.e. to investigate whether they are true issues in BC screening or related to study design. Since this thesis comprises the first published non-Danish psychometric investigation of the COS-BC it provides the first steps towards international harmonization of monitoring condition-specific psychosocial consequences of false-positive screening mammography.

The observations presented here have implications beyond the current setting as they also highlight areas to focus on in other programmes and countries adapting the COS questionnaires. This thesis provides support for the hypothesis of common cancer screening consequences, but the extent to which the results can be generalised outside Scandinavia remains to be determined. The consequences of false-positive cancer screening results operationalized as scale items in one context (e.g. LC screening) can be as useful as those identified in another context (e.g. BC screening), when applied therein. Development of a core item set for measurement of the psychosocial consequences of false-positive results across cancer screening programmes and countries has been previously proposed and the findings presented here support this proposal.

References

1. Lazarus RS, Folkman S. Stress, appraisal, and coping. New York: Springer; 1984.
2. The Council of the European Union, the Committee of ministers of member states. Recommendations on screening as a tool of preventive medicine. 1994. Available at: https://wcd.coe.int/com.instranet.InstraServlet?command=com.instranet.CmdBlobGet&InstranetImage=534532&SecMode=1&DocId=514336&Usage=2
3. World Health Organisation. International Agency for Research on Cancer. CANCERMondial. GLOBOCAN 2008. Available at: http://globocan.iarc.fr/
4. Jönsson P-E (redaktör). Bröstcancer. AstraZeneca Sverige, Södertälje: Trosa Tryckeri; 2009.
5. World Health Organisation. Breast cancer: prevention and control. Available at: http://www.who.int/cancer/detection/breastcancer/en/
6. IACR Handbooks of Cancer Prevention – Breast Cancer Screening: Lyon: IARC Press; 2002.
7. Paci E. EUROSCREEN Working Group. Summary of the evidence of breast cancer service screening outcomes in Europe and first estimate of the benefit and harm balance sheet. Journal of Medical Screening. 2012;19:5-13.
8. European Commission. Cancer screening in the European Union. Report on the implementation of the Council Recommendation on cancer screening (First report). European Communities; 2008. Available at: http://ec.europa.eu/health/ph_determinants/genetics/documents/cancer_screening.pdf
9. Juth N, Munthe C. The ethics of screening in health care and medicine: serving society or serving the patient? New York: Springer; 2011.
10. Hofvind S, Ponti A, Patnick J, Ascunce N, Njor S, Broeders M, et al. False-positive results in mammographic screening for breast cancer in Europe: a literature review and survey of service screening programmes. Journal of Medical Screening. 2012;19:57-66.
11. McCaffery KJ, Barratt AL. Assessing psychosocial/quality of life outcomes in screening: how do we do it better? Journal of Epidemiology and Community Health. 2004;58:968-70.
12. Brodersen J, McKenna SP, Doward LC, Thorsen H. Measuring the psychosocial consequences of screening. Health and Quality of Life Outcomes. 2007;5:3.
13. Brodersen J, Thorsen H, Cockburn J. The adequacy of measurement of short and long-term consequences of false-positive screening mammography. Journal of Medical Screening. 2004;11:39-44.
14. Brewer NT, Salz T, Lillie SE. Systematic review: the long-term effects of false-positive mammograms. Annals of Internal Medicine. 2007;146:502-10.

15. Bond M, Pavey T, Welch K, Cooper C, Garside R, Dean S, et al. Psychological consequences of false-positive screening mammograms in the UK. Evidence Based Medicine 2013;18:54-61.
16. Brett J, Bankhead C, Henderson B, Watson E, Austoker J. The psychological impact of mammographic screening. A systematic review. Psychooncology. 2005;14:917-38.
17. Brodersen J. Measuring psychosocial consequences of false-positive screening results – breast cancer as an example. Copenhagen: University of Copenhagen; 2006.
18. Streiner DL, Norman GR. Health measurement scales: A practical guide to their development and use. Oxford: Oxford University Press; 2008.
19. Brodersen J, Siersma VD. Long-term psychosocial consequences of false-positive screening mammography. Annals of Family Medicine. 2013;11:106-15.
20. Meystre-Agustoni G, Paccaud F, Jeannin A, Dubois-Arber F. Anxiety in a cohort of Swiss women participating in a mammographic screening programme. Journal of Medical Screening. 2001;8:213-9.
21. Olsson P, Armelius K, Nordahl G, Lenner P, Westman G. Women with false positive screening mammograms: how do they cope? Journal of Medical Screening. 1999;6:89-93.
22. Brett J, Austoker J. Women who are recalled for further investigation for breast screening: psychological consequences 3 years after recall and factors affecting re-attendance. Journal of Public Health Medicine. 2001;23:292-300.
23. Allen JD, Shelton RC, Harden E, Goldman RE. Follow-up of abnormal screening mammograms among low-income ethnically diverse women: findings from a qualitative study. Patient Education and Counseling. 2008;72:283-92.
24. Padgett DK, Yedidia MJ, Kerner J, Mandelblatt J. The emotional consequences of false positive mammography: African-American women's reactions in their own words. Women Health. 2001;33:1-14.
25. Solbjor M, Forsmo S, Skolbekken JA, Saetnan AR. Experiences of recall after mammography screening – a qualitative study. Health Care for Women International. 2011;32:1009-27.
26. Thorne SE, Harris SR, Hislop TG, Vestrup JA. The experience of waiting for diagnosis after an abnormal mammogram. The Breast Journal. 1999;5:42-51.
27. Raffle A, Gray M. Screening: evidence and practice. New York: Oxford University Press; 2007.
28. Morabia A, Zhang FF. History of medical screening: from concepts to action. Postgraduate Medical Journal. 2004;80:463-9.
29. National Cancer Institute. Screening and testing to detect cancer: Lung Cancer. Available at: http://www.cancer.gov/cancertopics/screening/lung
30. Wilson JMG, Jungner G. Principles and practice of screening for disease. Geneva: World Health Organisation; 1968. Available at: http://whqlibdoc.who.int/php/WHO_PHP_34.pdf
31. Andermann A, Blancquaert I, Beauchamp S, Dery V. Revisiting Wilson and Jungner in the genomic age: a review of screening criteria over the past 40 years. Bulletin of the World Health Organisation. 2008;86:317-9.
32. Strong K, Wald N, Miller A, Alwan A. Current concepts in screening for noncommunicable disease: World Health Organization Consultation Group Report

on methodology of noncommunicable disease screening. Journal of Medical Screening. 2005;12:12-9.
33. The Council of the European Union. Council recommendation of 2 December 2003 on cancer screening. Official Journal of the European Union. OJ 2003/878/EC; L 327: 34-38. Luxembourg: Office for Official Publications of the European Communities; 2003. Available at: http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2003:327:0034:0038:EN:PDF
34. Shapiro S. Periodic screening for breast cancer: the HIP randomized controlled trial. Health Insurance Plan. Journal of the National Cancer Institute Monographs. 1997;22:27-30.
35 Andersson I, Aspegren K, Janzon L, Landberg T, Lindholm K, Linell F, et al. Mammographic screening and mortality from breast cancer: the Malmö mammographic screening trial. British Medical Journal. 1988;297:943-8.
36. Frisell J, Eklund G, Hellstrom L, Lidbrink E, Rutqvist LE, Somell A. Randomized study of mammography screening- preliminary report on mortality in the Stockholm trial. Breast Cancer Research and Treatment. 1991;18:49-56.
37. Frisell J, Lidbrink E, Hellstrom L, Rutqvist LE. Follow up after 11 years – update of mortality results in the Stockholm mammographic screening trial. Breast Cancer Research and Treatment. 1997;45:263-70.
38. Tabar L, Fagerberg CJ, Gad A, Baldetorp L, Holmberg LH, Grontoft O, et al. Reduction in mortality from breast cancer after mass screening with mammography. Lancet. 1985;13:829-32.
39. Bjurstam N, Bjurneld L, Warwick J, Sala E, Duffy SW, Nystrom L, et al. The Gothenburg breast cancer screening trial. Cancer. 2003;97:2387-96.
40. Nystrom L, Rutqvist LE, Wall S, Lindgren A, Lindqvist M, Ryden S, et al. Breast cancer screening with mammography: overview of Swedish randomised trials. Lancet. 1993;341:973-8.
41. Andersson I, Janzon L. Reduced breast cancer mortality in women under age 50: updated results from the Malmo Mammographic Screening Program. Journal of the National Cancer Institute Monographs. 1997;22:63-7.
42. Bjurstam N, Bjorneld L, Duffy SW, Smith TC, Cahlin E, Eriksson O, et al. The Gothenburg breast screening trial: first results on mortality, incidence, and mode of detection for women ages 39-49 years at randomization. Cancer. 1997;80:2091-9.
43. Larsson LG, Andersson I, Bjurstam N, Fagerberg G, Frisell J, Tabar L, et al. Updated overview of the Swedish randomized trials on breast cancer screening with mammography: age group 40-49 at randomization. Journal of the National Cancer Institute Monographs. 1997;22:57-61.
44. Nystrom L, Andersson I, Bjurstam N, Frisell J, Nordenskjold B, Rutqvist LE. Long-term effects of mammography screening: updated overview of the Swedish randomised trials. Lancet. 2002;359:909-19.
45. Jorgensen KJ, Keen JD, Gotzsche PC. Is mammographic screening justifiable considering its substantial overdiagnosis rate and minor effect on mortality? Radiology. 2011;260:621-627.
46. Gotzsche PC, Nielsen M. Screening for breast cancer with mammography. Cochrane Database of Systematic Reviews. 2011;4:CD001877.

47. Gotzsche PC, Jorgensen KJ. Screening for breast cancer with mammography. Cochrane Database of Systematic Reviews. 2013;6:CD001877.
48. Autier P, Boniol M, Gavin A, Vatten LJ. Breast cancer mortality in neighbouring European countries with different levels of screening but similar access to treatment: trend analysis of WHO mortality database. British Medical Journal. 2011;343:d4411.
49. European Commission. European guidelines for quality assurance in breast cancer screening and diagnosis. 4. ed. Luxembourg: Office for Official Publications of the European Communities; 2006.
50. Olsson S, Andersson I, Karlberg I, Bjurstam N, Frodis E, Hakansson S. Implementation of service screening with mammography in Sweden: from pilot study to nationwide programme. Journal of Medical Screening. 2000;7:14-8.
51. Socialstyrelsen. Nationella riktlinjer för bröstcancersjukvård. Medicinsk och hälsoekonomiskt faktadokument. Lindesberg; Bergslagens Grafiska: 2007.
52. Socialstyrelsen. Screening för bröstcancer. Rekommendation och bedömningsunderlag. Remissversion. 2013. Available at: http://www.socialstyrelsen.se/SiteCollectionDocuments/nr-screening-brostcancer-rekommendation.pdf
53. Cecilia.Wattsgard@unilabs.com. Unilabs AB. Personal Communication.
54. Salz T, Richman AR, Brewer NT. Meta-analyses of the effect of false-positive mammograms on generic and specific psychosocial outcomes. Psychooncology. 2010;19:1026-34.
55. Lampic C, Thurfjell E, Bergh J, Sjoden PO. Short- and long-term anxiety and depression in women recalled after breast cancer screening. European Journal of Cancer. 2001;37:463-9.
56. Gilbert FJ, Cordiner CM, Affleck IR, Hood DB, Mathieson D, Walker LG. Breast screening: the psychological sequelae of false-positive recall in women with and without a family history of breast cancer. European Journal of Cancer. 1998;34:2010-4.
57. Cockburn J, Staples M, Hurley SF, De Luise T. Psychological consequences of screening mammography. Journal of Medical Screening. 1994;1:7-12.
58. Lidbrink E, Levi L, Pettersson I, Rosendahl I, Rutqvist LE, de la Torre B, et al. Single-view screening mammography: psychological, endocrine and immunological effects of recalling for a complete three-view examination. European Journal of Cancer. 1995;31:932-3.
59. Lowe JB, Balanda KP, Del Mar C, Hawes E. Psychologic distress in women with abnormal findings in mass mammography screening. Cancer. 1999;85:1114-8.
60. Brett J, Austoker J, Ong G. Do women who undergo further investigation for breast screening suffer adverse psychological consequences? A multi-centre follow-up study comparing different breast screening result groups five months after their last breast screening appointment. Journal of Public Health Medicine. 1998;20:396-403.
61. Aro AR, Pilvikki Absetz S, van Elderen TM, van der Ploeg E, van der Kamp LJ. False-positive findings in mammography screening induces short-term distress – breast cancer-specific concern prevails longer. European Journal of Cancer. 2000;36:1089-97
62. Lipkus IM, Halabi S, Strigo TS, Rimer BK. The impact of abnormal mammograms on psychosocial outcomes and subsequent screening. Psychooncology. 2000;9:402-10.
63. Scaf-Klomp W, Sanderman R, van de Wiel HB, Otter R, van den Heuvel WJ. Distressed or relieved? Psychological side effects of breast cancer screening in The Netherlands. Journal of Epidemiology and Community Health. 1997;51:705-10.

64. Sandin B, Chorot P, Valiente RM, Lostao L, Santed MA. Adverse psychological effects in women attending a second-stage breast cancer screening. Journal of Psychosomatic Research. 2002;52:303-9.
65. Ellman R, Angeli N, Christians A, Moss S, Chamberlain J, Maguire P. Psychiatric morbidity associated with screening for breast cancer. British Journal of Cancer. 1989;60:781-4.
66. Lampic C, Thurfjell E, Sjoden PO. The influence of a false-positive mammogram on a woman's subsequent behaviour for detecting breast cancer. European Journal of Cancer. 2003;39:1730-7.
67. Cockburn J, De Luise T, Hurley S, Clover K. Development and validation of the PCQ: a questionnaire to measure the psychological consequences of screening mammography. Social Science & Medicine. 1992;34:1129-34.
68. Brodersen J, Thorsen H. Consequences of Screening in Breast Cancer (COS-BC): development of a questionnaire. Scand Journal of Primary Health Care. 2008;26:251-6.
69. Brodersen J, Thorsen H, Kreiner S. Validation of a condition-specific measure for women having an abnormal screening mammography. Value in Health. 2007;10:294-304.
70. Brodersen J, Thorsen H, Kreiner S. Consequences of screening in lung cancer: development and dimensionality of a questionnaire. Value in Health. 2010;13:601-12.
71. Engel GL. The need for a new medical model: a challenge for biomedicine. Science. 1977;196:129-36.
72. Engel GL. The clinical application of the biopsychosocial model. American Journal of Psychiatry. 1980;137:535-44.
73. Hobart J, Cano S. Improving the evaluation of therapeutic interventions in multiple sclerosis: the role of new psychometric methods. Health Technology Assessment. 2009;13:1-200. Available at: http://www.hta.ac.uk/fullmono/mon1312.pdf
74. Definition of Objective Measurement. Institute for Objective Measurement. Available at: http://www.rasch.org/define.htm
75. Andrich D. Rasch models for measurement. Sage University paper series on quality application in the social sciences, series 07-068. Newbury Park; Sage Publications Inc.: 1988.
76. Rasch G. Probabilistic models for some intelligence and attainment tests. Studies in mathematical psychology, No. 1. Copenhagen; Danish Institute for Educational Research:1960.
77. Hagquist C, Bruce M, Gustavsson JP. Using the Rasch model in nursing research: an introduction and illustrative example. International Journal Nursing Studies. 2009;46:380-93.
78. Rothman M, Burke L, Erickson P, Leidy NK, Patrick DL, Petrie CD. Use of existing patient-reported outcome (PRO) instruments and their modification: the ISPOR good research practices for evaluating and documenting content validity for the use of existing instruments and their modification PRO task force report. Value in Health. 2009;12:1075-83
79. Polit DF, Beck CT. The content validity index: are you sure you know what's being reported? Critique and recommendations. Research in Nursing & Health. 2006;29:489-97.

80. Polit DF, Beck CT, Owen SV. Is the CVI an acceptable indicator of content validity? Appraisal and recommendations. Research in Nursing & Health. 2007;30:459-67.
81. Wild D, Grove A, Martin M, Eremenco S, McElroy S, Verjee-Lorenz A, et al. Principles of good practice for the translation and cultural adaptation process for patient-reported outcomes (PRO) measures: report of the ISPOR task force for translation and cultural adaptation. Value in Health. 2005;8(2):94-104.
82. Schuck P. Assessing reproducibility for interval data in health-related quality of life questionnaires: which coefficient should be used? Quality of Life Research. 2004;13:571-86.
83. U.S. Department of Health and Human Services Food and Drug Administration. Guidance for Industry. Patient Reported Outcome Measures: Use in medical product development to support labeling claims. 2009. Available at: http://www.fda.gov/downloads/Drugs/Guidances/UCM193282.pdf
84. Silverman E, Woloshin S, Schwartz LM, Byram SJ, Welch HG, Fischhoff B. Women's views on breast cancer risk and screening mammography: a qualitative interview study. Medical Decision Making. 2001;21:231-40.
85. Willis K. "I come because I am called": recruitment and participation in mammography screening in Uppsala, Sweden. Health Care for Women International. 2008;29:135-50.
86. Törnberg S, Nyström L. Värre med utredningsorsakad oro än förtidig död i bröstcancer? Läkartidningen 2009;45:3018.
87. Statens Beredning för Medicinsk Utvärdering. Gör hälsokontroller nytta? Vetenskap & praxis. 1993;4.
88. Irwig L, McCaffery K, Salkeld G, Bossuyt P. Informed choice for screening: implications for evaluation. British Medical Journal. 2006;332:1148-50.
89. Creswell JW. Research design: qualitative, quantitative, and mixed methods approaches. 2. ed. Thousand Oaks: Sage; 2003.
90. Swaine-Verdier A, Doward LC, Hagell P, Thorsen H, McKenna SP. Adapting quality of life instruments. Value in Health. 2004;7:27-30.
91. European group for quality of life and health measurement. European Guide to the Nottingham Health Profile. Brookwood Medical Publications, Surrey; 1993.
92. Lagerlund M. Factors affecting attendance at population-based mammography screening. Stockholm: Karolinska University; 2002.
93. Declaration of Helsinki – Ethical principles for medical research involving human subjects. Available at: http://www.wma.net/en/30publications/10policies/b3/
94. Graneheim UH, Lundman B. Qualitative content analysis in nursing research: concepts, procedures and measures to achieve trustworthiness. Nurse Education Today. 2004;24:105-12.
95. Andrich D, Sheridan B & Luo G. Interpreting RUMM2030. 4th edition. RUMM Laboratory Pty Ltd, Petrh; 2009.
96. Lipsey MW, Wilson D. Practical meta-analysis. London: SAGE; 2001.
97. Cohen J. Statistical power analysis for the behavioral sciences. 2. ed. Hillsdale: L. Erlbaum Associates; 1988.
98. Andrich D & Marais I. Advanced course in Rasch Measurement of Modern Test Theory, Learning guide 7 and 8. EDUC8606. The University of Western Australia; 2011.

99. Linacre L. Size vs. significance: Standardized chi-square fit statistic. Rasch Measurement Transactions. 2003;17:1:918. Available at: http://www.rasch.org/rmt/rmt171n.htm
100. Linacre L. Sample size and item calibration stability. Rasch Measurement Transactions. 1994;7:4:328. Available at: http://www.rasch.org/rmt/rmt74m.htm
101. Hobart JC, Riazi A, Lamping DL, Fitzpatrick R, Thompson AJ. Improving the evaluation of therapeutic interventions in multiple sclerosis: development of a patient-based measure of outcome. Health Technology Assessment. 2004;8:1-59.
102. Fayers PM, Machin D. Quality of life: assessment, analysis and interpretation. Chichester: Wiley; 2000.
103. Fitzpatrick R, Ziebland S, Jenkinson C, Mowat A. Transition questions to assess outcomes in rheumatoid arthritis. British Journal of Rheumatology. 1993;32:807-11.
104. Hosmer DW, Lemeshow S. Applied logistic regression. 2. ed. New York: Chichester: Wiley; 2000.
105. Hsieh HF, Shannon SE. Three approaches to qualitative content analysis. Qualitative Health Research. 2005;15:1277-88.
106. Cano SJ, Hobart JC. The problem with health measurement. Patient Prefer Adherence. 2011;5:279-90.
107. Brodersen J, Meads D, Kreiner S, Thorsen H, Doward L, McKenna SP. Methodological aspects of Differential Item Functioning in the Rasch Model. Journal of Medical Economics. 2007;10:309-24.
108. Ekelund G, Manjer J, Zackrisson S. Population-based screening for colorectal cancer with faecal occult blood test – do we really have enough evidence? International Journal of Colorectal Disease. 2010;25:1269-75.
109. Heckman BD, Fisher EB, Monsees B, Merbaum M, Ristvedt S, Bishop C. Coping and anxiety in women recalled for additional diagnostic procedures following an abnormal screening mammogram. Health Psychology. 2004;23:42-8.
110. Beauchamp TL, Childress JF. Principles of biomedical ethics. 5. ed. New York: Oxford University Press; 2001.
111. Ong GJ, Austoker J, Michell M. Early rescreen/recall in the UK national health service breast screening programme: epidemiological data. Journal of Medical Screening. 1998;5:146-55.
112. Clutton S, pakenham KI, Buckley B. Predictors of emotional well-being following a false positive breast cancer screening result. Psychology and Health. 1999;14:263-75.
113. Saltz T, DeFrank JT, Brewer NT. False positive mammograms in Europe: do they affect reattendance? Breast Cancer Research and Treatment. 2011;127:229-31.
114. Mathieu E, Barratt AL, McGeechan K, Davey HM, Howard K, Houssami N. Helping women make choices about mammography screening: an online randomized trial of a decision aid for 40-year-old women. Patient Education Counseling. 2010;81:63-72.
115. Mathieu E, Barratt A, Davey HM, McGeechan K, Howard K, Houssami N. Informed choice in mammography screening: a randomized trial of a decision aid for 70-year-old women. Archives of Internal Medicine. 2007;167:2039-46.
116. Hewitson P, Glasziou P, Irwig L, Towler B, Watson E. Screening for colorectal cancer using the faecal occult blood test, hemoccult. The Cochrane Database of Systematic Reviews. 2007;1:CD001216.

117. Raffle AE, Alden B, Quinn M, Babb PJ, Brett MT. Outcomes of screening to prevent cancer: analysis of cumulative incidence of cervical abnormality and modelling of cases and deaths prevented. British Medical Journal. 2003;326:1-5.
118. Idestrom M, Milsom I, Andersson-Ellstrom A. Women's experience of coping with a positive Pap smear: A register-based study of women with two consecutive Pap smears reported as CIN 1. Acta Obstetrica et Gynecologica Scandinavica. 2003;82:756-61.
119. Dillner J. Cervical cancer screening in Sweden. European Journal of Cancer. 2000;36:2255-9.
120. Socialstyrelsen. Screening för tjock- och ändtarmscancer. Rekommendation och bedömningsunderlag. Remissversion. 2013. Available at: http://www.socialstyrelsen.se/SiteCollectionDocuments/screening-tjockandtarmscancer-rekommendation.pdf
121. Wegwarth O, Gigerenzer G. Overdiagnosis and Overtreatment: Evaluation of what physicians tell their patients about screening harms. JAMA Internal Medicine. 2013.
122. Aberle DR, Adams AM, Berg CD, Black WC, Clapp JD, Fagerstrom RM, et al. Reduced lung-cancer mortality with low-dose computed tomographic screening. The New England Journal of Medicine. 2011;365:395-409.
123. Heleno B, Rasmussen JF, Brodersen J. Reduced lung-cancer mortality with CT screening. The New England Journal of Medicine. 2011;365:2036; author reply 7-8.

Populärvetenskaplig sammanfattning

Bröstcancer är den vanligaste dödsorsaken bland cancersjuka kvinnor. Därför rekommenderas regelbundna mammografiska hälsokontroller för tidig diagnostik och snar behandling för att förbättra prognosen. I Sverige rekommenderas kontroller för kvinnor mellan 40-74 år. Mammografiska hälsokontroller medför dock även risker, såsom till exempel behov av kompletterande bröstundersökning(ar) till följd av fynd i samband med hälsokontrollen för att slutligen fria kvinnan från misstankar om bröstcancer, dvs. falskt positiv mammografisk hälsokontroll. Det har uppskattats att bland 1000 kvinnor som deltar i 10 mammografiska hälsokontroller i Europa, kommer 200 att erfara falskt positiv hälsokontroll. Dessa kvinnor kan uppleva oro, rädsla eller ångest som följd av kompletterande undersökningar. Ett flertal studier har undersökt detta och visar på förekomst av kortsiktiga psykosociala konsekvenser. Det är däremot mer oklart hur kvinnor påverkas på sikt eftersom befintliga studier visat motstridiga resultat. Detta kan till exempel bero på att frågeformulär med oklar tillförlitlighet (reliabilitet och validitet) användes. Nyligen utvecklades ett frågeformulär i Danmark (Consequences of Screening – Breast Cancer, COS-BC) för att specifikt undersöka kort- och långsiktiga psykosociala konsekvenser av falskt positiv mammografisk hälsokontroll. Detta ger möjlighet till kartläggning av dessa konsekvenser på ett sätt som inte varit möjligt tidigare. Frågeformuläret behöver dock först översättas och anpassas till svensk kontext, samt testas avseende dess mätegenskaper (validitet och reliabilitet).

Trots flera publicerade studier är det fortfarande oklart hur kvinnor upplever falskt positiv mammografisk hälsokontroll. Genom att undersöka kvinnors upplevelser i sådan situation kan en djupare förståelse av fenomenet erhållas, inklusive vilka strategier som kvinnor använder för att hantera situationen. Vidare kan en kartläggning av faktorer som kan vara av betydelse för utveckling av konsekvenser göra det möjligt att identifiera kvinnor med försämrad förmåga att hantera falskt positiv mammografisk hälsokontroll. Sammantaget kan denna kunskap leda till utveckling av åtgärder som kan förebygga eller lindra psykosociala konsekvenser bland kvinnor

med ökad risk för utveckling av dessa oönskade effekter av mammografiska hälsokontroller.

I denna studie undersöktes psykosociala konsekvenser av falskt positiv mammografisk hälsokontroll i Sverige med hjälp av COS-BC, som testades avseende validitet och reliabilitet i svensk kontext, samt hur kvinnor hanterar att erfara falskt positiv mammografi.

Resultaten visade att; merparten av frågorna i en svensk version av COS-BC ansågs relevanta och heltäckande avseende konsekvenser av falskt positiv mammografisk hälsokontroll. Dock utvecklades två nya frågor, tre exkluderades pga bristande relevans och tvetydighet och elva frågor modifierades för att förbättra förståelsen.

Utvärdering av mätegenskaperna för den svenska COS-BC visade på god giltighet (validitet) och acceptabelt mätfel (reliabilitet) i fem domäner (nedstämdhet, ångest, beteende, sömn och existentiella aspekter). Dessa domäner användes därför för att undersöka förekomsten och långtidsutvecklingen av psykosociala konsekvenser av falskt positiv mammografisk hälsokontroll och dess potentiella bidragande orsaker. Resultaten bekräftade tidigare resultat, dvs kortsiktiga konsekvenser i anslutning till kompletterande bröstundersökningar. Dessutom sågs betydande konsekvenser inom samtliga fem domäner 6 och 12 månader efter avslutad utredning. Sannolikheten att uppleva dessa konsekvenser var mer är två gånger så höga hos kvinnor med falskt positiv jämfört med kvinnor med negativ (normal) mammografisk hälsokontroll. Samma mönster sågs även för samtliga övriga delar av COS-BC, förutom självundersökning av brösten. Dessa resultat bör dock tolkas med försiktighet, eftersom det inte fanns stöd för mätegenskaperna hos dessa domäner.

Att kontrolleras efter falskt positiv mammografi med tätare intervall än ordinarie hälsokontroll var den största bidragande orsaken till att uppleva långsiktiga konsekvenser av falskt positiv mammografisk hälsokontroll (>3 gånger högre sannolikhet). Andra bidragande orsaker avseende beteende, sömn, upplevelser av nedstämdhet och ångest samt tankar om existentiella aspekter var kvinnornas kulturella bakgrund, missnöje med informationen i samband med kompletterande bröstundersökningar och brist på socialt stöd.

Intervjuer med 13 kvinnor som upplevde kort- och/eller långsiktiga konsekvenser av falskt positiv mammografisk hälsokontroll visade att deras hantering av dessa konsekvenser upplevts som en berg- och dalbana av känslor och förnuft. Kvinnor beskrev att de befarade det värsta och levde i osäkerhet medan de väntade på diagnosen. Dessa upplevelser förekom till följd av känslor att vara hotad av en dödlig sjukdom. Å andra sidan upplevde kvinnorna trygghet hos familjen och genom att vara professionellt omhändertagna, vilket tillsammans med upplevelsen av syster-

skap och självuppmuntran frambringade styrka och hopp. Att vara medveten om ansvaret för familjen kom fram som en viktig aspekt. Erfarenheten av falskt positiv mammografisk hälsokontroll gav känslor av tacksamhet att vara noggrant undersökt, även om oron kvarstod. Vunnen insikt avseende mammografiska hälsokontroller och aspekter som är viktiga att värdesätta i livet kom också fram i intervjuerna.

Mammografiska hälsokontroller innebär ett ansvar hos verksamheten och beslutsfattare avseende kontinuerlig granskning och utvärdering av såväl fördelar som risker. Studiens resultat ger ny och fördjupad kunskap om en av riskerna med hälsokontrollerna; förekomst och långtidsutveckling av psykosociala konsekvenser av falskt positiv mammografi. Detta skapar även förutsättningar för stödåtgärder (exempelvis minimering av antalet tätare kontroller) med potential att förebygga utveckling av psykosociala konsekvenser hos kvinnor med falskt positiv mammografisk hälsokontroll, vilket även kan visa sig vara kostnadseffektivt. Vidare, att erbjuda personlig information, i synnerhet till riskgrupper (exempelvis utrikes födda kvinnor) kan utöver potentiell minskning av psykosociala konsekvenser även vara till gagn i ett längre perspektiv genom exempelvis ökad medvetenhet om bröstcancerscreeningen. Visserligen kan informationen om nackdelar med hälsokontrollerna potentiellt minska kvinnors deltagande i hälsokontrollerna, men denna typ av åtgärd beaktar kvinnornas integritet och rättighet till informerat samtycke att medverka i hälsokontrollerna. Å andra sidan, visar resultaten att kvinnor har förtroende för hälso- och sjukvården och avser att fortsatt delta i hälsokontroller trots erfarenheten av falskt positiv mammografi. Hög deltagarfrekvens i hälsokontrollerna är en av förutsättningarna för programmets effektivitet att upptäcka bröstcancer i tidigt stadium och därmed erbjuda tidig och förhoppningsvis mer effektiv behandling för att förbättra prognosen för individen och minska bröstcancerdödligheten i befolkningen. Därför finns det förutsättningar att resultaten i denna studie kan utgöra ett underlag för åtgärder som inte bara är fördelaktiga ur ett kvinnoperspektiv, utan även kan visa sig effektiva utifrån folkhälsoperspektivet.

Acknowledgements

My sincere gratitude goes to those without whom this thesis would have never seen the light of day.

Each and every woman who participated in the studies of which the thesis consists of – thank you for responding to the study questionnaires and sharing your experiences of false-positive screening mammography with me.

I am with beyond words of gratitude to my supervisors for the trust and freedom you have provided me with during the process of creating a researcher. I have undeniably tried to find my words to describe the tremendous support you have given me, but, no matter how much I have tried, the words will never fully mirror your support, nor will they ever reflect my feelings of appreciation. It was not a matter of that the story needed to be long, but it has been extremely hard to make it short.

Peter Hagell – psychometrics was a blurry phenomenon for me not so very long ago, your passion for psychometric analyses is contagious. Your crystalline way of thinking, open-mindedness, unpretentiousness and never ending enthusiasm make you an outstanding supervisor! Thank you for the challenges you set up that encouraged me to become a researcher. I cherish our meeting breaks conversing about Warszawa, feel-good films, and "self-cleaning" espresso machines.

Sophia Zackrisson – you have been a shoulder to cry on and a lifebuoy all the times when I thought there was no way out! Thank you for teaching me the huge importance of communicating research findings and how to do it in a comprehensible manner. Thank you for your excellent guidance through the field of mammographic screening.

Christine Wann-Hansson – you have guided me through the content analysis of the interview transcripts and challenged me to think in non-quantitative terms during my work with the thesis. Thank you for your questions, turning my way of thinking upside-down a lot of times, but always compelling me to create a new sound!

A special gratitude goes to, and in no particular order:

John Brodersen, my co-author in two papers – thank you for your help with adapting the COS-BC questionnaire to Swedish. Your cheerful way of being was a pleasant addition to our discussion meetings.

Gunbrith Borgström, my colleague and friend – over ten years ago you told me about the courses in Health Science and you knew long before me that I would become passionate about the new world the courses revealed. Thank you for believing in me! I appreciate your help in the enrolment of the study participants.

Eva Prahl – thank you for your priceless help with questionnaire distribution, transcribing the interviews, correcting the manuscripts, and limitless tiny little things a PhD student needs help with. Our talks about shoes, clothes, fashion, etc were always nice breaks.

Olle Ekberg, my boss – thank you for your never ending support! Working with you has always given me enormous satisfaction.

Lars Bååth, the executive of the Diagnostic Centre of Imaging and Functional Medicine Skåne University Hospital in Malmö, *Per Olof Iwars*, the health care manager at the department, and *Peter Leander*, previous executive of the department – thank you all for being so generous to provide me with the most optimal environment to work with the thesis.

Health care staff at the mammography facility Unilabs in Malmö – through countless days you have been patient with my presence in the middle of your busy daily work schedule, thank you for being tremendously helpful in the enrolment of the participants. My appreciation also goes to the executives of the facility for providing valuable data of the screening programme.

Women in the dual panel translation – thank you for your time and contribution in the translation of the study questionnaire.

Health care staff at the breast cancer surgery clinic Skåne University Hospital in Malmö – thank you for helping me to collect medical records.

My colleagues at the Diagnostic Centre of Imaging and Functional Medicine Skåne University Hospital in Malmö – working with you has always been encouraging and fun. A special acknowledgement goes to *Heléne Kjellström* and *Anna Johansson.*

My colleagues in the seminar group at the Department of Health Sciences, Lund University – a thousand words in appreciation for your constructive feedback.

Wojciech Jan Bołejko, my beloved husband, my soul mate – my respect to you for always standing by me and I know you always will. "Serce Twe busola ma"[1]. *Maximilian Bołejko*, my precious son, my heart – thank you for helping me with the language review. I was blessed the moments you both came into my life!

This thesis was supported by grants from the Diagnostic Centre of Imaging and Functional Medicine Skåne University Hospital in Malmö, Vårdakademin Skåne University Hospital in Malmö, the Faculty of Medicine Lund University, the Swedish Research Foundation, and the Foundation for Cancer Research at the Department of Oncology at the Malmö University Hospital.

1 Lyrics by Wojciech Młynarski