Study protocol

Official title: Breast Cancer Screening With MRI in Women Aged 50-75 Years With Extremely Dense Breast Tissue: the DENSE Trial

NCT number: NCT01315015

Document date: 29 November 2010, revision 7 June 2011

Protocol

This trial protocol has been provided by the authors to give readers additional information about their work.

Protocol for: Bakker MF, de Lange SV, Pijnappel RM, et al. Supplemental MRI screening for women with extremely dense breast tissue. N Engl J Med 2019;381:2091-102. DOI: 10.1056/NEJMoa1903986

Supplementary File

This supplement contains the following items:

1. Original protocol, this is also the final protocol.

   This is the protocol that was approved by the Ministry of Health, Welfare and Sport on 11th November 2011. For research within the breast cancer screening program in the Netherlands a permit is required under the Dutch Population Screening Act. The Ministry of Health, Welfare and Sport issued this permit after having been advised by the Netherlands Health Council. This advice includes Medical Ethical judgment of the protocol and replaces the Ethical Assessment by an Institutional Review Board (IRB).

2. The original statistical analysis plan is also the final statistical analysis plan and is included in the protocol (see paragraph 7. 'Sample size, feasibility' and paragraph 9. 'Outcome analysis, evaluation').

Ministry of Health, Welfare and Sport FILE NO. (leave blank)	
Application form to apply for a permit issued under the Dutch Population Screening Act. Please print and send form signed and dated to the Ministry of Health, Welfare and Sport, Public Health Department, PO box 20350, 2500 EJ, The Hague. The annexes and literature can be sent digitally together with the application form to: bevolkingsonderzoek@minvws.nl	
Part 1. General data	**Sending date: 29 November 2010 Sending date revision: 7 June 2011**
1. population-based screening program title	Breast cancer screening with MRI of women with high mammographic breast density between the age of 50 and 75
2. applicant	name: Dr Carla H. van Gils name contact person: Dr Carla H. van Gils correspondence address: PO box 85500, Str.6.131 postcode: 3508 GA city: Utrecht telefax: 088-755 5485
3. project leader	name: Dr Carla H. van Gils position: Senior University Lecturer Clinical Epidemiology of Cancer telephone number: 088-7553014 or 088-7559301 (Sec.) institution: Julius Center for Health Sciences and Primary Care correspondence address: PO box 85500 postcode: 3508 GA city: Utrecht telefax: 088-755 5485
4. type of population-based screening program	A permit is required because it is a population-based screening program O a. whereby ionising radiation is applied; ● b. on cancer; O c. on diseases or abnormalities which can neither be treated nor prevented. If category a: If the Nuclear Energy Act grants a permit for the X-ray machine, enclose a statement of the permit.

	If category c: Which exceptional circumstances would warrant the research? (max. 10 lines) The population-based screening program is also a medical scientific research: ● yes, then answer Part 7 as well O no
5. starting and end date of the population-based screening program	Starting date: 1 September 2011 End date: 1 September 2019
6. funding	If applicable, is the funding complete? ● yes, for the first screening round, contact person Dr C.H. van Gils O no, contact person:
7. signature	applicant: research or project leader:
Part 2. Justification for the population-based screening program (max. 300 lines)	
	1. Motivation, key figures Breast cancer is one of the most common cancers in women in the Netherlands, with an incidence of about 13,000 new invasive tumours per year. Each year, 3,180 women die as a result of this disease. In addition, breast cancer is sixth on the list of diseases with the greatest burden of disease for women (Source: www.nationaalkompas.nl). Over the past decades, new treatment options and the population-based mammography screening program for breast cancer have made an important contribution to the reduction of breast cancer mortality rates (together, reduction of 25%) (1). However, when we look at the current population-based screening program among women between the age of 50 and 75, still a third of the breast tumours are diagnosed between two screening rounds, the so-called interval cancers. These are tumours detected because they are palpable and/or cause complaints. Thus, these tumours did not benefit from the population-based screening program, are already at a more advanced stage when diagnosed, and have a significantly worse survival probability than tumours discovered during the population-based screening program (2). Women with extremely dense breast tissue >75% density, caused by a large amount of fibroglandular and stromal tissue in the breast, about 5% of all women in the screening age), are a high-risk group for breast cancer. Their chances at developing breast cancer is 4 to 6 times higher than those of women with low mammographic breast density (primarily fatty tissue in the breast) (3). On top of that, it is precisely this group, the highest-risk

group, where mammographic examination is of limited value because the fibroglandular and stromal tissue can easily mask the presence of a possible tumour. Various studies have shown that the sensitivity in the population-based screening program with mammography is significantly lower among women with very dense breasts (4-8). The study conducted by Kerlikowske et al. (6) indicates that the mammographic screening sensitivity in women aged 50 to 69 years is 65% among women with >75% density compared to 89% among women with <25% dense tissue. Or, in other words, the risk of an interval cancer for women with >75% density is almost 10 times higher than for women with <25% dense tissue. Boyd et al. even present a probability of almost 18 times higher for women with >75% density at getting interval cancers compared to women with <10% dense tissue (4). This dual effect of high breast density on the development of cancer on the one hand and on the sensitivity of mammography in this group on the other hand, means that the group with the highest risk at getting breast cancer is screened using a technique that has limited value to them, namely mammography.

Women with very dense breasts could benefit more from sensitive imaging techniques, such as digital mammography, ultrasound and magnetic resonance imaging (MRI). Digital mammography seems to be more sensitive than conventional mammography, but in a large study among screening participants, this effect seemed to be limited to women under the age of 50 with very dense breasts. There was no difference between the digital and conventional mammography for women above the age of 50, even for the group with very dense breasts (9). Ultrasound seems to result in higher detection rates among women with very dense breasts and a higher risk for breast cancer based on the Gail or Claus risk assessment models (10). However, the increased sensitivity goes hand in hand with a significantly higher number of false positives. Furthermore, the reproducibility of this technique depends a great deal on the person operating the machine, and the ultrasounds are preferably carried out by the radiologists themselves instead of technologists, which significantly raises the costs of this study.

To this day, the value of MRI for screening purposes has only been investigated in (young) women with a BRCA1 or 2 mutation or an otherwise very strong family history of breast cancer. In a number of studies where women were examined with (digital) mammography, ultrasound as well as MRI, the sensitivity of the MRI was significantly higher than that of the other techniques (11-14).

2. Objective of the population-based screening study

The proposed experimental population-based screening study entails that the participants of the current population-based screening program with the highest risk (>75% density) undergo an additional MRI examination if their mammographic screening does not show any abnormalities. The aim of this is to detect any possible tumours in an earlier stage in this high-risk group and to reduce the number of interval cancers in this group, to reduce the number of deaths due to breast cancer and to increase the quality of life of patients with breast cancer.

This could in particular be effective because the proposed population-based screening study focuses on the group with the highest risk, for whom, at the same time, the effectiveness of the mammography is the most limited.

The intended effect on a national level is as follows:
Every year, there are about 900,000 women in the Netherlands between the age of 50 and 75 who are screened using mammography as part of the Dutch population-based mammography screening program. About 5% of these women have extremely dense breasts (>75% density), so 45,000 women. It is expected that adding an MRI to the mammography screening in this high-risk group will decrease the number of interval cancers from 4.4 per 1000 screened women to 1.44 per 1000 screened women (1;6) (calculations are explained further in section 7.7). On a national level, this means that the number of interval cancers per screening round in this group decreases from 199 to 65.

The prognosis of patients with tumours discovered through screening is better than that of patients with interval cancers, with an estimated 5-year survival rate of 95% and 78% respectively (2). On the basis of these figures, we estimate that 23 deaths due to breast cancer could be prevented per screening round: 134 tumours (199 minus 65) detected now through screening instead of in the interval, and of which as a result thereof, 17% fewer breast cancer deaths in 5 years (95% minus 78%).

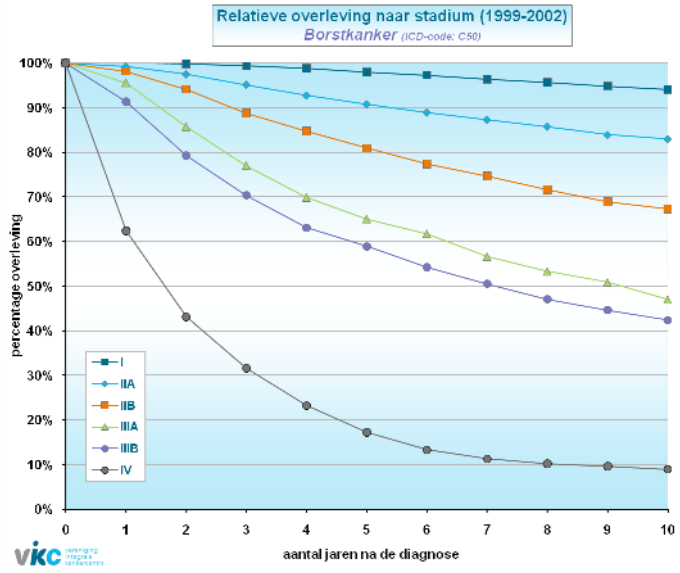
From previous research in the screening population of 50-70 years, we know that the average age of women with >75% density is around 53 years. If we calculate with an average life expectancy of another 27 years, an additional 621 extra life years are gained per round with the proposed population-based screening study.

In addition, this population-based screening study will have a favourable effect on the quality of life of breast cancer patients, considering a large part of the interval cancers (67%) will be diagnosed sooner through screening, which means an aggressive treatment will be required less frequently.

3. Natural course of the disease

About 13,000 women in the Netherlands are diagnosed with breast cancer annually (Source: www.ikcnet.nl 2009). Three quarters of these women are 50 years or older (± 9,750). Important predictive factors for the prognosis of breast cancer are tumour diameter, axillary lymph node status and histologic tumour grade (1;15). In the Netherlands, women between the age of 50 and 75 are screened every two years using mammography as part of the Dutch population-based mammography screening program. The population-based screening program detects about 5,000 breast tumours each year (51% of all tumours among women older than 50) (1). However, not all tumours are detected through the population-based screening program: Among the participants, about 30% of the tumours are discovered in the interval between the screening rounds (the so-called interval cancers) (1). This large number of interval cancers is a problem since the staging, and with that the prognosis, is less favourable compared to the

	tumours discovered through screening (1;2). In a patient monitoring study pilot in the Dutch screening organisation “Zuid-West”, data on the staging differences between screening tumours and interval cancers have been collected. Interval cancers are more frequently stage III (12%) or IV (15%) compared to screen-detected tumours (4% and 6% respectively) (1). A Finish study shows that the 5-year survival probability is 95% for women with screen-detected tumours and 78% for interval cancers (2). Women with very dense breasts have a higher risk of developing screen-detected tumours and even more so of developing interval cancer (4;5;7;8). Five per cent of all women who participate in the population-based screening program have very dense breasts (>75%) and 13% of the interval cancers is found in this group (6). Women with very dense breasts most likely have a higher chance at developing interval cancers because the large amount of fibroglandular and stromal tissue in the breast could mask the potential presence of a tumour during mammographic screening. There are no indications this would be caused by tumours growing faster in this group of women (4;16).
	4. Treatment options and prognosis The most commonly used treatments for breast cancer are: surgery, radiotherapy, chemotherapy, hormonal therapy and treatment through monoclonal antibodies. Most women with breast cancer receive a combination of the above-mentioned treatment methods. The choice and the sequence of the different treatments depends, among others, on the characteristics of the tumour, the stage of the disease, age and menopausal status. The breast cancer prognosis after treatment depends on the stage at diagnosis (see graph below). The 10-year survival probability varies from very good (>90%) at stage I to very bad at stage IV (<10%) (see graph). Treatment is not different for different density categories since there are no indicators of differences in prognosis of tumours in the same stage. The prognosis of interval cancers, which also occur more frequently among women with very dense breasts, is worse than that of tumours discovered through regular screening, because the stage at diagnosis is less favourable (1;2) (see *Section 3. Natural course of the disease*).



5/ 6. not applicable

7. Population-based screening program alternatives

Primary prevention
For breast cancer, no real strong risk factors are known, such as is the case with smoking and lung cancer. The known risk factors for only slightly increase the risk of breast cancer and heredity only plays a clear role in 5 to 10% of the women with breast cancer. Known risk factors are, for example: age above 50, various hormonal factors (early age of menarche, late menopause, not having children, etc.), being overweight after menopause, physical inactivity, alcohol use and having very dense breasts. Of all the risk factors, high breast density is, following heredity, associated with breast cancer the most. Reducing the density with medicines such as Tamoxifen or Raloxifene, could be a way to reduce the risk of developing breast cancer for women with very dense breasts (17). However, these medicines increase the risk of developing other conditions such as endometrial cancer and thromboembolism, and cause vasomotoric side effects. It is not yet known how long these medicines need to be taken for there to be a protective effect. As a result, these medicines are rarely used, even in the United States where these medicines have been approved by the FDA (18).

Curative treatment
In terms of curative treatment, no breakthroughs are expected in the near future that could significantly improve the diagnosis of breast cancer discovered in a relatively unfavourable stage (stage IIb, III, IV). This makes it all the more important to trace breast cancer in an early stage, so the prognosis is most favourable. The aim of this screening study is to contribute to early detection of breast cancer and with that, enhance the survival probability.

Part 3. Screening strategy
(max. 150 lines)

1. Target group

The target group of the proposed experimental population-based screening study are women at the age of 50 to 75, who participate in the current Dutch population-based mammography screening program. Women qualify for the examination if their mammographic breast density is higher than 75% (this is the case for about 5% of the screening participants) and if they have a negative (=without abnormalities) screening mammography. This is a mammography classified in the population-based screening program as BI-RADS (Breast Imaging Reporting Data System) category 1 or 2 according to the mammographic classification of the American College of Radiology (www.acr.org). The breast density is quantified using a method developed by N. Karssemeijer (19). This method has been validated against MRI for the GE Senographe mammograph (19).

Women with a contraindication to MRI with contrast media (Gadolinium) (for example, metal objects in the body such as a pacemaker), cannot take part in the trial.

2. Screening method, test characteristics

To this day, the value of MRI for screening purposes has only been investigated among (young) women with a BRCA1 or 2 mutation or an otherwise very strong family history of breast cancer.

A recent meta-analysis of these studies (20) shows a sensitivity of 94% (95%CI 90-97) for MRI and mammography combined and 39% (95%CI 37-41) for mammography only. The combination with MRI resulted in more false positives than mammography only (specificity 77% 95%CI 75-80% and 95% 95%CI 93-96% respectively). The specificity did, however, improve in the following screening rounds. Although survival data are still not available, the more favourable staging of the tumours discovered by adding the MRI predicts a significant reduction in breast cancer mortality. Based on these data, it is recommended that women with an increased risk of developing breast cancer (20-25% or greater) are screened annually with mammography and MRI from the age of 30. In the Dutch MRISC study on the effect of MRI screening among women with a family history of breast cancer, the results were also stratified into two groups of breast density (high and low) (21). The mammography and MRI combined led to detecting more tumours than mammography only among women with a low as well as women with a high breast density. Lehman et al. (22) investigated the value of MRI of the contralateral breast of women recently diagnosed with breast cancer. This research showed that MRI could identify tumours that were missed by mammography (3.1%) with high sensitivity in both dense (87%) and nondense (100%) breasts.

Screening with ultrasound also seems to result in higher breast cancer detection rates among women with an increased breast density and an increased risk -of developing breast cancer based on the Gail or Claus risk assessment models (10). Mammography combined with ultrasound had a

sensitivity of 77.5% compared to a sensitivity of 50% for mammography only. However, the increased sensitivity goes hand in hand with a significantly increased number of false positives (specificity 89.4% and 95.5% respectively). We selected MRI (in combination with mammography) for this experimental population-based screening study and not ultrasound as screening modality among women with very dense breasts for the following reasons:

- In several studies where women were examined using (digital) mammography, ultrasound as well as MRI, the sensitivity of MRI proved to be significantly higher than that of the ultrasound (11-14;23). For example, in the recent study (2010) by Kuhl et al. (11), the sensitivity of the combination of MRI + mammography was 100% compared to 48% for ultrasound + mammography. The positive predictive value of MRI is also considerably higher than that of ultrasound (48% versus 36%) (11), and the proportion breast cancers after biopsy too (25% versus 15%) (14). Based on this literature, the expected effectiveness of screening with MRI among women with very dense breasts is significantly higher than that of ultrasound.
- The reproducibility of ultrasound is lower and depends a great deal on the person operating the machine (24). For this reason, it is recommended that ultrasound screening of the breast is carried out by radiologists instead of technologists. Taking account of the fact that ultrasound screening of the breast is laborious (at least 15 minutes per breast) (24), the ultrasound screening costs could even be slightly higher than that of MRI with contrast (14).

In this experimental population-based screening study, we do not want to replace mammography by MRI, considering the sensitivity of the MRI and mammography combined are significantly higher than that of the separate modalities in studies among women with BRCA1 or 2 mutations (20). In addition, DCIS tumours are in general easier to detect by mammography (25;26).

3. Cut-off points, positivity criteria

Mammography and MRI results are typically classified in accordance with the American College of Radiology classification system, the Breast Imaging Reporting and Data System (BI-RADS) (www.acr.org; under Quality and Safety resources, BI-RADS atlas). A negative screening mammography is an inclusion criterion for this study, which means that per definition, all mammographic screenings have been classified as either BI-RADS 1 (=negative) or 2 (=benign finding). With the MRI assessment, the mammography results may be used, as well as those of previous examinations. The MRI assessments MRI BI-RADS 1 (=negative) and 2 (=benign finding) give no indication for further follow-up. These women are invited again after 2 years to take part in the population-based screening study. With MRI BI-RADS 3 (=probably benign abnormality), the MRI examination is repeated after six months. There is a positive finding with the MRI if it is classified as MRI BI-RADS 4 (=suspicious abnormality) or 5 (=highly suggestive of malignancy). These abnormalities will cytologically or histologically need to be confirmed. Women with a positive MRI result will undergo an ultrasound-guided or MRI-guided biopsy. The MRI BI-RADS 0

result (=incomplete examination) is merely a temporarily result, and completion of this examination will take place as soon as possible.

4. Technical realisation

MRI examinations are made no later than 8 weeks after participating in the population-based mammography screening program. For premenopausal women (women with at least 1 menstruation in the last 30 days), the MRI examination is scheduled between day 7 and 14 of the menstrual cycle as much as possible to avoid hormonal-induced enhancement.

All women in the MRI group will be scanned according to a “state-of-the-art” MRI protocol. Prior to the start of the trial, this protocol is standardised in the participating centres. The scans will be interpreted by experienced radiologists, using the under 3.3. referred to MRI BI-RADS criteria.

The MRI BI-RADS 3 results will be reassessed centrally *before* the results are reported back to the woman. With this double reading, by definition, the highest BI-RADS score is taken as the final result. The reason for reassessment of the BI-RADS 3 results is that this is typically the most difficult assessment category for the radiologists. This is because this is a very heterogeneous group of lesions, for which no clear classification category exists. All participating radiologists receive training to standardise the MRI procedures and assessment of MRI examinations.

5. Screening interval

The proposed experimental population-based screening study using mammography and MRI among women with very dense breasts will run parallel to the current Dutch breast cancer screening programme (2 year screening interval). This screening interval was chosen in the past based on a careful consideration of the costs and benefits. We continue with this, because there are no indications that women with very dense breasts in this age category develop other tumours (for example, faster-growing tumours) compared to women with fatty breasts. Besides, by adopting the current screening interval, the existing infrastructure can be used, which simplifies the research on the added value of MRI examinations.

6. Diagnosis

The MRI results lead to different strategies for further diagnostic work-up, as described in the Netherlands National Breast Cancer Dialogue (in Dutch: Nationaal Borstkanker Overleg Nederland, NABON) Mammary Carcinoma Directive (27):

- The MRI BI-RADS 0 result (=incomplete examination) is merely a temporarily result, and a completion of this examination will take place as soon as possible. A reason for a BI-RADS 0 result could be a technically incorrect execution of the MRI examination. The radiology department of the participating hospitals calls on the concerning woman to carry out a new MRI examination.
- When the MRI examination is negative (BI-RADS 1 or 2), further work-up is not required. Women with these scores will be invited again after 2 years for another mammography as part of the current

	population-based mammography screening program and for an MRI examination as part of this experimental population-based screening study. - With a BI-RADS 3 (=uncertain; probably benign abnormality) result, an MRI examination is repeated after six months. The radiology department will contact the woman for this examination. - A positive MRI examination (BI-RADS 4 or 5) requires a cytologic or histologic evaluation. This is done using an ultrasound-guided biopsy, or if the abnormality cannot be traced with a second-look ultrasound, using an MRI-guided biopsy. Cytological and histological evaluation is reviewed in the concerning centres by experienced mamma-pathologists to determine the subsequent clinical strategy. Within the framework of this study, all the material will also be evaluated centrally later. If an abnormality, found using MRI, turns out not to be malignant following further examination, the woman is invited again after 2 years for another mammography as part of the current population-based screening program and for an MRI examination as part of this experimental population-based screening study.
Part 4. Scientific appropriateness (max. 125 lines)	
	1. Efficacy The efficacy of adding MRI to the current breast cancer screening programme among women with very dense breasts is to this day unclear and is to be investigated in this experimental population-based screening study. There is, however, strong evidence from the literature that MRI could be worthwhile for these women. Different studies show that MRI can identify tumours missed by mammography and ultrasound (11-14). The value of MRI for screening purposes has been investigated among (young) women with a BRCA1 or 2 mutation or an otherwise very strong family history of breast cancer. Since young women have dense breasts more often, the results of these studies are important to this experimental population-based screening study. A recent meta-analysis of these studies (20) show a sensitivity of 94% (95%CI 90-97) for the MRI and mammography combined, and 39% (95%CI 37-41) for mammography only. The combination with MRI did result in more false positives than mammography only (specificity 77% 95%CI 75-80% and 95% 95%CI 93-96% respectively). Nevertheless, the specificity did improve in the following screening rounds. Although survival data are still not available, the more favourable staging of the tumours discovered by adding the MRI predicts a considerable reduction of breast cancer mortality rates. That is why the American Cancer Society advises to screen women with a high risk of developing breast cancer annually with MRI in addition to mammography. The first results of the ACRIN 6666 study among women with very dense breasts and an increased risk based on the Gail or Claus risk assessment models, where MRI is added after three years of screening with mammography and ultrasound support these results (23). This study

	shows that many more breast cancers at early stages are found by MRI compared to mammography and ultrasound and the combination of these two. This study also confirms that ultrasound screening or MRI leads to more false positive assessments and unnecessary biopsies than screening with mammography only. The drawback of the above-mentioned studies is the cross-sectional character, i.e. that all women in the studies have been offered all imaging techniques described in the study. Because of this set up, it is not clear whether MRI reduces the number of interval cancers, whether there is a difference in staging of the tumours found, and whether it does lead to a less aggressive treatment. In a randomised design like we propose in this experimental population-based screening study, we actually aim to get a better understanding of these clinically important outcome measures and the number of false positives (and unnecessary biopsies). The randomised design, the outcome measures chosen and the size of this experimental population-based screening study will generate the necessary weight of evidence on the efficacy of adding MRI to the current breast cancer screening programme among women with very dense breasts. **2. Other important implications** Investigating the value of adding MRI to the current breast cancer screening programme among women with very dense breasts is also important for younger women (45-49 years old). These women currently fall outside the Dutch population-based mammography screening program because it is not cost-effective. The screening is not cost-effective partly because, among other reasons, many women in this age group have very dense breasts so the sensitivity of mammography screening is limited. If this experimental population-based screening study reveals that MRI examinations are effective in detecting breast cancer at an early stage among women with very dense breasts, then this will also be important information for a potential cost-effective strategy for women of the age of 45 to 49. Another important effect is obtaining more insight into the relationship between density and breast cancer risk.
Part 5. Description of the population-based screening study (max. 300 lines)	
	1. Design, sites Described in Part 7. **2. Information, communication, complaints, report** Described in Part 7. **3. Objectives, feasibility, actual results** Objective of this experimental population-based screening study is to detect breast cancer in an early and better treatable stage among women with

	very dense breasts, which ultimately reduces the number of breast cancer-related mortalities. The feasibility is described in Section 7. The type of results referred to (Participation rates; Screening Results; Biopsy Rates; Detection Rates; Predictive Values; False Negatives and Variation in Seriousness of the Disease) will all be evaluated in this experimental population-based screening study. **4. Guidelines, working arrangements, quality control, evaluation** The quality of the MRI examinations and their assessment will be monitored by working according to the "state-of-the-art" guidelines and by creating working arrangements. The working arrangements will be concluded during consensus meetings and training sessions where radiologists from all participating centres will take part in. The working arrangements will be incorporated in a concise document that will be distributed to all the centres. This way, we can carry out the MRI reading and assessment thereof as standardised as possible. The data collected will be entered into an online central database (Research Online). This online database will incorporate different validation steps to monitor the quality of the data. The quality of the experimental population-based screening study will be monitored regularly through regular contact between the centres and the principal investigators. The principal investigators will also pay a regular visit to the centres to stay informed of the process made and any arising problems. Besides quality control by the researchers, there will also be an on-site monitoring, recommended by the "Kwaliteitsborging van mensgebonden onderzoek" (English: Quality Control on Research Involving Human Subjects) study group of the Netherlands Federation of University Medical Centres (28). Monitoring will consist of 1 visit per year per centre in accordance with the on-site monitoring guidelines for studies in the risk category: Minimum breach of insignificant risk (28).
Part 6. benefit and harm ratio (max. 175 lines)	
	1. Benefit The benefit refers to the decline in the number of breast cancer deaths by discovering breast cancer in the highest risk groups in an earlier stage than is currently the case. The early detection of breast tumours should, furthermore, lead to lesser invasive treatments and less burden of disease. **2. Harm** The harms primarily consist of false-positive MRI results and, as a result, further follow-up and possible interventions which turned out to be unnecessary. This could create stress and loss of trust. **3. Ratio** By selecting the group of women with the highest breast density, , MRI examination will take place in a group of women that will benefit from it the most, because of their high risk and very limited mammography value. The

	ratio of benefit and harm is determined by the sensitivity of MRI in this group on the one hand, and the number of false positives on the other hand.
Part 7. Medical-scientific research project description (max. 525 lines)	
	1. Projectleader Name contact person: Dr Carla H. van Gils position: Associate Professor Epidemiology institution: Julius Center for Health Sciences and Primary Care, UMC Utrecht correspondence address: PO box 85500 postcode: 3508 GA city: Utrecht telephone number: 088-7553014 telefax: 088-7555485
	2. Problem definition Women with very dense breast tissue (>75% of the breast consists of fibroglandular and stromal tissue) are 4-6 times more likely to develop breast cancer than women with little or no dense tissue (3). On top of that, the sensitivity of mammographic screening is very limited with these women (4-8), so in this group of women, breast cancer presents itself more often as interval cancer compared to the group with low breast density. Interval cancers are tumours that show itself in between two screening rounds; for example, because a woman has complaints. In general, interval cancers have a worse prognosis than tumours detected through screening. The women with the highest risk at developing breast cancer in the population-based screening program are momentarily screened using a method that has a merely limited value for them, and they could probably benefit more from more sensitive screening methods, such as MRI.
	3. Objective Our objective is to detect more breast tumours in an earlier stage among women who participate in the population-based mammography screening program and who have very dense breasts, eventually to reduce the number of breast cancer deaths.

	4. Research question 1) What is the effectiveness of screening every 2 years using mammography combined with MRI compared to mammography only in reducing the number of interval cancers (as proxy for breast cancer mortality) among women with extremely dense breasts? 2) What is the cost-effectiveness of each of these strategies? 3) What is the influence of the MRI examination on the quality of life of the participating women? 4) What is the degree of participation to the MRI study?
	5. Recent literature Women with extremely dense breasts (>75%) are 4-6 times more likely to develop breast cancer compared to women with <25% dense tissue. This relation has been indicated in the past many times before, both in premenopausal as well as postmenopausal women, and remains unchanged after statistical adjustment for differences in other breast cancer risk factors (3;29). The increased risk applies to all breast tumours, even for the more aggressive tumours (30). High breast density makes it substantially harder to detect tumours using mammography. This is because the fibroglandular and stromal tissue (dense tissue) and tumours block the same amount of X-rays. There is strong evidence that women with high mammographic density have a higher chance at developing both screen-detected tumours (tumours discovered by mammographic screening) and interval tumours. However, this effect is stronger for the interval cancers (4;6-8). Our research group already showed this difference in the past (8), but was more recently confirmed by a study conducted by Boyd et al. in the New England Journal of Medicine (4). They found that women with >75% density compared to women with <10% dense tissue had a higher risk of developing breast cancer (Odds ratio (OR)=4.7; 95% confidence interval (CI)=3.0-7.4). This applied to the screen-detected tumours (OR=3.5; 95%CI=2.0-6.2), and to the interval cancers, but the latter risk was much stronger (OR=17.8; 95%CI=4.8-65.9). In an accompanying editorial in the same journal, Kerlikowske et al. (6) confirmed again with data of the Breast Cancer Surveillance Consortium (BCSC) that the sensitivity of the screening programme is very much defined by the degree of breast density. With women with low breast density (<25%), the screening sensitivity was 0.91, 0.94 and 0.86 for women at the age of 40-49, 50-59 and 60-69 respectively. With women with very dense breasts, the sensitivity was 0.63, 0.66 and 0.64 respectively. There are no studies known on the use of MRI for breast cancer screening in this target population. As has been described extensively in *Part 3. Screening strategy*, MRI is in terms of sensitivity the most promising screening technique. However, MRI studies have so far been limited to women with a BRCA1 or BRCA2 mutation or women with a very strong family history of breast cancer. This strategy turned out to be cost-effective and the American Cancer Society and American College of Radiologists (www.acr.org) now recommends that women, who have a 20-25% lifetime risk of developing breast cancer throughout their lives, should get an MRI

	each year (in addition to mammography) starting from the age of 30 (31).
	6. Research design *Design* The study includes a randomised controlled trial (RCT) with rates of (extra) detected tumours, interval cancers and tumour stage distribution as the primary outcome measures. Many studies that evaluate the new screening methods use a so-called cross-sectional paired design whereby the participants receive both the 'old' as well as the 'new' test. This indicates how many more tumours are discovered by the new test compared to the old test, but it does not indicate in which stage the tumour with each of the tests could be detected. A similar major problem is that the proportion of interval cancers cannot be compared between the two tests. This is because all tumours detected by one of the two tests will be treated; only the tumours that are missed by both tests could appear as interval cancers. A decline in the number of interval cancers compared to the old method is important to show if the new screening test really is effective and does not cause merely more overdiagnosis (in other words, the detection of tumours that would never have become symptomatic during a woman's lifetime). For this reason, an RCT with parallel groups is necessary. For this research proposal a pre-randomisation design has been chosen. The researchers are of the opinion that the proposal meets the requirements stipulated by the Population Screening Act Committee (32): 1) It is reasonably likely that the study will lead to important new insights (requirement of *the importance*). The importance of this study is described extensively in paragraph 2.1. Women with extremely dense breast tissue are more likely to develop breast cancer, and mammographic screening has reduced sensitivity among these women. The hypothesis is that with an additional MRI examination, the sensitivity of the breast screening for these women will improve greatly. At this moment, it is not yet possible to make a reliable and founded statement on this (see paragraph 4.1). All factors that play a role in the consideration whether or not to introduce a new screening method are investigated (see paragraph 7.9). The results of this project may help ensure that women with extremely dense breast tissue will undergo breast cancer screening that is as sensitive as that in women with less dense breast tissue. 2) The research questions cannot be answered with a different research set up other than pre-randomisation (requirement of *subsidiarity*). The researchers are of the opinion that there are no reasonable alternatives for the pre-randomisation design. When the classic randomisation design is applied, there is also a great risk of contamination. Because the control group will also be informed of their high breast density and with this, the associated prognostic unfavourable combination of increased breast cancer risk with a lowered mammographic tumour detection, it is expected that a considerable percentage of women in the control group will request additional examination (for example, via the general practitioner). The

internal validity of the study will for this reason be compromised, and the effect will be diluted. This dilution will be the strongest if the women in the control group undergo an additional MRI examination, but also to a large degree when they undergo an ultrasound or other additional examination. This was also recently observed in a European prostate cancer screening trial, where the PSA test was the diagnostic intervention (33). The study results showed that there was a strong rise in the number of PSA requests in the control group.

With a pre-randomisation design, a setting is created that reflects the future practice the most. This way, an impression can be received of the acceptance rate of the MRI examination.

3) The disadvantages for and burden on the participants are negligible (requirement of *proportionality*).

The researchers are of the opinion that not informing the control group will not cause important adverse effects for or burden of these participants. Actually informing them means the participants are also informed of their risks they were not aware of until that moment. To the more than 20,000 women in the control group, no other screening method can be offered (of course except for the usual mammographic screening every 2 years) or a form of risk reduction, which probably triggers a lot of unrest.

To be able to answer our research question, we do not need any other information from the control group other than on the occurrence of breast tumours in between two screening rounds (interval cancers) or with later screening rounds or on mortality as a result of breast cancer. This data can be obtained by linking the screening data with the data from the Dutch Cancer Registry and Statistics Netherlands (CBS) as is done now routinely already as part of the evaluation of the Dutch population-based mammography screening program (1). This routine linkage is also described in the folder of the regular Dutch population-based mammography screening program. In that same folder, it describes how the women can object to such use. The data of both study arms are pseudonymized before processing and saving.

The women who are randomised to the intervention group are informed of the study, but this takes place after randomisation. Due to the pre-randomisation design, it is only possible to get permission to agree to the *result* of the randomisation. The researchers are of the opinion that the intervention group does not suffer any harm as a result thereof. After randomisation, the women invited for MRI receive an extensive information package on the study and its benefit-harm ratio.

Sampling frame

The study population consists of participants of the Dutch population-based mammography screening program aged between 50 and 75. They are from screening regions located in the operation area of the hospitals where the MRI examinations are carried out (see *Section 10. Organisation*). Women with >75% density and a negative mammographic screening (i.e. no reason for further diagnostic follow-up) are eligible for this.

	Selection of participants, inclusion and exclusion criteria The density is automatically quantified with the help of a full-automatic method developed for digital mammography(19). This result will be linked to the mammographic screening result. The women with extremely dense breasts and a negative screening result (BI-RADS 1 or BI-RADS 2) qualify for the study. The reason to focus on women between the ages of 50 and 75 and not on younger women with very dense breasts, is because the first group has a higher risk of developing breast cancer, and because the infrastructure for screening in this group already exists, so the value of the MRI can be examined relatively easily. If we can show that MRI has an added value to women with very dense breasts, this will be an important indication for a potential future cost-effective strategy for women of 45 to 49 years old. Women with contraindications for MRI cannot take part in the study. Examples of contraindications for contrast MRI are: allergy for gadolinium contrast media, presence of iron aneurysm clips, metal chips in the eye and pacemakers. Every participant fills out a safety questionnaire to find out if there are no contraindications to expose the participant to strong magnetic fields and/or contrast medium. *Randomisation and stratification* The DENSE trial applies the pre-randomisation design, whereby only the intervention group is informed about participation in the trial after randomisation (see paragraph 7.6 Research Method: design). Randomisation takes place if the breast density is extremely dense and a negative mammographic screening examination (BI-RADS mammography 1 or 2). Randomisation takes place between: Arm I (intervention): additional MRI examination Arm II (control; Dutch guideline): no further examination until the following screening round 2 years later (unless there is reason to do a breast examination in the meantime, because of complaints by the participant for instance). The participants are randomised by a computer-generated randomisation sequence, with stratification for screening centre. A screening centre is a radiologic partnership assessing the screening mammograms for a certain region. *Blinding* The proposed experimental population-based screening study is a pragmatic (in comparison to explanatory) trial whereby two different screening strategies, including possible placebo or external effects, are compared. The importance of blinding the outcome measure in a pragmatic trial depends on the ‘hardness’ with which the outcome measure can be determined. In this case, it concerns the presence of breast cancer (screening-detected and interval cancers), which will be assessed objectively with the help of histology. Thus, blinding in this trial is not necessary.

Starting and end date project
Within this experimental population-based screening study, 3 screening rounds will be carried out in a period of 6 years. Because of the capacity at the MRI centres, it might be necessary to include the screening rounds in stages, we pursue a longer period for the entire study, namely from 1 September 2011 to 1 September 2019. However, the research period per participant is still 6 years. For every participant, the intervention in a following screening round is the same as in her first screening round, so independent of, for example, a possible change in breast density. Eventually sensitivity analysis will be carried out to find out whether this has possibly influenced the results.

7. Sample size, feasibility

The sample size is calculated on a minimal detectable difference in the interval cancer percentage between the intervention group and the control group. The assumptions are based on the report of the National Evaluation Team for Breast Cancer Screening in the Netherlands 1990-2007 (1), a study by Kerlikowske et al. (6) and a study by Van Gils et al. (34).

Among those aged 49-50, there are 6.2 tumours diagnosed per 1000 screening examinations: 4.2/1000 by screening and 2/1000 in the interval between screenings (interval cancers) (1). In the table below, the expected occurrence of breast cancer in the different density categories is based on extrapolation of the data collected by Kerlikowske and et al. (6) to the Dutch setting. The programme sensitivity (calculated as the number of screen-detected tumours divided by the total of screen-detected and interval tumours) after a screening interval of 2 years is based on data of Van Gils et al. (34).

Breast density	Number of breast cancer diagnoses (screening + interval) per 1000 examinations*	Programme Sensitivity (Scr / (Scr+Int))***	Proportion interval cancers (=1-Sensitivity)
total	6.2**		
<25%	2.5	82%	18%
25-50%	5.6	74%	26%
50-75%	7.6	66%	34%
>75%	8.0	45%	55%

** relative occurrence of breast cancers between density categories based on data from Kerlikowske et al. (6) for women of 50-69 years of age*
*** from the National Evaluation Team for Breast Cancer Screening in the Netherlands 1990-2007 report (1)*
**** calculated from the article from Van Gils et al. (32) for a 2-year screening interval*

Within the group with >75% density, the estimated number of interval cancers per 1000 mammographic screening examinations is 4.4 (0.55*8.0). Our assumption is that the MRI is not hindered by fibroglandular and

	stromal tissue, which is why women with >75% density can reach the same sensitivity using an MRI as with mammography among women with <25% dense tissue (i.e. 82%, see table). By adding MRI, we also expect that the proportion of interval cancers decreases to 1.44 per 1000 screening examinationss (0.18*8.0). In the pre-randomisation design, it is necessary to take into account the fact that there are women who do not wish to take part after randomisation. Because this non-participation only occurs after randomisation, the non-participants must be included in the analysis according to the intention-to-treat principle. This means the required sample size becomes larger (35). A study by Berg et al. (23) on the added value of MRI in the ACRIN 6666 trial showed that 57.9% of the women contacted participated in the MRI substudy. For a number of women, money or not having insurance was the decisive factor. Without these reasons, the participation level would have been 66%. We expect to reach a participation of 66% because the cost of MRI examination is covered in our study. In the intervention group, the expected interval cancer rate is 1.44/1000 screening examinations among the 66% who participated in the MRI examination and 4.4/1000 among the 33% that did not participate in the MRI examination. As a result, the interval cancer rate in the intervention group is 2.49/1000 screening examinations. With a 1:4 ratio randomisation, 7,237 women are needed in the intervention group (of which 4,776 (66%) actually do take part in the MRI examination) and 28,948 women in the control group to show the difference between 4.4/1000 and 2.49/1000 as statistically significant (1-sided alpha=0.05) with a power of 80% after 1 screening round and the subsequent interval. This was calculated using the sample size software PASS. *Feasibility* The experimental population-based screening study will be integrated in the infrastructure of the Dutch population-based mammography screening program, which is national and entirely digital. We foresee to have to invite 7,237 women to be able to include 4,776 women (expected participation rate 66%) in the intervention group. The total study population (intervention and control group) will consist of 36,185 participants. The total target population of participants in the Dutch population-based mammography screening program with >75% density in the Netherlands is 90,000 in a period of 2 years, so we could expand our recruitment if the number of people participating is disappointing. We invite women living near the participating MRI centres (referred to in *Section 10. Organisation*). Every centre will have to make on average 400 extra MRI examinations per year for this experimental population-based screening study. The centres have already indicated their desire to participate.
	8. Effectiveness of treatment and practice options It is known that treating breast cancer in an early stage reduces the mortality rate. This is the motivation to conduct a population-based screening program on breast cancer in the Netherlands. The recently

	updated meta-analysis from Nelson et al. (36) for the US Preventive Taskforce 2009 report confirms again that mammographic screening reduces the number of deaths due to breast cancer. The relative risk of dying because of breast cancer is lower among women invited for mammographic screening compared to women in a control group who are not invited for screening (50-59 years: 0.89 95%CI 0.75-0.99; 60-69 years: 0.68 95%CI 0.54-0.87). Furthermore, it is known that with mammographic screening, breast tumours among women with extremely dense breasts are diagnosed at a later stage compared to breast tumours among women with low density. If we would be in the position to realise that same early detection among women with very dense breasts, it may be expected that this would reduce the number of deaths due to breast cancer. There are no indications thus far that breast tumours among women with very dense breasts have other intrinsic characteristics which would give them a more favourable or unfavourable prognosis (30;37).
	9. Outcome analysis, evaluation *Indicators of screening effectiveness* The extra number of tumours detected by MRI, but missed by mammography, shall be described as well as the recall rate, the positive predictive value (and the false positive percentages) of the MRI and the number of biopsies per positive MRI test, with the corresponding 95% confidence intervals. The primary outcome is the comparison of the interval cancer proportions between the study groups (Chi-square or Fisher's exact test in case of small numbers). Tumour size, stage and degree of the tumours will be compared between the arms as well as the distribution of histological and molecular subtypes. With respect to tumour size, the difference between means will be tested using the Student's T-test. If not normally distributed, the medians will be calculated and the differences will be tested using the Mann-Whitney U test. Differences in stages, grade, histological and molecular subtypes are tested using Chi-square or Fisher's exact test in case of small numbers. A problem already mentioned before, with the single consent version of the pre-randomisation design, is the dilution of the effect by non-participation in the intervention group after randomisation. Nevertheless, to be able to reach a true (nondiluted) estimation of the effect of the MRI, the analysis method of Cuzick et al. (35) will be used, which was specifically designed for these kinds of screening studies. *Estimation of mortality reduction* Calculating the actual mortality reduction requires a very long follow-up period and an even larger study population than proposed here. Breast cancer screening simulation models are a less expensive and faster approach to study the mortality reduction rate of this intervention. We use the MISCAN simulation programme that was specifically designed to build cancer screening models and also to analyse and explain the cancer

	screening trial results such as these, and to compare the cost-effectiveness of different screening measures (38;39). It is a validated model that has been applied numerous of times successfully in evaluating different screening strategies. It is also one of the most commonly used breast cancer models by the Cancer Intervention and Surveillance Modelling Network (CISNET) consortium sponsored by the National Cancer Institute (40). In MISCAN, the individual life courses are generated as a Markov process of stages and transitions. For this, the characteristics of the study population (demographic data, risk factors) are fed to the model. The results of the screening and the scenarios as applied in the proposed experimental population-based screening study are estimated for a cohort of 1 million women with very dense breasts. The effects on mortality and life years gained with the different screening scenarios are estimated at a period of 10 years after the start of the screening. The most important parameters in the model are the average duration of the preclinical stages that can be detected with screening, sensitivity of the screening test and improvement in prognosis after detection through screening. Outcomes of the model are the number of screen-detected tumours and interval cancers (including stage distribution), age-specific breast cancer incidence, stratified by stage, and age-specific breast cancer mortality. *Analysis of cost-effectiveness* The breast cancer simulation model will also be used to predict the costs and effects of different screening strategies for a situation in which the screening is carried out for 10 years. The costs of the diagnosis and treatment of breast cancer will be registered as part of the experimental population-based screening study. Costs and effects are calculated for a simulated cohort of 1 million women for a period of 10 years after the start of the screening. The costs are presented in euros. The effects are presented in terms of breast cancer mortality reduction and number of life years gained. Cost-effectiveness ratios (CER) are expressed in costs per life-year gained and incremental cost-effectiveness ratios (ICER) as extra costs per extra life-year gained. For the intervention group, the quality of life is assessed using the EuroQol-5D questionnaire. To estimate the number of quality-adjusted life years of the control group, we use the EuroQol-5D health status data of the general population of Dutch women of the age between 50 and 75 (41). We implement weighting factors with a reduction of 10%, 25%, 40% for respectively non-metastatic tumours, tumours with metastases in the regional lymph nodes or tumours with distant metastases. For this reason, the ICER is expressed in extra costs per quality adjusted life-year gained. *Analysis of quality of life* The effect of MRI screening on the quality of life of the women will be assessed using questionnaires. The EuroQol-5D questionnaire is a validated and a frequently-used questionnaire that is able to assess the quality of life with a few short questions, regarding physical as well as psychological consequences (41). In addition, the psychological consequences the experimental population-based screening study could cause will also get attention by using the 'Consequences of Screening---Breast Cancer' (COS-BC) questionnaire (42). The first part of the questionnaire (COS-BC-1) asks specifically about

	the fears and uncertainties during the screening procedure. The second part of the questionnaire (COS-BC-2) aims to get the woman to look back at the entire period of examinations and have her assess whether it has brought about positive and/or negative changes. The great advantage of the COS-BC questionnaire in comparison to other validated anxiety questionnaires is the fact that it is focussed specifically on breast cancer screening. With three short questions, the screen-specific items questionnaire uncover the burden experienced by the MRI examination (43). The group of women with MRI-detected breast cancer will be asked to answer the EORTC questionnaire once (44). The goal of the quality of life analysis is to investigate the (changing) quality of life in the intervention arm as a whole. Besides, whether differences can be observed between participants with false positive, true positive, false negative and true negative MRI results will be addressed. The moments at which time the different questionnaires are conducted is indicated in the flowchart (see 'Flowchart timing of questionnaires'). *Interim analyses* The first analyses are carried out after the first screening round, namely: 1. When all women in the intervention group have received an MRI, including any possible diagnostic work-up. At that point, the extra number of tumours detected by the MRI, and missed by the mammography, can be estimated, as well as the recall rate, the positive predictive value and the number of biopsies per positive MRI test. 2. Two years after the last participant has been screened, the difference in the proportion of interval cancers is analysed and the tumour size and stage between both groups are compared. An interim analysis on the primary outcome is not possible, because information on interval cancers is only available after all participants have been screened and followed for two years. For the false-positive percentage, interim monitoring will take place. From the abstract from Berg et al. (23) on the added value of MRI in the ACRIN 6666 trial, it is possible to calculate that the percentage 'unjustified' referrals in the entire investigated group is 17% and the percentage 'unjustified' biopsies is 9%. Based on this, we have decided the following: If after the first 750 MRI examinations the number of unjustified referrals is more than 25% or the number of unjustified biopsies is above 15%, consultation on modification of the intervention in relation to the referral criteria is necessary and suspending the trial could be considered. A Data Safety and Monitoring Board is installed to this end, in any case consisting of an independent biostatistician, radiologist and oncologist as its members.

	10. Organisation Steering group: *UMC Utrecht, Julius Center for Health Sciences and Primary Care:* Prof. P. Peeters (epidemiologist and coordinator of the Cancer Epidemiology group) Dr C. Van Gils (epidemiologist) Dr E. Monninkhof (epidemiologist) Dr A. de Wit (coordinator of the Medical Technology Assessment group) *UMC Utrecht Imaging Division:* Prof. W. Mali (radiologist) Prof. M. Van den Bosch (radiologist) Dr W. Veldhuis (radiologist) *Dutch Expert Centre for Screening (Dutch: Landelijk Referentie Centrum voor Bevolkingsonderzoek, LRCB)* Prof. G. Den Heeten (Director) Dr M. Broeders (epidemiologist) *Radboud University Medical Centre, Radiology Department:* Prof. N. Karssemeijer (medical physicist) *Screening organization Midden-West:* A. Bartels-Kortland MSc (Director) *Radiologists of participating centres where the MRI examinations will be made:* Prof. C. Boetes, UMC Maastricht A. Obdeijn MSc, Erasmus MC, Rotterdam Dr R. Pijnappel, Martini Hospital, Groningen Dr C. Loo, Netherlands Cancer Institute (NKI), Amsterdam Dr R. Mann, Radboud University Medical Centre, Nijmegen Dr W. Veldhuis, Prof. M. van den Bosch, UMC Utrecht *Advisors:* R. Reij MSc, RIVM, Centre for Population Study Prof. H. De Koning, Erasmus MC, Rotterdam Dr M. Tilanus-Linthorst, Erasmus MC, Rotterdam Prof. E. Van der Wall, UMC Utrecht, oncology Prof. P. Van Diest, UMC Utrecht, pathology *Breast Cancer Patient Association:* Ms R. van der Heide (former chairman)
	11. Risks The risks involved for the participants of this experimental population-based screening study are as follows: 1) A participant can experience side effects of the contrast medium used by the MRI examination. Millions of MRI examinations are carried out with

	Gadolinium-based contrast agent worldwide each year (45). In a recent study, 51 side effects were reported after 32,659 injections administered with gadolinium-based contrast agent (0.16%), of which 43 mild (0.13%, for example nauseous, lightheaded, headache), 6 moderately severe (0.02%, for example, shortness of breath, erythema) and 2 severe (0.006%, loss of consciousness, cardiac arrest) (46). Extrapolated to the 4,776 patients to be scanned for this study, 6 mild side effects are to be expected, 1 moderately severe and 0-1 severe side effects. These risks are further reduced by checking on allergies for the used contrast medium. There have been reports of Nephrogenic Systemic Fibrosis (NSF) among patients with severe kidney dysfunction (see FDA website: http://www.fda.gov/Drugs/DrugSafety/ucm223966.htm#aihp). NSF is not described with patients with normal kidney functions. Severe kidney dysfunction is therefore an exclusion criterion for this study. 2) With the MRI examination, there is a chance of a false positive result with uncertainty, followed by extensive further and possibly invasive examination. How often this will happen in this population is not known and is also subject of the trial. The best indication we have comes from a recently published abstract (23) on the MRI substudy of the ACRIN 6666 trial (10). This study shows that based on MRI, 17% of the participants underwent further examinations (imaging or biopsy) that were in retrospect unnecessary. Nine percent underwent, in retrospect, an unnecessary biopsy. Based on histopathological research, it turned out the abnormalities of these patients were benign. Participants will be informed thoroughly on the increased probability of false positives. Unnecessary invasive examination will be avoided as much as possible by looking back at the original mammographic examination, by assessing the MRI BI-RADS 3 results centrally and by scheduling a follow-up MRI after 6 months for this category and finally, by carrying out a target ultrasound before a possible biopsy. 3) There is a chance at being diagnosed with a relative indolent tumour that would never have been noticed if an MRI screening had not taken place (overdiagnosis). Overdiagnosis is inherent to screening: in the Dutch population-based screening program it is estimated that about 8% of the breast tumours discovered by screening would never have been diagnosed clinically in the absence of a screening programme (47). There are no estimates known yet on overdiagnosis in high-risk groups screened by MRI. We do not expect the overdiagnosis percentage to be higher in our study compared to the current population-based screening program because the ductal carcinoma in situ (DCIS) that plays a central role in the overdiagnosis problem, is actually easier to detect with mammography than with MRI (25;26). Besides, Kuhl et al. (48) have shown that the sensitivity of MRI for DCIS is especially high for the high-grade lesions that are most likely clinically more relevant. Overdiagnosis cannot be assessed for an individual, but our research is set up to get a better understanding of the degree of overdiagnosis on a population level. If the total number of tumours discovered in the MRI group (screening tumours) is greater than in the control group but the number of

	interval cancers does not decrease proportionally, this could be an indication of overdiagnosis.
	12. Information, consent The information and consent of the participants will be in accordance with the Central Committee on Research involving human subjects (in Dutch: Centrale Commissie Mensgebonden Onderzoek (CCMO)) regulations. Potential participants for the experimental population-based screening study will be recruited via the Dutch population-based mammography screening program. After randomisation, the intervention group will receive an information package through the mail about the study, consisting of the following documents: a cover letter from the screening organisation, information for the participants with the consent form, flow chart, MRI examination questionnaire and the 'General medical scientific research with human subjects brochure'. In the cover letter, the women are asked – even if they are not interested in participating – to log in to the website www.juliuscentrum.nl/dense-studie. The login codes are provided in the information for the participants. If the woman is not interested in the study, the procedure is merely unregistering with an option to indicate the reason thereof. Women who do wish to participate in the study, register on the website and indicate they are interested in participating. Then the woman enters her personal data so she can be contacted by telephone by the research team. During this telephone conversation, the study is explained again and the woman has the opportunity to ask any remaining questions. If a woman decides to take part in the study, the study team employee who spoke to her will sign the informed consent form. Then this employee will send the consent form to the MRI centre, where the participant will be asked to sign it before undergoing the MRI examination. If the woman indicates over the telephone to want more time to make a decision, a later time is agreed on to call again. An appointment is scheduled for the MRI examination for those women who do wish to participate in the study. The MRI appointment will be confirmed in writing. The information flowchart is also annexed to the information for the participants.
	13. Ethical assessment Not submitted for approval by Medical Ethical Committee.
	14. Funding This study is financially supported by: - Utrecht University Medical Center (UMCU) - Dutch Cancer Society (KWF) - The Netherlands Organisation for Health Research and Development (ZonMw) - Pink Ribbon Netherlands/A Sister's Hope - Bayer Schering Pharma (Bayer is not the sponsor or initiator of this research – that is UMC Utrecht). It has been contractually agreed with Bayer that all results and any possible intellectual properties as

	a result of this research is the property of the provider (UMCU). Bayer can not obstruct or postpone publication of the results. It is possible to see the contract on demand. Grant numbers: - UMCU DENSE - DCS-UU-2009-4348 - ZONMW-200320002-UMCU - Pink Ribbon-10074 - BSP-DENSE

Part 8. References
(summary of max. 100 lines)

(1) National evaluation of breast cancer screening in the Netherlands. National evaluation of breast cancer screening in the Netherlands 1990-2007. LETB XII. 2009. Rotterdam.

(2) Kauhava L, Immonen-Raiha P, Parvinen I, Helenius H, Kaljonen A, Rasanen O, et al. Lower costs of hospital treatment of breast cancer through a population-based mammography screening programme. Eur J Public Health 2004 Jun;14(2):128-33.

(3) McCormack VA, dos SS, I. Breast density and parenchymal patterns as markers of breast cancer risk: a meta-analysis. Cancer Epidemiol Biomarkers Prev 2006 Jun;15(6):1159-69.

(4) Boyd NF, Guo H, Martin LJ, Sun L, Stone J, Fishell E, et al. Mammographic Density and the Risk and Detection of Breast Cancer. N Engl J Med 2007 Jan 18;356(3):227-36.

(5) Kavanagh AM, Byrnes GB, Nickson C, Cawson JN, Giles GG, Hopper JL, et al. Using Mammographic Density to Improve Breast Cancer Screening Outcomes. Cancer Epidemiol Biomarkers Prev 2008 Oct 1;17(10):2818-24.

(6) Kerlikowske K. The mammogram that cried Wolfe. N Engl J Med 2007 Jan 18;356(3):297-300.

(7) Mandelson MT, Oestreicher N, Porter PL, White D, Finder CA, Taplin SH, et al. Breast density as a predictor of mammographic detection: comparison of interval- and screen-detected cancers. J Natl Cancer Inst 2000 Jul 5;92(13):1081-7.

(8) van Gils CH, Otten JD, Verbeek AL, Hendriks JH, Holland R. Effect of mammographic breast density on breast cancer screening performance: a study in Nijmegen, Netherlands. J Epidemiol Community Health 1998 Apr;52(4):267-71.

(9) Pisano ED, Hendrick RE, Yaffe MJ, Baum JK, Acharyya S, Cormack JB, et al. Diagnostic accuracy of digital versus film mammography: exploratory analysis of selected population subgroups in DMIST. Radiology 2008 Feb;246(2):376-83.

(10) Berg WA, Blume JD, Cormack JB, Mendelson EB, Lehrer D, Bohm-Velez M, et al. Combined screening with ultrasound and mammography vs mammography alone in women at elevated risk of breast cancer. JAMA 2008 May 14;299(18):2151-63.

(11) Kuhl C, Weigel S, Schrading S, Arand B, Bieling H, Konig R, et al. Prospective

Multicenter Cohort Study to Refine Management Recommendations for Women at Elevated Familial Risk of Breast Cancer: The EVA Trial. J Clin Oncol 2010 Mar 20;28(9):1450-7.

(12) Lord SJ, Lei W, Craft P, Cawson JN, Morris I, Walleser S, et al. A systematic review of the effectiveness of magnetic resonance imaging (MRI) as an addition to mammography and ultrasound in screening young women at high risk of breast cancer. Eur J Cancer 2007 Sep;43(13):1905-17.

(13) Warner E, Plewes DB, Hill KA, Causer PA, Zubovits JT, Jong RA, et al. Surveillance of BRCA1 and BRCA2 mutation carriers with magnetic resonance imaging, ultrasound, mammography, and clinical breast examination. JAMA 2004 Sep 15;292(11):1317-25.

(14) Weinstein SP, Localio AR, Conant EF, Rosen M, Thomas KM, Schnall MD. Multimodality screening of high-risk women: a prospective cohort study. J Clin Oncol 2009 Dec 20;27(36):6124-8.

(15) Tabar L, Chen HH, Duffy SW, Yen MF, Chiang CF, Dean PB, et al. A novel method for prediction of long-term outcome of women with T1a, T1b, and 10-14 mm invasive breast cancers: a prospective study. Lancet 2000 Feb 5;355(9202):429-33.

(16) Tilanus-Linthorst MM, Obdeijn IM, Hop WC, Causer PA, Leach MO, Warner E, et al. BRCA1 mutation and young age predict fast breast cancer growth in the Dutch, United Kingdom, and Canadian magnetic resonance imaging screening trials. Clin Cancer Res 2007 Dec 15;13(24):7357-62.

(17) Visvanathan K, Chlebowski RT, Hurley P, Col NF, Ropka M, Collyar D, et al. American society of clinical oncology clinical practice guideline update on the use of pharmacologic interventions including tamoxifen, raloxifene, and aromatase inhibition for breast cancer risk reduction. J Clin Oncol 2009 Jul 1;27(19):3235-58.

(18) Ropka ME, Keim J, Philbrick JT. Patient decisions about breast cancer chemoprevention: a systematic review and meta-analysis. J Clin Oncol 2010 Jun 20;28(18):3090-5.

(19) van Engeland S, Snoeren PR, Huisman H, Boetes C, Karssemeijer N. Volumetric breast density estimation from full-field digital mammograms. IEEE Trans Med Imaging 2006 Mar;25(3):273-82.

(20) Warner E, Messersmith H, Causer P, Eisen A, Shumak R, Plewes D. Systematic review: using magnetic resonance imaging to screen women at high risk for breast cancer. Ann Intern Med 2008 May 6;148(9):671-9.

(21) Kriege M, Brekelmans CT, Obdeijn IM, Boetes C, Zonderland HM, Muller SH, et al. Factors affecting sensitivity and specificity of screening mammography and MRI in women with an inherited risk for breast cancer. Breast Cancer Res Treat 2006 Nov;100(1):109-19.

(22) Lehman CD, Gatsonis C, Kuhl CK, Hendrick RE, Pisano ED, Hanna L, et al. MRI evaluation of the contralateral breast in women with recently diagnosed breast cancer. N Engl J Med 2007 Mar 29;356(13):1295-303.

(23) Berg W, Zhang Z, Cornack J, Jong R, Barr R, Lehrer D, et al. Supplemental yield and performance characteristics of screening MRI after combined ultrasound and

mammography: ACRIN 6666. American College of Radiology Imaging Network. 2009. http://www.acrin.org/

(24) Gordon PB. Ultrasound for breast cancer screening and staging. Radiol Clin North Am 2002 May;40(3):431-41.

(25) Kriege M, Brekelmans CT, Boetes C, Besnard PE, Zonderland HM, Obdeijn IM, et al. Efficacy of MRI and mammography for breast-cancer screening in women with a familial or genetic predisposition. N Engl J Med 2004 Jul 29;351(5):427-37.

(26) Leach MO, Boggis CR, Dixon AK, Easton DF, Eeles RA, Evans DG, et al. Screening with magnetic resonance imaging and mammography of a UK population at high familial risk of breast cancer: a prospective multicentre cohort study (MARIBS). Lancet 2005 May 21;365(9473):1769-78.

(27) Nationaal Borstkanker Overleg Nederland: Richtlijn mammacarcinoom 2008. http://www.oncoline.nl/

(28) Netherlands Federation of University Medical Centres. Kwaliteitsborging van mensgebonden onderzoek. Houten: Badoux Printing House; 2010 Jun.

(29) Vachon CM, van Gils CH, Sellers TA, Ghosh K, Pruthi S, Brandt KR, et al. Mammographic density, breast cancer risk and risk prediction. Breast Cancer Res 2007;9(6):217.

(30) Ziv E, Tice J, Smith-Bindman R, Shepherd J, Cummings S, Kerlikowske K. Mammographic density and estrogen receptor status of breast cancer. Cancer Epidemiol Biomarkers Prev 2004 Dec;13(12):2090-5.

(31) Saslow D, Boetes C, Burke W, Harms S, Leach MO, Lehman CD, et al. American Cancer Society guidelines for breast screening with MRI as an adjunct to mammography. CA Cancer J Clin 2007 Mar;57(2):75-89.

(32) Netherlands Health Council. Wet bevolkingsonderzoek: prenatale screening en risicoperceptie. The Hague: Netherlands Health Council; 1999: 1999/04WBO.

(33) Otto SJ, van der Cruijsen IW, Liem MK, Korfage IJ, Lous JJ, Schröder FZ, de Koning HJ. Effective PSA contamination in the Rotterdam section of the European Randomized Study of Screening for Prostate Cancer. Int J Cancer 2003 105:394-399.

(34) van Gils CH, Otten JD, Hendriks JH, Holland R, Straatman H, Verbeek AL. High mammographic breast density and its implications for the early detection of breast cancer. J Med Screen 1999;6(4):200-4.

(35) Cuzick J, Edwards R, Segnan N. Adjusting for non-compliance and contamination in randomized clinical trials. Stat Med 1997 May 15;16(9):1017-29.

(36) Nelson HD, Tyne K, Naik A, Bougatsos C, Chan BK, Humphrey L. Screening for breast cancer: an update for the U.S. Preventive Services Task Force. Ann Intern Med 2009 Nov 17;151(10):727-42.

(37) Ma H, Luo J, Press MF, Wang Y, Bernstein L, Ursin G. Is there a difference in the association between percent mammographic density and subtypes of breast cancer? Luminal A and triple-negative breast cancer. Cancer Epidemiol Biomarkers Prev

2009 Feb;18(2):479-85.

(38) de Gelder R., Bulliard JL, de WC, Fracheboud J, Draisma G, Schopper D, et al. Cost-effectiveness of opportunistic versus organised mammography screening in Switzerland. Eur J Cancer 2009 Jan;45(1):127-38.

(39) Habbema JD, van Oortmarssen GJ, Lubbe JT, van der Maas PJ. The MISCAN simulation program for the evaluation of screening for disease. Comput Methods Programs Biomed 1985 May;20(1):79-93.

(40) Mandelblatt JS, Cronin KA, Bailey S, Berry DA, de Koning HJ, Draisma G, et al. Effects of mammography screening under different screening schedules: model estimates of potential benefits and harms. Ann Intern Med 2009 Nov 17;151(10):738-47.

(41) Hoeymans N, van LH, Westert GP. The health status of the Dutch population as assessed by the EQ-6D. Qual Life Res 2005 Apr;14(3):655-63.

(42) Brodersen J, Thorsen H. Consequences of Screening in Breast Cancer (COS-BC): development of a questionnaire. Scandinavian Journal of Primary Health Care. 2008; 26: 251-256.

(43) Rijnsburger AJ, Essink-Bot ML, van Dooren S, et al. Impact of screening for breast cancer in high-risk women on health-related quality of life. Br J Cancer 2004; 91: 69–76.

(44) European Organisation for Research and Treatment of Cancer (EORTC) Study Group on Quality of Life. http://groups.eortc.be/qol.

(45) Colletti PM. Nephrogenic systemic fibrosis and gadolinium: a perfect storm. AJR Am J Roentgenol 2008 Oct;191(4):1150-3.

(46) Abujudeh HH, Kosaraju VK, Kaewlai R. Acute adverse reactions to gadopentetate dimeglumine and gadobenate dimeglumine: experience with 32,659 injections. AJR Am J Roentgenol 2010 Feb;194(2):430-4.

(47) de Koning HJ, Draisma G, Fracheboud J, de BA. Overdiagnosis and overtreatment of breast cancer: microsimulation modelling estimates based on observed screen and clinical data. Breast Cancer Res 2006;8(1):202.

(48) Kuhl CK, Schrading S, Bieling HB, Wardelmann E, Leutner CC, Koenig R, et al. MRI for diagnosis of pure ductal carcinoma in situ: a prospective observational study. Lancet 2007 Aug 11;370(9586):485-92.