Lessons Learned from Setting Up a Prospective, Longitudinal, Multicenter Study with Women at High Risk for Breast Cancer

Susana I.S. Patuleia[1,2], Sophie C. Hagenaars[3], Cathy B. Moelans[1], Margreet G.E.M. Ausems[4], Carla H. van Gils[5], Rob A.E.M. Tollenaar[3], Paul J. van Diest[1], Wilma E. Mesker[3], and Elsken van der Wall[2]

ABSTRACT

Women identified with an increased risk of breast cancer due to mutations in cancer susceptibility genes or a familial history of breast cancer undergo tailored screening with the goal of detecting tumors earlier, when potential curative interventions are still possible. Ideally, screening would identify signs of carcinogenesis even before a tumor is detectable by imaging. This could be achieved by timely signaling of altered biomarker levels for precancerous processes in liquid biopsies. Currently, the Nipple Aspirate Fluid (NAF) and the Trial Early Serum Test BREAST cancer (TESTBREAST), both ongoing, prospective, multicenter studies, are investigating biomarkers in liquid biopsies to improve breast cancer screening in high-risk women. The NAF study focuses on changes over time in miRNA expression levels both in blood and NAF samples, whereas the TESTBREAST study analyzes changes in protein levels in blood samples at sequential interval timepoints. These within-subject changes are studied in relation to later occurrence of breast cancer using a nested case–control design. These longitudinal studies face their own challenges in execution, such as hindrances in logistics and in sample processing that were difficult to anticipate. This article offers insight into those challenges and concurrently aims to provide useful strategies for the set-up of similar studies.

See related commentary by Sauter, p. 429

Introduction

Screening of women at high risk of breast cancer

Women identified with mutations in breast cancer susceptibility genes or with a family history of breast cancer have a moderate or strong increased lifetime risk (LTR) of developing breast cancer. The LTR can go up to 72% (1) and, therefore, demands adequate screening programs to spare these women the physical and psychosocial sequelae of breast cancer. Current screening practices in high-risk women, especially in genetic mutation carriers, comprise adapted, more intensive programs than the regular nationwide screening for women at a population risk. The tailored screening starts at an earlier age, is more frequent, is mostly combined with a clinical breast examination, and often includes breast MRI depending on the risk group and age (2–4). Advantages of breast MRI over the widely used mammography are its higher sensitivity (5) and absence of radiation and hence, the possible radiation-induced side effects (6, 7). However, whereas the improvement of the applied imaging techniques has led to a better detection of smaller and *in situ* tumors, the fact remains that it detects cancer when it has already been developed. Another issue is the timing of screening. Even though the screening is performed at regular intervals (biannually or annually), around 3%–17% of the detected breast tumors are diagnosed between these scheduled screening moments (i.e., interval cancers; refs. 8–11). Finally, another drawback is the postponement of screening by means of mammography and MRI during pregnancy and lactation. These imaging techniques are not indicated during these periods, due to the potentially harmful effects of mammography and the decreased specificity and sensitivity of both the techniques (12). Even though breast ultrasound can represent an alternative for these women according to a few guidelines (13), it is still not ideal given its low positive predictive value and high false-positive rates (12). Taken all of this together, it is clear that there remains an urgent need to improve current screening protocols.

As scientific knowledge on the biology of cancer is still accumulating, extending screening with the analysis of biological tumor markers in so-called liquid biopsies becomes an increasingly realistic approach. In these biofluids, tumor-derived material shed by cancerous cells, such as DNA, RNA, and proteins, can be found. Serial monitoring of these markers bears the potential to reveal early carcinogenesis by a noninvasive approach, as has already been shown for lung cancer in the MILD trial (14, 15) and for *BRCA2*-associated prostate cancer in the IMPACT trial (16), among others (17–21). The IMPACT trial has even led to the implementation of biomarker monitoring in current Dutch screening practice (22, 23). Such biomarkers could also be used as an indicator to anticipate the next screening moment and hence, decrease the occurrence of interval cancers. These advantages would also apply to a non–high-risk population, for instance, for women with dense breast tissue, for whom mammography is less sensitive (24), and the additional value of supplemental MRI is being studied in the DENSE study (25). Moreover, as there is currently no international consensus regarding the best age for *BRCA1* and *BRCA2* mutation carriers to undergo prophylactic mastectomy (26), biomarker monitoring could play an additional role to personalize this delicate decision.

[1]Department of Pathology, University Medical Center Utrecht, Utrecht University, Utrecht, the Netherlands. [2]Department of Medical Oncology, University Medical Center Utrecht, Utrecht University, Utrecht, the Netherlands. [3]Department of Surgery, Leiden University Medical Center, Leiden University, Leiden, the Netherlands. [4]Department of Genetics, University Medical Center Utrecht, Utrecht University, Utrecht, the Netherlands. [5]Department of Epidemiology of the Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, Utrecht University, Utrecht, the Netherlands.

Note: S.I.S. Patuleia and S.C. Hagenaars contributed equally to this article.

Corresponding Author: Elsken van der Wall, University Medical Center Utrecht, PO Box 85500, Utrecht 3508 GA, the Netherlands. Phone: 318-8755-7409; E-mail: E.vanderWall@umcutrecht.nl

Cancer Epidemiol Biomarkers Prev 2021;30:441–9

doi: 10.1158/1055-9965.EPI-20-0770

The Nipple Aspirate Fluid and Trial Early Serum Test BREAST cancer studies

The exploratory work to identify a panel of predictive biomarkers in liquid biopsies as an additional screening tool requires an appropriate study design. The prospective, longitudinal follow-up of an ample-sized cohort of the targeted population at risk with repeated liquid sampling before cancer onset and at the moment of cancer diagnosis, is the optimal design. Two studies, which have both received strong support of patient advocates, have been designed accordingly. These include the Dutch Nipple Aspirate Fluid (NAF) study (Dutch trial register number NL8661; ref. 27), initiated by the University Medical Center Utrecht (Utrecht, the Netherlands), and the Trial Early Serum Test BREAST cancer (TESTBREAST) study (Dutch trial register number NL8724; ref. 28), initiated by the Leiden University Medical Center (Leiden, the Netherlands). The set-up of these studies allows to identify prediagnostic changes in biomarker levels between serial samples from the same study subject. A certain proportion of the study subjects will develop breast cancer after they were recruited (cases). The serial comparison of samples facilitates investigating whether the biomarker levels were already altered before the tumor was diagnosed and if so, how long beforehand these changes could already be measured (**Fig. 1**). These changes over time in the cases could then be compared with those determined in study subjects who have not developed breast cancer (controls).

Both are long-term, ongoing studies: the NAF study started in 2008 and the TESTBREAST study was initiated in 2011. In the TESTBREAST study, blood samples from high-risk women are prospectively collected during regular screening appointments. The last sample is collected if and when an event occurs, that is, invasive breast cancer or carcinoma *in situ* (ductal or lobular). The NAF study uses the same set-up, but besides blood samples, it also collects NAF samples from these high-risk women. The additional hypothesis tested in the NAF study is whether NAF samples mirror the breast microenvironment best, as this "liquid biopsy" is directly derived from the ducts and/or lobules including those that harbor the cancer cells. The biomarkers investigated in both studies differ: while the NAF study focuses on miRNA expression levels, the TESTBREAST study analyzes protein changes over time. The overarching aim of both research groups is similar, that is, to identify a combination of biomarkers for early breast cancer detection.

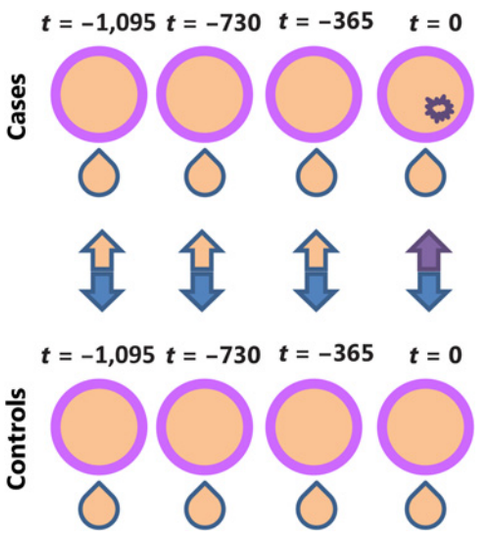

Figure 1.

Nested case–control analysis. In the described longitudinal prospective cohort, healthy, high-risk women were included and serial samples are collected over time. The subjects who develop breast cancer after they are recruited in the cohort become the "cases." "Controls" are all the cohort members who did not develop breast cancer. In this figure, the circles represent the study visits over time; the droplets underneath represent sample acquisition. Any "*t*" (in days) stands for the number of days before the event. In the cases group, $t = 0$ represents the time of the event; in the control group, timepoint $t = 0$ represents the moment at which the most recent sample was acquired. The arrows represent the comparisons between cases and controls that allow a nested case–control analysis.

For this article, both research groups convened to reflect on the challenges that go hand in hand with conducting such a prolonged, longitudinal, multicenter biomarker study. Our goal was to provide an overview of these challenges and its possible solutions to support future researchers who intend to design similar cohort studies.

Lesson 1. Study Phases and Cohort Size

The first and predominant challenge is that, to obtain a longitudinal series of prediagnostic samples from a sufficient number of breast cancer cases, a very large, long-running biobank needs to be constructed. Even in cohorts of women at high risk for breast cancer, only about 1%–6% of the study participants develop an event over a period of 4–8 years (9–11). Therefore, much effort is needed to set up the biobank and to acquire and store the great number of samples, whereas final analysis will only be performed in a nested case–control design on a limited number of cases and controls to answer the research question. Since the start, the focus of the study should be on setting up a large, longitudinal cohort of healthy, high-risk women. The required time to reach a large number of inclusions and a sufficient follow-up period makes this "inclusion and biobanking phase" the most time-consuming part of the study phases. When enough cases have been developed to provide sufficient statistical power to test the predictive value of changes in biomarker levels in relation to cancer occurrence, the next phase can be initiated. This "sample and data analysis phase" includes selecting matching controls based on the characteristics of the cases, retrieving samples of the selected study subjects from the biobank, analyzing biological samples in the laboratory, and performing subsequent statistical analyses. In the meantime, the "inclusion and biobanking phase" becomes a "follow-up and biobanking phase," allowing for more women to develop events for the subsequent validation studies.

A few potential pitfalls should be taken into consideration. The first pitfall is to only engage experts at specific phases during the study. Early commitment and involvement of several experts lets the study run smoothly from the start and avoids making incorrect assumptions or decisions. We recommend involving all necessary experts who need to be engaged from the start of the study. A list of suggested experts for the presented study design and cohort is presented in **Table 1**.

Another pitfall is to mainly focus on the number of inclusions in the cohort, while it is also relevant to regularly monitor more aspects of the cohort, such as study drop-outs. Specifically for biomarker studies, issues such as the number of successfully, serially acquired samples for analysis should be monitored closely. Some samples may be absent because of missed visits, lost by unexpected flaws, processing errors, and/or laboratory technical failures. These issues and recommendations on how to avoid them will be further elucidated in lessons 3, 4, 7, and 9. Finally, predicting the time frame of the study is challenging. From our experience, more than 5 years are needed for the initial "inclusion and biobanking phase." This depends on the extent of the abovementioned hurdles, which cannot be anticipated beforehand. We

Table 1. List of suggested team members.

Core team	Consulting experts	Additional essential members
• PIs	• Clinical geneticist(s)	• Screening doctors and nurses
• PhD student	• Epidemiologist and methodologist	• Local PIs, research nurses, and laboratory staff inside centers
• Project coordinator	• Statistician	• CRA
• Research nurses	• Biobank coordinator	
• Biomedical expert		
• Data manager		
• Laboratory technician		

Abbreviations: CRA, clinical research associate; PI, principal investigator.

recommend to monitor the participation rate and to be prepared for a prolonged time frame of the study.

Lesson 2. Funding

The lack of fast results in prolonged longitudinal studies makes funding less appealing, both to governmental and private sponsors. Preliminary data and achieving milestones are usually the foundations to apply for the next grant. However, given the long "inclusion and biobanking phase," which was described in lesson 1, limited data will be generated for a long period of time. This is a relevant and delicate challenge, because progress of the study is highly dependent on getting continuous, long-term funding for one project. Discontinuation of funding is prejudicial at any phase of the project. Essential budgetary costs also refer to the salaries of the team members involved in running the study (**Table 1**); intermittently downsizing the team due to financial shortcomings endangers, among others, adequate data acquisition and, in the end, its interpretation.

To ensure continuous funding, principal investigators (PI) should apply for several grants throughout the course of the study. Each grant application should focus on one of the study aims. Infrastructural grants are ideal to achieve biobanking aims in the "inclusion and biobanking phase," whereas research grants, including high-risk pilot grants and proof-of-concept grants, are more suited for aims focused on acquiring results as part of the "sample and data analysis phase." Private sponsors that share the vision that a prolonged study is required to obtain translational results, may be instrumental.

Lesson 3. The Inclusion and Exclusion Criteria are Dynamic

Inclusion criteria

The underlying causes for having an increased risk of developing breast cancer are quite diverse, as these include, among others, pathogenic variants of one of the breast cancer genes (e.g., *BRCA1, BRCA2, CHEK2, PALB2*, and *ATM*), a personal and/or family history of breast cancer, and a medical history of radiotherapy in the thoracic field. Inclusion criteria can be defined on the basis of a threshold LTR percentage or on the basis of one specific high-risk group for the cohort. The pitfall of applying LTR as the main inclusion criterion lies in the inclusion of a wide range of high-risk subgroups. Note that if the predictive value of the biomarker panel varies per subgroup, but it is analyzed as the total group, there will be a diluted effect, and one may miss biomarkers that are relevant for one subgroup but not for others. If there are upfront strong ideas that relevant biomarkers may differ per high-risk subgroup, it is advisable to calculate the necessary sample size per subgroup.

One can also choose to only include mutation carriers of a certain age to make the study financially efficient in terms of costs for sample collection, processing, and storage. However, this would mean that the biomarker panel could only be tested for this group and not give the possibility to explore whether it is also usable for other high-risk women, who are in need of such an early detection screening tool as well. Thus, including several high-risk categories has the advantage that one large high-risk breast cancer biobank can be established that allows biomarker testing in a wider scope. The increased LTR for developing breast cancer is well defined for most subgroups, for example, in women with an established genetic mutation. Yet, the risk may be dynamic due to increasing age, a new breast cancer diagnosis in the family, new scientific developments that lead to the discovery of additional gene mutations, expansion of the tested gene mutation panel (29), and new algorithms for risk assessment (30) that result in renewed LTR estimates and screening guidelines. For women with a 50% chance of having a *BRCA1* or *BRCA2* gene mutation, their LTR lowers if genetic testing of the study subject and/or a family member is negative for mutation(s). For these cases, the study protocol should state whether they should be excluded. If so, the research team should decide beforehand whether the already acquired samples will then be used for other research purposes, and describe this possibility in the informed consent form. Arguably, follow-up of these women could still be continued, as a lower risk does not rule out the chance of developing breast cancer. The study team should be aware of the possibility of this decreased LTR and keep this particular screening information of study subjects up to date, which can be facilitated by keeping a close collaboration with the clinical geneticists involved.

Exclusion criteria

A criterion to consider for exclusion in a biomarker study is a personal history of breast cancer. Even though these women have an increased risk of developing breast cancer in the contralateral breast (31), the past treatment might by itself influence the biomarker pattern of choice. For the same reason, a medical history of other malignancies could be considered to be an exclusion criterion, as was adopted by the TESTBREAST study team (**Table 2**). Given the novelty component that comes with the NAF samples and, therefore, lack of literature-based arguments for the implementation of these criteria, these were not adopted by the NAF study team. Nevertheless, this information will be considered in the statistical analysis.

Commonly used study exclusion criteria are pregnancy and lactation. One of the reasons why these criteria were implemented in the NAF study was to avoid milky sample collection; as such, a study visit can only be planned more than 3 months after completing breastfeeding. Still, samples collected between 3 and 24 months after the end of lactation may contain milk components. These components may influence biomarker analyses when these samples are being compared

Table 2. Inclusion and exclusion criteria of the NAF and TESTBREAST studies.

	NAF study	TESTBREAST study
Inclusion criteria	• Female ≥18 years of all ethnic backgrounds • A >20% LTR of developing breast cancer, including germline *BRCA1* or *BRCA2* mutations and previous DCIS/invasive breast cancer	• Female ages between 25 and 75 years • Screening indication due to familiar or genetically enhanced risk of developing breast cancer or LTR >15%
Exclusion criteria	• Bilateral ablative breast surgery • Bilateral breast reduction with nipple graft • Pregnancy or lactation • Active breast infection • Disseminated breast cancer	• Previous invasive breast cancer • Other malignancies <10 years (other than basal cell carcinoma)

Abbreviations: *BRCA*, breast cancer gene; DCIS, ductal carcinoma *in situ*.

with NAF samples collected more than 24 months after the end of breastfeeding, as are most of the samples. Other reasons were the potential rare side effects of the oxytocin nose spray, such as uterus contractions (32). Although oxytocin is used in every study visit in a very low dosage to increase the success rate of bilateral NAF harvesting (which currently is 65.8% for the entire study period and was previously reported to be 62%–73%; refs. 33–35), a total absence of effect on the uterine muscle and on biomarker patterns cannot be guaranteed. As these are long-term studies and young participating women might become pregnant during the study period, these criteria should be defined as "temporary exclusion criteria" to allow study continuation afterwards. From another perspective, because there are currently no alternatives for screening during that period, measurement of blood-based biomarkers in pregnant high-risk women may be even more valuable. This could, however, be limited by the fact that pregnant women do not undergo routine screening in this period, so they would have to visit the hospital solely to donate blood samples for research purposes, which might be an extra burden for some. Finally, in countries where chemoprevention is advised (3, 36), this should be considered as an exclusion criterion due to the woman's lowered LTR and the possible influence of chemoprevention on the biomarker panel.

To sum up, the study protocol should identify which measures should be taken when a woman's LTR lowers and this should be clearly stated in the study information form. Herein, the subsequent consequences regarding participation in the study and handling of already acquired data and biobank material should be included. Finally, PIs should define criteria that may influence the biomarker of choice and add these to the list of exclusion criteria in the study protocol.

Lesson 4. Delayed Inclusion Moment and Discontinuation of Participation

Both studies aimed to pose as little as possible extra burden on the study participants with regard to hospital visits. This means, in theory, that the first opportunity to inform potential candidates about the study is during their regular scheduled screening appointment. By Dutch law, it is required that individuals who are asked to participate in a scientific study should have a reasonable amount of time to consider this request, in principle prohibiting start of the study on the day that information about the study is provided. Consequently, by holding on to having sample acquisition coincide with a regular screening appointment, the actual moment of study inclusion could be delayed for 6–12 months (i.e., when the next screening appointment is scheduled). This observation led to an adapted protocol in the TESTBREAST study, where research nurses now attempt to inform potential participants about the study in advance by post or e-mail, so that inclusion in the study can start at the moment of the already scheduled screening visit.

Given the long time period until publication of study results, discontinuation of participation among enrolled women is prone to occur. Therefore, efficient strategies to keep compliance should be applied. An essential starting point is having dedicated research nurses who clearly describe the aim of the study to participants, and the importance of collecting serial samples. Also, a movie portraying a study visit and/or explaining the set-up of the study provides further clarification to participating women. To keep study subjects informed about the course of the study, one should consider sending regular study newsletters, frequently update the study website, and organize annual meetings where the study team is present. The latter provides study participants ample opportunity to ask questions, which in our studies has proven to be very well appreciated. A combination of abovementioned strategies leads to a clear realization of the importance of their study participation. This is not only relevant for the participating women, but also for team members in participating hospitals.

Despite these various strategies, drop-outs are almost an implicit component of long-term clinical studies. The main reasons for women to stop participating in the NAF study, in rank order (from highest to lowest), were: (i) reaching the end of the study according to protocol (established as 10 years in the NAF study; 18%), (ii) repeated unsuccessful aspiration (17%), (iii) preventive bilateral mastectomy (8.6%), (iv) loss to follow-up (7.4%), (v) no screening appointments at the hospital (6.7%), and (vi) development of breast cancer (6%), among other reasons. Planned prophylactic mastectomies, specifically for *BRCA1* or *BRCA2* mutation carriers, are unavoidable drop-outs that can be expected. For the NAF study, a specific drop-out reason was repeated unsuccessful NAF acquisition (i.e., 0 μL), which can be related to parameters such as menopausal status, spontaneous nipple discharge, breast size, bilateral oophorectomy, and previous use of hormone replacement therapy or antihormonal treatment (34). Other reasons were logistic issues, such as relocation of participants to another area and hospital, no longer having time to come to the outpatient clinic, or the end of hospital screening (i.e., return to regular out-of-hospital nationwide screening). These problems could theoretically be surpassed if blood draws could be taken at any location (e.g., at a blood draw clinic or at the general practitioner's office). However, in practice, women may overlook or postpone having their blood drawn. To avoid long waiting times as a potential barrier, the TESTBREAST study team established a rule of priority for blood sampling at the laboratory. This resulted in an increased willingness of up to 86% of the study population to have their blood samples taken. All of these factors should be considered when estimating the number of required inclusions.

To summarize, to safeguard the inclusion rate, the study team should ideally inform potential candidates about the study before the next scheduled clinical screening visit. For subsequent study visit invitations, the research team should create and maintain a recall system to invite study subjects timely and sample acquisition should be facilitated by implementing a fast track for obtaining blood samples. On top of that, the core study team should keep study subjects and other participating team members informed about the course of the study. These strategies diminish the number of unnecessary drop-outs and missed study visits.

Lesson 5. Logistics

Given the high number of participants and an even higher number of study appointments in this trial design, it can be challenging to maintain continuous registration of the incoming data. For example, in the NAF study, a two-center study in which every study visit takes about 1 hour, up to 12 study visits are planned every week. Study visit registration aspects include digital processing of a questionnaire and collection, processing, registration, and biobanking of NAF and blood samples which are aliquoted into 22 vials (three vials per breast for NAF, 10 vials for serum, and six vials for plasma). Questionnaire data comprise lifestyle and hormonal factors. Sample registration data comprise NAF sample color, including bloody appearance, consistency, and volume per aliquot. NAF color can vary between study subjects, between breasts from the same woman, and even within a breast. NAF color registration provides insight into the color variation and permits exploration of its association with breast cancer risk. NAF samples should be labeled as bloody or nonbloody before analysis, as bloody NAF samples could lead to alterations in biomarker analyses compared with nonbloody NAF. TESTBREAST study visits comprise a questionnaire and a blood draw, which, in contrast to the NAF study, does not need to be performed by trained study team members, but can be done by regularly trained hospital personnel. Because the TESTBREAST is a nine-center study with one to four visits per participant per year (as established by hospital-specific screening guidelines and commitment by study participants), the number of incoming questionnaires and blood samples to be processed and stored is high. Therefore, it is key to set up a well-established administrative process. To guarantee a good cohesive administration that is appropriately maintained according to data-monitoring rules, annual monitoring by an independent clinical research associate (CRA) throughout the complete course of the study is advisable. In addition, both a data manager and a(n) (online) data management system (e.g., ProMISe for the TESTBREAST study) are essential to keep the incoming information up to date and facilitate the traceability of samples (37). Recent developments of online databases also allow integration of online questionnaires. This ensures completeness of the questionnaires and automated integration of the information directly into the database.

A turnover of study team members is expected during a long-term study. Such a turnover can occur in team members that are daily and routinely involved in the study, like research nurses and the coordinator of the clinical study. In the Netherlands, coordination of a clinical study is usually performed by an MD-PhD student. As they ought to finish their research project within 4 years, several PhD students are usually involved in such a project before it reaches its end. A way to maintain consistent working procedures is to establish and frequently update standardized operating procedures (SOP), which are supervised by a CRA. Also, contact information of team members should be updated and departing team members should train new incoming members about the SOPs. In multicenter studies, a fixed contact person on-site is responsible for the correct local execution of the study. A site training of a specific study technique (such as the nipple fluid aspiration) might be necessary in some studies. Structured meetings with the local study coordinators are recommended to secure proper collection of data.

Lesson 6. Sample Processing

A difficulty of long-term, prospective studies is to guarantee optimal quality of biological material for multiple years until the time of analysis. For instance, a prolonged time in the freezer is accompanied with the risk of sample quality loss and loss of sample by evaporation (38). Therefore, correct sample handling during collection and storage is essential. Sample collection and processing aspects to be considered include consistent use of the same type of collection tubes and buffers, a predefined time until sample processing and freezing, and a defined monitored temperature at which samples are stored. In addition, aliquoting samples is a good strategy to later avoid freeze–thaw cycles, which possibly influences sample quality. Finally, pilot sample analysis should not be postponed until the end of the study, as it allows timely sample quality monitoring and assessment of natural temporal biomarker fluctuations (**Fig. 1**). In this context, pilot testing to ensure that the collected samples are adequate for biomarker analyses should be performed. As an example, in the NAF study, pilot testing has shown that, regardless of the RNA concentration at the start, miRNA qPCR can always be measured given the high sensitivity of this technique (39).

Working with a number of hospitals over a long period of time comes together with the potential pitfall that different types of collection tubes and a variety of buffers are used across centers or even within centers, which could hamper a comparable sample analysis. It is important to maintain consistency among and within participating hospitals. Recommended approaches include providing sets of collection tubes, labels, laboratory forms, and questionnaires to all centers, as has been part of the standard procedure in the NAF and TESTBREAST studies.

Finally, prolonged processing time of samples can lead to the alteration of biomarker characteristics and cause the final results to deflect (40). Therefore, continuous presence of a team member in the laboratory is required to process new samples within the defined time limit. The standards used for sample processing in the NAF and TESTBREAST studies are depicted in **Table 3**.

Lesson 7. Sampling at Events

For both longitudinal cohorts of the NAF and the TESTBREAST studies, multiple samples from study subjects are prospectively acquired; the last sample is collected if and when an event occurs. As the primary aim of the studies is to compare data from liquid biopsies obtained at the time of an event with those acquired before diagnosis, sampling at the time of an event is valuable. This sample is relevant in the investigational setting, because biomarker levels at the time of the event best reflect the carcinogenetic signature and, as such, function as the reference levels in the paired analysis (**Fig. 1**).

As participating women follow an intensive regular screening program, an event may be detected during such an appointment, which is usually combined with a study visit. In those cases, sampling at the time of an event is guaranteed. However, for interval cancers, chances are that liquid biopsy collection is missed, because neither the treating physician nor the study participant is sufficiently aware of the

Table 3. Sample processing in the NAF and TESTBREAST studies.

	NAF study	TESTBREAST study
Sample processing time	<60 minutes	<4 hours
Centrifuging speed	15 minutes at 300 × *g* (1,200 rpm) for serum Mini-centrifugation for NAF	10 minutes at 1,000 × *g*
Storage temperature	−80°C	−80°C
Sample volume	10 × 600 μL for serum 6 × 600 μL for plasma 6 × 0–30 μL for NAF	550 μL and 4 × 500 μL for serum
Type of collection tubes	BD Vacutainer SST II Advance for serum BD Vacutainer K2E (EDTA) for plasma Brooks FluidX 0.7 mL external threat tube for NAF	BD Vacutainer SST II Advance for serum
Analysis	RT-qPCR	Mass spectrometry methods

Abbreviation: RT-qPCR, reverse transcription quantitative PCR.

importance to inform the study team: the physician because he/she might not be informed about study participation and the participant may neglect to inform the study team about her diagnosis, because of her overwhelming current situation. An active, regular search by the study team to rule out or confirm new breast-related pathology is not efficient, is labor-intensive, and has a high chance of still leading to discovering the events too late (i.e., after they started treatment). A practical, automated solution is to generate a "study alert notification" in electronic platforms such as the electronic health records or in the nationwide online registry of pathology reports (in the Netherlands, this has the acronym PALGA; ref. 41), so that new events will be reported to the study team before the start of treatment. Such interconnection between the cohort and tissue repositories and cancer registries is also essential to ensure a complete follow-up, provided that this interconnection is included in the informed consent form. Probably the most practical approach is instructing study participants to inform the research team every time they have additional hospital visits for breast-related issues. Therefore, participants must be well-informed, well-instructed, and perhaps, regularly reminded that when they develop an event they should contact the study team for final sampling. Still, understandably, this is a delicate time for a woman to consider an additional hospital visit, especially in the context of a clinical study.

Finally, there is a chance that women develop an event after having completed the study period. According to the Medical Research Ethics Committee guidelines, the protocol of a clinical study is required to have a specified delimited study period. Absence of such a limit could lead to endless clinical studies with the threat of losing aim and perspective. Still, we recommend to add a prolonged time margin to the study protocol, which would allow women to participate for a longer period and to have an additional sample taken at the time of an event, provided that this sample is acquired within a reasonable amount of time after the end of the study period.

By combining the abovementioned approaches, chances of acquiring a sample at the time of an event are increased, leading to a more complete intrasubject sample series in the biobank and, thereby, achieving the main aim of the study.

Lesson 8. Biobanking

As a consequence of the high number of samples in these specific studies, storage room might become an issue. As explained in lesson 5, around one to 22 aliquots per study visit per study subject were acquired throughout the TESTBREAST and the NAF studies, respectively. To give an impression, in the NAF study, more than 15,000 vials were stored in 11 years' time for 555 study subjects. In the TESTBREAST study, more than 3,000 samples were acquired between 2011 and 2019 from more than 930 participants (some blood sample series for the same women surpass the 20 visits). Therefore, it is essential to timely ensure substantial biobank storage, which will enable holding the samples safely for many years. In parallel, a well-defined biobanking system that allows for continuous sample registration in a database and consensual labeling throughout the study should be operational. Labels should be cold and moisture resistant, printed, and include standardized information, ideally the name of the team member who acquired and handled the sample, study subject number, visit number, sample type, volume, aliquot number, sample date, and a QR code. Specifically for the NAF study, breast side of the NAF sample should be added to the label.

Lesson 9. Sample Handling at the Time of Analysis

As highlighted in lesson 1, only a relatively small part of the cohort is expected to develop an event. Depending on the ratio of sample volume attained and technical volume required, samples could perhaps only be used once. For instance, in the NAF study, given the small volume acquired (10–50 μL), the majority or the complete volume may be needed for a single experiment. Thus, application of a trustworthy, familiar technique that has already been optimized to analyze the presence of biomarkers provides reassurance.

The exploration of improved, cheaper, new, faster, and potentially better techniques developed during the course of the study should be tested. This is inherent to technological developments and cannot be anticipated upon. Therefore, sufficient sample volume has to be acquired to allow for such technical exploration and comparisons. We recommend considering including technical testing in the list of secondary aims of the study protocol. Obtaining more volume is understandably less of a limitation for blood samples than for the low-volume NAF samples. As highlighted in lesson 8, a downside of obtaining higher volumes in general is that even more samples will have to be stored in the biobank, which demands for more storage.

Lesson 10. Adjustment of the Biomarker of Choice

The biomarker of choice to investigate in the high-risk cohort often derives from case–control studies. Specifically for the TESTBREAST

study, the focus of research is based on previous promising results in proteomic expression profiling, showing very high sensitivity and specificity for the detection of breast cancer (42–44). However, preliminary case–control studies may also lead to alteration of the initially chosen biomarker due to insufficient diagnostic accuracy in interim analyses or in other studies. For instance, in the NAF study, gene methylation was the subject of investigation at first, but when interim analyses did not reveal a sufficient AUC to justify further translation into a clinical test (45), the biomarker of choice changed to miRNA. Furthermore, alterations may occur due to the discovery of new classes of biomarkers. Therefore, it is important to create the ideal conditions to be able to switch to another biomarker during the study by, among others, obtaining broad informed consent. Nevertheless, it is difficult to define what the ideal conditions are to be able to switch to new biomarkers, since this anticipates future developments. The use of buffers that allow different types of isolations, of, for example, DNA, RNA, and proteins, is advisable. Storing sufficient amounts of biofluids in different aliquots allows even more types of analyses than initially anticipated for the project. Moreover, storing different blood (half) products, such as whole blood, plasma, and serum is recommended, but, as a downside, leads to increased preprocessing costs. When switching to another biomarker or technical platform, one has to be aware that new markers and platforms will have to be technically validated, and the validity of the "old" data has to be established and might even have to be discarded.

Lesson 11. Nested Case–Control Analysis

There are a few issues that hamper the nested case–control analysis. First, the envisioned time frames of sampling (**Fig. 1**) are seldom exactly reached. The main reasons include delayed screening appointments at the hospital, which are due to hospital logistics, a woman's preference, pregnancy (a contraindication for imaging), or lactation (leads to reduced imaging sensitivity; ref. 12). Because study visits are preferentially combined with hospital screening visits, a change of the

Table 4. Summary of the challenges, lessons learned, and recommendations for prospective, longitudinal, multicenter studies of women with a high risk of developing breast cancer.

Challenges	Lessons learned and recommendations
1. Study phases and cohort size	• Engage all necessary experts from the start of the study. • Consider in the sample size calculation of the complete cohort, matters like study withdrawals, lowered LTRs, and incomplete sample series. Sample size customization during the study and an extended study time frame may be necessary.
2. Funding	• Divide the study aims in parts to apply for several infrastructural and research funding grants throughout the years.
3. The inclusion criteria and exclusion criteria are dynamic	• A woman's LTR may lower: describe which measures to take in the study protocol and in the study information form. • Investigate which factors may influence the biomarker of choice and add these to the list of exclusion criteria in the study protocol.
4. Delayed inclusion moment and discontinuation of participation	• Inform potential candidates about the study before the scheduled screening appointment. This allows the first study visit to take place together with the upcoming scheduled screening moment, avoiding study inclusion delay. • Provide regular study updates to study subjects and team members.
5. Logistics	• Ensure a data management system that allows data overview and sample traceability. • Monitor the study administration yearly. • Generate SOPs and use them to train new members of the study team.
6. Sample processing	• Keep consistency by using the same collection tubes and buffers. • Assure sample quality during multiple years by choosing the best buffer, aliquoting samples, defining the maximum time until sample processing, time until freezing, and storage temperature. • Periodically test sample quality. • Compare serial samples to assess natural temporal biomarker fluctuation.
7. Sampling at events	• Keeping study subjects involved and having dedicated research nurses are relevant strategies to increase sampling at events. • Add a time margin in the protocol for sampling at events. • Awareness among study subjects about the relevance of notifying the study team in case of an event is crucial. • Interconnection of the cohort with tissue repositories and cancer registries is essential to ensure a complete follow-up, provided that such interconnection is included in the informed consent form.
8. Biobanking	• Ensure the availability of substantial biobanking storage. • Establish a well-defined biobanking system with consensual labeling.
9. Sample handling at the time of analysis	• Use a trustworthy, optimized, familiar technique. • Compare promising new, cheap, and evolved techniques to perform the primary analysis with the most optimal one. Include such a technical analysis in the secondary aims of the study.
10. Adjustment of the biomarker of choice	• Get informed consent as broad as possible to allow analysis of various and emerging biomarker classes.
11. Nested case–control analysis	• Keep a real-time overview of the number of events, including the timing and number of successfully acquired samples. Envisioned time frames of sampling and completeness of sample series are seldom reached. This can be overcome in the statistical analysis by applying linear mixed models.

latter affects the timepoints of the study. Consequently, in practice, there will be variation in intervals of sample acquisition within the same subject and between study subjects. Second, another naturally occurring issue is incompleteness of sample series. This is because of missed study visits or unsuccessful sample collection. The statistical approach that allows data analysis with individual repeated measurements in case of missing data points and flexible time schedules, includes the use of linear mixed models (46). A solution could be to cluster data sampled at several timepoints into categorized subgroups (e.g., <6, 6–12, and >12 months before the event, as clustered by Weber and colleagues; ref. 47). Still, to optimally allow paired comparison analysis, it is of importance to continuously keep a real-time overview of the number of events and also to keep the number of successfully acquired samples and timing of sample acquisition up to date. If incomplete series are timely noticed, adjusting the sample sizes accordingly is still possible.

Discussion

Biological markers in liquid biopsies reflecting carcinogenesis could be of great additional value to further improve the screening program for breast cancer, specifically in high-risk women. To find a suitable panel of biomarkers for early breast cancer detection, a longitudinal, prospective cohort should be set up to allow serial collection of samples from high-risk women until an event occurs. This permits serial analysis of changes in biomarker levels over time and especially how long beforehand biomarkers can signal breast cancer onset. Nevertheless, this study design is also accompanied with pitfalls, which were encountered by two independent research teams setting up similar long-lasting studies. Some of the pitfalls can be circumvented by defining crucial elements in the protocol before the start of the study. Also, proper SOPs should be written to contribute to uniformity in study execution, because change of personnel is unavoidable in view of the lengthy time of the study. Moreover, substantial biobank storage needs to be established timely and study subject withdrawal and missing data should be monitored closely. However, one of the greatest challenges is to have a sustained flow of funding, for instance, by research foundations or generous private donors. On a general note, all the advances in knowledge and technology, such as database management, online questionnaires, QR codes, and laboratory technologies, cannot be anticipated at the time of the study set-up, but should be actively tailored throughout the course of the study. **Table 4** provides a summary of these different pitfalls, lessons learned by us, and recommendations that may be of use for other research groups setting up similar long-lasting cohort studies.

Authors' Disclosures

No disclosures were reported.

Acknowledgments

The authors are greatly indebted to all women who participated and are participating in the NAF and TESTBREAST studies. W.E. Mesker received funding from "Het Bollenstreekfonds," A Sister's Hope, Nuts-Ohra and Pink Ribbon (grant no. 2011.VW06.C725). P.J. van Diest received funding from a private donor (Mr. H) and a private foundation (Foundation M).

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received May 25, 2020; revised July 30, 2020; accepted October 9, 2020; published first October 20, 2020.

References

1. Kuchenbaecker KB, Hopper JL, Barnes DR, Phillips KA, Mooij TM, Roos-Blom MJ, et al. Risks of breast, ovarian, and contralateral breast cancer for BRCA1 and BRCA2 mutation carriers. JAMA 2017;317:2402–16.
2. Smith RA, Andrews KS, Brooks D, Fedewa SA, Manassaram-Baptiste D, Saslow D, et al. Cancer screening in the United States, 2017: a review of current American Cancer Society guidelines and current issues in cancer screening. CA Cancer J Clin 2017;67:100–21.
3. Familial breast cancer: classification, care and managing breast cancer and related risks in people with a family history of breast cancer; [about 46 screens]. Available from: https://www.nice.org.uk/guidance/cg164/chapter/Recommendations#surveillance-and-strategies-forearly-detection-of-breast-cancer.
4. Screening buiten het landelijk bevolkingsonderzoek; [about 2 screens]. Available from: https://www.oncoline.nl/borstkanker.
5. Riedl CC, Luft N, Bernhart C, Weber M, Bernathova M, Tea MK, et al. Triple-modality screening trial for familial breast cancer underlines the importance of magnetic resonance imaging and questions the role of mammography and ultrasound regardless of patient mutation status, age, and breast density. J Clin Oncol 2015;33:1128–35.
6. Nelson HD, Pappas M, Cantor A, Griffin J, Daeges M, Humphrey L. Harms of breast cancer screening: systematic review to update the 2009 U.S. Preventive Services Task Force Recommendation. Ann Intern Med. 2016;164:256–67.
7. Drooger JC, Hooning MJ, Seynaeve CM, Baaijens MH, Obdeijn IM, Sleijfer S, et al. Diagnostic and therapeutic ionizing radiation and the risk of a first and second primary breast cancer, with special attention for BRCA1 and BRCA2 mutation carriers: a critical review of the literature. Cancer Treat Rev 2015;41:187–96.
8. Heemskerk-Gerritsen BAM, Jager A, Koppert LB, Obdeijn AI, Collee M, Meijers-Heijboer HEJ, et al. Survival after bilateral risk-reducing mastectomy in healthy BRCA1 and BRCA2 mutation carriers. Breast Cancer Res Treat 2019;177:723–33.
9. 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;351:427–37.
10. Obdeijn IM, Loo CE, Rijnsburger AJ, Wasser MN, Bergers E, Kok T, et al. Assessment of false-negative cases of breast MR imaging in women with a familial or genetic predisposition. Breast Cancer Res Treat 2010;119:399–407.
11. Saadatmand S, Geuzinge HA, Rutgers EJT, Mann RM, de Roy van Zuidewijn DBW, Zonderland HM, et al. MRI versus mammography for breast cancer screening in women with familial risk (FaMRIsc): a multicentre, randomised, controlled trial. Lancet Oncol 2019;20:1136–47.
12. Carmichael H, Matsen C, Freer P, Kohlmann W, Stein M, Buys SS, et al. Breast cancer screening of pregnant and breastfeeding women with BRCA mutations. Breast Cancer Res Treat 2017;162:225–30.
13. Evans A, Trimboli RM, Athanasiou A, Balleyguier C, Baltzer PA, Bick U, et al. Breast ultrasound: recommendations for information to women and referring physicians by the European Society of Breast Imaging. Insights Imaging 2018;9:449–61.
14. Boeri M, Verri C, Conte D, Roz L, Modena P, Facchinetti F, et al. MicroRNA signatures in tissues and plasma predict development and prognosis of computed tomography detected lung cancer. Proc Natl Acad Sci U S A 2011;108:3713–8.
15. Montani F, Marzi MJ, Dezi F, Dama E, Carletti RM, Bonizzi G, et al. miR-Test: a blood test for lung cancer early detection. J Natl Cancer Inst 2015;107:djv063.
16. Page EC, Bancroft EK, Brook MN, Assel M, Hassan Al Battat M, Thomas S, et al. Interim results from the IMPACT study: evidence for prostate-specific antigen screening in BRCA2 mutation carriers. Eur Urol 2019;76:831–42.
17. Casabonne D, Benavente Y, Seifert J, Costas L, Armesto M, Arestin M, et al. Serum levels of hsa-miR-16-5p, hsa-miR-29a-3p, hsa-miR-150-5p, hsa-miR-155-5p and hsa-miR-223-3p and subsequent risk of chronic lymphocytic leukemia in the EPIC study. Int J Cancer 2020;147:1315–24.

18. Widschwendter M, Zikan M, Wahl B, Lempiainen H, Paprotka T, Evans I, et al. The potential of circulating tumor DNA methylation analysis for the early detection and management of ovarian cancer. Genome Med 2017;9:116.
19. Cohen JD, Li L, Wang Y, Thoburn C, Afsari B, Danilova L, et al. Detection and localization of surgically resectable cancers with a multi-analyte blood test. Science 2018;359:926–30.
20. Blyuss O, Burnell M, Ryan A, Gentry-Maharaj A, Marino IP, Kalsi J, et al. Comparison of longitudinal CA125 algorithms as a first-line screen for ovarian cancer in the general population. Clin Cancer Res 2018;24:4726–33.
21. Lin XJ, Chong Y, Guo ZW, Xie C, Yang XJ, Zhang Q, et al. A serum microRNA classifier for early detection of hepatocellular carcinoma: a multicentre, retrospective, longitudinal biomarker identification study with a nested case-control study. Lancet Oncol 2015;16:804–15.
22. van Asperen CJ, Bangma CH. Screening naar prostaatkanker bij mannen met een BRCA2 mutatie: adviezen voor de praktijk gebaseerd op resultaten van internationale studie. Tijdschrift voor Urologie 2020;10:36–9.
23. Ausems MGEM, Kiemeney LALM. Genetisch onderzoek bij prostaatkanker: nieuwe ontwikkelingen. Tijdschrift voor Urologie 2020;10:30–5.
24. Wanders JOP, Holland K, Karssemeijer N, Peeters PHM, Veldhuis WB, Mann RM, et al. The effect of volumetric breast density on the risk of screen-detected and interval breast cancers: a cohort study. Breast Cancer Res 2017; 19:67.
25. Bakker MF, de Lange SV, Pijnappel RM, Mann RM, Peeters PHM, Monninkhof EM, et al. Supplemental MRI screening for women with extremely dense breast tissue. N Engl J Med 2019;381:2091–102.
26. Kurian AW, Munoz DF, Rust P, Schackmann EA, Smith M, Clarke L, et al. Online tool to guide decisions for BRCA1/2 mutation carriers. J Clin Oncol 2012; 30:497–506.
27. Netherlands Trial Register: Trial NL8661. NAF study. Early detection of hereditary breast cancer by monitoring microRNA expression in nipple aspirate fluid; [about 4 screens]. Available from:https://www.trialregister.nl/trial/8661.
28. Netherlands Trial Register: Trial NL8724. TESTBREAST study. Early detection of breast cancer in women with a familial or genetic predisposition of developing breast cancer using biomarkers in serum; [about 3 screens]. Available from: https://www.trialregister.nl/trial/8724.
29. Rosenthal ET, Evans B, Kidd J, Brown K, Gorringe H, van Orman M, et al. Increased identification of candidates for high-risk breast cancer screening through expanded genetic testing. J Am Coll Radiol 2017;14:561–8.
30. Evans DG, Howell A. Breast cancer risk-assessment models. Breast Cancer Res 2007;9:213.
31. Giardiello D, Steyerberg EW, Hauptmann M, Adank MA, Akdeniz D, Blomqvist C, et al. Prediction and clinical utility of a contralateral breast cancer risk model. Breast Cancer Res 2019;21:144.
32. MacDonald E, Dadds MR, Brennan JL, Williams K, Levy F, Cauchi AJ. A review of safety, side-effects and subjective reactions to intranasal oxytocin in human research. Psychoneuroendocrinology 2011;36:1114–26.
33. Suijkerbuijk KP, van der Wall E, Meijrink H, Pan X, Borel Rinkes IH, Ausems MG, et al. Successful oxytocin-assisted nipple aspiration in women at increased risk for breast cancer. Fam Cancer 2010;9:321–5.
34. de Groot JS, Moelans CB, Elias SG, Hennink A, Verolme B, Suijkerbuijk KP, et al. Repeated nipple fluid aspiration: compliance and feasibility results from a prospective multicenter study. PLoS One 2015;10:e0127895.
35. Suijkerbuijk KP, van der Wall E, van Diest PJ. Oxytocin: bringing magic into nipple aspiration. Ann Oncol 2007;18:1743–4.
36. Nelson HD, Fu R, Zakher B, Pappas M, McDonagh M. Medication use for the risk reduction of primary breast cancer in women: updated evidence report and systematic review for the US Preventive Services Task Force. JAMA 2019;322: 868–86.
37. Schiermeier Q. Data management made simple. Nature 2018;555:403–5.
38. Gast MC, van Gils CH, Wessels LF, Harris N, Bonfrer JM, Rutgers EJ, et al. Influence of sample storage duration on serum protein profiles assessed by surface-enhanced laser desorption/ionisation time-of-flight mass spectrometry (SELDI-TOF MS). Clin Chem Lab Med 2009;47:694–705.
39. Moelans CB, Patuleia SIS, van Gils CH, van der Wall E, van Diest PJ. Application of nipple aspirate fluid miRNA profiles for early breast cancer detection and management. Int J Mol Sci 2019;20:5814.
40. Yin P, Lehmann R, Xu G. Effects of pre-analytical processes on blood samples used in metabolomics studies. Anal Bioanal Chem 2015;407:4879–92.
41. Casparie M, Tiebosch AT, Burger G, Blauwgeers H, van de Pol A, van Krieken JH, et al. Pathology databanking and biobanking in The Netherlands, a central role for PALGA, the nationwide histopathology and cytopathology data network and archive. Cell Oncol 2007;29:19–24.
42. de Noo ME, Deelder A, van der Werff M, Ozalp A, Mertens B, Tollenaar R. MALDI-TOF serum protein profiling for the detection of breast cancer. Onkologie 2006;29:501–6.
43. van der Werff MP, Mertens B, de Noo ME, Bladergroen MR, Dalebout HC, Tollenaar RA, et al. Case-control breast cancer study of MALDI-TOF proteomic mass spectrometry data on serum samples. Stat Appl Genet Mol Biol 2008;7:1–12.
44. Velstra B, van der Burgt YE, Mertens BJ, Mesker WE, Deelder AM, Tollenaar RA. Improved classification of breast cancer peptide and protein profiles by combining two serum workup procedures. J Cancer Res Clin Oncol 2012;138:1983–92.
45. de Groot JS, Moelans CB, Elias SG, Jo Fackler M, van Domselaar R, Suijkerbuijk KP, et al. DNA promoter hypermethylation in nipple fluid: a potential tool for early breast cancer detection. Oncotarget 2016;7:24778–91.
46. Krueger C, Tian L. A comparison of the general linear mixed model and repeated measures ANOVA using a dataset with multiple missing data points. Biol Res Nurs 2004;6:151–7.
47. Weber DG, Brik A, Casjens S, Burek K, Lehnert M, Pesch B, et al. Are circulating microRNAs suitable for the early detection of malignant mesothelioma? Results from a nested case-control study. BMC Res Notes 2019;12:77.