

Examples of Funded Grants in Healthcare Delivery Research

Overview

The National Cancer Institute (NCI) frequently receives requests for examples of funded grant applications. Several investigators and their organizations agreed to let the Healthcare Delivery Research Program (HDRP) post excerpts of their healthcare delivery research grant applications online.

About

We are grateful to the investigators and their institutions for allowing us to provide this important resource to the community. We only include a copy of the SF 424 R&R Face Page, Project Summary/Abstract (Description), Project Narrative, Specific Aims, and Research Strategy; we do not include other SF 424 (R&R) forms or requisite information found in the full grant application (e.g., performance sites, key personnel, biographical sketches). To maintain confidentiality, we have redacted some information from these documents (e.g., budgets, social security numbers, home addresses, introduction to revised application).

Copyright Information

The text of the grant applications is copyrighted. Text from these applications can only be used for nonprofit, educational purposes. When using text from these applications for nonprofit, educational purposes, the text cannot be changed and the respective Principal Investigator, institution, and NCI must be appropriately cited and credited.

Accessibility

Individuals using assistive technology (e.g., screen reader, Braille reader) who experience difficulty accessing any information should send an email to the Healthcare Delivery Research Program (NCIHDRP@mail.nih.gov).

424 R&R and PHS-398 Specific Table Of Contents

| | |
|---|----|
| Examples of Funded Grants in Healthcare Delivery Research | 1 |
| Table Of Contents | 2 |
| SF 424 R&R Face Page | 3 |
| Abstract | 4 |
| Narrative | 5 |
| Specific Aims | 6 |
| A. SIGNIFICANCE | 7 |
| B. INNOVATION | 8 |
| C. APPROACH | 9 |
| References | 14 |

SF 424 R&R Face Page

PI: Sprague, Brian L

Grant Number: 1 R03 CA223725-01

Title: Multi-center evaluation of digital breast tomosynthesis with synthesized two-dimensional mammography for breast cancer screening

FOA: PAR16-416

FOA Title: NCI SMALL GRANTS PROGRAM FOR CANCER RESEARCH (NCI OMNIBUS R03)

Organization: UNIVERSITY OF VERMONT & ST AGRIC COLLEGE

Department: Surgery

Senior/Key Personnel: Brian Sprague

Organization: The University of Vermont and State Agricultural College

Role Category: PD/PI

Abstract

Digital mammography (DM) combined with digital breast tomosynthesis (DBT) has emerged as a promising primary breast cancer screening technology with improvements in specificity and sensitivity compared to screening with DM alone. However, the total radiation dose of a combined DM/DBT study is often twice that of screening with DM alone, raising concerns about the cumulative dose over a lifetime of screening. In 2013, the FDA approved synthetic “2D-like” (s2D) images, derived from data acquired during the DBT acquisition, as a replacement for the conventional “dose” DM imaging in DBT screening. Our preliminary single site data and early reader studies suggest that s2D/DBT screening may further reduce recall and biopsy rates compared to DM/DBT, while providing high invasive breast cancer detection rates but reduced detection of ductal carcinoma in situ (DCIS). In addition, the mammographic density assignments from s2D/DBT screening changed significantly with less patients deemed “dense”, which could impact risk assessment and density-based clinical decisions regarding supplemental screening since breast density notification is mandated by laws in more than half of US states. We propose a multi-center observational study with adequate power to comprehensively characterize multiple aspects of s2D/DBT screening performance. We will leverage a natural experiment at early adopter radiology facilities participating in the University of Vermont and University of Pennsylvania PROSPR Research Centers that have transitioned from DM/DBT to s2D/DM for routine screening. Our specific aims are 1) To determine the clinical screening performance (e.g., sensitivity and specificity) of breast cancer screening with s2D/DBT; 2) To evaluate the rates of specific malignant and benign lesions detected by breast cancer screening with s2D/DBT; and 3) To determine the distribution of mammographic breast density assessments on s2D/DBT exams. For all aims, we will compare s2D/DBT screening performance to the performance of DM/DBT screening at the same facilities prior to their adoption of s2D/DBT. We anticipate a sample size of approximately 86,000 s2D/DBT exams and 98,000 DM/DBT exams. The University of Vermont and University of Pennsylvania PROSPR Research Centers have prospectively collected breast cancer screening data since 2011 in a standardized fashion to permit pooled analyses. Our registries contain exam-level breast imaging common data elements, including high quality information on indication of exam (i.e., screening vs. diagnostic), mammographic views, assessments, and breast density, linked to pathology databases and statewide cancer registries thereby permitting investigation of both benign and malignant diagnoses. The results of this project will provide needed evidence for women evaluating screening options, providers assessing whether to adopt s2D/DBT, health care payers considering reimbursement rates, and policy makers debating breast cancer screening recommendations in the context of this new technology.

Narrative

Digital breast tomosynthesis (“3D mammography”) is a new breast imaging technology that improves breast cancer screening performance when added to conventional 2D digital mammography, though there are concerns regarding the extra radiation dose. Our multi-center study will evaluate the clinical performance of a new approach in which digital breast tomosynthesis is used without conventional 2D mammography.

Specific Aims

Digital breast tomosynthesis (DBT) is disseminating widely into clinical practice for primary breast cancer screening.¹⁻³ This new x-ray technique images the breast with multiple low-dose exposures obtained across an arc, which are reconstructed into a series of thin images of the breast.^{4,5} We and others have shown that DBT reduces the frequency of false-positive exams and increases invasive breast cancer detection when added to conventional 2D digital mammography (DM).^{6,7} This evidence has resulted in accelerated adoption of DBT technology in clinical practice, expanded health insurance reimbursement for the additional costs of DBT, and provides the basis for ongoing policy debates about optimal breast cancer screening strategies.

However, implementation of DBT is changing since the FDA approved in 2013 the use of DBT with synthesized two-dimensional (s2D) mammography to reduce the total radiation dose when compared to imaging with combined DM/DBT. Nearly all evidence to date regarding the performance of DBT has been based on imaging exams in which DBT is performed concurrently with DM.⁶⁻⁸ The goal of this proposal is to evaluate the clinical performance of DBT with s2D (s2D/DBT), without conventional 2D DM imaging. With concurrent DM/DBT the interpreting radiologist has access to both the conventional 2D DM views as well as the quasi-3D DBT image set. By exchanging the “dose” DM images with s2D images reconstructed from the DBT acquisition, the radiation dose is significantly reduced but the interpreting radiologist is presented only with s2D images that may appear quite different from a conventional DM image. Our preliminary single site data and early reader studies suggest that s2D/DBT screening may further reduce recall and biopsy rates compared to DM/DBT, while providing high invasive breast cancer detection rates but reduced detection of ductal carcinoma in situ (DCIS).^{9,10} In addition, there may be a significant change in mammographic density ratings when s2D/DBT is used,⁹ potentially impacting risk assessment and density-based clinical decisions regarding supplemental screening, as density notification is now mandated by law in more than half of US states.¹¹ In this application we propose a multi-center assessment of the clinical performance of DBT with synthetic 2D images (s2D/DBT), with comparison to the performance metrics at these same facilities when using DM/DBT prior to their adoption of s2D/DBT. We will analyze existing data from two comprehensive breast imaging registries within the National Cancer Institute’s PROSPR (Population-based Research Optimizing Screening Through Personalized Regimens) consortium: the University of Vermont and the University of Pennsylvania PROSPR Research Centers. These centers include data from diverse radiology facilities that have been early adopters of s2D/DBT screening, thereby offering a unique opportunity to efficiently assess the performance of this new breast cancer screening test. Our specific aims are:

- 1) To determine the clinical performance characteristics of breast cancer screening with s2D/DBT.** We hypothesize that compared to DM/DBT screening performance, s2D/DBT will have lower recall and higher specificity, with comparable overall cancer detection, sensitivity, and false negative rates.
- 2) To evaluate the rates of specific malignant and benign lesions detected by breast cancer screening with s2D/DBT.** We hypothesize that compared to DM/DBT detection rates, s2D/DBT will have elevated invasive breast cancer detection, lower DCIS detection, and lower benign biopsy rates.
- 3) To determine the distribution of mammographic breast density assessment on s2D/DBT exams.** We hypothesize that compared to the density distribution for DM/DBT screening exams, the prevalence of heterogeneously and extremely dense breasts will be lower on s2D/DBT screening exams.

Since 2011, the University of Vermont and University of Pennsylvania PROSPR Research Centers have prospectively collected breast imaging and outcomes data in a standardized fashion to permit pooled analyses. The registries contain exam-level breast imaging common data elements, including high quality information on indication of exam (i.e., screening vs. diagnostic), mammographic views, results, and breast density, linked to regional and statewide cancer registries and pathology databases thereby permitting investigation of both benign and malignant diagnoses. Radiologist and imaging facility IDs are available for all exams to permit consideration of within-radiologist and within-facility effects. **Our study will include approximately 86,000 s2D/DBT screening exams and 98,000 DM/DBT exams interpreted by 22 radiologists at 7 screening facilities, and for the first time provide adequate cancer outcomes follow-up and statistical power to examine specificity, sensitivity, false negative rates, and the distribution of malignant and benign lesions.** We have previously demonstrated our ability to collaborate productively addressing breast cancer screening questions, including the performance of DM/DBT and variation in breast density assessment.^{6,12}

The results of this project will provide needed evidence as to whether there are clinically significant differences in the performance of s2D/DBT breast cancer screening compared to DM/DBT. Our findings will be informative for women evaluating screening choices, providers assessing whether to adopt s2D/DBT, health care payers considering reimbursement rates, and policy makers debating breast cancer screening recommendations in the context of this new technology.

A. SIGNIFICANCE

Among US women, breast cancer is the most common cancer and the second most common cause of cancer mortality.¹³ Due to advances in screening and treatment, the mortality rate for breast cancer in the US has decreased steadily over the past 25 years.^{14,15} However, there remains widespread controversy and uncertainty regarding optimal screening regimens, due in part to concerns regarding the harms and limitations of mammography screening, including false-positive exams, benign biopsies, and missed cancers.¹⁶⁻²⁰ The United States Preventive Services Task Force has controversially advocated reduced screening intensity to mitigate screening harms.^{3,21} At the same time, digital breast tomosynthesis (DBT) has emerged as a new screening modality that could substantially improve the benefit-to-harm ratio for screening.¹ Evidence on the screening performance of DBT is accumulating; however, this evidence is almost exclusively based on the performance of DBT imaging in combination with conventional 2D digital mammography. A new technique is now FDA-approved to create synthetic 2D images reconstructed from the DBT views without the requirement of the 2D digital mammogram. This substantially reduces the radiation dose to the patient and is therefore highly attractive to health care providers and patients, yet there is very little evidence to date regarding the screening performance of this new approach.

A.1. Digital breast tomosynthesis (DBT) has emerged as a promising primary screening technology.

Digital mammography has been the standard of care for breast cancer screening of the general population since it supplanted film mammography more than 10 years ago.²² In 2011, a new technology called digital breast tomosynthesis (DBT; often referred to as “3D mammography”) was FDA-approved for use in combination with digital mammography (DM) and has since steadily disseminated into clinical practice for primary breast cancer screening.¹⁻³ This novel x-ray technique images the breast with multiple low-dose exposures obtained across an arc, which are then reconstructed into a series of thin images or “slices” of the breast.^{4,5} The radiologist can scroll through many thin mammographic images, thereby minimizing the masking effects of overlapping tissue that are problematic on conventional digital mammography views.²³ If DBT can “bend the curve” by improving both the sensitivity and specificity of breast cancer screening, the balance of benefits and harms for patients could improve dramatically.

Early observational studies suggest that DBT decreases recall rates and increases invasive cancer detection rates when added to conventional DM.^{7,8,24-31} The largest study to date (173,000 DBT screens) used aggregate data from 13 breast centers and found that combined DM/DBT screening was associated with an absolute reduction in recall rate of 16 per 1000 (95% CI, 14-18) and an increase in cancer detection rate of 1.2 per 1000 (95% CI, 0.8-1.6) compared to DM alone.⁷ While early data are promising, the USPSTF concluded in early 2016 that there is insufficient evidence to evaluate the benefits and harms of DBT for screening.³ The supporting AHRQ evidence review concluded that studies are needed from community practice that comprehensively capture interval cancers during at least one year of follow-up so that false-negative rates, sensitivity, and specificity can be accurately determined, as well as studies evaluating cancer characteristics to assess if DBT increases detection of indolent vs. aggressive cancer types.²⁴ Nevertheless, in January 2015 the Centers for Medicare & Medicaid Services enacted add-on billing codes to pay for the higher costs of adding DBT to digital mammography (DM) breast cancer screening.³²

A.2. Synthetic two-dimensional image reconstruction offers reduced radiation dose to patients.

Although it is beneficial to recall and cancer detection rates, the addition of DBT to DM screening effectively doubles the radiation dose to patients.^{33,34} While this dose still falls within the Mammography Quality Standards Act limit of 3.0 mGy,³⁵ mammography screening should be performed at the lowest possible dose that maintains optimal benefits.³⁶ Synthesized two-dimensional (s2D) mammography, generated via slab reconstruction from the DBT image set, was approved by the FDA in May 2013 as an alternative to combined DM/DBT screening. When s2D mammography is used in the place of DM, the dose received by patients is only from the DBT image acquisition; this was estimated to be approximately 45% lower than that of a combination DM/DBT study.³⁷ There have been a few reader studies and limited single institution clinical studies that compare the performance of s2D mammography versus digital mammography/DBT.^{9,10,37-41} Results have been mixed: two clinical studies and one reader study showed similar performance between both combination modalities;^{10,37,39,40} two clinical studies observing reduced recall rates and comparable cancer detection,^{9,41} and one reader study that showed lower sensitivity and comparable specificity in s2D mammography and DBT (referred to as s2D/DBT) versus DM/DBT.³⁸ The clinical studies to date have been limited by small sample sizes at single institutions, in addition to various threats to validity. In one study, symptomatic patients were included such that the cohort was not a true screening population.³⁷ In another analysis, DM was used to select cancer cases for reader interpretation, which may have led to selection bias.^{10,40}

A.3. Current breast cancer screening performance lags behind expert-recommended benchmarks.

There is widespread interest in evaluating the extent to which DBT technology can improve breast cancer screening performance metrics. Recently published screening performance statistics based on digital mammography found that in a large national sample of radiologists, only 59% had recall rates below the expert-recommended threshold of 12%, and 62% of radiologists did not meet the recommended range of 20-40% for positive predictive value of biopsy recommendation.⁴² In contrast, 92% of radiologists had a cancer detection rate above the expert-recommended threshold. These findings suggest a major need for improved specificity in breast cancer screening. DBT offers the promise of improved specificity by providing image sets across the 3D architecture of the breast, thereby enabling the radiologist to resolve suspicious areas that would appear on 2D projection images due to overlapping tissue. Evidence to date on combined DM/DBT exams supports this premise, with recall rates reduced by approximately 20%.^{6,7} Other aspects of DM/DBT performance remain unclear, including the effects on positive predictive value for biopsy recommendation, the rate of benign biopsies, the rate of interval cancers, and the distribution of specific malignant and benign diagnoses. Furthermore, the potential impact of replacing “dose” DM views with s2D is poorly understood. We hypothesize that s2D/DBT screening performance will differ from combined DM/DBT because the reconstructed 2D image appears different than the conventional 2D DM image. Areas of subtle architectural distortion reconstructed from the tomosynthesis acquisition or distinct, high optical density ligaments may be enhanced in the s2D reconstruction algorithm and may appear more conspicuous.⁴³ While this may cause subtle malignancies to appear more conspicuous on s2D than on DM resulting in improved detection, other reconstructed normal structures such as dense ligamentous structures may undergo over-enhancement on reconstruction. This over-enhancement may make some normal structures appear more conspicuous and “pop” with focal density mimicking calcification-like densities thereby resulting in false positive recalls. Our preliminary data suggests that s2D/DBT screening may permit further reductions in recall rate while maintaining high cancer detection.⁹ Notably, there was a maintenance of invasive cancer detection with a reduction in recall due specifically to asymmetries, suggesting that radiologists are better able to resolve these findings when using s2D/DBT compared to DM/DBT. However, there was also a reduction in calcific lesions and the detection of DCIS on s2D/DBT suggesting that radiologists may not be able to detect or correctly characterize some calcific lesions.⁴³ Anecdotal evidence from radiologists suggests an increased tolerance for calcification-like densities due to the over-enhancement effect, thereby potentially leading to a decline in both benign and malignant lesions that present primarily via calcifications (e.g., atypical ductal hyperplasia, DCIS).

B. INNOVATION

S2D/DBT screening is a new technique for breast cancer screening that offers the potential benefits of DBT imaging without the extra radiation associated with combined DM/DBT exams. We propose the first multi-center study to evaluate this new approach as implemented in clinical practice.

B.1. We will examine multiple aspects of s2D/DBT screening performance that impact patient care.

In addition to the core screening performance metrics assessed in Aim 1, we will investigate additional aspects of performance that affect patient management in Aims 2 and 3. An increasing concern regarding breast cancer screening is overdiagnosis. Concerns regarding breast cancer overdiagnosis are focused primarily on ductal carcinoma in situ (DCIS; stage 0 breast cancer). While DCIS overdiagnosis cannot be directly quantified, various sources of evidence suggest that as few as half of detected DCIS cases would actually progress to invasive breast cancer if left untreated.^{17,44-47} Thus, the relative detection of invasive breast cancer vs. DCIS on screening may have a substantial impact on the benefits and harms of screening.

Detection of benign disease is also often perceived as a limitation or harm of screening, as the goal of screening is to advance the detection of cancer. Nevertheless, certain benign lesions are considered high-risk and frequently surgically excised. Optimal management of benign lesions remains an area of uncertainty.⁴⁸ We seek to understand the extent to which s2D/DBT changes the rate of benign lesion detection and the distribution of specific low-risk and high-risk benign lesions. Overall, these results will provide evidence as to the extent to which s2D/DBT influences the balance of aggressive vs. indolent screen-detected lesions.

Finally, mammographic breast density has emerged as an important factor in breast cancer screening. Breast density refers to the relative amounts of radiodense stromal and epithelial tissue vs. the radiolucent fat tissue appearing on a mammogram.⁴⁹ Radiologists qualitatively rate breast density during mammogram interpretation using four categories defined by the American College of Radiology’s Breast Imaging Reporting and Data System (BI-RADS): “almost entirely fat”, “scattered fibroglandular densities”, “heterogeneously dense”, and “extremely dense”⁵⁰. Breast density has been recognized as a major risk factor for breast cancer^{49,51} and also impairs mammography performance.⁵²⁻⁵⁵ To ensure that women with dense breasts are aware of the limitations

of mammography and their increased breast cancer risk, over 50% of US states have enacted laws mandating the disclosure of breast density directly to women.^{11,56} In many states these notifications are required to include language advising women with dense breasts to discuss supplemental screening tests with her providers.^{57,58} The Food and Drug Administration is also considering an amendment to its regulations issued under the Mammography Quality Standards Act that would require density reporting to patients.⁵⁹ These legislative and regulatory initiatives have generated controversy because of the large number of women affected and the lack of evidence or consensus in the medical community regarding appropriate supplemental screening strategies for women with dense breasts. Breast density notification laws generally define “dense breasts” as those rated on mammograms as “heterogeneously dense” or “extremely dense”. Approximately 40% of US women aged 40-74 have dense breasts based on this definition.⁶⁰ The 2016 evidence review for the USPSTF concluded that there is insufficient evidence to support supplemental screening strategies (including DBT) at this time for women with dense breasts.⁶¹ Density assessment is a subjective measure and we have recently shown high variability in the use of density categories across radiologists in clinical practice.¹² The replacement of conventional 2D DM images with s2D images is likely to alter radiologists’ perception of breast density. Our preliminary data suggests that s2D/DBT screening exams are less likely than DM/DBT exams to be assessed as high density.⁹ Our study proposes to investigate this potential shift in density assessment in a multi-center study with a large number of radiologists.

B.2. Our study provides a unique opportunity to efficiently assess s2D/DBT screening performance.

The University of Vermont and University of Pennsylvania PROSPR Research Centers were established in 2011 within the newly formed PROSPR consortium, at a time in which DM/DBT just began to disseminate into clinical practice. The participating radiology facilities within our PROSPR centers included very early adopters of DM/DBT screening, and many of these facilities are now early adopters of s2D/DBT screening. Our research centers established prospective breast cancer screening data collection using common data elements to permit valid pooling of data across sites. As a result, we are uniquely positioned to assess the performance of s2D/DBT screening. The high quality breast imaging data elements are collected directly from radiology facilities, including indication for exam (to accurately identify screening exams), mammographic views, mammographic breast density, and patient characteristics. We also link to regional and statewide cancer registries and pathology databases to ensure high capture of false negative exams as well as an ability to look at both malignant and benign diagnoses. Our multi-center approach increases the sample size and improves generalizability due to our diverse patient populations; the University of Vermont PROSPR Research Center includes a predominantly white rural population, whereas the University of Pennsylvania medical center covers an urban population that is nearly 50% African-American.

C. APPROACH

We propose a multi-center study of s2D/DBT screening using data from the University of Vermont and University of Pennsylvania PROSPR Research Centers. We will include data from radiology facilities that have transitioned from DM/DBT to s2D/DBT screening. Screening outcomes will be assessed via linkage to cancer registries and pathology databases. Statistical analyses will be performed to assess s2D/DBT screening performance, with comparisons to DM/DBT screening performance prior to the adoption of s2D/DBT.

C.1. Study Population

C.1.a. The University of Vermont PROSPR Research Center. The University of Vermont PROSPR Research Center collects integrated patient, radiology, pathology, and outcomes data for all women undergoing breast imaging in the state of Vermont (approximately 80,000 women per year).^{62,63} All women undergoing breast imaging exams in one of Vermont’s 15 breast imaging facilities complete a health questionnaire. Passive consent to use their data for research is collected on this form via an opt-out checkbox. Approximately 95% of subjects participate (do not opt out). Combined DM/DBT screening in Vermont began at a handful of facilities in 2012, and 11 of the 15 facilities are now using DBT as part of routine screening. s2D/DBT screening with no concurrent DM began at one facility in 2014 and is now used uniformly in DBT screening (i.e., routine use for all patients receiving DBT exams) at 6 facilities, including three facilities within the University of Vermont medical center and three community practice facilities located at small hospitals. These six facilities screen a combined annual volume of approximately 38,000 women, with mammograms interpreted by 16 radiologists.

C.1.b. The University of Pennsylvania PROSPR Research Center. The University of Pennsylvania PROSPR Research Center collects integrated breast cancer screening data for women undergoing screening in Penn Medicine’s integrated health network (approximately 74,000) per year. A waiver of informed consent was secured to allow inclusion of data from all women undergoing breast cancer screening at the six breast imaging radiology facilities in the Penn Medicine network. Combined DM/DBT screening in the Penn Medicine

network began in 2011, with all six facilities now using DBT as part of routine screening. Routine s2D/DBT screening with no concurrent DM was adopted uniformly at the University of Pennsylvania medical center in 2015, with an annual volume of approximately 11,000 women and mammograms interpreted by 6 radiologists.

C.1.c. Eligibility criteria and study population characteristics. Eligibility will be restricted to women undergoing breast cancer screening mammography at radiology facilities within our PROSPR Research Centers that performed DM/DBT screening for at least 1 year and subsequently adopted s2D/DBT screening on a uniform basis. The study period will include screening exams performed between 2011 and 2017. Women with a prior personal history of breast cancer will be excluded since mammography protocols for breast imaging of survivors frequently involves additional views to assess the site of the prior breast cancer location.

Characteristics of women screened at the included facilities for this study are provided in **Table 1**.

C.2. Available data

The University of Vermont and University of Pennsylvania PROSPR Research Centers obtain data from participating radiology facilities, pathology facilities, and statewide cancer registries. The strength of this integrated data at both centers has been demonstrated in numerous publications. The University of Vermont PROSPR Research Center has particular strengths in evaluation of breast cancer risk,^{60,64-68} utilization of screening,^{63,69-72} and screening performance.^{42,73-77} The University of Pennsylvania has particular strengths in evaluating screening performance^{6,7,27,55,74,78-80} and mammographic and MRI-based image biomarkers including breast density.⁸¹⁻⁹² Our centers have recently collaborated on a number of analyses addressing breast cancer etiology,⁹³ implementation of breast cancer screening recommendations in primary care,⁹⁴⁻⁹⁶ breast density assessment,¹² development of novel conceptual models for the evaluation of cancer screening,^{97,98} and evaluation of DM/DBT screening.⁶

C.2.a. Patient data. Data on patient characteristics are obtained from radiology information systems and patient electronic medical records. Women undergoing breast imaging at participating facilities complete a questionnaire that includes clinically relevant breast cancer risk factors (family history of breast cancer, reproductive and menstrual history, etc.), patient demographics (age, race, education, etc.), and medical history (prior cancer diagnoses, date of last mammogram, breast symptoms, etc.).

C.2.b. Radiology exam data. Radiology data elements recorded prospectively by the interpreting radiologist include indication for exam (e.g., screening), breast density, assessments, and recommendations. These data are transmitted electronically from the facilities' radiology information systems.

C.2.c. Benign and malignant diagnoses. The University of Vermont PROSPR Research Center receives pathology data on all benign and malignant breast specimens evaluated at pathology facilities in Vermont on a standardized abstraction form accompanied by a copy of the facility's pathology report. The University of Pennsylvania PROSPR Research Center accesses pathology data electronically within Penn Medicine electronic health records. Pathology data available for the study include specimen type and diagnosis, including behavior, histology, grade, size, estrogen and progesterone hormone receptor status, HER2 status, and OncotypeDX recurrence scores. Consolidated data on breast cancer diagnoses among Vermont and Pennsylvania residents are obtained from the Vermont and Pennsylvania Cancer Registries. This provides additional validated data for cancer cases, including stage at diagnosis, and ensures the capture of cancer diagnosis data for cases screened at participating radiology facilities but diagnosed with cancer elsewhere.

C.2.d. Pooling data from the two centers. Through the PROSPR consortium, the University of Vermont and University of Pennsylvania PROSPR Research Centers developed common data elements for submitting breast cancer screening and outcomes to the PROSPR central data repository. The data collection instruments at both PROSPR centers align closely with national standards, including the BI-RADS atlas,⁹⁹ the College of American Pathologists breast specimen examination protocols,^{100,101} and the North American Association of Central Cancer Registries data standards.¹⁰² For this study the University of Pennsylvania PROSPR Research Center will prepare an exam-level analytic dataset containing the required common data elements to address

Table 1. Characteristics of the screening population at participating facilities.

| | Vermont % | University of Pennsylvania % |
|-----------------------|--------------|------------------------------------|
| Age | | |
| <40 | 2.3 | 2.4 |
| 40-49 | 21.0 | 26.7 |
| 50-59 | 31.5 | 32.0 |
| 60-69 | 28.3 | 26.4 |
| 70+ | 16.9 | 12.5 |
| Race | | |
| White | 96.2 | 42.2 |
| Black | 0.5 | 48.1 |
| Asian | 1.2 | 3.4 |
| Other | 2.1 | 4.9 |
| Breast density | | |
| Almost entirely fatty | 16.2 | 14.7 |
| Scattered densities | 51.2 | 56.7 |
| Heterogeneously dense | 27.8 | 26.1 |
| Extremely dense | 4.7 | 2.5 |

the Specific Aims. The dataset will be securely transferred to the University of Vermont for pooling with corresponding exam-level data from the Vermont PROSPR Research Center.

C.2.e. Preliminary data. Preliminary data from the University of Pennsylvania Research Center based on 5,366 s2D/DBT screening exams and 15,571 DM/DBT screening exams suggests that s2D/DBT screening results in lower recall rates (7.1% vs. 8.8%, $P < 0.001$) and comparable overall cancer detection (5.0 vs. 5.5 per 1000, $P = 0.72$).⁹ Recall for calcifications and asymmetries were reduced for s2D/DBT compared to DM/DBT ($p < 0.05$). The DCIS detection rate was lower with s2D/DBT (0.9 vs. 1.5 per 1000) though this finding did not reach statistical significance ($P = 0.30$). We also observed that fewer s2D/DBT exams were categorized as “dense” (28.6% vs. 32.5%).

We have previously published screening performance estimates for DM/DBT using data from the Vermont and Pennsylvania PROSPR Research Centers, combined with data from the Dartmouth PROSPR Research Center (**Table 2**).⁶ This demonstrates our ability to pool data across our centers and also provides estimates of DM/DBT performance metrics to inform our power calculations.

C.3. Primary exposures of interest: DM/DBT and s2D/DBT

All facilities from both the University of Vermont and University of Pennsylvania PROSPR Research Centers that use DBT in screening use the Hologic® Selenia® Dimensions® 2D/3D Mammography™ system (Hologic; Bedford, MA), which was the first tomosynthesis unit approved by the FDA and remains the most common tomosynthesis-enabled mammography machine in the country. All facilities performing s2D/DBT screening use Hologic’s C-View™ software to reconstruct synthesized 2D images from the DBT image set. All screening exams routinely include bilateral craniocaudal and mediolateral oblique images. All analyses will be restricted to screening exams performed at radiology facilities (N=7) that performed DM/DBT screening for at least 1 year and then switched to s2D/DBT. Two cohorts will be constructed based on the type of mammography screening performed: 1) DM/DBT screening performed prior to the adoption of s2D/DBT; and 2) s2D/DBT with no DM. Notably, some facilities used DM/DBT combined with s2D for a small period of time during the transition to s2D/DBT; these exams will be excluded from the study.

C.4. Primary outcomes of interest

C.4.a. Aim 1: Clinical breast cancer screening performance metrics. We will follow the ACR BI-RADS 5th edition manual for defining all performance metrics, including the cancer detection rate, abnormal interpretation (recall) rate, positive predictive value-1 (PPV1), false negative rate, sensitivity, and specificity.⁹⁹ Radiologists at each participating facility prospectively record a BI-RADS assessment for each screening exam: negative (1), benign (2), probably benign (3), needs additional evaluation (0), suspicious (4), highly suggestive of malignancy (5). A positive screening exam is defined as assessment categories 0, 3, 4, or 5; recall rate is positive exams divided by total number of exams. Cancer outcomes are defined based on diagnoses within 12 months following the screening exam. Cancer detection rate is defined as true positive exams divided by the total number of exams. False negative rate is defined as false negative exams divided by the total number of exams. Sensitivity is calculated by dividing the number of true positive exams by the total number of exams associated with cancer (true positive plus false negative), and specificity is calculated by dividing the number of true negative exams by the total number of exams without cancer (true negative plus false positive). Positive predictive value is the number of true positive exams divided by the total number of positive exams.

C.4.b. Aim 2: Rates of specific screen-detected malignant and benign lesions. Linkage of screening exam data with pathology databases and cancer registries permits the determination of cancer detection rates for specific malignant and benign lesions. We will determine the rates of invasive breast cancers, DCIS, and benign lesion detection via positive screening exams. Benign lesions will be further subdivided as non-proliferative changes, fibroepithelial proliferations, proliferative lesions without atypia, and atypical lesions. Exploratory analyses will examine invasive breast cancers according to ER/PR/HER2 subtypes.

C.4.c. Aim 3: Mammographic breast density. Radiologists at participating facilities prospectively record breast density assessments in one of four categories according to BI-RADS lexicon.⁹⁹ We will conduct analyses using these four categories, as well as a dichotomized variable representing dense (heterogeneously or extremely dense) vs. non-dense (almost entirely fatty or scattered fibroglandular density), which corresponds to the definition of dense breasts in density notification laws.

C.5. Statistical analyses and power

Table 2. DM/DBT screening performance based on data from Vermont, Pennsylvania, and Dartmouth PROSPR Research Centers⁶

| | |
|-----------------------|---------------|
| Recall rate | 8.7% |
| Benign biopsy rate | 14.1 per 1000 |
| Cancer detection rate | 5.9 per 1000 |
| Invasive cancer | 4.2 per 1000 |
| DCIS | 1.7 per 1000 |
| PPV1 | 6.4% |
| False negative rate | 0.6 per 1000 |
| Sensitivity | 90.9% |
| Specificity | 91.3% |
| % with dense breasts | 34.0% |

We estimate data will be available on 98,850 DM/DBT and 86,850 s2D/DBT exams performed at 7 facilities and interpreted by 22 radiologists during 2011-2017. Descriptive statistics will first be used to compare patient characteristics between women undergoing DM/DBT vs. s2D/DBT. We anticipate minimal differences given the study conditions; however, patient variables known to be associated with screening performance (e.g., age, screening interval, etc.) will be incorporated as covariates in regression models to control confounding.

C.5.a. Aim 1: Clinical breast cancer screening performance metrics. General estimating equations (GEE) will be used to estimate performance metrics (cancer detection rate, recall rate, biopsy rate, PPV, false negative rate, sensitivity, and specificity) by type of screening exam (DM/DBT or s2D/DBT). To control for correlations between exams within women, radiologists and facilities, we will use the approach described by Miglioretti and Heagerty for modeling marginal (population average) effects from data with multilevel, non-nested clustering.¹⁰³ While confounding by patient characteristics is not anticipated, we will investigate the influence of including covariates in the model to adjust for patient factors known to be associated with screening performance, including age, race/ethnicity, breast density, family history of breast cancer, and screening interval. Interactions between the patient characteristics and type of imaging will be examined to determine if their effects on accuracy are different for DM/DBT and s2D/DBT.

It is difficult to determine a priori how adjustment for correlations within women, radiologists and facilities will affect power because it can either increase or decrease GEE variance estimates.¹⁰³ However, because both DM/DBT and s2D/DM exams will be obtained from all participating radiologists and facilities, as well as from most women, it is unlikely that adjustment will increase the variances for comparisons of the two imaging modalities. With 98,850 DM/DBT exams and 86,850 s2D/DBT exams, we will have 80% power to detect absolute differences of 0.4% in the recall rate and 0.8 per 1000 in the cancer detection rate. We anticipate having a total of 1,114 breast cancers (true positives plus false negatives) and 174,546 non-cancer outcomes (true negatives plus false positives), which will provide 80% power to detect absolute differences of 5.6% in sensitivity and 0.4% in specificity between DM/DBT and s2D/DBT.

C.5.b. Aim 2: Rates of specific screen-detected malignant and benign lesions. Using similar methods to Aim 1, we will use descriptive statistics to describe the detection of invasive breast cancer, DCIS, and benign lesions via s2D/DBT screening and compare to DM/DBT statistics using logistic regression. We will have 80% power to detect absolute decreases of 0.8 per 1000 in invasive breast cancer detection, 0.5 per 1000 in DCIS detection, and 1.5 per 1000 in benign lesion detection.

C.5.c. Aim 3: Mammographic breast density. We will determine the percent of exams rated in each of the four BI-RADS density categories for s2D/DBT and compare to the density distribution for DM/DBT exams using chi-square testing and logistic regression with adjustment for patient factors associated with breast density, including age, race/ethnicity, and body mass index. Analyses will also be conducted based on the dichotomized dense vs. non-dense categories. We will have 80% power to detect an absolute difference as small as 0.6% in the proportion of exams with dense breasts between s2D/DBT and DM/DBT screening.

C.6. Potential problems and alternate strategies

The gold standard for evaluation of new screening strategies is a randomized controlled trial. However, our observational study design makes efficient use of a natural experiment that has limited threats to validity for the following reasons: 1) both types of exams come from the same set of facilities and radiologists; 2) the demographic and risk characteristics of the patient population served at each facility were unlikely to change substantially during the study period; and 3) s2D/DBT adoption occurred uniformly at the facility level (i.e., s2D/DBT was not targeted towards certain patient subgroups). Furthermore, our comprehensive exam-level data on patient risk factors, demographics and screening history permits statistical control of potentially confounding factors. Our study offers strong generalizability through the use of data from routine clinical practice (i.e., outside of a trial environment) and permits the rapid and cost-efficient ascertainment of our study aims by using existing data. Notably, the American College of Radiology Imaging Network (ACRIN) is initiating a trial (TMIST; Tomosynthesis Mammographic Imaging Screening Trial) to compare DBT screening to digital mammography.¹⁰⁴ Current plans for this trial permit DBT sites to perform either combined DM/DBT screening or s2D/DBT screening, with the primary aim of comparing DBT performance to DM-alone screening. Given the current low level of adoption of s2D/DBT it is anticipated that the vast majority of DBT sites will employ DM/DBT, and the trial does not aim to evaluate the performance of s2D/DBT in comparison to DM/DBT. Current sample size plans include approximately 44,000 women in each arm (DBT vs. DM) so power would be very limited to evaluate s2D/DBT even if there is moderate uptake by enrolling centers. Thus, our study will contribute important and timely information about how the use of s2D impacts DBT screening performance.

C.7. Summary

Screening with s2D/DBT offers the potential to maintain the benefits of DBT technology without exposing the

patient to the extra radiation of “dose” 2D DM imaging. Evidence regarding the screening performance of s2D/DBT is urgently needed to inform health care providers who are considering this new technology but are anxious about discontinuing 2D DM imaging, which has been the standard of care in breast cancer screening for more than a decade. The transition to s2D/DBT screening could have multiple impacts on performance metrics and patient care, including changes in recall rate, detection rates for malignant and benign lesions detected, and the assessment of breast density. Our multi-center study leverages high quality clinical data available from academic and community practice facilities that have been early adopters of s2D/DBT. Our multi-disciplinary team has the required expertise in radiology, pathology, and epidemiology and a track record of successful collaboration. Our findings will be informative for women considering breast cancer screening options, providers assessing whether to adopt s2D/DBT, health care payers considering reimbursement rates, and policy makers debating breast cancer screening recommendations in the context of this new technology.

References

1. Houssami N, Miglioretti DL. Digital Breast Tomosynthesis: A Brave New World of Mammography Screening. *JAMA oncology* 2016.
2. Houssami N, Skaane P. Overview of the evidence on digital breast tomosynthesis in breast cancer detection. *Breast* 2013; **22**(2): 101-8.
3. Siu AL. Screening for Breast Cancer: U.S. Preventive Services Task Force Recommendation Statement. *Ann Intern Med* 2016; **164**(4): 279-96.
4. Sechopoulos I. A review of breast tomosynthesis. Part I. The image acquisition process. *Medical physics* 2013; **40**(1): 014301.
5. Sechopoulos I. A review of breast tomosynthesis. Part II. Image reconstruction, processing and analysis, and advanced applications. *Medical physics* 2013; **40**(1): 014302.
6. Conant EF, Beaber EF, Sprague BL, et al. Breast cancer screening using tomosynthesis in combination with digital mammography compared to digital mammography alone: a cohort study within the PROSPR consortium. *Breast Cancer Res Treat* 2016; **156**(1): 109-16.
7. Friedewald SM, Rafferty EA, Rose SL, et al. Breast cancer screening using tomosynthesis in combination with digital mammography. *JAMA* 2014; **311**(24): 2499-507.
8. Ciatto S, Houssami N, Bernardi D, et al. Integration of 3D digital mammography with tomosynthesis for population breast-cancer screening (STORM): a prospective comparison study. *Lancet Oncol* 2013; **14**(7): 583-9.
9. Zuckerman SP, Conant EF, Keller BM, et al. Implementation of Synthesized Two-dimensional Mammography in a Population-based Digital Breast Tomosynthesis Screening Program. *Radiology* 2016; **281**(3): 730-6.
10. Gilbert FJ, Tucker L, Gillan MG, et al. Accuracy of Digital Breast Tomosynthesis for Depicting Breast Cancer Subgroups in a UK Retrospective Reading Study (TOMMY Trial). *Radiology* 2015; **277**(3): 697-706.
11. Haas JS, Kaplan CP. The Divide Between Breast Density Notification Laws and Evidence-Based Guidelines for Breast Cancer Screening: Legislating Practice. *JAMA Intern Med* 2015; **175**(9): 1439-40.
12. Sprague BL, Conant EF, Onega T, et al. Variation in Mammographic Breast Density Assessments Among Radiologists in Clinical Practice: A Multicenter Observational Study. *Ann Intern Med* 2016; **165**(7):457-64.
13. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2015. *CA Cancer J Clin* 2015; **65**(1): 5-29.
14. Edwards BK, Noone AM, Mariotto AB, et al. Annual Report to the Nation on the status of cancer, 1975-2010, featuring prevalence of comorbidity and impact on survival among persons with lung, colorectal, breast, or prostate cancer. *Cancer* 2014; **120**(9): 1290-314.
15. Munoz D, Near AM, van Ravesteyn NT, et al. Effects of screening and systemic adjuvant therapy on ER-specific US breast cancer mortality. *J Natl Cancer Inst* 2014; **106**(11): dju289.
16. Marmot MG. Sorting Through the Arguments on Breast Screening. *JAMA* 2013; **309**(24): 2553-4.
17. Esserman L, Shieh Y, Thompson I. Rethinking screening for breast cancer and prostate cancer. *JAMA* 2009; **302**(15): 1685-92.
18. Bleyer A, Welch HG. Effect of three decades of screening mammography on breast-cancer incidence. *N Engl J Med* 2012; **367**(21): 1998-2005.
19. Myers ER, Moorman P, Gierisch JM, et al. Benefits and Harms of Breast Cancer Screening: A Systematic Review. *JAMA* 2015; **314**(15): 1615-34.
20. 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**(4): 256-67.
21. U.S. Preventive Services Task Force. Screening for breast cancer: U.S. Preventive Services Task Force recommendation statement. *Ann Intern Med* 2009; **151**(10): 716-26, W-236.
22. Stout NK, Lee SJ, Schechter CB, et al. Benefits, Harms, and Costs for Breast Cancer Screening After US Implementation of Digital Mammography. *J Natl Cancer Inst* 2014; **106**(6): dju092.
23. Kopans DB. Digital breast tomosynthesis from concept to clinical care. *AJR Am J Roentgenol* 2014; **202**(2): 299-308.
24. Melnikow J, Fenton JJ, Miglioretti D, Whitlock EP, Weyrich MS. Screening for Breast Cancer with Digital Breast Tomosynthesis, AHRQ Publication No. 14-05201-EF-2: Center for Healthcare Policy and Research, University of California-Davis and Kaiser Permanente Research Affiliates Evidence-based

- Practice Center, Kaiser Permanente Center for Health Research, 2015.
25. Skaane P, Bandos AI, Gullien R, et al. Comparison of digital mammography alone and digital mammography plus tomosynthesis in a population-based screening program. *Radiology* 2013; **267**(1): 47-56.
 26. Haas BM, Kalra V, Geisel J, Raghu M, Durand M, Philpotts LE. Comparison of tomosynthesis plus digital mammography and digital mammography alone for breast cancer screening. *Radiology* 2013; **269**(3):694-700.
 27. McCarthy AM, Kontos D, Synnestvedt M, et al. Screening outcomes following implementation of digital breast tomosynthesis in a general-population screening program. *J Natl Cancer Inst* 2014; **106**(11).
 28. Rose SL, Tidwell AL, Bujnoch LJ, Kushwaha AC, Nordmann AS, Sexton R, Jr. Implementation of breast tomosynthesis in a routine screening practice: an observational study. *AJR Am J Roentgenol* 2013; **200**(6): 1401-8.
 29. Greenberg JS, Javitt MC, Katzen J, Michael S, Holland AE. Clinical performance metrics of 3D digital breast tomosynthesis compared with 2D digital mammography for breast cancer screening in community practice. *AJR Am J Roentgenol* 2014; **203**(3): 687-93.
 30. Lourenco AP, Barry-Brooks M, Baird GL, Tuttle A, Mainiero MB. Changes in recall type and patient treatment following implementation of screening digital breast tomosynthesis. *Radiology* 2015; **274**(2): 337- 42.
 31. Sharpe RE, Jr., Venkataraman S, Phillips J, et al. Increased Cancer Detection Rate and Variations in the Recall Rate Resulting from Implementation of 3D Digital Breast Tomosynthesis into a Population-based Screening Program. *Radiology* 2016; **278**(3): 698-706.
 32. Centers for Medicare & Medicaid Services. Policy and payment changes to the Medicare Physician Fee Schedule for 2015. 2014. <http://www.cms.gov/newsroom/mediareleasedatabase/fact-sheets/2014-Fact-sheets-items/2014-10-31-7.html> (accessed April 24, 2015).
 33. Skaane P, Bandos AI, Gullien R, et al. Prospective trial comparing full-field digital mammography (FFDM) versus combined FFDM and tomosynthesis in a population-based screening programme using independent double reading with arbitration. *European radiology* 2013; **23**(8): 2061-71.
 34. Svahn TM, Houssami N, Sechopoulos I, Mattsson S. Review of radiation dose estimates in digital breast tomosynthesis relative to those in two-view full-field digital mammography. *Breast* 2015; **24**(2): 93-9.
 35. U.S. Food and Drug Administration. Mammography Quality Standards Act and Program Document Archives. April 23, 2013 2013. <http://www.fda.gov/Radiation-EmittingProducts/MammographyQualityStandardsActandProgram/DocumentArchives/ucm123313.htm> (accessed September 3, 2013).
 36. Miglioretti DL, Lange J, van den Broek JJ, et al. Radiation-Induced Breast Cancer Incidence and Mortality From Digital Mammography Screening: A Modeling Study. *Ann Intern Med* 2016; **164**(4): 205-14.
 37. Skaane P, Bandos AI, Eben EB, et al. Two-view digital breast tomosynthesis screening with synthetically reconstructed projection images: comparison with digital breast tomosynthesis with full-field digital mammographic images. *Radiology* 2014; **271**(3): 655-63.
 38. Gur D, Zuley ML, Anello MI, et al. Dose reduction in digital breast tomosynthesis (DBT) screening using synthetically reconstructed projection images: an observer performance study. *Acad Radiol* 2012; **19**(2): 166-71.
 39. Zuley ML, Guo B, Catullo VJ, et al. Comparison of two-dimensional synthesized mammograms versus original digital mammograms alone and in combination with tomosynthesis images. *Radiology* 2014; **271**(3): 664-71.
 40. Gilbert FJ, Tucker L, Gillan MG, et al. The TOMMY trial: a comparison of TOMosynthesis with digital MammographY in the UK NHS Breast Screening Programme--a multicentre retrospective reading study comparing the diagnostic performance of digital breast tomosynthesis and digital mammography with digital mammography alone. *Health technology assessment* 2015; **19**(4): i-xxv, 1-136.
 41. Aujero MP, Gavenonis SC, Benjamin R, Zhang Z, Holt JS. Clinical Performance of Synthesized Two-dimensional Mammography Combined with Tomosynthesis in a Large Screening Population. *Radiology* 2017: 162674.
 42. Lehman CD, Arao RF, Sprague BL, et al. National Performance Benchmarks for Modern Screening Digital Mammography: Update from the Breast Cancer Surveillance Consortium. *Radiology* 2016: 161174.
 43. Zuckerman S, Maidment A, Weinstein S, McDonald E, Conant EF. Imaging with synthesized two-dimensional mammography: differences, advantages, and pitfalls compared to digital mammography.

American Journal of Roentgenology (in press).

44. Ernster VL, Barclay J. Increases in ductal carcinoma in situ (DCIS) of the breast in relation to mammography: a dilemma. *J Natl Cancer Inst Monogr* 1997; (22): 151-6.
45. Duffy SW, Agbaje O, Tabar L, et al. Overdiagnosis and overtreatment of breast cancer: estimates of overdiagnosis from two trials of mammographic screening for breast cancer. *Breast Cancer Res* 2005; **7**(6): 258-65.
46. Welch HG, Woloshin S, Schwartz LM. The sea of uncertainty surrounding ductal carcinoma in situ--the price of screening mammography. *J Natl Cancer Inst* 2008; **100**(4): 228-9.
47. Yen MF, Tabar L, Vitak B, Smith RA, Chen HH, Duffy SW. Quantifying the potential problem of overdiagnosis of ductal carcinoma in situ in breast cancer screening. *Eur J Cancer* 2003; **39**(12): 1746-54.
48. Menes TS, Kerlikowske K, Lange J, Jaffer S, Rosenberg R, Miglioretti DL. Subsequent Breast Cancer Risk Following Diagnosis of Atypical Ductal Hyperplasia on Needle Biopsy. *JAMA oncology* 2017; **3**(1): 36-41.
49. Boyd NF, Guo H, Martin LJ, et al. Mammographic density and the risk and detection of breast cancer. *N Engl J Med* 2007; **356**(3): 227-36.
50. American College of Radiology. ACR BI-RADS® - Mammography. 4th Edition. ACR Breast Imaging Reporting and Data System, Breast Imaging Atlas. Reston, VA: American College of Radiology; 2003.
51. Tice JA, Cummings SR, Smith-Bindman R, Ichikawa L, Barlow WE, Kerlikowske K. Using clinical factors and mammographic breast density to estimate breast cancer risk: development and validation of a new predictive model. *Ann Intern Med* 2008; **148**(5): 337-47.
52. Kerlikowske K, Grady D, Barclay J, Sickles EA, Ernster V. Effect of age, breast density, and family history on the sensitivity of first screening mammography. *JAMA* 1996; **276**(1): 33-8.
53. Buist DS, Porter PL, Lehman C, Taplin SH, White E. Factors contributing to mammography failure in women aged 40-49 years. *Journal of the National Cancer Institute* 2004; **96**(19): 1432-40.
54. Carney PA, Miglioretti DL, Yankaskas BC, et al. Individual and combined effects of age, breast density, and hormone replacement therapy use on the accuracy of screening mammography. *Ann Intern Med* 2003; **138**(3): 168-75.
55. Pisano ED, Gatsonis C, Hendrick E, et al. Diagnostic performance of digital versus film mammography for breast-cancer screening. *N Engl J Med* 2005; **353**(17): 1773-83.
56. Are You Dense Inc. 24 Density Reporting Bills and A Growing Number of Introduced Bills in 2016. 2016. http://areyoudenseadvocacy.org/worxcms_published/news_page200.shtml (accessed March 16, 2016).
57. Dehkordy SF, Carlos RC. Dense breast legislation in the United States: state of the states. *J Am Coll Radiol* 2013; **10**(12): 899-902.
58. Ray KM, Price ER, Joe BN. Breast density legislation: mandatory disclosure to patients, alternative screening, billing, reimbursement. *AJR Am J Roentgenol* 2015; **204**(2): 257-60.
59. U.S. Food and Drug Administration. Mammography Quality Standards Act; Regulatory Amendments. Accessed at <http://www.gpo.gov/fdsys/pkg/FR-2015-06-18/pdf/2015-14352.pdf> on July 31, 2015. *Federal Register* 2015; **80**(117): 35020.
60. Sprague BL, Gangnon RE, Burt V, et al. Prevalence of mammographically dense breasts in the United States. *J Natl Cancer Inst* 2014; **106**(10).
61. Melnikow J, Fenton JJ, Whitlock EP, et al. Supplemental Screening for Breast Cancer in Women With Dense Breasts: A Systematic Review for the U.S. Preventive Services Task Force. *Ann Intern Med* 2016; **164**(4): 268-78.
62. Onega T, Beaber EF, Sprague BL, et al. Breast cancer screening in an era of personalized regimens: A conceptual model and National Cancer Institute initiative for risk-based and preference-based approaches at a population level. *Cancer* 2014.
63. Sprague BL, Bolton KC, Mace JL, et al. Registry-based study of trends in breast cancer screening mammography before and after the 2009 U.S. Preventive Services Task Force recommendations. *Radiology* 2014; **270**(2): 354-61.
64. Engmann NJ, Golmakani MK, Miglioretti DL, Sprague BL, Kerlikowske K, Breast Cancer Surveillance C. Population-Attributable Risk Proportion of Clinical Risk Factors for Breast Cancer. *JAMA oncology* 2017.
65. Ahern TP, Sprague BL, Bissell MC, et al. Family history of breast cancer, breast density, and breast cancer risk in a U.S. breast cancer screening population. *Cancer Epidemiol Biomarkers Prev* 2017.
66. Henderson LM, Hubbard RA, Sprague BL, Zhu W, Kerlikowske K. Increased Risk of Developing Breast Cancer after a False-Positive Screening Mammogram. *Cancer Epidemiol Biomarkers Prev* 2015; **24**(12):

67. Kerlikowske K, Gard CC, Sprague BL, Tice JA, Miglioretti DL, Breast Cancer Surveillance C. One versus Two Breast Density Measures to Predict 5- and 10-Year Breast Cancer Risk. *Cancer Epidemiol Biomarkers Prev* 2015; **24**(6): 889-97.
68. Kerlikowske K, Zhu W, Tosteson AN, et al. Identifying women with dense breasts at high risk for interval cancer: a cohort study. *Ann Intern Med* 2015; **162**(10): 673-81.
69. Bolton KC, Mace JL, Vacek PM, et al. Changes in breast cancer risk distribution among Vermontwomen using screening mammography. *J Natl Cancer Inst* 2014; **106**(8): dju157.
70. Wernli KJ, Arao RF, Hubbard RA, et al. Change in Breast Cancer Screening Intervals Since the 2009 USPSTF Guideline. *J Womens Health (Larchmt)* 2017.
71. Beaver EF, Tosteson AN, Haas JS, et al. Breast cancer screening initiation after turning 40 years of age within the PROSPR consortium. *Breast Cancer Res Treat* 2016; **160**(2): 323-31.
72. Hubbard RA, O'Meara ES, Henderson LM, et al. Multilevel factors associated with long-term adherence to screening mammography in older women in the U.S. *Prev Med* 2016; **89**: 169-77.
73. Miglioretti DL, Zhu W, Kerlikowske K, et al. Breast Tumor Prognostic Characteristics and Biennial vs Annual Mammography, Age, and Menopausal Status. *JAMA oncology* 2015; **1**(8): 1069-77.
74. Sprague BL, Stout NK, Schechter C, et al. Benefits, harms, and cost-effectiveness of supplemental ultrasonography screening for women with dense breasts. *Ann Intern Med* 2015; **162**(3): 157-66.
75. Kemp Jacobsen K, O'Meara ES, Key D, et al. Comparing sensitivity and specificity of screening mammography in the United States and Denmark. *Int J Cancer* 2015; **137**(9): 2198-207.
76. Henderson LM, Miglioretti DL, Kerlikowske K, Wernli KJ, Sprague BL, Lehman CD. Breast Cancer Characteristics Associated With Digital Versus Film-Screen Mammography for Screen-Detected and Interval Cancers. *AJR Am J Roentgenol* 2015; **205**(3): 676-84.
77. Houssami N, Abraham LA, Onega T, et al. Accuracy of screening mammography in women with a history of lobular carcinoma in situ or atypical hyperplasia of the breast. *Breast Cancer Res Treat* 2014; **145**(3): 765-73.
78. McCarthy AM, Yamartino P, Yang J, Bristol M, Conant EF, Armstrong K. Racial differences in false-positive mammogram rates: results from the ACRIN Digital Mammographic Imaging Screening Trial (DMIST). *Med Care* 2015; **53**(8): 673-8.
79. McDonald ES, Oustimov A, Weinstein SP, Synnestvedt MB, Schnall M, Conant EF. Effectiveness of Digital Breast Tomosynthesis Compared With Digital Mammography: Outcomes Analysis From 3 Years of Breast Cancer Screening. *JAMA oncology* 2016.
80. Pisano ED, Hendrick RE, Yaffe MJ, et al. Diagnostic accuracy of digital versus film mammography: exploratory analysis of selected population subgroups in DMIST. *Radiology* 2008; **246**(2): 376-83.
81. Ashraf A, Gaonkar B, Mies C, et al. Breast DCE-MRI Kinetic Heterogeneity Tumor Markers: Preliminary Associations With Neoadjuvant Chemotherapy Response. *Translational oncology* 2015; **8**(3): 154-62.
82. Ashraf AB, Daye D, Gavenonis S, et al. Identification of intrinsic imaging phenotypes for breast cancer tumors: preliminary associations with gene expression profiles. *Radiology* 2014; **272**(2): 374-84.
83. Ashraf AB, Gavenonis SC, Daye D, Mies C, Rosen MA, Kontos D. A multichannel Markov random field framework for tumor segmentation with an application to classification of gene expression-based breast cancer recurrence risk. *IEEE transactions on medical imaging* 2013; **32**(4): 637-48.
84. Keller BM, Chen J, Daye D, Conant EF, Kontos D. Preliminary evaluation of the publicly available Laboratory for Breast Radiodensity Assessment (LIBRA) software tool: comparison of fully automated area and volumetric density measures in a case-control study with digital mammography. *Breast Cancer Res* 2015; **17**(1): 117.
85. Keller BM, Nathan DL, Wang Y, et al. Estimation of breast percent density in raw and processed full field digital mammography images via adaptive fuzzy c-means clustering and support vector machine segmentation. *Medical physics* 2012; **39**(8): 4903-17.
86. Keller BM, Oustimov A, Wang Y, et al. Parenchymal texture analysis in digital mammography: robust texture feature identification and equivalence across devices. *Journal of medical imaging* 2015; **2**(2): 024501.
87. Kontos D, Ikejimba LC, Bakic PR, Troxel AB, Conant EF, Maidment AD. Analysis of parenchymal texture with digital breast tomosynthesis: comparison with digital mammography and implications for cancer risk assessment. *Radiology* 2011; **261**(1): 80-91.
88. Mahrooghy M, Ashraf AB, Daye D, et al. Pharmacokinetic Tumor Heterogeneity as a Prognostic Biomarker for Classifying Breast Cancer Recurrence Risk. *IEEE transactions on bio-medical engineering* 2015; **62**(6):

1585-94.

89. Pertuz S, McDonald ES, Weinstein SP, Conant EF, Kontos D. Fully Automated Quantitative Estimation of Volumetric Breast Density from Digital Breast Tomosynthesis Images: Preliminary Results and Comparison with Digital Mammography and MR Imaging. *Radiology* 2015; 150277.
90. Wu S, Weinstein SP, Conant EF, Kontos D. Automated fibroglandular tissue segmentation and volumetric density estimation in breast MRI using an atlas-aided fuzzy C-means method. *Medical physics* 2013; **40**(12): 122302.
91. Wu S, Weinstein SP, DeLeo MJ, 3rd, et al. Quantitative assessment of background parenchymal enhancement in breast MRI predicts response to risk-reducing salpingo-oophorectomy: preliminary evaluation in a cohort of BRCA1/2 mutation carriers. *Breast Cancer Res* 2015; **17**: 67.
92. Zheng Y, Keller BM, Ray S, et al. Parenchymal texture analysis in digital mammography: A fully automated pipeline for breast cancer risk assessment. *Medical physics* 2015; **42**(7): 4149-60.
93. Wood ME, Sprague BL, Oustimov A, et al. Aspirin use is associated with lower mammographic density in a large screening cohort. *Breast Cancer Res Treat* 2017.
94. Haas JS, Barlow WE, Schapira MM, et al. Primary Care Providers' Beliefs and Recommendations and Use of Screening Mammography by their Patients. *J Gen Intern Med* 2017.
95. Schapira MM, Sprague BL, Klabunde CN, et al. Inadequate Systems to Support Breast and Cervical Cancer Screening in Primary Care Practice. *J Gen Intern Med* 2016; **31**(10): 1148-55.
96. Haas JS, Sprague BL, Klabunde CN, et al. Provider Attitudes and Screening Practices Following Changes in Breast and Cervical Cancer Screening Guidelines. *J Gen Intern Med* 2016; **31**(1): 52-9.
97. Onega T, Beaber EF, Sprague BL, et al. Breast cancer screening in an era of personalized regimens: a conceptual model and National Cancer Institute initiative for risk-based and preference-based approaches at a population level. *Cancer* 2014; **120**(19): 2955-64.
98. Beaber EF, Kim JJ, Schapira MM, et al. Unifying Screening Processes Within the PROSPR Consortium: A Conceptual Model for Breast, Cervical, and Colorectal Cancer Screening. *J Natl Cancer Inst* 2015; **107**(6).
99. American College of Radiology. ACR BI-RADS® - Mammography. 5th Edition. ACR BI-RADS Atlas: Breast Imaging Reporting and Data System. Reston, VA: American College of Radiology; 2013.
100. Lester SC, Bose S, Chen YY, et al. Protocol for the examination of specimens from patients with invasive carcinoma of the breast. *Archives of pathology & laboratory medicine* 2009; **133**(10): 1515-38.
101. Lester SC, Bose S, Chen YY, et al. Protocol for the examination of specimens from patients with ductal carcinoma in situ of the breast. *Archives of pathology & laboratory medicine* 2009; **133**(1): 15-25.
102. Thornton ML, editor. Standards for Cancer Registries Volume II: Data Standards and Data Dictionary, Record Layout Version 15, 19th ed. Springfield, IL: North American Association of Central Cancer Registries; 2015.
103. Miglioretti DL, Heagerty PJ. Marginal modeling of nonnested multilevel data using standard software. *Am J Epidemiol* 2007; **165**(4): 453-63.
104. American College of Radiology Imaging Network. PROTOCOL ACRIN PA 4006: Comparison of full- field digital mammography with digital breast tomosynthesis image acquisition in relation to screening call-back rate. 2015. http://www.acrin.org/4006_protocol.aspx (accessed April 24, 2015).