

UCSF

UC San Francisco Previously Published Works

Title

“It’s personalized, but it’s still bucket based”: the promise of personalized medicine vs. the reality of genomic risk stratification in a breast cancer screening trial

Permalink

<https://escholarship.org/uc/item/44q242hw>

Journal

New Genetics and Society, 41(3)

ISSN

1463-6778

Authors

James, Jennifer Elyse

Joseph, Galen

Publication Date

2022-07-03

DOI

10.1080/14636778.2022.2115348

Copyright Information

This work is made available under the terms of a Creative Commons Attribution License, available at <https://creativecommons.org/licenses/by/4.0/>

Peer reviewed



Published in final edited form as:

New Genet Soc. 2022 ; 41(3): 228–253. doi:10.1080/14636778.2022.2115348.

“It’s personalized, but it’s still bucket based”: The promise of personalized medicine vs. the reality of genomic risk stratification in a breast cancer screening trial

Jennifer Elyse James¹, Galen Joseph²

¹Institute for Health and Aging, University of California, San Francisco

²Department of Humanities and Social Sciences, University of California, San Francisco

Abstract

Adaptive pragmatic clinical trials offer an innovative approach that integrates clinical care and research. Yet, blurring the boundaries between research and clinical care raises questions about how clinicians and investigators balance their caregiving and research roles and what types of knowledge and risk assessment are most valued. This paper presents findings from an ethnographic ELSI (Ethical, Legal, Social Implications) study of an innovative clinical trial of risk-based breast cancer screening that utilizes genomics to stratify risk and recommend a breast cancer screening commensurate with the assessed risk. We argue that the trial demonstrates a fundamental tension between the promissory ideals of personalized medicine, and the reality of implementing risk stratified care on a population scale. We examine the development of a Screening Assignment Review Board in response to this tension which allows clinician-investigators to negotiate, but never fully resolve, the inherent contradiction of ‘precision population screening’.

Keywords

personalized medicine; risk-stratification; risk-based screening; breast cancer screening; pragmatic adaptive trial design; genomics; ethnography; embedded ELSI

Introduction

The promissory tropes of precision and personalized medicine are well trodden in the medical literature and popular media, and extensively critiqued in the social science and STS literature (Kerr et al. 2021; Adams, Murphy, and Clarke 2009; Tutton 2012; Erikainen and Chan 2019). Oncology has been at the forefront of advances in precision medicine as well as these dialogues about anticipated futures, with a primary focus on cancer diagnosis and treatment (Kerr et al. 2021; Kerr et al. 2018; Bourret, Keating, and Cambrosio 2011; Keating and Pace 2018; Cambrosio et al. 2018). Only more recently have clinician scientists and their social science observers turned to the possibility of using genomic testing of healthy populations to inform screening and prevention, moving precision medicine from the

realm of patient care (e.g. in oncology) to population level screening and the primary care context (Khoury and Galea 2016; Armstrong 2019; James et al. 2021).

This article examines the complexities and challenges involved in the implementation of a breast cancer risk algorithm in the context of an ongoing national breast cancer screening trial in the United States. The WISDOM trial (Women Informed to Screen Depending on Measures of Risk), funded by the US government's Patient Centered Outcomes Institute (PCORI) and National Cancer Institute (NCI), is an adaptive pragmatic clinical trial using genomics and other measures of risk to test a risk-based approach to breast cancer screening (Clinical Trial: [NCT02620852](#)). With the goal of supplanting the long-standing paradigm of annual mammography, this approach to breast cancer screening integrates genomic risk factors with other biological (e.g., breast density) and demographic (e.g., age, race) risk factors into an algorithm that estimates risk and makes corresponding recommendations for screening schedule and modality for study participants. It does so through an innovative trial design that merges this experimental risk algorithm and clinical care in a single endeavor. Thus, WISDOM is at the forefront of translating genomics to preventive medicine at the population level, one of the key promises of precision medicine (Abrahams 2008). Interestingly, the WISDOM trial has not been framed as precision medicine study; rather, the investigators speak to the advancement of a personalized approach to screening, using genetics and genomics.

WISDOM exemplifies a trend described by Nelson and colleagues (2014) away from trials as "testing machines" (the traditional randomized controlled trial), and toward trials as "clinical experimental systems" which are characterized by "the management of heterogeneity, the flexibility of protocols, the institutions needed to execute the trials, and type of information that can be gleaned from clinical trial participants." They argue that this characterization is "useful for understanding the tensions surrounding the implementation of hybrid trial designs and new research practices that attempt to satisfy both experimental and testing aims, and for understanding the increasingly dense connections between the laboratory and the clinic that are prominent in emerging forms of translational research" (2014, 75). WISDOM's adaptive, pragmatic study design both supports the integration of rapidly emerging genomic data—polygenic risk and risk associated with high and moderate penetrance gene variants—and creates substantial tension between the optimal provision of clinical care based on clinical judgment and the responsibilities of research (Bourret, Keating, and Cambrosio 2011; Montgomery 2017). The tension is fueled by and fuels role ambiguity (Wolf et al. 2018) for many of the study's investigators who are also clinicians, conducting research across the increasingly blurry boundary between research and care in translational genomics (Wolf et al. 2018). As such, WISDOM raises critical questions about the practical implementation of precision medicine at the population scale, and the inherent conflict between the algorithm-based risk assessment being tested and the personalization promised as a core feature of precision medicine (Erikainen and Chan 2019).

In what follows, we describe the WISDOM trial, and then we examine the emergence of a "Screening Assignment Review Board", an "organizational mechanism" and "work routine" the clinician-investigators¹ created to manage the tensions and contradictions inherent in the trial design and implementation (Crabu 2021). This review board is an

organizational strategy that provides a space to negotiate how the trial is implemented, and adaptations to the risk algorithm and risk thresholds, all of which have implications for statistical analyses of trial outcomes, and thus for the knowledge the trial will produce. We argue that WISDOM demonstrates a fundamental tension between the promissory ideals of personalized medicine, and the reality of implementing risk stratified care on a population scale. The Screening Assignment Review Board (SARB), developed in response to this tension and enabled by WISDOM's pragmatic adaptive design, allows clinician-investigators to grapple with this tension and to manage conflicts over individual cases, but never fully resolves the inherent contradiction of population screening. Our ethnography and the case of the WISDOM trial demonstrate the challenges of translating genomic risk prediction to clinical practice.

WISDOM: A Pragmatic, Adaptive Trial of Personalized Breast Cancer Screening

In 2009, the US Preventive Services Task Force introduced changes to its breast cancer screening guidelines, recommending that annual mammography for all women age forty to seventy-five be replaced by biennial screening for women ages fifty to seventy-five, and that screening in the forties should be individualized by taking patient context into account, including the patient's values regarding specific benefits and harms of screening (Keating and Pace 2018). In 2015, the American Cancer Society changed its guidelines to recommend annual mammography starting at age forty-five and biennial screening for women over the age of fifty-five (Oeffinger et al. 2015). But these guidelines continue to spark debate and scientific controversies about the benefits, harms and effectiveness of annual screening. On the one hand, many in the radiology and OB/GYN communities argue that annual mammography starting at forty reduces interval cancers (those that arise between screenings) and should not be changed. On the other hand, some epidemiologists and clinicians maintain that annual mammography results in false positives and unnecessary treatment, and that a more targeted screening approach could reduce false positives and over-diagnosis without increasing interval cancers (Esserman 2017).

The WISDOM trial aims to inform these intense policy debates about mammography use and to create a new risk-based paradigm of breast cancer screening by demonstrating that comprehensive risk-based screening is as safe, less morbid and has greater healthcare value than annual mammography (Shieh et al. 2017; Esserman 2017) (see Figure 1). WISDOM's risk-based screening approach, and the application of precision medicine to the screening context in general, is intended to move away from the "one size fits all" approach by differentiating among individuals at high, average, and low risk of breast cancer and recommending screening schedules and modalities commensurate with risk: those at higher risk are screened more and those at lower risk are screened less. WISDOM is the first effort in the United States to recommend breast cancer screening according to individual genomic risk, as opposed to population characteristics. The European Union and Canada have recently begun similar trials.²

¹We refer throughout to trial personnel who are both active clinical providers (primary care providers, breast surgeons, oncologists, genetic counselors etc.) as "clinician-investigators." This is a term some use to describe themselves and their dual role as both clinical care providers and researchers.

This debate has proved confusing for women and their providers and may be especially harmful to members of vulnerable populations who are less likely to have a regular provider with whom to discuss conflicting recommendations and personal/family medical history (Aragon et al. 2011; Martin and Wingfield 2012; Anderson et al. 2013; 2014). Further, advocates for a targeted approach, such as the principal investigator of WISDOM, believe this approach will provide better health care value (Shieh et al. 2017; Esserman 2017). The idea of value-based care feeds concerns and fears among some patients (including WISDOM participants we interviewed) that corporate interests –the various for-profit actors within the US healthcare system—are simply aiming to save money by rationing care.

The WISDOM trial is built upon the infrastructure of the Athena Breast Health Network. Athena is a collaborative network across the five University of California Medical Centers and Sanford Health, an integrated healthcare system in the upper Midwest region of the US. The Athena network established a cohort of women who agreed to have their clinical data used in research, with the goal of enabling rapid translation of emerging data into evidence-based clinical care. The WISDOM trial, which began within the Athena network, and has since expanded across the US, aims to recruit 100k women. WISDOM first opened recruitment at two of the Athena sites, progressively opening at all six sites during the first eighteen months of the trial. After two years of lower than expected (and less racially and ethnically diverse than expected) accrual, WISDOM began to enroll anyone meeting eligibility criteria (age 40–74; identify as female³; no previous breast cancer diagnosis), first in California, and then across the US, regardless of healthcare system; additional sites at other healthcare centers across the US followed. As shown in Figure 1, participants are randomized, or if they have a strong preference may choose to be in either the annual mammography or risk-based screening arm.

Participants in both arms of WISDOM are evaluated for breast cancer risk using the Breast Cancer Surveillance Consortium (BCSC) risk calculator which incorporates age, race, if a first degree relative has had breast cancer, breast density, and history of biopsies (Shieh et al. 2017). Those in the risk-based screening arm are assessed with BCSC in combination with genomic screening of nine high and moderate penetrance breast cancer genes (ATM, BRCA1, BRCA2, CDH1, CHEK2, PALB2, PTEN, STK11, and TP53), and >200 single nucleotide polymorphisms (SNPs). The SNPs are combined to produce a Polygenic Risk Score (PRS), which is then used as a multiplier to the BCSC score to produce a five-year breast cancer risk estimate. Women in the RBS arm are assigned to one of five screening schedules according to their combined BCSC and PRS, with thresholds set by the study investigators. (See Figure 2.) Based on these risk thresholds, participants are recommended to delay screening (for average-risk women below age 50); discontinue routine screening (for low-risk women over age seventy); have a mammogram every two years; have annual mammograms; or have annual mammograms plus annual MRI. All

²Since the WISDOM trial was funded in 2015, trial investigators have collaborated with groups in other countries on similar efforts. The My Personal Breast Screening (MyPeBS) trial in the European Union (“Home Page” n.d.) and the Personalized Risk Assessment for Prevention and Early Detection of Breast Cancer (PERSPECTIVE) trial in Canada (“Personalized Breast Cancer Risk Assessment Study” 2019) Both also use Polygenic Risk Scores to predict breast cancer risk and offer screening recommendations.

³This is the current inclusion criteria for WISDOM. Our team has been engaged in discussions with the trial investigators and staff about broadening this language to better reflect best practices involved in the care of the gender-diverse population that undergoes routine mammographic screening.

screening recommendations fall within at least one current US breast cancer screening guideline (USPSTF, ACS, ACR, etc.). As participants age, and as they report specific changes (on annual surveys) such as family members diagnosed with breast cancer or personal medical changes such as biopsies, their risk may be adjusted. Each participant's risk is re-calculated annually, and this new information may change the clinical screening recommendation participants receive each year. WISDOM is an on-going study and thus outcome data, including details of how many participants were assessed at low or elevated risk, remains blinded or embargoed.

WISDOM's Pragmatic Adaptive Design

The WISDOM trial's pragmatic, adaptive design reflects an integration of clinical care and research, following a trend in translational research away from the traditional randomized controlled trial (Nelson et al. 2014; Montgomery 2017; Alberto Cambrosio et al. 2018; Yee et al. 2012). As Cambrosio et al argue, "the relation between research and care needs to be historicized by examining its repeated re-emergence and re-definition, and the shifting relations between these two components" (Cambrosio et al. 2018, 213). In historical terms, the boundary between clinical care and research was established relatively recently, as part of the human subject protection regime put in place after World War II to protect individuals and patients from abuse (Wolf et al. 2018). The *Belmont Report* firmly established boundaries between research and clinical care, with the purpose of the former to produce generalizable knowledge for the benefit of future patients, and the latter to provide optimal care to patients today (Largent, Joffe, and Miller 2011). With the rapid pace of genomic translation, this blurred boundary between clinical and research activities has become particularly pronounced (Wolf et al. 2015). As Wolf and colleagues contend, "translational genomics reflects a growing challenge to the traditional view that research and clinical care are distinct activities that should be governed by separate norms and rules" (Wolf et al. 2018, 546).

One of the most important reasons to maintain a distinction between research and clinical care is the distinction between the patient-physician relationship and the subject-investigator relationship (Brody and Miller 2013; Kass et al. 2013) Montgomery and colleagues argue, "When experiment and provision of care become one and the same thing, the status of the patient is rendered ambiguous; it is no longer clear whether the priority is to treat the individual in their best interests or to extract maximum information value from them in the interests of the experiment" (2017, 249). These arguments are largely made with treatment settings in mind, where the treating physician is also conducting research. WISDOM presents a different scenario, where the research is not based in the clinic, and providers are not recruiting their own patients for research. As we explore below, this nevertheless becomes a struggle for WISDOM clinician-investigators.

The structure of pragmatic trials explicitly aims to bridge clinical practice and research and to inform "real-world" choices about health and health care. They aim to move from "the bench to the bedside" quickly, simultaneously generating evidence about the intervention as well as how it works in practice. Califf and Sugarman (2015) define pragmatic trials as: "Designed for the primary purpose of informing decision-makers regarding the comparative

balance of benefits, burdens and risks of a biomedical or behavioral health intervention at the individual or population level” (438). Breast cancer screening is a prime example of such biomedical/behavioral health interventions, and WISDOM aims to inform national policy on breast cancer screening.

Some pragmatic trials, like WISDOM, have the added complexity of being adaptive.⁴ As defined by Chow, “an adaptive design is one that allows adaptations in trial procedures and/or statistical procedures after initiation of the trial without undermining the validity and integrity of the trial” (2005, 575). The adaptive design allows for changes in various aspects of study implementation throughout the trial, which may involve integration of new scientific evidence or responses to participants’ experiences in the trial. While adaptive trial designs vary (e.g. aspects of the trial that are adapted and timing of the adaptations), some of the advantages of such designs include the ability to incorporate emerging data from outside the trial and to react to and change procedures depending on evidence emerging within the trial. The downside is generally more complex and difficult statistical analyses to account for all the changes. In WISDOM, the adaptive design is operationalized by integrating new scientific evidence into the risk assessment algorithm as it emerges (e.g. new SNP data for the PRS), potentially changing screening recommendations for individual participants during the trial. It also permits changes to such key elements in the trial as the thresholds at which certain levels are risk are set, highlighting the arbitrariness of such thresholds and the subjective assessments that undergird evidence-based medicine (Wynants et al. 2019).

METHODS

The analysis presented here is based on multiple empirical methods grounded in an ethnographic approach. Our team of ethnographers spent five years observing thousands of meetings of more than twenty ongoing WISDOM trial working groups, which meet regularly to address and implement various aspects of the study (e.g., risk thresholds for screening recommendations, statistical methods, processes for return of genomic results, etc). In fieldnotes, we documented how study procedures developed and evolved in WISDOM; how genomic screening was implemented; how decisions were made; and the experiences of variously positioned participants in the process. Our team meets every two weeks to share our field notes and discuss themes and processes emerging across working groups. We also conducted key informant interviews with WISDOM staff, clinicians and investigators; study collaborators such as patient advocates, payors and policy makers; and experts working in academic and commercial genomics. To elucidate the experience of receiving genomic risk information and screening recommendations in the context of the trial, we audio recorded telephone genetic test results disclosure sessions in which WISDOM participants learned that they were at elevated risk for breast cancer. We also conducted interviews with trial participants at low, average, and high risk for breast cancer. By combining participant perspectives with ethnographic observations and key informant interviews, we were able to develop a multifaceted understanding of the trial implementation. The primary data for this paper are ethnographic observations of two key

⁴The WISDOM model comes out of an adaptive approach to developing new breast cancer treatments (I-SPY) previously developed and led by the WISDOM principal investigator Laura Esserman (Nelson 2014; Yee et al 2012).

WISDOM working groups, the SARB and the Risk Thresholds Working Group, along with interviews with WISDOM trial clinician-investigators. For our analysis, we have drawn on social constructionist approaches to discourse analysis, wherein we examine how language is used to construct the cultural and social world and its subjects—in this case the world of the WISDOM clinical trial, and translational genomic research more broadly along with subjectivities of clinician, researcher, research participant, patient (Jørgensen and Phillips 2002).

This study was approved by the UCSF Institutional Review Board. WISDOM staff, clinicians, and investigators were informed about our study via emailed information sheets, presentations at virtual meetings, and physical information sheets and presentations at in-person meetings. All participants were informed that they could request to have us leave any meeting or event at any time. Verbal consent was obtained at the start of each interview.

Findings

Implementing the risk algorithm—The WISDOM trial was designed to be highly automated (necessary to achieve the accrual goal of 100,000 participants) and set up so that participants who are determined by the algorithm to be at low or average risk do not engage with study personnel or have their records or questionnaires reviewed manually. Instead, they receive their risk assessment and screening recommendation via an email directing them to the study portal. As the study launched however, a WISDOM BHS elected to conduct manual reviews of each participant's risk calculation and screening assignment to ensure that the algorithm was working properly, both from an IT perspective and a Risk Threshold perspective. This review enabled the investigators to find any bugs in the system as well as to adjust the risk algorithm for a variety of circumstances and participant health histories they had not anticipated in the original trial protocol. For example, prior to the trial, investigators had not considered that there might be participants who had already taken breast cancer risk-reducing medications such as tamoxifen or raloxifene. The trial recommends these medications to those at the highest risk but did not have a mechanism in place to adjust a participant's calculated risk based on their prior use of the medications. The risk algorithm was amended to take into account this risk criterion.

This manual review process also identified participants whose trial screening recommendation varied, at times drastically, from the screening schedule that would typically be recommended in a clinical setting. At a time when clinician-investigators were first learning the WISDOM model and implementation process, they were often reviewing cases of participants they had seen in clinic, or whose medical records they could easily access. As described above, when WISDOM first began, it was only recruiting from within the Athena Network where clinician-investigators were in practice. This situation presented a double-edged sword when it came to implementing the study algorithm. While being able to verify family or personal breast history could be useful, we observed that it was especially challenging for clinician-investigators when a woman who was under the care of their site's high-risk breast clinic and who had been advised to have an annual mammogram and annual MRI was categorized as low risk by the study algorithm and assigned to screen every two years (thus moving her from the highest to the lowest risk screening regimen).

For most participants in the trial, the assigned screening schedule appeared to match the level of risk a clinician would assign. For example, a sixty-five year-old woman with no family history, low breast density and a low genetic-risk profile would receive a screening recommendation of biennial mammogram, a schedule that is both within national guidelines and within the scope of usual care at many institutions (though it might require a change if she had been screening annually, as many women in the US continue to do). When a forty-five year-old woman learns through her participation in WISDOM that she has a BRCA mutation, she would be recommended to screen with alternating MRI and mammogram every six months, and be referred for consultation on prophylactic surgery. While this is likely a shift from her prior screening regimen, the recommendations mirror what she would likely be advised to do in a clinical setting. However, questions arose when the algorithmic recommendation and the clinical practice conflicted. As one clinician-investigator described, “Everybody feels a little hedgy telling a woman who has three relatives with breast cancer, you know, ‘You can go in the low-risk bucket because we looked at all your other stuff and we couldn’t find anything.’ I completely understand that.” Members of the study team appreciated the need for an improved model of breast cancer screening and were united in a desire to expand the use of genetic testing to assess risk, yet WISDOM’s application of precision medicine to the screening context diverges from standard clinical practice in ways that were troubling – particularly when women were given a recommendation to screen less often.

As described above, the WISDOM risk algorithm combines two forms of risk assessment. First, it utilizes a breast cancer risk model called the Breast Cancer Surveillance Consortium Risk Calculator (Tice et al. 2019). This model, which one of the WISDOM clinician-investigators was instrumental in developing, is thought to be superior to other models in that it incorporates breast density as a risk factor, which many breast cancer risk models do not. However, this model uses family history as a dichotomous variable: yes or no, does a woman have a *first degree* relative with family history?⁵ This was worrisome to many clinician-investigators on the trial. For example, in BCSC, a woman with a paternal grandmother and four paternal aunts with breast cancer would be considered to have no family history because all those relatives are *second degree*. Further, BCSC would calculate equally the risk of a woman whose mother was diagnosed with stage I breast cancer at age sixty-eight and a woman whose mother and three sisters were all diagnosed with late-stage breast cancer in their forties.

Second, the BCSC score is combined with the polygenic risk score (PRS) to produce the WISDOM risk assessment score. The PRS is the more experimental aspect of WISDOM and the piece of the risk assessment that is most distinct from risk assessment in the clinical setting. For many clinician-investigators, the addition of PRS was a double-edged sword; while it offers information beyond what would be available clinically, clinician-investigators often expressed concern that PRS may not fully or accurately reflect the risk in any individual participant. For some, a participant with a concerning family history but a low

⁵WISDOM is using the Breast Cancer Surveillance Consortium Risk Calculator, Version 2. Version 3, which WISDOM plans to integrate into its risk algorithm, was expected to be released in 2021 but has been repeatedly delayed. It will include an expanded definition of family history to allow for second degree relatives and more than one first degree relative.

PRS, the genetic information was reassuring. Others were unsure about the relationship between family history and PRS and if a low PRS should change the weight given to family history. Further, WISDOM uses a race/ethnic specific PRS, and the team often expressed concern that it may be less accurate for individuals of non-European descent (James et al. 2021). Individual clinician-investigators were encountering cases for which they were unsure if the WISDOM risk algorithm produced the most appropriate screening recommendation. To review such cases as they emerged, the study team developed a Screening Assignment Review Board.

The Screening Assignment Review Board—It is important to note that the Screening Assignment Review Board (SARB) -- or even the need to review the screening recommendations of low and average-risk participants -- was not articulated in the original study design. The SARB emerged organically and reflects the tension between the idealized notion of personalized medicine and the real-world implementation of a pragmatic clinical trial of risk-based population screening. The SARB offered a space for clinician-investigators to interrogate the complexity of individual participant risk within the confines of the trial, rather than responding to each in an ad hoc manner. The clinician-investigators aimed to develop a process to discuss these cases collectively and systematically in order to determine if the algorithm should be changed. Operating in a similar manner to a molecular tumor board, this new organization was created to “frame the epistemic activities of the bioclinical collectives” (Cambrosio et al. 2018, 210).

The first SARB was held as an open session at a study retreat. Twice per year, the Athena Network brings together approximately 120 clinicians, researchers, and stakeholders to provide study updates, build relationships among dispersed staff and investigators, and hold scientific seminars on relevant issues. Attendees range from investigators and staff working on the WISDOM trial, to patient advocates, representatives from participating insurance companies, partners, funders, and clinicians and researchers working in breast health across the University of California system. The initial SARB session, held during the Spring 2017 Athena retreat, highlighted the tension between clinical care and research. With approximately thirty people in the room, many were less familiar with the details of the WISDOM study protocol and voiced their disagreement with the recommendations produced by the study algorithm. In general, many were inclined to “bump up” study recommendations – to put women assigned to lower-risk screening regimens into higher-risk screening regimens, usually due to family history that was perceived as concerning. Many clinicians in the room also wanted to offer additional clinical follow-up to these participants. These clinicians, who were less familiar with the details of WISDOM study protocol and implementation, did not recognize that it would not be feasible to offer follow-up care to such a large number of participants, particularly the many who are not affiliated with the health systems conducting the study. These clinicians were focused on providing ideal clinical care, even if that required altering the trial protocol. Additionally, at the time this meeting was held, the structure of the group had not been determined, raising questions about who on the team should have final authority in determining if a participant’s screening assignment could be altered.

After the retreat, the SARB was reorganized (See Figure 3) and more fully articulated. First, the process of identifying cases to discuss at the SARB was formalized and systematized. Through the manual review process, clinician-investigators elucidated certain categories of participants to “flag” for review. The “low-risk review” category included participants whom the trial algorithm had categorized as low risk, but who, in a clinical setting, might be considered high risk due to their family history. Trial investigators created a set of “low-risk review” criteria to automatically generate a report of participants who will then be reviewed by the clinician-investigators. These criteria are based on the US Preventive Services Task Force and National Comprehensive Cancer Network guidelines for genetic testing and involve combinations of breast and ovarian cancer in first- and second-degree relatives, with age of diagnosis playing a key role.

Next, fewer people were invited to participate in the SARB: only Breast Health Specialists who spoke with participants about their risk, and members of the Risk Thresholds working group were included. At least one primary care provider who was also a member of the Risk Thresholds group was required to attend each meeting. While implementing voting mechanisms was considered, they ultimately decided to operate by consensus, with the opinions of Risk Thresholds committee members carrying more weight. For the first few years of the SARB, screening assignment changes were elevated to the Risk Thresholds group and ultimately to the study PI. Subsequently the SARB’s decisions were only reviewed by the Risk Thresholds group when clear consensus could not be reached. In practice, this has meant that many screening assignments the SARB decided to change, ultimately stayed the same after veto by someone higher up. Importantly, elevating the decision to the Risk Thresholds group is an opportunity for the study to adjust the trial’s risk thresholds, thus changing the screening recommendations for many participants, not only the case that was flagged for review. For example, after seeing many cases of participants in their forties who were recommended to delay having a mammogram until age fifty but who had an extensive family history of breast cancer in first- and second-degree relatives, diagnosed at a young age, the risk thresholds group decided that these participants would automatically be moved to biennial screening if there was no confirmed genetic mutation in the family associated with hereditary breast and ovarian cancer syndrome. This decision was a compromise; some in the group felt strongly that those with a concerning family history should be screened at least annually while others felt that since these participants had been tested for the nine hereditary breast and ovarian cancer genes on the WISDOM panel, they should no longer be considered at elevated risk. Thus, the SARB process allowed the group to come to a consensus on how to move forward.

As this group established a rhythm, case reviews became more routine. They established a “tendency to stick with the screening recommendation whenever possible,” as one clinician-investigator put it. In their first meeting after the retreat, the group discussed the fact that many people who had been at the retreat meeting were less comfortable with the idea of biennial mammograms as the standard of care, despite it being the standard in other countries and increasingly recommended by US health organizations (e.g., US Preventive Services Task Force).⁶ One clinician-investigator noted the trial aims to address the very concerns raised about appropriate screening intervals and modalities for women in various circumstances.

While discussing a particularly difficult case during a SARB meeting around that time, one clinician-investigator spoke to the embodiment of the clinical-research process as she described approaching the case as different versions of herself, noting that “Scientist Laurie” wants to put blinders on and pretend she never saw the patient in the clinic; just accept that the study recommendation is the study recommendation and, unfortunately, for some patients, the study algorithm may not produce the most ideal screening schedule. In contrast, “Clinician Laurie” wants to acknowledge and act upon the risk she knows about from her clinical interaction with the patient. She acknowledged that the study algorithm included more than was tested for in the clinic due to the addition of polygenic risk, but she did not think that alone accounted for this participant’s particularly strong family history.

As the SARB developed, three guiding principles emerged that the team often repeated, becoming refrains, as they made decisions about particularly difficult cases: “trust the model”; “women do not have to follow the study’s recommendation”; and “women in every ‘risk bucket’ will develop cancer.” (See Table 1).

“Trust the model”: The WISDOM trial is testing the idea that risk-based screening is just as safe as annual mammography, yet retains flexibility in the inputs and criteria deployed in the risk thresholds. As an adaptive, pragmatic clinical trial, WISDOM is not strictly bound by a particular algorithm or set of risk thresholds; this is what creates a need and possibility for the SARB. Within the SARB, clinician-investigators have the space to challenge the specific elements used to construct the risk thresholds, yet to test the model, they must implement it with some degree of consistency. This idea that they must “trust the model” is a theme the clinician-investigators come back to time and time again. They remind themselves that they need to trust in the risk algorithm the study has developed and continues to refine in order to fully implement and test it. But, as one clinician-investigator described, “I think that throwing – not throwing away the family history but trusting that the algorithms that have been built through WISDOM are adequately assessing a clinical picture has been a work in progress.” Notice that she started to say they were throwing out family history and then caught herself. This highlights that, at times, it can feel like they are disregarding what was once the primary factor in determining screening recommendations for women who are not mutation carriers in high-risk breast cancer genes like BRCA. The cases where clinical care and trial recommendations are misaligned are in many ways the most critical for the trial; these are the cases that determine if risk-based screening is as effective as standard of care.

When the clinician-investigators encourage themselves and others to trust the model, they are attempting to both build confidence and to remind each other of the purpose of the trial. The trial can only be successful if they are able to trust the model enough to test it. In fact, unlike the other guiding principles, this idea of trusting the model was discussed

⁶The issue of the most appropriate standard of care to serve as the comparator in WISDOM has been debated from early on in the study. Over the last decade, many guidelines and prominent medical organizations have shifted their standard of care away from annual mammography beginning at age forty. However, annual mammography still remains the standard for most professional radiology groups; even within medical practices and organizations that routinely recommend biennial mammography to their patients, radiology centers continue to remind the same patients to have a mammogram annually. Thus, WISDOM has elected to continue with annual mammography as the comparator.

as the primary approach this group should take when in doubt. In one example, early on in the tenure of the SARB, they discussed the case of a 57-year-old participant whose mother and four out of five maternal aunts had been diagnosed with breast cancer, with several diagnoses occurring in their thirties. The team went back and forth discussing several elements of her case: they'd feel more strongly if she was 40, but at 57 she had passed most of the risk period without developing cancer; she'd had an oophorectomy which may have lowered her breast cancer risk, but she'd also taken hormone replacement therapy which could elevate her risk although they didn't have details on timing or duration of the hormone replacement therapy; her breast density was high but not the highest category; they did not know whether anyone in her family had had genetic testing. Eventually, many agreed that it could be argued both ways and that there was not a clear right or wrong answer. Someone asked what they do in that case, to which one clinician-investigator responded, "When we are on the fence, we default to trusting the study algorithm." This did indeed become the default. The burden of proof shifted from needing to affirm that the screening recommendation produced by the algorithm was the best schedule for each individual to instead demonstrating why the study model should *not* be trusted – and whether a new rule or risk threshold was needed to account for the new scenario. Yet, there are still situations where members of the team can't quite trust the model; the next two refrains are utilized to provide reassurance when a team member is still uncomfortable with a screening recommendation.

"Women do not have to follow the study's recommendation": Fundamental to WISDOM's design as a pragmatic trial, is the idea that participants are free to follow or not to follow WISDOM's screening recommendation. Ideally, to test WISDOM's primary hypothesis (that risk-based screening is as safe as annual mammography), participants will follow the study's screening recommendation. However, a secondary outcome of the trial is whether women will follow the recommendations: is risk-based screening feasible at scale and, thus, could it be implemented nationally as the new screening paradigm?

Given this element of the study design, SARB members would often remind each other that "women do not have to follow the study recommendation." Simultaneously committing to utilize the model while also taking comfort in knowing that some women will not follow the screening recommendations eases the burden for clinician-investigators who are not fully convinced of the validity of WISDOM's risk assessment in all cases. This refrain calls into question an implicit statistical assumption that most participants will follow the study recommendations and, to some degree, undermines the exhortation to trust in the model. In such cases, the SARB members don't actually have to trust the model; instead, they can hope that the participant and her providers will revert to standard clinical practice. If for example, they know that the participant is a patient at their own institution or a patient in their own high-risk clinic, they can feel confident that the participant will in fact be screened according to standard clinical practice rather than the WISDOM study's recommendation.

In one case, a 44-year-old participant was given the recommendation to wait until age fifty to start screening. Clinician-investigators at her site were concerned about her family history. As they reported to the SARB, her mother was diagnosed with breast cancer at 55 and her sister was diagnosed with breast cancer at 48. She also had a maternal aunt with

ovarian cancer at 50. While ovarian cancer is not used in the WISDOM risk calculations, a family history of ovarian cancer can be suggestive of Hereditary Breast and Ovarian Cancer syndrome, so this type of family history can be concerning. At the time, this participant was being seen at a high-risk breast clinic and was undergoing an annual MRI and an annual mammogram, alternating every six months. Additionally, the clinician presenting this case included her risk estimate from another breast cancer risk predictor called IBIS (International Breast Cancer Intervention Study) also called the Tyrer-Cuzick Breast Cancer Risk Evaluation Tool. Similar to BCSC, IBIS is a risk assessment model commonly used clinically to assess Breast Cancer risk (Warwick et al. 2014). IBIS is often critiqued by WISDOM clinician-investigators for *overestimating* risk.

WISDOM clinician-investigators use IBIS or other risk assessment models in an attempt to show why a participant may be *truly* high risk, in contrast with the WISDOM algorithm's estimate. This reliance on other risk models to validate, explicate, or contradict the WISDOM model demonstrates a lack of confidence in the WISDOM model. When a patient is shown to be at high risk based on both clinical judgment and other risk calculators, many clinician-investigators prefer to make a screening recommendation in line with the other models. While this is often countered by arguments highlighting why the WISDOM model may offer a superior risk estimate (e.g. genetic testing is not routinely offered in primary care; IBIS does not take into account PRS), those rationales are not always sufficient to assuage concerns or augment trust in the WISDOM model.

In the case of one 44-year-old woman with an extensive family history, her IBIS score indicated a lifetime risk of 31.2%, well above the 20% threshold set by National Comprehensive Cancer Network (and followed by many payors) for the recommendation to screen with both mammogram and MRI alternating every six months. The clinicians who presented this case were uncomfortable with the idea that this participant might not screen for six years, but others on the team felt that genetic testing, including negative results on the nine gene panel and her low Polygenic Risk Score, indicated that she was low risk. The investigator presenting the case ended the conversation by saying "I don't think she'll follow it; I'm assuming she will stick with the specialist here at the clinic" noting that it was such a "dramatic switch." This clinician-investigator could feel comfortable leaving WISDOM's recommendation in place only because she was confident that the participant would be unlikely to follow it and her clinician would continue to recommend a more aggressive screening plan.

"Women in every risk bucket will develop cancer": Screening does not prevent cancer. The goal of WISDOM, and of screening in general, is to detect cancers at the point when the cancer can be treated effectively. As one clinician-investigator said,

"It's personalized but it's still bucket-based... You have to go back to population principles and remember that we're putting people in buckets; we're not predicting whether or not they will ever get breast cancer. And the low-risk bucket should have, at the end of the day, a lower incidence of breast cancer and the high-risk bucket should have a higher incidence of breast cancer. But we still can't tell any one woman in any bucket that they will or will not get breast cancer. It's putting

you with a peer group, which is a very population way of thinking about it rather than personalized.”

This clinician-investigator knows that some women in every “risk bucket” will be diagnosed with breast cancer. The WISDOM hypothesis is that fewer women in the low-risk “bucket” will be diagnosed, but it is still expected to happen. As researchers, looking at numbers in a dataset or formulas in a statistics program, the team members can accept this. It feels harder when they are looking at the details of a participant’s family and medical history and thinking about the implications for her life – and even harder if they know the participant personally as a patient. By repeating the refrain “*women in every risk bucket will develop cancer*,” clinician-investigators are re-training themselves to approach these decisions within a population context, rather than the clinical context. They often need to remind themselves that no screening regimen will prevent cancer – whether it’s annual for everyone or it’s a risk-stratified system. The goal of creating the different screening categories (instead of screening all women the same way) is to reduce the harms caused by unnecessary screening, such as false positives, unnecessary biopsies and the anxiety that go along with both. Yet, despite an intellectual commitment to these ideals, this tension between research – a desire to stick to the model to see if it works – and clinical – a desire to make sure that each participant receives the best care possible – persists. Many of the clinician-investigators approach these decisions as clinicians: they want to assess what is the right screening for the “patient” in front of them. This is often in conflict with their role as researcher, where the goal of each decision should be to generate knowledge to help the largest number of patients in the future.

A member of the team presented the case of a 41-year-old woman who received the recommendation to have biennial mammograms. It should be noted that this is actually an elevated screening recommendation for her age; for women in their forties, those who are low risk are recommended to begin screening at age 50. However, this participant’s mother was diagnosed with breast cancer at 63 and her sister was diagnosed with breast cancer at 41. Most concerning for those on SARB was that her relatives with breast cancer were reported to have had genetic testing, but with negative results. Meaning, this participant was not a “true negative”; with no variant identified as the cause of the cancer in her family, her results could not determine whether she had the risk factor(s) causing the cancer in her family. While most clinician-investigators were comfortable with the recommendation to have biennial mammograms as they felt reassured that she was not being asked to forgo screening until age 50, one clinician-investigator felt concerned and asked, “are we sure we are screening for all mutations? Are there things we have yet to identify? Are we confident?” Another clinician-investigator, a leader in the group, responded “Absolutely not. There are things we don’t know about. We will miss things. We will have cancer come up in women with less than 1% risk.”

The team is well aware that participants will develop breast cancer, and that the model does not perfectly predict who will get it. In fact, the number of women who get breast cancer during the study is a primary outcome of the trial. Yet, herein lies the contradiction of the SARB. The clinician-investigators hypothesize the combination of PRS and BCSC improves risk prediction over the current model of age and gender that is the basis for annual

mammography. As one clinician-investigator involved in establishing the risk thresholds explained in an interview, “Cancer screening is inherently risk-based. We just only used two risk factors, being female and age, to drive it in the past. In my mind, we’re just trying to improve on that.” He continued, “In the end, you need to personalize things not just based on risk, but also on other values of the patient. We do that every day in internal medicine.” WISDOM clinician-investigators acknowledge that participants will get breast cancer, some of which may not be caught as early as it may have been with alternative screening models. There is inherent risk that a risk-based screening recommendation may not be the ideal screening schedule for an individual participant. So, why not trust the model, with the acknowledgement that participants do not have to follow it, for every participant in the trial?

Limits of the guiding principles: a case example: The cases we described briefly above were ones in which the SARB members relied on the three refrains to negotiate acceptance of the WISDOM screening recommendation, even when it differed from the screening schedule the participant had previously or would typically be offered in the clinic. However, there are other cases that arise in which the team is unable to reach a consensus to trust the model. As one clinician-investigator reflected on the process, “I think over time the Breast Health Specialists have become more comfortable and we’re really focusing more now on those cases that are really outliers, where there’s something about the case that just makes you wonder, ‘is there a mistake somewhere’.” It is important to note, that these “outliers” tend to be in cases when the recommendation is for biennial screening or delaying the start of screening. From our observations, clinician-investigators, with very few exceptions, are comfortable when the risk algorithm leads to a more intensive screening recommendation than what would be recommended clinically. The cases where a “mistake” seems possible, tend to be those that offer less screening than usual care. Next, we describe one such case and how the clinician-investigators utilized the refrains in their discussions but did not ultimately follow their ethos.

A Breast Health Specialist presented the case of a 52-year-old woman who was set to receive a recommendation for biennial mammograms. This participant had, as another Breast Health Specialist described it, a “striking” family history. Her mother had been diagnosed with breast cancer at the age of 45, considered “early onset” for breast cancer; early breast cancers can be more aggressive or a sign of Hereditary Breast and Ovarian Cancer syndrome. This participant also had a sister who was diagnosed with breast cancer at 56, a maternal aunt who was diagnosed with breast cancer at 35, a paternal aunt with breast cancer at 65 and two cousins with breast cancer. Additionally, her father had colon cancer, her sister had lymphoma, and her brother had pancreatic cancer. This is the type of history that the clinicians, especially the genetic counselors, find very concerning. This participant also had other risk factors for breast cancer: heterogeneously dense breasts and a history of breast biopsies, which are taken into account in the BCSC model, as well as a few concerning lifestyle factors, such as a history of smoking, which are not.

When considering her BCSC score alone, this participant fell into the top 2.5% of risk for her age (WISDOM’s cut off for annual mammogram and a risk reduction consultation). Yet,

when combined with her Polygenic Risk Score (the experimental component of the study), her risk estimate of 3.12% put her below the top 2.5% of risk for women her age.

On paper, this participant is not dissimilar from many cases discussed on the SARB, yet the extensive nature of the family history caused many on the team to have more difficulty following the guiding principles. First, several members of the team who identified themselves as having a tendency to trust the model, used this designation as a justification for making this case a rare exception to the guiding principle of “sticking with” the model. One clinician-investigator invoked this argument saying, “I tend to be pretty conservative with changing [screening recommendations] but...” to preface her recommendation to shift the participant to annual mammography. Another said, “this is happening...rarely enough that if people feel more comfortable screening at every year, that’s fine.”

Another argument for changing the participant’s screening was the recognition that concern about an aggressive cancer might be weighted more heavily than an early cancer. One clinician-investigator asked, “Is the concern that she’ll get an *aggressive* cancer or that she’ll get *early* cancer? Those two things need to be disentangled.” This clinician-investigator was pointing out that moving her to annual screening could help identify an aggressive cancer, but if the SARB members were concerned about cancer at a young age (which may be slow growing or very responsive to treatment), changing her screening may not change her mortality risk. However, this idea of that the *type* of cancer the participant might be at risk for, and thus the larger role of screening in the prevention of breast cancer mortality (as opposed to *morbidity*, preventing breast cancer itself), was not enough to reassure the members of the team who felt that the (experimental) PRS may be too heavily weighted. The SARB ultimately made the decision to recommend this participant be moved to annual screening, along with a risk-reduction and hereditary cancer consultation with a Breast Health Specialist.

Discussion

The promise of personalized medicine is the right treatment, for the right patient, at the right time (Abrahams 2008). This approach is mirrored within the personalized *screening* framework of WISDOM, and clinician-investigators repeatedly invoked this “mantra” in interviews and ethnographic observations. On the one hand, personalized screening “refers to a medical model using characterisation of individuals’ phenotypes and genotypes (e.g. molecular profiling, medical imaging, lifestyle data)... to determine the predisposition to disease and/or to deliver timely and targeted prevention”; yet, personalized medicine also relates to the “broader concept of patient-centered care, which takes into account that, in general, healthcare systems need to better respond to patient needs” (Erikainen and Chan 2019, 313). Our research demonstrates that this latter aspect of personalized medicine, patient-centered care, is what many clinician-investigators feel is missing from the WISDOM model, despite the trial’s framing of personalized medicine. The term “personalized” is routinely used by WISDOM team members to refer to the trial arm that utilizes genomic data to stratify risk, and also as a description of a general approach to care that aims to tailor interventions to the values and preferences of the patient. The clinician-investigators desire a personalized model that does both. Similarly, in interviews

with WISDOM participants, many understood “personalized” to mean more personal care, rather than the use of genomic information to tailor the care offered.

Due to this conflation, as well as the shift to *precision* within oncology, some have argued that it may be more appropriate to use the term *stratified* rather than personalized to “evade the kinds of “overly optimistic” associations relating to the “personalization” concept” (Erikainen and Chan 2019, 317). According to Erikainen and Chan (2019), both the World Health Organization and the UK Academy of Medical Sciences, have argued for the use of *stratification* due to the “possibly overambitious promise” of personalization, as well as noting that, in the population context, this type of medicine often involves the “grouping of patients based on risk of disease” (317). As noted above, the clinician-investigators have recognized that WISDOM, while more personalized than broad recommendations of annual mammography for all, is still “bucket based.” WISDOM is testing a “risk-based” or risk-stratified approach to screening, and the lack of accounting for patient-provider judgment and interaction in the trial model is challenging for many on the team.

The SARB offers the clinician-investigators an “organizational mechanism” and “work routine” within the trial to acknowledge and weigh clinical judgment on individual participant cases (Crabu 2021). There, the clinician-investigators are able to utilize their clinical expertise to evaluate the appropriateness of the screening recommendation determined by the WISDOM risk algorithm. The SARB, through its deployment of the three refrains – “trust the model,” “women do not have to follow the recommendations,” and “women in all ‘buckets’ will develop cancer”— is engaged in several forms of work. First, the SARB as an intervention, does the “bridging work” (Stefan Timmermans and Buchbinder 2012; S. Timmermans, Tietbohl, and Skaperdas 2017) to reconcile the WISDOM risk algorithm with clinical practice. The SARB is “both epistemic and organizational” (Cambrosio et al. 2018) in that it both generates risk thresholds to address previously unconsidered risk elements (and individual participant circumstances) and offers an alternative to ad hoc deviation from the trial protocol. Finally, through the refrains that articulate the guiding principles, the SARB encourages the clinician-investigators to center the population goals of the trial, while at the same time enabling them to assuage (to some degree) their concerns about their obligation to the “patient” – even in this trial where the participants are, for the most part, not actually their patients.

Although WISDOM is a multi-sited trial operating online, outside of a clinical setting, and with thousands of participants the clinician-investigators have never met, they still have trouble releasing themselves from the clinical mindset. The refrain of “trust the model” is deployed as a reminder to the team to center the population (and the future), but the clinician-investigators struggle not to prioritize the individual participant (a.k.a. patient) today. In many respects, the SARB represents what one might expect if risk-stratified screening—the imagined new paradigm that will result from WISDOM—is implemented broadly. The clinician-investigators are weighing the risk assessment produced by the algorithm with clinical judgment, the patient’s prior screening schedule, and (assumptions about) patient values and preferences. The WISDOM algorithm becomes one data point to guide decision making, rather than the sole driver of the screening schedule. They are bringing the real-world elements of clinical screening decision making to the trial context

but attempting to do so without undermining the ideals or statistical demands and necessities of the trial.

As an adaptive pragmatic trial, the criteria and inputs that make up the WISDOM risk model have not remained constant; trial investigators recognize that if they were still testing the risk thresholds and algorithm designed in 2015 at the inception of the trial, it would no longer be relevant by the time trial results were analyzed and ready for implementation. However, this study design, and in particular the use of a mechanism like the SARB to potentially make adaptations on the level of individual participants, raises novel questions not faced in many RCTs. With the potential to shift the risk algorithm and incorporate new knowledge about how to understand and count risk, how strictly should WISDOM enforce its own risk algorithm? What are the implications for scientific knowledge production when allowing for flexibility and real-world considerations? In some respects, the creation of the SARB represents an emergent bureaucracy and speaks to the limitations of current regulatory structures, such as institutional review boards and data science and monitoring boards, to content with the challenges of pragmatic research. Yet, questions remain about what level of oversight or compliance should be levied by this model as well as where research regulation ends, and clinical regulation and oversight begin.

We contend that what complicates the trial *is* in fact the adaptive and pragmatic design and the unclear role of clinical judgment within this type of trial. WISDOM tests the idea that risk-stratified medicine, in the form of a population approach to genomic screening and risk algorithms, can better determine how often women should be screened for breast cancer than a one-size-fits-all approach. Yet, importantly, current breast cancer screening guidelines still often include a parameter that screening decisions should be made in collaboration with the primary care provider (Keating and Pace 2018). Further, despite guidelines shifting away from annual mammograms, the widely held belief that “early detection saves lives” has been deeply ingrained in both patients and providers. A recent study demonstrated that more than 80% of clinics still recommend annual mammography starting at age 40 (Patel, Lee, and Marti 2021). The SARB demonstrates the unique challenges of implementing a personalized approach to population wide breast cancer screening with a technology that only allows for risk-stratification. It is still unclear what role primary care providers, and more specifically, *clinical judgment*, should play in risk stratified breast cancer screening and to what extent an algorithm, such as the one WISDOM uses and continues to refine, can or should replace or supplement clinical judgment. While the SARB represents a structural solution to the role ambiguity felt by clinician-investigators, it only offers an avenue to weigh clinical judgment, but does not fully resolve the tensions between research and clinical care as it cannot substitute for the patient-provider relationship.

The SARB model presents one avenue to bring clinical judgement into the clinical trial setting, but further research will be needed to assess the efficacy of this model compared to others that may be in use in similar ‘precision population screening’ trials. Such research could advance pragmatic approaches that center real-world settings, including patient values and the patient-provider relationship, in evaluations of new technologies and interventions. Importantly, our analysis demonstrates the value of embedded ELSI research for producing a

multifaceted understanding of the challenges involved in translating genomic technology to clinical practice.

Acknowledgements:

We would like to thank the WISDOM trial investigators, staff, and advocates for their support of this project and our ongoing collaboration. We are grateful to the WISDOM Bioethics Working Group for their invaluable work shaping our understanding of these issues. Ms. Leslie Riddle and Ms. Mikaella Caruncho contributed greatly to the data collection and analysis of the ongoing WISDOM study. We are thankful to Dr. Barbara Koenig for her contributions as a principal investigator of the WISDOM ELSI study and her ongoing support and guidance.

References

- Abrahams Edward. 2008. "Right Drug—Right Patient—Right Time: Personalized Medicine Coalition." *Clinical and Translational Science* 1 (1): 11–12. 10.1111/j.1752-8062.2008.00003.x. [PubMed: 20443813]
- Adams Vincanne, Murphy Michelle, and Clarke Adele E.. 2009. "Anticipation: Technoscience, Life, Affect, Temporality." *Subjectivity* 28 (1): 246–65. 10.1057/sub.2009.18.
- Anderson Britta L., Pearlman Mark, Griffin Jennifer, and Schulkin Jay. 2013. "Conflicting and Changing Breast Cancer Screening Recommendations: Survey Study of a National Sample of Ob-Gyns after the Release of the 2009 USPSTF Guidelines." *Journal for Healthcare Quality: Official Publication of the National Association for Healthcare Quality* 35 (4): 25–35. 10.1111/jhq.12009. [PubMed: 23590634]
- Anderson Britta L., Urban Renata R., Pearlman Mark, and Schulkin Jay. 2014. "Obstetrician-Gynecologists' Knowledge and Opinions about the United States Preventive Services Task Force (USPSTF) Committee, the Women's Health Amendment, and the Affordable Care Act: National Study after the Release of the USPSTF 2009 Breast Cancer Screening Recommendation Statement." *Preventive Medicine* 59 (February): 79–82. 10.1016/j.ypmed.2013.11.008. [PubMed: 24246966]
- Aragon Robert, Morgan John, Wong Jan H., and Lum Sharon. 2011. "Potential Impact of USPSTF Recommendations on Early Diagnosis of Breast Cancer." *Annals of Surgical Oncology* 18 (11): 3137–42. 10.1245/s10434-011-1915-9. [PubMed: 21947591]
- Armstrong David. 2019. "Diagnosis: From Classification to Prediction." *Social Science & Medicine* (1982) 237 (September): 112444. 10.1016/j.socscimed.2019.112444. [PubMed: 31374408]
- Bourret Pascale, Keating Peter, and Cambrosio Alberto. 2011. "Regulating Diagnosis in Post-Genomic Medicine: Re-Aligning Clinical Judgment?" *Social Science & Medicine* (1982) 73 (6): 816–24. 10.1016/j.socscimed.2011.04.022. [PubMed: 21664021]
- Brody Howard, and Miller Franklin G.. 2013. "The Research-Clinical Practice Distinction, Learning Health Systems, and Relationships." *Hastings Center Report* 43 (5): 41–47. 10.1002/hast.199.
- Califf Robert M, and Sugarman Jeremy. 2015. "Exploring the Ethical and Regulatory Issues in Pragmatic Clinical Trials." *Clinical Trials* (London, England) 12 (5): 436–41. 10.1177/1740774515598334. [PubMed: 26374676]
- Cambrosio A, Keating P, É Vignola-Gagné, Besle S, and Bourret P. 2018. "Extending Experimentation: Oncology's Fading Boundary between Research and Care." *New Genetics and Society*. 10.1080/14636778.2018.1487281.
- Cambrosio Alberto, Keating Peter, Etienne Vignola-Gagné Sylvain Besle, and Bourret Pascale. 2018. "Extending Experimentation: Oncology's Fading Boundary between Research and Care." *New Genetics and Society* 37 (3): 207–26. 10.1080/14636778.2018.1487281.
- Chow Shein-Chung, Chang Mark, and Pong Annpey. 2005. "Statistical Consideration of Adaptive Methods in Clinical Development." *Journal of Biopharmaceutical Statistics* 15 (4): 575–91. 10.1081/BIP-200062277. [PubMed: 16022164]
- Crabu Stefano. 2021. "Organizing the Precision Clinic: Arranging Expertise, Knowledge and Technologies in Cancer Precision Medicine Clinical Trials." *New Genetics and Society* 40 (1): 58–72. 10.1080/14636778.2021.1882839.

- Erikainen Sonja, and Chan Sarah. 2019. "Contested Futures: Envisioning 'Personalized,' 'Stratified,' and 'Precision' Medicine." *New Genetics and Society* 38 (3): 308–30. 10.1080/14636778.2019.1637720. [PubMed: 31708685]
- Esserman Laura J. 2017. "The WISDOM Study: Breaking the Deadlock in the Breast Cancer Screening Debate." *Npj Breast Cancer* 3 (1): 34. 10.1038/s41523-017-0035-5. [PubMed: 28944288]
- "Home Page." n.d. MyPeBS. Accessed October 3, 2020. <https://mypebs.eu/>.
- James Jennifer Elyse, Riddle Leslie, Koenig Barbara Ann, and Joseph Galen. 2021. "The Limits of Personalization in Precision Medicine: Polygenic Risk Scores and Racial Categorization in a Precision Breast Cancer Screening Trial." *PLOS ONE* 16 (10): e0258571. 10.1371/journal.pone.0258571. [PubMed: 34714858]
- Jørgensen Marianne W., and Phillips Louise J.. 2002. *Discourse Analysis as Theory and Method*. SAGE.
- Kass Nancy E., Faden Ruth R., Goodman Steven N., Pronovost Peter, Tunis Sean, and Beauchamp Tom L.. 2013. "The Research-Treatment Distinction: A Problematic Approach for Determining Which Activities Should Have Ethical Oversight." *Hastings Center Report* 43 (s1): S4–15. 10.1002/hast.133.
- Keating Nancy L., and Pace Lydia E.. 2018. "Breast Cancer Screening in 2018: Time for Shared Decision Making." *JAMA* 319 (17): 1814. 10.1001/jama.2018.3388. [PubMed: 29715344]
- Kerr A, Ross E, Jacques G, and Cunningham-Burley S. 2018. "The Sociology of Cancer: A Decade of Research." *Sociology of Health & Illness* 40 (3): 552–76. [PubMed: 29446117]
- Kerr Anne, Choon Key Chekar Emily Ross, Swallow Julia, and Cunningham-Burley Sarah. 2021. *Personalised Cancer Medicine: Future Crafting in the Genomic Era*. Wellcome Trust–Funded Monographs and Book Chapters. Manchester (UK): Manchester University Press. <http://www.ncbi.nlm.nih.gov/books/NBK567274/>.
- Khoury Muin J., and Galea Sandro. 2016. "Will Precision Medicine Improve Population Health?" *JAMA* 316 (13): 1357–58. 10.1001/jama.2016.12260. [PubMed: 27541310]
- Largent Emily A., Joffe Steven, and Miller Franklin G.. 2011. "Can Research and Care Be Ethically Integrated?" *The Hastings Center Report* 41 (4): 37–46. 10.1002/j.1552-146x.2011.tb00123.x.
- Martin Nicolle, and Wingfield John. 2012. "USPSTF Screening Recommendations for Breast Cancer: The Potential Impact on the African American Community." *Journal of Health Care for the Poor and Underserved* 23 (2 Suppl): 91–97. 10.1353/hpu.2012.0072. [PubMed: 22643557]
- Montgomery Catherine M. 2017. "From Standardization to Adaptation: Clinical Trials and the Moral Economy of Anticipation." *Science as Culture* 26 (2): 232–54. 10.1080/09505431.2016.1255721.
- Nelson Nicole C., Keating Peter, Cambrosio Alberto, Aguilar-Mahecha Adriana, and Basik Mark. 2014. "Testing Devices or Experimental Systems? Cancer Clinical Trials Take the Genomic Turn." *Social Science & Medicine* 111 (June): 74–83. 10.1016/j.socscimed.2014.04.008. [PubMed: 24768778]
- Oeffinger Kevin C., Fontham Elizabeth T. H., Etzioni Ruth, Herzig Abbe, Michaelson James S., Ya-Chen Tina Shih Louise C. Walter, et al. 2015. "Breast Cancer Screening for Women at Average Risk: 2015 Guideline Update From the American Cancer Society." *JAMA* 314 (15): 1599–1614. 10.1001/jama.2015.12783. [PubMed: 26501536]
- Patel Neal S., Lee Mark, and Marti Jennifer L.. 2021. "Assessment of Screening Mammography Recommendations by Breast Cancer Centers in the US." *JAMA Internal Medicine* 181 (5): 717–19. 10.1001/jamainternmed.2021.0157. [PubMed: 33720303]
- "Personalized Breast Cancer Risk Assessment Study." 2019. Cancer Care Ontario. October 28, 2019. <https://www.cancercareontario.ca/en/guidelines-advice/cancer-continuum/screening/personalized-breast-cancer-risk-assessment-study>.
- Shieh Yiwey, Eklund Martin, Madlensky Lisa, Sawyer Sarah D., Thompson Carlie K., Allison Stover Fiscalini Elad Ziv, et al. 2017. "Breast Cancer Screening in the Precision Medicine Era: Risk-Based Screening in a Population-Based Trial." *JNCI: Journal of the National Cancer Institute* 109 (5). 10.1093/jnci/djw290.
- Tice Jeffrey A., Bissell Michael C. S., Miglioretti Diana L., Gard Charlotte C., Rauscher Garth H., Dabbous Firas M., and Kerlikowske Karla. 2019. "Validation of the Breast Cancer Surveillance

- Consortium Model of Breast Cancer Risk.” *Breast Cancer Research and Treatment* 175 (2): 519–23. 10.1007/s10549-019-05167-2. [PubMed: 30796654]
- Timmermans S, Caroline K. Tietbohl, and Skaperdas Eleni. 2017. “Narrating Uncertainty: Variants of Uncertain Significance (VUS) in Clinical Exome Sequencing.” 10.1057/S41292-016-0020-5.
- Timmermans Stefan, and Buchbinder Mara. 2012. “Expanded Newborn Screening: Articulating the Ontology of Diseases with Bridging Work in the Clinic.” *Sociology of Health & Illness* 34 (2): 208–20. 10.1111/j.1467-9566.2011.01398.x. [PubMed: 21929648]
- Tutton Richard. 2012. “Personalizing Medicine: Futures Present and Past.” *Social Science & Medicine* (1982) 75 (10): 1721–28. 10.1016/j.socscimed.2012.07.031. [PubMed: 22906526]
- Warwick Jane, Birke Hanna, Stone Jennifer, Warren Ruth ML, Pinney Elizabeth, Brentnall Adam R., Duffy Stephen W., Howell Anthony, and Cuzick Jack. 2014. “Mammographic Breast Density Refines Tyrer-Cuzick Estimates of Breast Cancer Risk in High-Risk Women: Findings from the Placebo Arm of the International Breast Cancer Intervention Study I.” *Breast Cancer Research* 16 (5): 451. 10.1186/s13058-014-0451-5. [PubMed: 25292294]
- Wolf Susan M, Amendola Laura M, Berg Jonathan S, Chung Wendy K, Clayton Ellen Wright, Green Robert C, Harris-Wai Julie, et al. 2018. “Navigating the Research–Clinical Interface in Genomic Medicine: Analysis from the CSER Consortium.” *Genetics in Medicine* 20 (5): 545–53. 10.1038/gim.2017.137. [PubMed: 28858330]
- Wolf Susan M., Branum Rebecca, Koenig Barbara A., Petersen Gloria M., Berry Susan A., Beskow Laura M., Daly Mary B., et al. 2015. “Returning a Research Participant’s Genomic Results to Relatives: Analysis and Recommendations.” *The Journal of Law, Medicine & Ethics* 43 (3): 440–63. 10.1111/jlme.12288.
- Wynants Laure, van Smeden Maarten, McLernon David J., Timmerman Dirk, Steyerberg Ewout W., Van Calster Ben, and on behalf of the Topic Group ‘Evaluating diagnostic tests and prediction models’ of the STRATOS initiative. 2019. “Three Myths about Risk Thresholds for Prediction Models.” *BMC Medicine* 17 (1): 192. 10.1186/s12916-019-1425-3. [PubMed: 31651317]
- Yee Douglas, Haddad Tufia, Albain Kathy, Barker Anna, Benz Christopher, Boughey Judy, Buxton Meredith, et al. 2012. “Adaptive Trials in the Neoadjuvant Setting: A Model to Safely Tailor Care While Accelerating Drug Development.” *Journal of Clinical Oncology* 30 (36): 4584–86. 10.1200/JCO.2012.44.1022. [PubMed: 23169510]

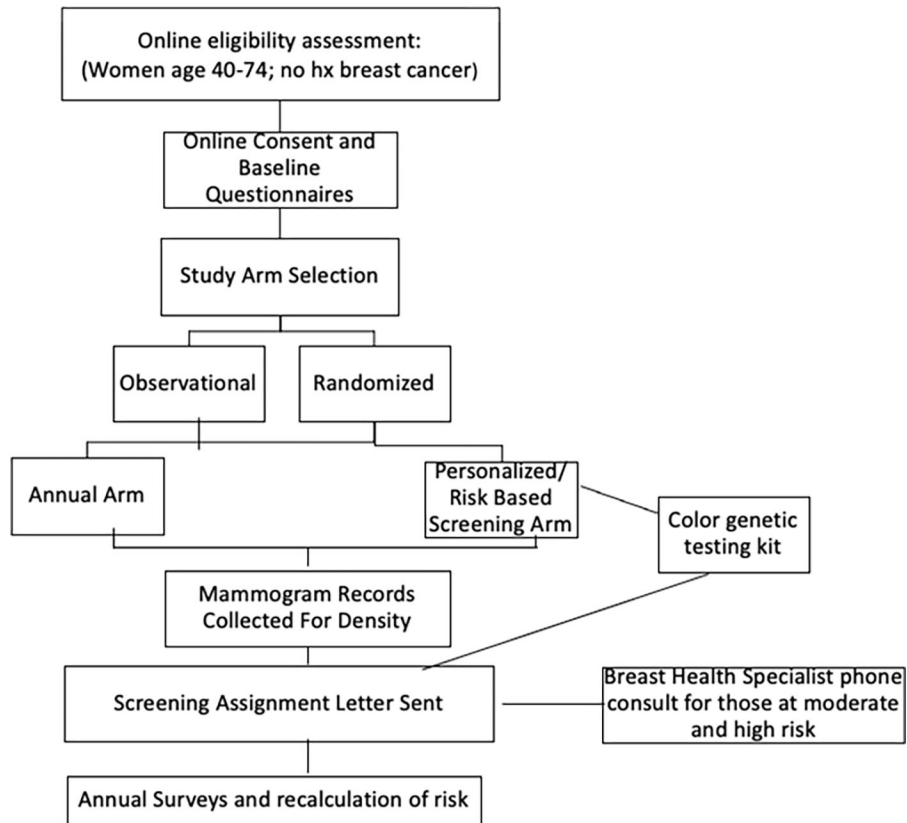
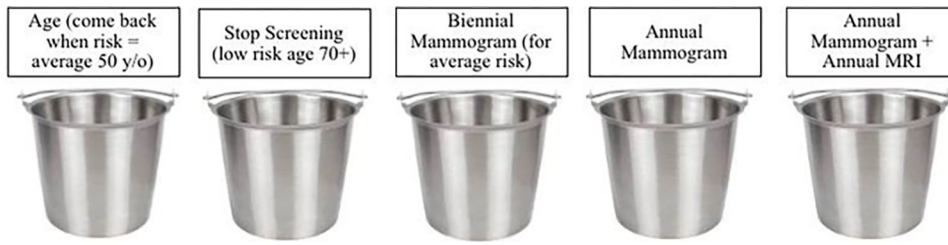


Figure 1.
WISDOM Trial flow diagram.



These screening recommendation categories are referred to colloquially by the WISDOM team as “risk buckets.”

Figure 2.
WISDOM Risk Based Screening Recommendations

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

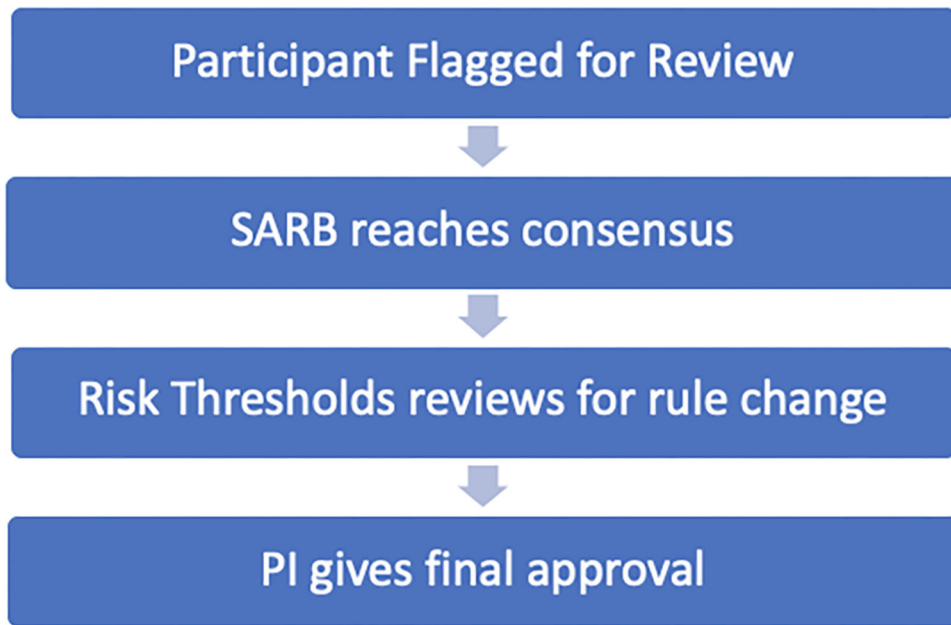


Figure 3.
Screening Assignment Review Board Processes

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 1.

Screening Assignment Review Board Guiding Principles and Case Examples

<p>Guiding Principal: <i>Trust the Model</i></p> <p><u>Definition:</u> In order to test the model effectively, the SARB should default to keeping the screening recommendation generated by the risk algorithm</p> <p><u>Representative Quote:</u> “When we are on the fence, we default to trusting the study algorithm”</p>	<p><u>Case example:</u></p> <ul style="list-style-type: none"> • Age: 57 • WISDOM screening Recommendation: Biennial • Considerations: <ul style="list-style-type: none"> ○ Family History: Mother & 4/5 maternal aunts with BC, some occurring early (before age 40); unknown if family has had genetic testing ○ Other risk factors: Oophorectomy (lowers risk); Hormone Replacement Theory (raises risk) ○ Density: C • Decision: Stay with WISDOM recommendation of Biennial mammogram
<p>Guiding Principal: <i>Women do not have to follow the study’s recommendation</i></p> <p><u>Definition:</u> The SARB is not deciding how the participant <i>must</i> screen; the participant will be free to follow the recommendation or not based on her own or her doctor’s risk assessment.</p> <p><u>Representative Quote:</u> “I don’t think she’ll follow it; I’m assuming she will stick with the specialist”</p>	<p><u>Case example:</u></p> <ul style="list-style-type: none"> • Age: 44 • WISDOM Screening Recommendation: Screen at age 50 • Considerations: <ul style="list-style-type: none"> ○ Family History: Mother diagnosed at 55, Sister diagnosed at 48, Maternal aunt with ovarian cancer at 50 ○ Current Screening: Annual Mammogram and Annual MRI ○ Risk Assessment: IBIS lifetime risk estimate 31.2%; Negative for 9 breast cancer gene variants and low PRS • Decision: Stay with WISDOM screening recommendation to wait until age 50 to have a mammogram, with the assumption that participant will not follow it.
<p>Guiding Principal: <i>Women in every ‘risk bucket’ will develop cancer.</i></p> <p><u>Definition:</u> No screening schedule can prevent cancer. Instead, the WISDOM hypothesis is that fewer women who are identified at low risk will develop cancer than those assessed to be high risk and the risk model is designed to work on a population level.</p> <p><u>Representative Quote:</u> “It’s personalized but it’s still bucket-based... You have to go back to population principles and remember that we’re putting people in buckets; we’re not predicting whether or not they will ever get breast cancer.”</p>	<p><u>Case example:</u></p> <ul style="list-style-type: none"> • Age: 41 • WISDOM Screening Recommendation: Biennial mammogram • Considerations: <ul style="list-style-type: none"> ○ Family history: Mother diagnosed at 63; sister diagnosed at 41; family had genetic testing with negative results. ○ Concern that genetic test is not adequate (with only nine genes) and that the study could be “missing” an unknown genetic risk factor. • Decision: Stay with WISDOM screening recommendation of biennial mammogram