Who Gains When Medicine Becomes More Precise?

Evidence from Genomic Testing in Breast Cancer*

Jasmin Moshfegh

August 2025

First Draft: November 2023

Abstract

I study the equity and efficiency consequences of the entry of a precision medicine test in the context of breast cancer treatment. Using Medicare claims data linked to cancer registry data from the Surveillance, Epidemiology, and End Results Program (SEER), I leverage variation in the timing of a patient's breast cancer surgery relative to the technology adoption date of the provider, in an event-study framework. On average, I find that patients who saw a provider in the year after the provider adopted the technology are more likely to avoid unnecessary treatments. More specifically, patients are less likely to receive chemotherapy at no increased risk in the 5-year cancer mortality rate. However, this welfare-improving impact is concentrated among white patients. Conditional on seeing the same provider, Black patients are 40 percent less likely to be tested. Adjusting for differences in medical characteristics (e.g., cancer stage, hormone receptor status) only explains about half of the within-provider Black-white gap in testing, underscoring the role of provider discretion. Overall, the results highlight the possibility of precision medicine to improve healthcare efficiency at a potential risk of widening health disparities.

*Stanford University. Email: moshfegh@stanford.edu. I am grateful to Laurence C. Baker, Petra Persson, and Maya Rossin-Slater for their advice and guidance. I would also like to thank Dominique Araya, Alex Chan, David Chan, Mark Duggan, Caroline Hoxby, Suhani Jalota, Johan Ljungkvist, Alexa Pohl, Maria Polyakova, Keisha Solomon, Kurt Sweat, and seminar participants at Stanford, IHEA Economics of Genomics and Precision Medicine Seminar, the 2nd TSE Health Economics Conference, the Stockholm School of Economics, the Stanford-Surgery Policy Improvement Research & Education Center (S-SPIRE), and the NBER Trends and Patterns in Health Disparities meeting. This research is supported by an R36 award ('The Causes and Consequences of the Diffusion of Precision Medicine: Evidence from Innovations in Breast Cancer Medicine') and grant T32HS026128 from the Agency for Healthcare Research and Quality. The content is solely the responsibility of the authors and does not necessarily represent the official views of the Agency for Healthcare Research and Quality. This research is also supported by the Sweden-America Foundation and the Stanford Institute for Economic Policy and Research George P. Shultz Dis-

sertation Fund. This study used the linked SEER-Medicare database. The interpretation and reporting of these data are the sole responsibility of the authors. The authors acknowledge the efforts of the National Cancer Institute; Information Management Services (IMS), Inc.; and the Surveillance, Epidemiology, and End Results (SEER) Program tumor registries in the creation of the SEER-Medicare database. The collection of cancer incidence data used in this study was supported by the California Department of Public Health pursuant to California Health and Safety Code Section 103885; Centers for Disease Control and Prevention's (CDC) National Program of Cancer Registries, under cooperative agreement 1NU58DP007156; the National Cancer Institute's Surveillance, Epidemiology and End Results Program under contract HHSN261201800032I awarded to the University of California, San Francisco, contract HHSN261201800015I awarded to the University of Southern California, and contract HHSN261201800009I awarded to the Public Health Institute. The ideas and opinions expressed herein are those of the author(s) and do not necessarily reflect the opinions of the State of California, Department of Public Health, the National Cancer Institute, and the Centers for Disease Control and Prevention or their Contractors and Subcontractors. Additional data for this project were accessed using the Stanford Center for Population Health Sciences Data Core. The PHS Data Core is supported by a National Institutes of Health National Center for Advancing Translational Science Clinical and Translational Science Award (UL1TR003142) and from Internal Stanford funding. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH.

1 Introduction

Many high-stakes decisions are made with limited information. Doctors deciding on the optimal course of treatment, judges deciding whether to let a defendant await trial at home or in jail, and banks deciding whether to extend credit are some salient examples of decisions made with imperfect information (Manski, 2018; Kleinberg et al., 2018; Blattner and Nelson, 2021; Fuster et al., 2022). The availability of new data – and new approaches to leverage that data – can improve decision-making and raise efficiency. Yet the same innovations can widen gaps across demographic groups, depending on who receives the underlying signal and how well that signal works for each group. Whether the explosive growth of medical signals and algorithms generated from new uses of genetic information - often referred to as precision medicine - shifts that balance towards greater efficiency without sacrificing equity remains an open empirical question.

This paper offers one of the first large-scale, quasi-experimental looks at the efficiency–equity trade-off created by a precision-oncology innovation in the context of breast cancer. Each year in the United States, over 200,000 women are diagnosed with breast cancer. For a large subset of these patients, physicians frequently face a complicated decision of whether or not to treat the patient with chemotherapy. While treating a patient with chemotherapy reduces the risk of the cancer recurring, treatment itself is both physically and financially costly.[1] In the early 2000s, a technology was developed to aid in this decision. By assessing the expression of 21 genes present in tumor tissue, and using a proprietary algorithm to predict the risk of the cancer recurring, the test outputs a risk score intended to refine the classic medical classification problem of whom to treat – and whom not to treat.

Between 2004 and 2018, close to 800,000 individuals with breast cancer enrolled in Medicare were treated by oncologists who had adopted the technology at different times, or who had not yet adopted the technology.[2] Using detailed cancer registry data from the Surveillance, Epidemiology, and End Results (SEER) Program linked to Medicare claims data, I leverage within-provider variation in the timing of a patient's breast cancer diagnosis[3] relative to the provider's initial adoption of

[1] According to a 2022 survey of breast cancer patients in the US conducted by Breastcancer.org, 47 percent of respondents reported their breast cancer-related out-of-pocket costs were a significant or catastrophic burden (Breastcancer.org, 2022).

[2] This corresponds to about 160,000 patients in a 20 percent random sample of SEER-Medicare claims.

[3] More specifically, I leverage within-provider variation in the timing of a patient's breast cancer surgery which occurs shortly after diagnosis.

the technology to estimate the impact of increased patient access to the technology in an event-study framework. The findings suggest substantial changes in treatment decisions. On average, patients who saw an oncologist three years after the oncologist adopted the technology are more likely to be tested and approximately 15 percent less likely to receive chemotherapy. If chemotherapy use prior to technology adoption represents overtreatment, or unnecessary use, the reduction should not be associated with an increased risk of long-term mortality or cancer recurrence. Consistent with prior overtreatment, I find no increase in five-year cancer mortality or cancer metastasis. Patients are also less likely to experience common side effects associated with chemotherapy treatments, such as nausea, as measured by a decline in the use of anti-nausea prescriptions.[4]

To quantify the efficiency gains from adoption, I leverage insights from prior literature measuring productivity in healthcare (Abaluck et al., 2016; Abaluck et al., 2020) and Chan et al. (2022), in particular. More specifically, the previous results are viewed through the lens of a classification problem in which providers first observe patient risk types with noise and based on this information classify patients as low-risk or high-risk. The classification in turn determines whether or not a patient receives treatment. Imperfect information results in diagnostic errors or inefficiencies, such as false positives which occur when a low-risk patient is classified as high-risk and receives treatment. Under the assumption that the true distribution of patient risk-types does not change before and after provider technology adoption, the event-study design can be used to gauge the impact of adoption on the true and false negative rates. The estimates show that genomic testing adoption increases the likelihood of a true negative, without a corresponding increase in the likelihood of a false negative, meaning physicians target chemotherapy more accurately – a clear efficiency gain attributable to the technology.

To what extent is the efficiency gain distributed equally across Black and white patients? Over the decade after Medicare coverage began, I find that Black women were significantly less likely than white women to be tested, regardless of income (proxied by Medicaid dual eligibility) or geography—echoing, but also sharpening, mixed findings in the clinical literature (see e.g., Dinan et al., 2015a; Roberts et al., 2016; Peethambaram et al., 2017; Zhang et al., 2020; Chen et al., 2023). I next explore the drivers of the Black-white gap in testing. First, disparities in testing may reflect

[4] I focus on one of the most common classes of drugs used in conjunction with chemotherapy - anti-nausea drugs, although there are many other treatments that cancer patients may undergo to treat or avoid the side effects of chemotherapy.

allocation or *access* - Black patients may be less likely to see adopting providers. I find that Black patients are 40 percent less likely to be tested, conditional on seeing the *same* oncologist. Second, disparities in testing may reflect differences in medical appropriateness for the technology (e.g., due to differences in cancer characteristics such as staging).[5] If by contrast, a within-provider gap exists for two medically identical patients, this could be suggestive of racial differences in provider testing behavior. Using the detailed cancer registry data from SEER, I find that controlling for measures of medical appropriateness - and more specifically, *objectively* defined criteria for testing stated in major clinical guidelines, including but not limited to cancer staging, hormone status, and lymph node spread - accounts for only about 50 percent of the within-provider racial gap in testing - underscoring the role of provider discretion.

Overall, these results highlight how the extent to which different population groups eventually gain from the arrival of a new technology depends on various distinct margins. Conceptually, a Black–white gap in test uptake can arise on four sequential margins – when the technology 1) is more appropriate in one group because it targets a disease more prevalent in one group ("differential appropriateness"), 2) conditional on appropriateness is accessed to a greater extent by one group ("differential access or allocation"), 3) conditional on appropriateness and access, is used more intensively in one group ("differential testing or discrimination"), and conditional on all of the above, 4) is of greater value in one group ("differential treatment effects"). When this is true, and groups are defined along the lines of socioeconomic status or race, the arrival of a new technology gives rise to, or worsens, an existing health disparity. A back-of-the-envelope calculation in the setting of my study highlights the role of differential appropriateness and testing as driving the overall gap in testing, as opposed to differences in access.

This paper makes four key contributions. First, this paper builds on a growing literature on the impacts and value of new sources - and uses of - information on decision-making, with a particular focus on genomic information.[6] The findings of this paper suggest the potential for large efficiency gains from the use of genomic algorithms to guide physician decision-making, specifically driven by a reduction in *overtreatment* as a result of baseline (human) diagnostic inaccuracy.

[5]This would be true if Black patients generally present with more advanced stages of cancer making chemotherapy a necessity, and hence a test that guides in the decision to forgo it, irrelevant.

[6]Some recent examples in the space of genetic/genomic information include Berndt et al. (2019) and Conner et al. (2025), in the space of AI and machine learning include Kleinberg et al. (2018), Mullainathan and Obermeyer (2022), and Agarwal et al. (2023), and in the space of algorithms include Boswell Dean et al. (2024) and Raymond (2023).

Second, this paper contributes to the literature on the causes of racial disparities in health, and in particular the role of within- versus between-provider variation (Chandra et al., 2022). With respect to within-variation, a growing body of literature investigates racial disparities for patients seeing the *same* provider (see e.g., Chandra and Staiger, 2010; Corredor-Waldron et al., 2024). One of the main challenges in establishing disparities in healthcare among patients who see the same provider is the presence of unobserved (to the econometrician) differences in (medical) patient characteristics that may render differences in care appropriate. One of the strengths of this study is the use of rich cancer data and objective measures of appropriate care (as stipulated by well-known guidelines) in measuring disparities. This in turn, allows for a more accurate measurement of the role of provider discretion in explaining disparities in care.

Third, this paper contributes to a literature on the impacts of new medical treatments on mortality and productivity (see e.g., Cutler and Huckman, 2003; Cutler, 2004; Chandra and Skinner, 2012, and Cutler (2008) and Lichtenberg (2014) for examples in the context of cancer)[7] and its potential distributional effects (see e.g., Lleras-Muney and Lichtenberg, 2005; Cutler et al., 2006; Glied and Lleras-Muney, 2008; Aizer and Stroud, 2010; Cutler et al., 2012; Jayachandran et al., 2010; Alsan et al., 2021; Bögl et al., 2024; Fadlon et al., 2024; Callison et al., 2025). Building on Cutler et al. (2012) in particular, this paper highlights the distinct pathways by which a disparity may arise, some of which are unique to information technologies, and gauges the role of each pathway.

Fourth, this paper provides evidence on the performance of the 21-gene recurrence in the real-world setting, using quasi-experimental variation in access to testing to quantify the (heterogenous) impact of the test on treatments. As such, it complements the evidence from clinical trials, which tend to carry the highest form of internal validity but which often have limited external validity given selection into participation, with minorities and Black patients often underrepresented (Unger et al., 2016; Nazha et al., 2019; Alsan et al., 2022; Alsan et al., 2023).[8] The findings from this study also complement medical studies documenting trends in the use of testing, and the association between testing and chemotherapy use in the real-world setting.[9] In this regard, the contribution of

[7] See Daysal et al. (2024) for a recent example in the context of breast cancer.

[8] In particular, the largest randomized clinical trial of the 21-gene recurrence score, the Trial Assigning Individualized Options for Treatment (TAILORx), randomized over 10,200 women with early-stage breast cancer across more than 1,000 sites in the United States, Australia, Canada, Ireland, New Zealand, and Peru (National Cancer Institute, 2018). Results from the trial suggested that standard of care would lead to overtreating 25 percent of patients with a low risk of recurrence, while undertreating 43 percent of patients with high risk (Sparano et al., 2015; Sparano et al., 2018).

[9] More specifically, a large body of prior medical literature has shown that test use is associated with an overall

this study is to provide causal estimates of the effects of increased access to the 21-gene recurrence score on chemotherapy use, complementary treatments, and metastasis and mortality in a large nationally representative population.

This paper is organized as follows. Section 2 provides background on breast cancer treatments and genomic testing. Section 3 describes the data and provides descriptive statistics. Section 4 describes the empirical strategy and the impacts of increased access to genomic testing. Section 5 describes the efficiency implications of technology diffusion. Section 6 describes the implications of technology diffusion on racial disparities, and Section 7 concludes.

2 Setting

Breast cancer is the second most common cancer in women in the United states, representing about 1 in 3 of all new female cancers each year (American Cancer Society, 2023a). It is also the second leading cause of cancer death in women. According to the American Cancer Society's 2023 estimates, there will be 297,790 new cases of invasive breast cancer diagnosed in women, of which 43,700 will result in death. On the upside, the probability of surviving breast cancer has increased considerably in the past decades. Estimates from the Surveillance, Epidemiology, and End Results (SEER) database for cancers diagnosed between 2012-2018 suggest an overall 5-year relative survival rate of 91 percent, ranging from 30 percent to 99 percent depending on the spread of the cancer (American Cancer Society, 2023b). Breast cancer is a disease affecting women across the age distribution with a median age at diagnosis of 62 years.

2.1 Breast cancer treatment

Breast cancer is often discovered after an individual experiences symptoms (e.g. discovering a lump in the breast), or through an abnormal finding during a screening mammogram. To confirm that the abnormal finding is a breast cancer, diagnosis is made following a (surgical) biopsy. If a patient is found to have breast cancer, several tests are often performed to determine the type of the cancer.

reduction in chemotherapy use, although results have been mixed, partly reflecting different patient populations and study periods (Schaafsma et al., 2021; Dinan et al., 2015b; Su et al., 2016; Carlson and Roth, 2013; Ademuyiwa et al., 2011; Losk et al., 2021; Choi et al., 2020; Eiermann et al., 2013; Geffen et al., 2011; Paik et al., 2006; Bhutiani et al., 2018; Marchionni et al., 2007). At least in the shorter-run, no evidence of increased mortality has been documented (Roberts et al., 2017; Petkov et al., 2016). Most relevant to this study is Dinan et al. (2020) which uses an approach to address selection bias.

There are different types of breast cancer, were "type" refers to the protein/hormone and gene expression of the cancer cell. For instance, many cancers are characterized by cancer cells that have hormone receptors (also known as hormone receptor positive cancers) and some cancers have the HER2 gene (also known as HER2+ cancers). The treatment a patient receives, in turn, if any, depends on the type of the breast cancer.

Most women with breast cancer will have some type of surgery to remove the tumor. This may then be followed by treatment with "systemic therapies" or drugs. Common options are chemotherapy, hormone therapy, targeted drug therapy, or immunotherapy. The overall treatment decision is generally guided by the type of the cancer and the stage of the cancer. So for instance, a patient with breast cancer and cells that have hormone receptors, would be treated with hormone therapy, possibly in addition to other therapies, and a patient with HER2+ cancer would be treated with targeted drug therapies such as Trastuzumab (brand name: Herceptin).

Different systemic therapies carry different risks and benefits. While potentially very effective in treating cancer, chemotherapy is notorious for its dreaded side effects, which include hair loss, nausea, fatigue, as well as potential long-term damages to the heart and nerves. To avoid or treat these side effects, patients may receive other treatments and drugs, such as anti-emetic drugs to combat nausea. By contrast, hormone therapy does not carry the same side effects profile.

2.2 Genomic testing

In a subset of breast cancer patients, genomic testing and the 21-gene recurrence score assay, can help determine which patients would benefit from chemotherapy treatment, and which patients could safely forgo it. In lay terms, this is the subset of patients who respond to hormone therapy, have HER2 negative cancer, and who have early-stage invasive breast cancers.[10] In this set of patients, hormone therapy may suffice.

The test was developed in 2004 by a private commercial lab located in Redwood City, California.[11] It assesses the expression of 21 genes in a tumor sample that is removed from the patient

[10] In clinical terms, today, these are patients with ER/PR positive cancers, HER2- cancers, stages I-IIIa. Upon the inclusion of the technology in the NCCN guideline in 2008, the test was strongly recommended for node-negative patients with hormone receptor-positive (i.e. patients with cancer cells that have hormone receptors and who therefore may respond to hormonal therapies), HER2-negative tumors (i.e. patients with cancers cells that have HER2 receptors and may therefore respond to HER2 targeted treatments like Trastuzumab) that are 0.6-1.0 cm and moderately/poorly differentiated or with unfavorable features, or >1cm.

[11] Although several other tests are available on the market today, the use of other similar tests is negligible in the Medicare covered population.

during surgery or biopsy. The expression of these genes, 16 which are cancer-related and 5 which are reference genes[12], are informative of 1) the likelihood of cancer recurrence, and 2) the likelihood that the patient will respond to chemotherapy. The test uses this information and an algorithm to produce a single score, ranging from 0 to 100, also known as the "recurrence score". The score reflects the risk of the breast cancer coming back (recurring) in the distant future if the patient is treated with hormone therapy alone and how likely the patient is to benefit from getting chemo after surgery. Individuals with low scores, can "safely" forgo chemotherapy, while individuals with high scores are likely to benefit from chemotherapy.

The typical treatment pathway may look like the following. 1) providers/patients discover a lump, 2) a biopsy is made for initial diagnosis and to determine the characteristics of the cancer (e.g. HER2-, hormone+, early stage), 3) a surgery is performed to remove the tumor and to further diagnose the patient, and 4) a decision is made about the course of treatment and in particular, if chemotherapy should be used in addition to hormone therapy:

Simplified pathway to chemotherapy

1. Abnormal finding
2. (Surgical) biopsy for initial diagnosis
3. Surgery to remove tumor and for further diagnosis
4. Begin hormone therapy and potentially chemotherapy to treat the cancer and reduce the risk of recurrence

Genomic testing typically occurs on tumor tissue obtained from the surgery, or in some cases after the (surgical) biopsy. The major clinical guidelines for oncology in the US, the American Society of Clinical Oncology (ASCO) and National Comprehensive Cancer Network (NCCN) guidelines, began recommending the use of the technology in 2007 (Harris et al., 2007), and January 2008, respectively.[13] The guidelines include score cutoffs to determine patients at low, intermediate, and

[12]The reference genes are: Beta-actin, GAPDH, RPLPO, GUS, and TFRC. The cancer-related genes are HER2-related(GRB7 and HER2), Estrogen-related (ER, PGR, Bcl2, SCUBE2), Invasion-related (Stromelysin 3, Cathepsin L2), and Proliferation-related (Ki-67, STK15, Survivin, Cyclin B1, MYBL2).

[13]How were providers and patients making chemotherapy decisions prior to the inclusion of the technology in the guidelines? From the 2007 version of the NCCN guidelines: "The decision to add chemotherapy to hormonal therapy should be individualized, especially in those with a favorable prognosis and in women age 60 y where the incremental benefit of chemotherapy may be smaller. Available data suggest sequential or concurrent hormonal therapy with RT

high risk of recurrence. These cutoffs have changed over time as the evidence base has evolved. In the first release of the NCCN guidelines in 2008, patients were deemed to be "low risk" if they had scores below 18.[14]

To test a patient, a provider orders a specimen kit from the manufacturing laboratory, and submits a sample obtained during surgery or biopsy to the lab. The lab performs the analysis using the technology, and submits a report to the physician within two weeks.

2.3 Medicare coverage policies relevant to cancer

Medicare is the largest US public health insurance program covering individuals aged 65 and older, those who are disabled, and those with end-stage renal disease. Different parts of Medicare cover different types of healthcare services. Medicare Part A primarily covers inpatient hospital stays, Part B primarily covers physician services, outpatient care, as well as medical devices, and Part D covers prescription drugs. Part C refers to the Medicare Advantage (MA) program, representing private health plans that Medicare has contracted with. Most elderly patients in the US are enrolled in "traditional" Medicare; typically these are individuals with Part A and Part B coverage, and to an increasing extent Part D coverage since 2006, although MA enrollment has increased rapidly in the last decade.[15]

Medicare started covering the technology in 2006 under the Part B schedule.[16] Although most Part B services are subject to a deductible and 20 percent coinsurance, provided on having a breast cancer diagnosis, a tested Medicare patient faces no co-insurance, i.e. the test comes at no cost to the patient. Medicare generally reimburses the lab for the cost of the test under the Clinical Laboratory Fee Schedule (CLFS). In 2018, this rate was set at $\$3,873$. In some instances, and in particular, when the test is performed as part of an inpatient stay our as part of certain outpatient

is acceptable."

[14]Moreover, "intermediate" was defined as a score between 18 and 30, and "high" was defined as a score greater than or equal to 31. By 2020, the cutoffs were contingent on the node status of the cancer. The cutoffs were the same for node positive patients. For node negative patients they were 26, 26-30 and greater to or equal to 31. The greatest change in the recommended use of the test occurred after the release of new findings from the ongoing clinical trial associated with the technology in 2022. As a result, the intermediate risk-group was removed for several patient groups, including postmenopausal women, and the cutoffs were set at, less than 26 or greater than 26.

[15]More specifically, MA enrollment increased from roughly 16 percent of all Medicare beneficiaries in 2006 to 34 percent in 2018 (Kaiser Family Foundation, 2019).

[16]In practice, NHIC - a Medicare contractor - issued a Local Coverage Determination (LCD) in 2006, but because the laboratory manufacturing the test (at that point, Genomic Health) is located at a single site (Redwood city, California), the LCD effectively became a national policy.

stays, the laboratory receives reimbursement from the hospital, and the hospital bills Medicare. Physicians are required to order the test, but face no direct cost and receive no reimbursement for doing so.

Chemotherapy drugs are primarily covered under the Part B Schedule,[17] which means patients face a 20 percent coinsurance. Given that there is no maximum out-of-pocket (OOP) for services on the Part B schedule, Medicare patients can be particularly vulnerable to the costs of cancer treatments, potentially facing OOP costs that are between 23.7 percent of household income (Narang and Nicholas, 2017). Most other self-administered drugs are covered under Part D. This includes most hormone therapy drugs. Anti-emetic medications taken to combat nausea during chemotherapy treatment are covered by both Part B and Part D, depending on the timing of the prescription in relation to the chemotherapy treatment. For drugs that are on the Part B schedule, providers purchase them from wholesalers and get reimbursement by Medicare.

3 Data and Descriptive Statistics

The primary source of data is a 100 percent sample of administrative Medicare claims data linked to the Surveillance, Epidemiology and End Results (SEER) program for the years 2003-2019 for all breast cancer cases. The SEER program is a program that the National Cancer Institute (NCI) uses to support cancer surveillance activities. Established in 1973, the SEER program collects data from population-based cancer registries. The number of registries participating in the program has grown over time; in 1973 the program covered less than 10 percent of the US population, and by 2022 the program consisted of 18 registries representing 48 percent of the US population. The SEER data file includes detailed information about the cancer. More specifically, this data file includes one record per tumor diagnosed, and each record includes information such as the stage, histology, and site of the cancer.

The Medicare claims data is organized into different files. For each individual, I observe at the daily level, the diagnosis and treatments individuals receive, and the doctors who treat them. The analysis relies on the following Medicare files: 1) the master beneficiary summary File (MBSF) which tracks beneficiary characteristics such as the age of the beneficiary and the zip code of residence,

[17] Some chemotherapy drugs are covered under Part D. However, the use of these drugs is negligible in the setting of this study.

2) the carrier file which contains claims submitted by non-institutional providers such as physicians and clinical laboratories, 3) the outpatient file which contains claims submitted by hospitals for non-inpatient care, 4) the MEDPAR file which contains information about inpatient stays, and 5) the Part D event file which contains information about prescription drugs.

The SEER-Medicare linked data does not include granular and unmasked data on where individuals live. Therefore, for some of the descriptive analysis on technology diffusion in the appendix, a 20 percent sample of (unlinked) Medicare claims data is used, which includes information on the zip code of the residence of Medicare beneficiaries. This data spans the years 2007-2019.

3.1 Sample construction

Surgery sample. To identify the sample of patients potentially eligible for the technology, the carrier and outpatient files are used to identify all women with a claim for a breast cancer surgical procedure between Jan 1, 2004 and Dec 31, 2018. Next, the following set of exclusions are made. First, individuals are excluded if they were enrolled in HMO plans, or were not continuously enrolled in Medicare Parts A and B in the 12 months prior to, and 12 months, proceeding the surgery (unless they died). This restriction is made to be able to follow treatment decisions in the year following surgery and to measure comorbidities in the year before surgery. Second, individuals without a primary breast cancer diagnosis according to SEER are excluded. Third, men with breast cancer are excluded, as well as individuals with missing information on race, and an age at surgery below 40 or exceeding 80. This restriction is made to limit the analysis to patients who are not too old to be considered for the technology/chemotherapy. Fourth, individuals without a lobular and ductal breast cancer histology are excluded as well as individuals with non-incident cancers, as the testing and chemotherapy decision facing a patient with a recurrent cancer is different from that of a newly diagnosed patient. This sample of N=163,613 is the breast cancer surgery population. This population is a superset of the "true" at-risk population. That is, some patients in the sample are not necessarily eligible for the technology, in particular patients with advanced stages of cancer.

An alternative surgery sample is constructed using the unlinked Medicare claims data, to characterize technology diffusion by various measures of socioeconomic status. The sample construction follows a similar process as for the main surgery sample, excluding steps that involve data that is not included in the unlinked Medicare data. To identify the sample of patients potentially

eligible for the technology, the carrier and outpatient files are used to identify all individuals with a claim for a breast cancer surgical procedure between Jan 1, 2007 and Dec 31, 2019. Next, the following set of exclusions are made. First, individuals are excluded if they were enrolled in HMO plans, or were not continuously enrolled in Medicare Parts A and B in the 12 months prior to, and 12 months, proceeding the surgery (unless they died). Second, men with breast cancer are excluded, as well as individuals with missing information on race, and an age at surgery below 40 or exceeding 80.[18] Third, individuals with non-incident cancers are excluded. This sample of N=95,740 is the breast cancer surgery population identified in the unlinked Medicare claims data.

Event-study sample. Surgical patients are matched to oncologists by identifying any oncologist[19] seen by the patient for an office visit within [-3,3] months of the breast cancer surgery. The sample is also restricted to patients who saw providers with at least ten patients overall (this corresponds to the average number of patients per provider), excluding the index patient (i.e., the first patient the provider tests). A flow chart of the sample construction process is found in Figure A.1. The final sample for the event study analysis consists of N=90,389 patients and N= 3,481 providers.

3.2 Measurement

Treatments and outcomes. For each patient in the analysis sample, the Medicare files are used to determine if the patient was tested, the treatments the patient received, and all-cause mortality. The SEER file is used to determine cancer-specific, as well as breast-cancer specific mortality. All outcomes are measured at various time points relative to surgery. Testing is identified using the Carrier file.[20] Since the Carrier file includes files submitted by commercial laboratories, all patients

[18]This study includes Medicare beneficiaries under age 65, who qualify for coverage due to disability. While this group is not fully representative of the general Medicare population (see e.g., Enewold et al., 2020), their inclusion allows for a more complete view of patients receiving care within the program and increases statistical power. This is noted as a consideration when interpreting the generalizability of the findings.

[19]"Oncologists" are identified as physicians with a CMS speciality code in hematology/oncology (83), medical oncology (90), or surgical oncology (91). Radiation oncology (92) and gynecologist/oncologist (98) are not included as these providers are unlikely to order the test. Providers from other specialties may use the test as well, but the analysis is restricted to oncologists since this subset is the most relevant provider population to adopt the technology

[20]The early arrival of the 21-gene assay allows for an analysis of both the short-term and long-term implications of technological progress. This makes this technology a particularly important one to study and it also makes the study sufficiently powered. Moreover, coding for laboratory tests predates the entry of precision medicine. More specifically, laboratory tests were historically billed using CPT codes that reflected the individual steps or components of a molecular service, such as "DNA extraction" (also known as "stacking codes"). This made it difficult for payers to know what tests they were paying for. Amid increasing pressure to reform the system of coding for molecular pathology tests - the AMA implemented several changes, of which the major reform was the replacement of the process-oriented

who were tested can be identified by finding all claims submitted by Genomic Health (CLIA lab number 05D1018272) to Medicare.[21] It is also possible that a patient is tested as part of an inpatient stay or a hospital outpatient visit. In the latter case, the test is generally bundled with other services and falls under the Hospital Outpatient Prospective Payment System. The hospital bills Medicare, and the laboratory collects payment from the hospital. Importantly, those tests would not be identifiable in the data. However, according to the company's annual report from 2011, such patients represent less than 1 percent of the company's total breast cancer testing population.[22] Nevertheless, measurement error in test receipt may lead to an under count of the true testing rate. To the extent that this is true, this would lead to an attenuation bias as some of the patients in the control group may be treated. Moreover, because the identification strategy leverages within-provider variation, this concern is limited unless providers use different billing rules for different patients.

The main treatment of interest is chemotherapy, as this is the treatment directly impacted by the use of the test. I also consider if there are spillover effects to other common cancer treatments, such as hormone therapy and radiation therapy. Since chemotherapy is often associated with a range of side effects, I also consider treatments for the management of nausea, which is a common side effect of chemotherapy, and the use of immunostimulants, which are primarily used to reduce the risk of infection due to a low white blood cell count in patients treated with chemotherapy. To identify the treatments patients receive, the carrier, outpatient, MEDPAR, and Part D files are used. See Appendix Table A.1 for a list of codes used.

The main measure of patient health is mortality. All-cause mortality is measured using the master beneficiary summary file. The SEER file is used to measure cancer-specific as well as breast-cancer specific mortality. As a proxy for the spread of cancer, the carrier, outpatient, and MEDPAR files are used to identify claims with diagnosis codes indicating a secondary cancer or metastasis. I refer to this as a *proxy* for recurrence, as coding for metastasis may be measured with noise.

Patient characteristics. The NCI comorbidity index Klabunde et al., 2007) and the Charl-

CPT codes to gene-specifc CPT codes starting in 2012. As such it is often not possible to identify the receipt of a particular genomic/genetic test in most sources of administrative claims data in the United States. By contrast the 21-gene recurrence assay is identifiable in claims data over time. The reason for this is that the 21-recurrence assay was manufactured by a single lab - identifiable in the claims data - and for which the 21-gene assay was its main product.

[21] Later in the sample period Genomic Health also launched a separate test for prostate cancer. To avoid including these tests, an individual's testing date is defined as the first date at which a patient had a claim submitted by Genomic Health for a covered breast cancer diagnosis code.

[22] Source: https://www.sec.gov/Archives/edgar/data/1131324/000095012311024832/f58029e10vk.htm

son Comorbidity index (Charlson et al., 1987; Deyo et al., 1992; Romano et al., 1993; Quan et al., 2005) is measured in the year prior to the surgery using the carrier, outpatient, and inpatient files. The NCI comorbidity index builds on the Charlson comorbidity index, but is specifically adapted to be suitable for cancer cases. The master summary beneficiary file is used to identify a patient's age, race, original reason for Medicare enrollment, and more specifically whether or not the patient was originally eligible due to disability or old-age, where the former represents a much more vulnerable population, and a patient's dual status or enrollment in Medicaid in addition to Medicare. For each individual, the metropolitan status of an individual's county of residence is determined according to the Rural-Urban Continuum Codes in 2013, which define metropolitan counties as counties in metro areas.[23]

To characterize the socioeconomic environment, the National Neighborhood Data Archive (NaNDA) is used (Melendez et al., 2020). The NaNDA contains data on zip code level neighborhood deprivation and affluence measures, which are indices based on zip-code level statistics from the American Community Surveys.[24] More specifically, the neighborhood socioeconomic disadvantage is originally constructed using information from the American Community Survey (2008-2012) as an average of five census indicators (proportion non-Hispanic Black, proportion of female headed families with children, proportion of households with public assistance income or food stamps; proportion of families with income below the federal poverty level; proportion of population age 16+ unemployed) and it ranges from 0 to 1.0. The neighborhood affluence is an average of three census indicators (proportion of households with income greater than $75K, proportion of population age 16+ employed in professional or managerial occupations; proportion of adults with Bachelor's Degree or higher) ranging from 0 to 1.0. Using this data, patient zip codes are characterized according to their 2008-2012 1) neighborhood disadvantage index, 2) neighborhood affluence index, 3) proportion of people with income past 12 months below poverty level, and 4) proportion of households with public assistance income. Zip codes are then classified into deciles of each of the

[23]More specifically, metropolitan counties are defined as counties in 1) Counties in metro areas of 1 million population or more, 2) Counties in metro areas of 250,000 to 1 million population, or 3) Counties in metro areas of fewer than 250,000 population. Non-metropolitan counties are counties with 4) Urban population of 20,000 or more, adjacent to a metro area, 5) Urban population of 20,000 or more, not adjacent to a metro area, 6) Urban population of 2,500 to 19,999, adjacent to a metro area, 7) Urban population of 2,500 to 19,999, not adjacent to a metro area, 8) Completely rural or less than 2,500 urban population, adjacent to a metro area, or 9) Completely rural or less than 2,500 urban population, not adjacent to a metro area according to the 2013 Rural-Urban Continuum Codes (RUCC).

[24]The NaNDA contains measures at the ZCTA level. A crosswalk from the USDS mapper is used to obtain the measures at the 5-digit zip level. Source: https://udsmapper.org/zip-code-to-zcta-crosswalk/

above measures.

Provider technology adoption. A provider's adoption date is defined as the first date the provider used the technology for any breast cancer patient in the Medicare population. Since providers may have used the technology in a non-Medicare patient, there will be measurement error in the adoption date, and more specifically, the adoption date is likely to be biased upwards for all providers.

3.3 Descriptive statistics

This section presents aggregate time series of technology diffusion in the at-risk population between 2007-2019 and summary statistics.

3.3.1 Technology diffusion

Figure 1, panel (a), shows the share of patients tested within three months of their breast cancer surgery over time in the surgery sample. From the figure it can be seen that technology adoption increased from roughly 0 percent in 2004 to close to 40 percent in 2018.[25] Over this time period, the share of patients treated with chemotherapy within one year of surgery also declined, potentially reflecting the impact of the technology. Over this time-period, five year mortality due to any cause, cancer, and breast cancer, specifically, have declined. Overall, these descriptive patterns illustrate that the technology had a big footprint - almost four in ten breast cancer patients in the sample are tested by 2018.

To what extent is adoption evenly distributed across space, and in particular across more vs less disadvantaged areas? Figure A.2 suggests an uneven distribution. This figure shows the predicted share of patients tested by different measures of zip code level socioeconomic characteristics.[26] For instance, individuals living in zip codes in the bottom decile of the neighborhood affluence index have an adoption rate of roughly 20 percent while individuals living in the top decile have an adoption rate of roughy 25 percent, representing a difference of almost 25 percent in uptake.

To what extent was technology diffusion even across different patient groups, and race in particular? Figure A.3 illustrates large differences in the share of patients getting tested by race

[25] Not all patients in the sample will be indicated for the test, hence it is not possible to infer under-adoption.

[26] More specifically, the predictions are based off of regressions of testing on the decile of a given measure of zip-code level SES, controlling for age at surgery indicators, comorbidity indicators, and HER2 status.

and a proxy for patient income, namely, enrollment in Medicaid. In 2018, ca 35 percent of white patients were getting tested, while the corresponding share for Black patients was 27 percent (Figure A.3, panel a). Compared to low-income Medicare patients, i.e., patients who are also eligible for Medicaid, relatively high-income Medicare patients are more likely to be tested (Figure A.3, panel b). By contrast, diffusion was evenly distributed across metropolitan and non-metropolitan counties (Figure A.4, panel b).

To gauge to what extent the racial gap in testing is driven by differences in income, Figure A.3 panels (c) and (d) condition on a patient's Medicaid status. From these panels it can be seen that the racial gap exists independent of Medicaid status, although the descriptives are noisy. Overall, Figure A.3 suggests an uneven diffusion of the technology by a patient's race and income, independently of each other. To gauge the extent to which this pattern is a function of where individuals live, and in light of the gradients observed in Figure A.2, Figure A.5 depicts these gradients by race. From the figure it can be seen that Black patients are less likely to be tested than white patients, regardless of where they live. This suggests the racial gap in testing is not solely explained by differences in where people live.

3.3.2 Sample statistics

The empirical strategy leverages variation in when a patient undergoes breast cancer surgery relative to when the provider adopts the technology. This design hinges on sufficient variation in the timing of adoption by providers. Figure A.6 illustrates considerable variation in the timing of adoption. It increases rapidly following Medicare's decision to cover the technology in 2006 and reaches a peak in 2008, coinciding with the release of a major clinical guideline recommending the use of the technology in clinical practice. Summary statistics of the surgery sample and event-study sample are displayed in Table 1, columns (1) and (2), respectively. In the event-study sample (column 2), the median age at surgery is approximately 71 years, the majority of patients are white (87 percent), and 16 percent have Medicaid coverage. Almost 18 percent were tested and 29 percent were treated with chemotherapy post surgery.

4 Impacts of Genomic Testing Adoption

In this section, I show how an oncologist's decision to adopt genomic testing impacts a breast cancer patient's probability of undergoing testing, receiving certain treatments, and intermediate-run mortality. To do so, I leverage variation in the timing of technology adoption across providers relative to when a patient undergoes breast cancer surgery similar to previous literature looking at the effect of physician guideline awareness (Abaluck et al., 2020). More specifically, I estimate event-study models that compare outcomes for patients who have their breast cancer surgery after their oncologist adopted the technology to patients who had it before, relative to the overall time trend of the outcome which is informed by all breast-cancer treating oncologists, including those who have not adopted the technology by the end of the time period:

$$Y_i = \sum_{r=-3, r\neq -1}^{r=5} \mathbb{1}(r(i) = r)\beta_r + \gamma_{d(i)} + \eta_{t(i)} + \delta X' + \epsilon_i \tag{1}$$

where $Y_i \in \{0, 1\}$ is the outcome of interest for patient i (e.g., chemotherapy treatment following surgery), $\mathbb{1}(\mathrm{r(i)} = \mathrm{r})$ is a function that returns the year of patient $i's$ surgery relative to the technology adoption date of the oncologist (i.e., a patient's surgery year − a provider's adoption year). $\gamma_{d(i)}$ are provider fixed effects and $\eta_{t(i)}$ are year of surgery fixed effects. X' includes a range of patient controls including a patient's disability status, NCI comorbidity tertiles, patient age (binned in 5-year intervals), and various measures of cancer characteristics, including indicators for the tumor being well-differentiated, neoadjuvant chemotherapy use, localized cancer, HER2 positive treatment use, and the cancer being classified as metastatic within 90 days of surgery. These patient characteristics may be correlated with the likelihood that a patient receives chemotherapy. [27] Controls also include indicators for the health service areas associated with the zip code of patient residence, as well as an indicator for the metropolitan status of the county, to account for geographic differences in treatment intensities.

The main coefficient of interest is β_r which represents the marginal (percentage point) change in Y_i for a patient that had her surgery r years after her oncologist had adopted the technology, relative to another patient who had her surgery one year before the same oncologist adopted the

[27] More specifically, the older the patient, and the greater the number of comorbidities, the less likely is it that the patient will benefit from being treated with chemotherapy. Conversely, less localized, poorly differentiated, HER2 positive cancers are more likely to benefit from chemotherapy.

technology, adjusting for the underlying time trend in the outcome. Given that breast cancer surgery is measured with noise, and to avoid undercounting testing and treatment among individuals, testing is measured within 3 months of surgery, although the vast majority of individuals who are tested have a testing claim within the first of month of surgery. Similarly, chemotherapy is measured within one year of surgery.

I also estimate the associated difference-in-differences model, to obtain a summary measure of the impact of technology adoption (β):

$$Y_i = \gamma_{d(i)} + \eta_{t(i)} + \beta \text{Adopted}_{d(i)} + \delta X' + \varepsilon_i \tag{2}$$

where again $\gamma_{d(i)}$ are provider fixed effects and $\eta_{t(i)}$ are year of surgery fixed effects, and $\text{Adopted}_{d(i)}$ is an indicator equal to 1 if the oncologist that the patient saw in relation to the breast cancer surgery had adopted the technology prior to that point.

Two key assumptions are critical for a causal interpretation in this setting. The first is that the timing of adoption is uncorrelated with time-varying determinants of patient outcomes, and in particular the likelihood of receiving chemotherapy. The second is that technology adoption did not coincide with the adoption of other de-escalation or chemo-reducing treatment styles or technologies.

To gauge the validity of the first assumption, I predict the likelihood that a patient receives chemotherapy within one year of their surgery, based on a patient's age, NCI comorbidity tertile, a patient's disability status, an indicator for living in a metropolitan area, and various measures of cancer characteristics, including indicators for the tumor being well-differentiated, neoadjuvant chemotherapy use, localized cancer, HER2 positive treatment use, and the cancer being classified as metastatic within 90 days of surgery. These predicted values are plotted against the timing of provider adoption in Figure A.13. This figure illustrates a smoothly evolving predicted probability of receiving chemotherapy in support of the identifying assumption, with some caveats. More specifically, although most of these characteristics are statistically significant, there are likely many other factors that explain the likelihood of a chemotherapy receipt. More discussion is provided in the robustness section. Moreover, the absence of pre-trends in the proceeding event studies (e.g., Figure 2) provides further support for the identifying assumption.

4.1 Chemotherapy

Figure 2 presents results from estimating the event study models in Equation (1). The figure suggests that patients who had breast cancer surgery three years after the oncologist had adopted the technology were almost 20 percentage points more likely to be tested than a patient who saw the same oncologist a year prior to technology adoption (panel a).[28] Once, a provider adopts the technology, there is more or less a level-shift in the use of the technology, suggesting awareness of the technology rather than experience and learning are governing the adoption process. Moreover, patients are almost 5 percentage points less likely to be treated with chemotherapy within a year of the surgery (panel b). This represents an almost 15 percent (5/35) decline in chemotherapy use following technology adoption. The lack of pre-trends, and the clear level-shift in chemotherapy use at the time point of technology adoption, supports the notion that the change in chemotherapy use is a function of technology adoption. Figure A.7 documents that the overall reduction in chemotherapy use is driven by a reduction in the use of specific chemotherapy drugs. In particular, the figure suggests a shift away from more (cardio)toxic chemotherapy drugs like Doxorubicin (an Anthracycline) to less (cardio)toxic chemotherapy drugs like Docetaxel, and a marked shift away from Anthracyclines in general (Figure A.8, panel e).

4.2 Prevention and treatment of side effects

I next gauge the possible health spillover effects of technology adoption. Since chemotherapy is a physically costly treatment to undergo, Figure 3 displays results from estimating Equation (1) for plausibly related health effects that patients may avoid if they avoid chemotherapy. Indeed, patients have fewer days of anti-nausea medications prescribed (panel a) and are less likely to have a prescription for immunosupportive medications (panel b) - both medications that patients undergoing chemotherapy often take. It is important to note that nausea and a loss of white blood cells represent only two margins of health that may be impacted by chemotherapy treatment, and the absence thereof. The true health spillovers of forgoing chemotherapy are likely much higher, and not necessarily measurable.

[28] Note, a patient who sees a provider who has not yet adopted the technology can get tested by another provider.

4.3 Spillovers to other (non-targeted) treatments

Although the test targets chemotherapy use, it is possible that the information revealed by the test has implications for other commonly used cancer treatments. For instance, a reduction in chemotherapy use may be accompanied by an increase in alternative treatments. From Figure A.9 it can be seen that, technology adoption is associated with an increase in the use of radiation therapy. More specifically, patients undergoing surgery with a provider who adopted the technology in the past five years, are about 8 percent (5/63) more likely to be treated with radiation therapy. This appears to be driven by a shift from more aggressive surgery (complete and radical mastectomies) to less aggressive surgery (lumpectomies), where the latter case is more likely to also rely on radiation (see Figure A.10). Overall, this suggests that the adoption of genomic testing not only affected targeted treatments, but also non-targeted ones.

4.4 Patient mortality and proxies for recurrence

The sharp reduction in the likelihood of receiving chemotherapy for patients seeing an adopting provider raises the question of whether these treatments were unnecessary. On the one hand, chemotherapy can be physically, mentally, and financially costly in the short run. On the other hand, it carries the benefit of reducing the risk of the cancer recurring. Truly unnecessary treatments are treatments that don't increase the long-run likelihood from dying from cancer. Figure 4, displays the overall probability of dying due to breast cancer (panel a) or any cancer (panel b), within 5 years of the breast cancer surgery. From the figure it can be seen that there is no statistically significant increase in five-year mortality following technology adoption. To gauge whether there is an increase in metastasis or the treatment of a recurrent cancer (that may not necessarily have lead to death), Figure A.11 displays the probability of having a metastasis or being treated with chemotherapy between 2-5 years relative to the surgery. From the figure it can be seen that patients are, if anything, *less* likely to experience a metastasis or new chemotherapy episode. This could be consistent with an overall better targeting of chemo to patients, such that previously undertreated patients receive treatment.

4.5 Heterogeneity by race and SES

Overall, the main results suggest technology adoption was associated with an increase in testing and a reduction in chemotherapy use, without any associated health costs. I next, consider whether these impacts are distributed evenly across patients. Figure A.12 shows heterogeneity in the results for different patient subpopulations. From the figure, it can for instance be seen that white patients and patients of higher SES are relatively more likely to be tested, compared to Black patients and patients of lower SES, although the results are noisy.

4.6 Robustness

Patient sorting. One potential concern with the empirical design is that patients respond to the adoption decision of providers. If patient sorting to adopting providers is correlated with the likelihood of the patient needing chemotherapy, the estimated impact on chemotherapy of seeing an adopting provider is partly due to the technology, and partly due to changing patient characteristics. Smoothly evolving patient characteristics, speak against patient sorting, although sorting on unobservable patient characteristics, it not verifiable. To gauge patient sorting, the previous event study is estimated and presented in Figure A.14. Overall, the patient characteristics displayed in Figure A.14 evolve smoothly before and after adoption, although the NCI comorbidity score appear to be trending downward. This however, if anything, would bias against finding a chemotherapy reduction, as individuals with lower comorbidity scores are more likely to be able to tolerate chemotherapy treatment.

Conditioning instead of controlling for medical characteristics. One of the identifying assumptions of the event-study model in (1) is that the timing of adoption is uncorrelated with time-varying determinants of patients outcomes, and in particular the likelihood of receiving chemotherapy. In particular, if providers shifted towards seeing earlier-stage patients at the same time as they adopted the technology, this would cause the assumption to break. Since, the SEER data is very rich in cancer characteristics, one can relax this assumption and directly estimate impacts in a sample with fixed cancer characteristics (as opposed to controlling for it). The trade-off is that some of the most detailed characteristics are not available for all SEER registries, leading to a loss in sample size. Despite this, the results in Figure A.15 provide support for one of the main identifying

assumptions of the event-study model in (1).

Placebo checks. Another concern is that the adoption of the technology was concurrent with other events that might have impacted the use of chemotherapy. Although it is not possible to completely rule out the possibility of other concurrent changes, I estimate whether providers are more likely to use two other technologies and services. The first is immunohistochemistry (IHC) testing, which is commonly used for diagnostic purposes in breast cancer, including determining hormone positive receptors. The second placebo is the sentinel lymph node biopsy (SLNB) which is a surgical procedure used to determine whether breast cancer (or another solid tumor) has spread to the lymph nodes. Over time the SLNB has replaced full axillary lymph node dissection (ALND), particularly in early-stage patients.

A concurrent adoption of IHC and SLNB with the technology is worrying, as both IHC and SLNB have the potential to improve diagnosis, and thereby change chemotherapy decisions. Smoothness in the use of these two services is explored in Figure A.16. From Figure A.16, panel (a) it can be seen that there is no increased likelihood of using IHC testing around the time point of genomic testing adoption. Similarly, Figure A.16, panel (b), reveals no increased likelihood of SLNB use.

Control vs treated indications. As an alternative identification strategy, I leverage variation in provider adoption dates in addition to variation in a patients appropriateness for testing. More specifically, since the technology is primarily indicated for patients with early stage cancers, patients with metastatic cancers can be used as a (imperfect) control group. Both early stage and late stage patients may undergo chemotherapy in conjunction with hormone therapy. Intuitively, the empirical design compares the probability that a patient with an early stage cancer receives chemotherapy after her provider has adopted the technology, compared to before, relative to a patient with a late stage cancer. The corresponding event studies are given by:

$$Y_i = \sum_{r=-3, t\neq -1}^{r=4} \mathbb{1}(\text{r(i)} = \text{r})\mathbb{1}(\text{Treated}_i)\beta_m + \mathbb{1}(\text{Treated}_i)\tau + \gamma_{d(i)} + \eta_{t(i)} + \delta X' + \epsilon_i \tag{3}$$

where Y_i is the outcome variables for individual i (e.g. and indicator equal to one if patient i gets

chemotherapy within one year of the surgery), $\mathbb{1}(\text{r(i)} = \text{r})$ is a function that returns the year of $i's$ surgery relative to the technology adoption date of the oncologist, and $\mathbb{1}(\text{Treated}_i)$ is an indicator equal to one if the patient is flagged as having an early-stage cancer and hence would be appropriate for the technology. $\gamma_{d(i)}$ are provider fixed effects and $\eta_{t(i)}$ are year of surgery fixed effect. X' include time-invariant patient characteristics and time-varying provider characteristics previously defined. The main coefficient of interest is β_m which measures the increased/decreased probability of e.g., technology and chemotherapy receipt for patient i who had a late stage cancer and who saw provider j, t years after provider j first started using the technology, relative to a patient with an early-stage cancer. The results in Figure A.17 support the overall results.

Alternative estimators. To address recent debates about the interpretation of staggered difference-in-difference designs (Goodman-Bacon, 2021; Sun and Abraham, 2021; Borusyak and Jaravel, 2018; Callaway and Sant'Anna, 2021; De Chaisemartin and d'Haultfoeuille, 2022), some of the main results are replicated using the Sun and Abraham estimator (Sun and Abraham, 2021), in which the control group is set to be the never-treated. The results in Figure A.18 corroborate the main findings.

5 Efficiency Implications

5.1 Setup

This section sketches out a simple model of physician treatment choice in the presence of genomic testing following Chan et al. (2022). Oncologists $j \in J$ choose between treating or not treating a patient $i \in N$. Denote the choice set (treating vs not treating) as $t \in \{0, 1\}$. There are two types of patients - low-risk and high-risk patients, denoted as $s_i \in \{0, 1\}$, and treatment is only appropriate for high-risk patients - i.e., $u_i(0, 0) > u_i(1, 0)$ and $u_i(1, 1) > u_i(0, 1)$, where the utility function of the patient takes the form $u_i(t, s)$. However, oncologist only observe a noisy signal of a patient's actual risk-type in the absence of a testing technology:

$$\tilde{s}_i = s_i + \epsilon$$

where $\epsilon \sim N(0,1)$.[29] Oncologists only care about treating patients appropriately, hence the utility function of an oncologist takes the form of $u_j(t,s)$, with $u_j(0,0) > u_j(1,0)$ and $u_j(1,1) > u_j(0,1)$. The signal, $\tilde{s}_i$ is increasing in s_i, hence treatment, t_{ij}, is appropriate whenever the observed signal exceeds the oncologist's decision threshold, τ_j.[30] In this framework, genomic testing can be modeled as a factor that can reduce noise in the decision-making process

$$\tilde{s}_i = s_i + \frac{\epsilon}{\gamma(s_i)}$$

where $\gamma(s_i) \in [1, \infty)$ and higher levels of $\gamma(s_i)$ imply higher levels of effectiveness.

5.2 Classification problem

With binary patient types and treatments, a physician's treatment choice can be characterized by four outcomes, illustrated in the classification matrix in panel (a) of Figure 5. Noise results in diagnostic errors. Some low-risk patients are classified as high-risk (false positives or type I errors) - while other high-risk patients are classified as low-risk (false negatives or type II errors):

- True positive (TP): a high-risk patient ($s_i = 1$) is classified as high-risk ($\tilde{s}_i = 1$)
- True negatives (TN): a low-risk patient ($s_i = 0$) is classified as low-risk ($\tilde{s}_i = 0$)
- False positives (FP): a low-risk patient ($s_i = 0$) is classified as high-risk ($\tilde{s}_i = 1$)
- False negatives (FT): a high-risk patient ($s_i = 1$) is classified as low-risk ($\tilde{s}_i = 0$)

The overall performance of the classification problem can be visualized using the Receiver Operating Characteristic (ROC) curve which shows the relationship between the True Positive Rate, $\text{TPR} = \frac{\text{TP}}{\text{TP+FN}}$ and the False Positive Rate, $\text{FPR} = \frac{\text{FP}}{\text{FP+TN}}$ (Figure 5, panel (b)). Points along the ROC curve represents different combinations of TPR and FPR that result from differences in treatment thresholds τ_j. Hence, as discussed in Chan et al., 2022, the ROC curve can be viewed as a production possibility frontier. Movements along the curve may represent differences in preferences for true positives and false positives, while an increase in productivity can be thought of as an outward shift of the ROC curve to the left, as would be the case if the technology is effective. As

[29] For simplicity, I assume no variation in diagnostic skill across oncologist. In principle, this could be included as a scaling factor of the error term (see e.g., Currie et al., 2024)

[30] That is, the treatment decision can be modelled as $t_{ij} = \begin{cases} 1, & \tilde{s}_{ij} \geq \tau_j \\ 0, & \tilde{s}_{ij} < \bar{\tau}_j \end{cases}$

visualized in panel (c) of Figure 5, technology adoption can 1) increase the true positive rate at the expense of also increasing the false positive rate, 2) increase the true positive rate without increasing and possibly decreasing the false positive rate, 3) decrease the true positive rate while increasing the false positive rate, and 4) decrease the true positive rate while decreasing or not affecting the false positive rate. The second case represents a productivity gain and the third case represents a productivity loss. The first and fourth case are ambiguous and depend on the extent to which a gain (loss) in the true positive rate is valued relative to a gain (loss) in the false positive rate.

5.3 Identification

Measuring the TPR and FPR in this setting is challenging, as neither classified nor actual risk-type is observable in the data. However, identification of the *change* in the TPR and FPR is possible under a set of assumptions. First, under the assumption that the actual risk distribution in the patient population does not change before and after technology adoption (see Appendix A.5, for derivations) the change in the TPR and FPR can be reduced to the following, where S is the share of patients that are truly high-risk and P is the share of patients that are classified as high-risk:

$$\Delta\text{TPR} = \frac{-\Delta\text{FN}}{\text{S}} \tag{4}$$
$$\Delta\text{FPR} = \frac{\Delta\text{P} + \Delta\text{FN}}{1 - \text{S}}$$

Second, under the assumption that providers always treat high-risk patients (this follows from the model of physician choice), and never treat low-risk patients, a patient who is observed to be treated must have been classified as high-risk by the provider. This implies that the share of patients classified as high-risk in the population, P, equals the share of patients treated in the population. This object is observable and hence so is its change. More specifically, define ΔP to be the change in the share of the population treated with chemotherapy within one year of a breast cancer surgery.

Third, under the assumption that an *untreated* patient's recurrence or death in the future was indicative of a patient's actual risk type being high, ΔFN equals the share of patients in the population who were not treated and who did experience a recurrence. This object is also observable and hence so is its change. More specifically, define ΔFN as the change in the share of the population who were not treated with chemotherapy within one year of surgery and who died of any cancer

within five years of the surgery.

The remaining object - the population share of patients with high-risk cancers S is unobservable. The change in the true positive and false positive rates (TPR and FPR) can thus be expressed as functions of S. The first assumption is essentially also one of the identifying assumptions of the difference in difference design. Hence, the DiD estimator can be used to retrieve estimates of ΔP and ΔFN.

5.4 Results

Figure 6 illustrates the results from estimating Equation (1) using the true negative indicator (i.e. no chemotherapy within one year, and no cancer death within 5 years) and the false negative indicator (i.e. no chemotherapy within one year and a cancer death within 5 years) as the outcome variable. From the figure it can be seen that technology adoption is associated with an approximate 5 percentage point increase in the probability of a true negative . Importantly, technology adoption is not associated with any impact on the probability of a false negative. Hence, the estimates suggest that the performance of the classification problem improves following adoption. This illustrates the efficiency gain from genomic testing adoption.

Table 2 provides estimates of the inputs needed for estimating the change in the true positive rate (TPR) and false positive rate (FPR). The numbers correspond to estimates from equation (2). From the table it can be seen that technology adoption is associated with an almost 10 percentage point increase in the likelihood of getting tested and a one percentage point reduction in the likelihood of getting treated with chemotherapy. The latter represents the estimate of ΔP. Technology adoption is also associated with a similar reduction in the probability of a true negative and no statistically significant effect on the probability of a false negative. The former represents our estimate of ΔFN. Note that the estimates are in general smaller in magnitude compared to those of the event studies, as the DD estimates represent a weighted average of per-period coefficients.

Using these estimates and Equation (4), Figure 7 displays the quadrant of Figure 5, panel (c), in which the impacts of genomic testing adoption are estimated to lie. More specifically, the figure shows estimates from Equation (4) for varying levels of the share, S, of patients that are truly high-risk, ranging from 20 percent to 80 percent. $S = 50$ is displayed as well - it represents the empirically estimated share of patients who are likely too high-risk to be tested. Overall, under

the assumptions of the model, the figure suggests that technology adoption increased productivity primarily by decreasing the false positive rate, without an associated decrease in the true positive rate.

Do the productivity impacts of genomic testing adoption vary by race? From Table 2 it was seen that adoption was associated with an increase in the likelihood of getting tested, a decrease in the likelihood of getting chemotherapy, and an increase in the likelihood of a true negative at no increased risk of a false negative. Do similar results hold up for Black patients when compared to white patients? To test for this, Equation (2) is fully interacted with race and the results are presented in Table 3. From the table it can be seen that Black patients who have their surgery with a provider after the provider has adopted the technology are almost five percentage points less likely to be tested than white patients, conditional on seeing the same provider. The point estimate for chemotherapy use is positive, but statistically insignificant. The sample size and relatively small effect size, make it difficult to rule out relatively small disparities. However, it is likely that large disparities in chemotherapy treatment can be rejected. Ultimately, the downstream consequences of lower levels of testing post-adoption are difficult to assess with the current sample sizes and without assumptions about who the marginal tested patient is. This is further studied in the next section.

6 Equity Implications

The previous section suggests genomic testing adoption was associated with substantial productivity gains. To what extent were these gains born equally for patients of different race? The previous descriptives on diffusion suggest a racial gap in testing, that exists independent of area and income (see e.g. Figure A.3, Figure A.5). Moreover, Figure A.12 is suggestive of a racial gap in testing among patients who see the *same* provider. Table 3 formally establishes a lower likelihood of getting tested for Black patients relative to white patients following provider adoption of the technology. What explains the within-provider racial gap in testing? And what is the relative role of healthcare access, medical appropriateness, and provider decision-making in determining the racial gap in testing?

6.1 Mechanisms behind the within-provider racial gap in testing

What explains the within-provider racial gap in testing? In particular, is the gap in testing reflective of racial differences in medical appropriateness or differences in provider testing decisions?

Differences in patient medical appropriateness. Black patients may be less likely to be tested due to unobserved (to the econometrician) patient characteristics that make them less likely to benefit from the technology. As an example, if black patients are more likely to have advanced stages of cancer that would require chemotherapy than white patients, testing will be less beneficial in Black patients.

To gauge the extent to which the gap in testing is a function of differences in appropriateness, I construct a subset of the event-study sample. More specifically, I focus on registries and years for which more detailed information on cancer is available. Starting with the event study sample, I implement the following restrictions. First, I limit the sample to 2010-2018, to allow for direct measurement of HER2 status. Second, I restrict the sample to registries with non-missing information on characteristics that clinical guidelines state are relevant for the decision to test. This essentially amounts to dropping the Massachusetts and Texas registries. Third, I restrict the sample to adopting providers, and surgeries that happen after the provider has adopted the technology. This allows me to compare the testing decision for two, based on cancer etiology, very similar Black and white patients, who both saw a provider who had adopted the test at the time point of their surgery.

Next, the probability of being tested within 3 months of a breast cancer diagnosis is estimated in the previously defined sample using the following regression model:

$$Y_i = \beta \text{Black}_i + \gamma_{d(i)} + \delta X_i' + \epsilon_i \tag{5}$$

where Y_i is the probability of being tested within 3 months of the breast cancer surgery, $\gamma_{d(i)}$ are provider fixed effects, and X_i' are patient-level characteristics measured at the time point of diagnosis (e.g., age at diagnosis). The racial gap in testing (i.e., β) is then estimated in a sample of provider-patient pairs, were the provider has already adopted the technology at the time point of a patient's surgery.

To gauge the extent to which the within-provider racial gap is due to differences in medical appropriateness between Black and white patients, aspects of medical appropriateness are successively

(and separately) controlled for in Table 4. From the table it can be seen that controlling for staging reduces the racial gap in testing by 13 percent, while controlling for whether or not the cancer is hormone positive, reduces the gap by 36 percent. While a large fraction of the gap is explained by differences in medical appropriateness (49 percent), a large fraction of the gap remains unexplained after having controlled for these particular aspects of medical appropriateness. It is possible that other aspects of medical appropriateness, not captured here, explain differences in testing across Black and white patients. However, the characteristics selected here are the ones listed in the NCCN Breast Cancer guidelines as relevant in deciding whether or not to test a patient.

Differences in provider testing decisions. If providers use different testing threshold for Black patients relative to white patient, and in particular, if they use a higher threshold for Black patients, this would render a racial gap in testing. One approach to gauge this is using outcomes-based tests (see e.g., Knowles et al. (2001) and Anwar and Fang (2006) for early applications in the criminal justice setting, Arnold et al. (2018) for a more recent one, and (Chandra and Staiger (2010) for a related application in the healthcare setting). The outcomes-based approaches rely on the idea that a difference in testing thresholds across groups will render the returns to testing different across groups. In this case, part of the value, or returns to testing, lies in the possibility to forego chemotherapy. Conditional on testing, do Black patients have a higher likelihood of foregoing chemotherapy (consistent with a higher testing threshold)?

To gauge differential returns to testing, the likelihood of foregoing chemotherapy is estimated in a subset of the event-study sample, namely those that were tested, using the following regression:

$$Y_i = \alpha \text{ Black }_i + \gamma_{d(i)} + X' + \epsilon_i \tag{6}$$

where Y_i is an indicator equal to one if the patient was not treated with chemotherapy within one year of their surgery, $\gamma_{d(i)}$ represent provider fixed effects, and X' include NCI comorbidity tertiles, an indicator for being recommended for the test according to NCCN guidelines, and age indicators. The result from this regression reveals an $\alpha = -0.22$, statistically significant at the ten percent level of significance. Hence, if anything, Black patients have *lower* returns to testing than white patients.

This result should however be interpreted with caution. First, although the likelihood of

foregoing chemotherapy is a benefit of the test, this "benefit" is also a function of provider and patient decision-making, which complicates the outcomes-test. The negative α could, for instance, be consistent with providers applying a higher threshold to testing Black patients, and simultaneously applying a lower threshold to *treating* Black patients with chemotherapy. Second, as previous studies have highlighted, if the risk distribution of Black patients differ from those of white patients, this also poses a threat to any outcomes-based test (also known as the infra-marginality critique, see e.g. Corbett-Davies et al. (2017)). Although this study can leverage detailed cancer registry data, controlling for cancer characteristics only partially accounts for this problem.

6.2 Decomposing the overall racial gap in testing

Appendix A.6 provides a conceptual framework for how different forces govern the extent to which different population groups eventually gain from the arrival of a new technology. In particular, a disparity in the gains from innovation across two groups emerges and amplifies when the technology 1) is more appropriate in one group because it targets a disease more prevalent in one group ("differential appropriatenes"), 2) conditional on appropriateness is accessed to a greater extent by one group ("differential access or allocation"), 3) conditional on appropriateness and access, is used more intensively in one group ("differential testing or discrimination"), and conditional on all of the above, 4) is of greater value in one group "differential treatment effect"). When this is true, and groups are defined along the lines of socioeconomic status or race, the arrival of a new technology gives rise to, or worsens, an existing health disparity.

Abstracting from differences in treatment effects, and focusing on the determinants of testing, I do a simple empirical approximation of the role of differences in access, appropriateness, and testing. The results are displayed in table Table 5. The table shows 1) the estimated share of Black and white patients who have access to an adopting provider, 2) conditional on access, the estimated share of Black and white patients who are medically appropriate based on clinical guidelines, and 3) conditional on access and appropriateness, the estimated share of Black and white patients who are tested. Appendix A.6 describes how the product of these probabilities constitutes the estimated conditional probability of getting tested for a Black and white patient, and how a log-transformation allows for backing out the relative role of each factor. Table 5 shows the results from these calculations, and it highlights the role of differential appropriateness and testing as driving

the overall gap in testing, as opposed to differences in access. It is important to note that these calculations are performed in a sample of patients that were matched to an oncologist. To the extent that there are racial differences in access to any oncologist, this could magnify the role of access.

7 Discussion and Conclusion

The last two decades have seen a rapid rate of entry of genetic and genomic technologies designed to personalize medicine. This paper evaluated the efficiency and equity consequences of the diffusion of genomic testing in the context of breast cancer in the Medicare population. The findings suggest substantial efficiency gains. More specifically, patients who were diagnosed with breast cancer after their oncologist had adopted the technology, were less likely to receive chemotherapy, at no intermediate-term increased mortality or recurrence risk. Patients were also as a result less likely to experience chemotherapy related side-effect. These effects reflect only a portion of the true value of forgoing chemotherapy, which to the patient aside from health costs also include the financial burden of chemotherapy treatments and productivity loss for those still in the workforce. From a societal perspective, there are substantial gains from reduced healthcare spending on chemotherapy, the management of its side effect, and the freeing up of healthcare resources deployed to chemotherapy treatment.

However, the gains from innovation appear to have been concentrated to certain patient populations, with Black patients less likely to be tested, primarily for reasons other than costs and access. Existing disparities in staging contribute to the technology being less relevant to the average Black patient. To the extent that disparities in screening correlate with the increased likelihood of Black patients presenting with later stage disease (Smith-Bindman et al., 2006), addressing the causes of these patterns may render a substantial fraction of Black patients eligible for the technology. However, controlling for staging and other medical characteristics, Black patients are still less likely to be tested.

As such, this paper contributes to our understanding of the role of medical advancements in precision medicine on the healthcare system. While such innovation can generate a lot of efficiency gains to the healthcare system, by for instance improving diagnostic accuracy, it may also generate or exacerbate health disparities. Addressing these disparities may require a more nuanced understanding of its root causes, that include but are not limited to differences in affordability or

healthcare access, but that also include a better understanding of potential racial differences in the ways providers are treating patients as well as existing disparities in medical care such as screenings.

As a case in point, although today Black women with breast cancer have a 40 percent higher death rate than white women, this racial divergence is a relatively recent phenomenon emerging in the 1980s. The divergence has in part been attributed to medical advancements including the arrival of mammography and hormone therapy, which due to access and differences in medical appropriateness, tended to benefit white patients more Jatoi et al. (2022).

As precision medicine and oncology to an increasing extent target diseases that occur in earlier stages, as opposed to end-of-life, and population health improves not only as a result of reductions in mortality but also morbidity, the case for reducing existing racial disparities in preventive medicine and screening is becoming increasingly important. In the absence of interventions that address these disparities, current and future technologies that target early stages of disease, such as the much anticipated "liquid biopsies", may end up disproportionately benefiting certain groups in society more. In the context of precision technologies that *spare* patients from treatments, the effects can be two-fold and regressive - patients not only fail to realize the health benefits stemming from the technology, they also fail to realize the financial savings from forgoing costly treatments.

To conclude, as new precision technologies continue to develop and diffuse through the healthcare, policies that ensure the equitable diffusion of these technologies are critical in ensuring that the gains from innovation are realized across patient groups, regardless of socioeconomic status and race.

References

Abaluck, Jason, Leila Agha, Chris Kabrhel, Ali Raja, and Arjun Venkatesh, "The Determinants of Productivity in Medical Testing: Intensity and Allocation of Care," *American Economic Review*, 2016, *106* (12), 3730–3764.

_ , _ , David C Chan Jr, Daniel Singer, and Diana Zhu, "Fixing Misallocation with Guidelines: Awareness vs. Adherence," Report, National Bureau of Economic Research 2020.

Ademuyiwa, Foluso O, Austin Miller, Tracey O'Connor, Stephen B Edge, Mangesh A Thorat, George W Sledge, Ellis Levine, and Sunil Badve, "The Effects of Oncotype DX Recurrence Scores on Chemotherapy Utilization in a Multi-Institutional Breast Cancer Cohort," *Breast Cancer Research and Treatment*, 2011, *126* (3), 797–802.

Agarwal, Nikhil, Alex Moehring, Pranav Rajpurkar, and Tobias Salz, "Combining Human Expertise with Artificial Intelligence: Experimental Evidence from Eadiology," Report, National Bureau of Economic Research 2023.

Aizer, Anna and Laura Stroud, "Education, Knowledge and the Evolution of Disparities in Health," Report, National Bureau of Economic Research 2010.

Alsan, Marcella, Maya Durvasula, Harsh Gupta, Joshua Schwartzstein, and Heidi L Williams, "Representation and Extrapolation: Evidence from Clinical Trials," Report, National Bureau of Economic Research 2022.

_ , Romaine A Campbell, Lukas Leister, and Ayotomiwa Ojo, "Investigator Racial Diversity and Clinical Trial Participation," Report, National Bureau of Economic Research 2023.

_ , Vincenzo Atella, Jay Bhattacharya, Valentina Conti, Iván Mejía-Guevara, and Grant Miller, "Technological Progress and Health Convergence: The Case of Penicillin in Postwar Italy," *Demography*, 2021, *58* (4), 1473–1498.

American Cancer Society, "Key Statistics for Breast Cancer," 2023. Accessed Oct 10, 2023. https://www.cancer.org/cancer/types/breast-cancer/about/how-common-is-breast-cancer.html.

_ , "Survival Rates for Breast Cancer," 2023. Accessed Oct 10, 2023. https://www.cancer.org/cancer/types/breast-cancer/understanding-a-breast-cancer-diagnosis/breast-cancer-survival-rates.html.

Anwar, Shamena and Hanming Fang, "An Alternative Test of Racial Prejudice in Motor Vehicle Searches: Theory and Evidence," *American Economic Review*, 2006, *96* (1), 127–151.

Arnold, David, Will Dobbie, and Crystal S Yang, "Racial Bias in Bail Decisions*," *The Quarterly Journal of Economics*, 2018, *133* (4), 1885–1932.

Berndt, Ernst R, Dana P Goldman, and John Rowe, *Economic Dimensions of Personalized and Precision Medicine*, University of Chicago Press, 2019.

Bhutiani, Neal, Michael E Egger, Nicolás Ajkay, Charles R Scoggins, Robert CG Martin II, and Kelly M McMasters, "Multigene Signature Panels and Breast Cancer Therapy: Patterns of Use and Impact on Clinical Decision Making," *Journal of the American College of Surgeons*, 2018, *226* (4), 406–412. e1.

Blattner, Laura and Scott Nelson, "How Costly is Noise? Data and Disparities in Consumer Credit," *arXiv preprint arXiv:2105.07554*, 2021.

Borusyak, Kirill and Xavier Jaravel, *Revisiting Event Study Designs*, SSRN, 2018.

Breastcancer.org, "Survey Finds Nearly Half of People With Breast Cancer Report Serious Financial Burdens," 2022. Accessed July 15, 2025. https://www.breastcancer.org/managing-life/covering-cost-of-care/cost-of-care-report/survey-results.

Bögl, Sarah, Jasmin Moshfegh, Petra Persson, and Maria Polyakova, "The Economics of Infertility: Evidence from Reproductive Medicine," Report, National Bureau of Economic Research 2024.

Callaway, Brantly and Pedro HC Sant'Anna, "Difference-in-Differences with Multiple Time Periods," *Journal of Econometrics*, 2021, *225* (2), 200–230.

Callison, Kevin, Michael E Darden, and Keith F Teltser, "Innovation and Health Inequality: Evidence from Direct-Acting Antivirals," Report, National Bureau of Economic Research 2025.

Carlson, Josh J and Joshua A Roth, "The Impact of the Oncotype Dx Breast Cancer Assay in Clinical Practice: a Systematic Review and Meta-Analysis," *Breast Cancer Research and Treatment*, 2013, *141* (1), 13–22.

Chaisemartin, Clément De and Xavier d'Haultfoeuille, "Difference-in-Differences Estimators of Intertemporal Treatment Effects," Report, National Bureau of Economic Research 2022.

Chan, David C, Matthew Gentzkow, and Chuan Yu, "Selection with Variation in Diagnostic Skill: Evidence from Radiologists," *The Quarterly Journal of Economics*, 2022, *137* (2), 729–783.

Chandra, Amitabh and Douglas O Staiger, "Identifying Provider Prejudice in Healthcare," Report, National Bureau of Economic Research 2010.

_ **and Jonathan Skinner**, "Technology Growth and Expenditure Growth in Health Care," *Journal of Economic Literature*, 2012, *50* (3), 645–680.

_ **, Pragya Kakani, and Adam Sacarny**, "Hospital Allocation and Racial Disparities in Health Care," *Review of Economics and Statistics*, 2022, pp. 1–39.

Charlson, Mary E, Peter Pompei, Kathy L Ales, and C Ronald MacKenzie, "A New Method of Classifying Prognostic Comorbidity in Longitudinal Studies: Development and Validation," *Journal of Chronic Diseases*, 1987, *40* (5), 373–383.

Chen, Shuyi, Christopher Thacker, Shengxuan Wang, Katelyn A Young, Rebecca L Hoffman, and Joseph A Blansfield, "Adherence Disparities and Utilization Trends of Oncotype DX Assay: A National Cancer Database Study," *Journal of Surgical Research*, 2023, *286*, 65–73.

Choi, In Sil, Jiwoong Jung, Byoung Hyuck Kim, Sohee Oh, Jongjin Kim, Jin Hyun Park, Jeong Hwan Park, and Ki-Tae Hwang, "The 21-gene Recurrence Score Assay and Prediction of Chemotherapy Benefit: A Propensity Score-Matched Analysis of the SEER Database," *Cancers*, 2020, *12* (7), 1829.

Conner, Peter, Liran Einav, Amy Finkelstein, and Petra Persson, "Targeting Precision Medicine: Evidence from Prenatal Screening," *Journal of Political Economy*, 2025, *133* (2), 604–651.

Corbett-Davies, Sam, Camelia Simoiu, and Sharad Goel, "The Problem of Infra-marginality in Outcome Tests for Discrimination," *The Annals of Applied Statistics*, 2017.

Corredor-Waldron, Adriana, Janet Currie, and Molly Schnell, "Drivers of Racial Differences in C-Sections," Report, National Bureau of Economic Research 2024.

Currie, Janet, W Bentley MacLeod, and Kate Musen, "First Do No Harm? Doctor Decision Making and Patient Outcomes," Report, National Bureau of Economic Research 2024.

Cutler, David, Angus Deaton, and Adriana Lleras-Muney, "The Determinants of Mortality," *Journal of Economic Perspectives*, 2006, *20* (3), 97–120.

Cutler, David M, *Your Money or Your Life: Strong Medicine for America's Health Care System*, Oxford University Press, 2004.

_ , "Are We Finally Winning the War on Cancer?," *Journal of Economic Perspectives*, 2008, *22* (4), 3–26.

_ **and Robert S Huckman**, "Technological Development and Medical Productivity: the Diffusion of Angioplasty in New York State," *Journal of Health Economics*, 2003, *22* (2), 187–217.

_ **, Ellen Meara, and Seth Richards-Shubik**, "Induced Innovation and Social Inequality: Evidence from Infant Medical Care," *Journal of Human Resources*, 2012, *47* (2), 456–492.

Daysal, N Meltem, William N Evans, Mikkel Hasse Pedersen, and Mircea Trandafir, "Do Medical Treatments Work for Work?: Evidence from Breast Cancer Patients," Report, National Bureau of Economic Research 2024.

Dean, Emma Boswell, Mallick Hossain, and Daniel Kaliski, "Does Algorithmic Risk Assessment Affect Physician Prescribing? Understanding Aggregate Versus Micro-Level Impacts," Report, Social Science Research Network 2024.

Deyo, Richard A, Daniel C Cherkin, and Marcia A Ciol, "Adapting a Clinical Comorbidity Index for Use with ICD-9-CM Administrative Databases," *Journal of Clinical Epidemiology*, 1992, *45* (6), 613–619.

Dinan, Michaela A, Lauren E Wilson, Shelby D Reed, Jennifer J Griggs, and Edward C Norton, "Association of 21-Gene Assay (OncotypeDX) Testing and Receipt of Chemotherapy in the Medicare Breast Cancer Patient Population Following Initial Adoption," *Clinical Breast Cancer*, 2020, *20* (6), 487–494. e1.

_ **, Xiaojuan Mi, Shelby D Reed, Bradford R Hirsch, Gary H Lyman, and Lesley H Curtis**, "Initial Trends in the Use of the 21-Gene Recurrence Score Assay for Patients with Breast Cancer in the Medicare Population, 2005-2009," *JAMA Oncology*, 2015, *1* (2), 158–166.

_ , _ , _ **, Gary H Lyman, and Lesley H Curtis**, "Association Between Use of the 21-Gene Recurrence Score Assay and Receipt of Chemotherapy Among Medicare Beneficiaries with Early-Stage Breast cancer, 2005-2009," *JAMA Oncology*, 2015, *1* (8), 1098–1109.

Eiermann, W, M Rezai, S Kümmel, T Kühn, M Warm, K Friedrichs, A Schneeweiss, S Markmann, H Eggemann, and J Hilfrich, "The 21-Gene Recurrence Score Assay Impacts Adjuvant Therapy Recommendations for ER-Positive, Node-Negative and Node-Positive Early Breast Cancer Resulting in a Risk-Adapted Change in Chemotherapy Use," *Annals of Oncology*, 2013, *24* (3), 618–624.

Enewold, Lindsey, Helen Parsons, Lirong Zhao, David Bott, Donna R Rivera, Michael J Barrett, Beth A Virnig, and Joan L Warren, "Updated Overview of the SEER-Medicare Data: Enhanced Content and Applications," *JNCI Monographs*, 2020, *2020* (55), 3–13.

Fadlon, Itzik, Parag Agnihotri, Christopher Longhurst, and Ming Tai-Seale, "Equity and Efficiency in Technology Adoption: Evidence from Digital Health," Report, National Bureau of Economic Research 2024.

Fuster, Andreas, Paul Goldsmith-Pinkham, Tarun Ramadorai, and Ansgar Walther, "Predictably Unequal? The Effects of Machine Learning on Credit Markets," *The Journal of Finance*, 2022, *77* (1), 5–47.

Geffen, DB, S Abu-Ghanem, N Sion-Vardy, R Braunstein, M Tokar, S Ariad, B Delgado, M Bayme, and M Koretz, "The Impact of the 21-Gene Recurrence Score Assay on Decision Making About Adjuvant Chemotherapy in Early-Stage Estrogen-Receptor-Positive Breast Cancer in an Oncology Practice with a Unified Treatment Policy," *Annals of Oncology*, 2011, *22* (11), 2381–2386.

Glied, Sherry and Adriana Lleras-Muney, "Technological Innovation and Inequality in Health," *Demography*, 2008, *45*, 741–761.

Goodman-Bacon, Andrew, "Difference-in-Differences with Variation in Treatment Timing," *Journal of Econometrics*, 2021, *225* (2), 254–277.

Harris, Lyndsay, Herbert Fritsche, Robert Mennel, Larry Norton, Peter Ravdin, Sheila Taube, Mark R Somerfield, Daniel F Hayes, and Robert C Bast Jr, "American Society of Clinical Oncology 2007 Update of Recommendations for the Use of Tumor Markers in Breast Cancer," *Journal of Clinical Oncology*, 2007, *25* (33), 5287–5312.

Jatoi, Ismail, Hyuna Sung, and Ahmedin Jemal, "The Emergence of the Racial Disparity in US Breast-Cancer Mortality," *New England Journal of Medicine*, 2022, *386* (25), 2349–2352.

Jayachandran, Seema, Adriana Lleras-Muney, and Kimberly V Smith, "Modern Medicine and the Twentieth Century Decline in Mortality: Evidence on the Impact of Sulfa Drugs," *American Economic Journal: Applied Economics*, 2010, *2* (2), 118–146.

Kaiser Family Foundation, "An Overview of Medicare," 2019. Accessed Oct 10, 2023. https://www.kff.org/medicare/issue-brief/an-overview-of-medicare/.

Klabunde, Carrie N, Julie M Legler, Joan L Warren, Laura-Mae Baldwin, and Deborah

Schrag, "A Refined Comorbidity Measurement Algorithm for Claims-Based Studies of Breast, Prostate, Colorectal, and Lung Cancer Patients," *Annals of Epidemiology*, 2007, *17* (8), 584–590.

Kleinberg, Jon, Himabindu Lakkaraju, Jure Leskovec, Jens Ludwig, and Sendhil Mullainathan, "Human Decisions and Machine Predictions," *The Quarterly Journal of Economics*, 2018, *133* (1), 237–293.

Knowles, John, Nicola Persico, and Petra Todd, "Racial bias in Motor Vehicle Searches: Theory and Evidence," *Journal of Political Economy*, 2001, *109* (1), 203–229.

Lichtenberg, Frank R, "Has Medical Innovation Reduced Cancer Mortality?," *CESifo Economic Studies*, 2014, *60* (1), 135–177.

Lleras-Muney, Adriana and Frank R Lichtenberg, "Are the More Educated More Likely to Use New Drugs?," *Annales d'Économie et de Statistique*, 2005, pp. 671–696.

Losk, Katya, Rachel A Freedman, Alison Laws, Olga Kantor, Elizabeth A Mittendorf, Zhenying Tan-Wasielewski, Lorenzo Trippa, Nancy U Lin, Eric P Winer, and Tari A King, "Oncotype DX Testing in Node-Positive Breast Cancer Strongly Impacts Chemotherapy Use at a Comprehensive Cancer Center," *Breast Cancer Research and Treatment*, 2021, *185* (1), 215–227.

Manski, Charles F, "Reasonable Patient Care Under Uncertainty," *Health Economics*, 2018, *27* (10), 1397–1421.

Marchionni, Luigi, Renee F Wilson, Spyridon S Marinopoulos, Antonio C Wolff, Giovanni Parmigiani, Eric B Bass, and Steven N Goodman, "Impact of Gene Expression Profiling Tests on Breast Cancer Outcomes," *Evidence Report/Technology Assessment*, 2007, (160), 1–105.

Melendez, Robert, Philippa Clarke, Anam Khan, Iris Gomez-Lopez, Mao Li, and Megan Chenoweth, "National Neighborhood Data Archive (NaNDA): Socioeconomic Status and Demographic Characteristics of Census Tracts, United States, 2008–2017," *Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor]*, 2020, pp. 05–19.

Mullainathan, Sendhil and Ziad Obermeyer, "Diagnosing Physician Error: A Machine Learning Approach to Low-Value Health Care," *The Quarterly Journal of Economics*, 2022, *137* (2), 679–727.

Narang, Amol K and Lauren Hersch Nicholas, "Out-of-Pocket Spending and Financial Burden

Among Medicare Beneficiaries with Cancer," *JAMA Oncology*, 2017, *3* (6), 757–765.

National Cancer Institute, "The TAILORx Breast Cancer Trial," 2018. Accessed Oct 10, 2023. https://www.cancer.gov/types/breast/research/tailorx.

Nazha, Bassel, Manoj Mishra, Rebecca Pentz, and Taofeek K Owonikoko, "Enrollment of Racial Minorities in Clinical Trials: Old Problem Assumes New Urgency in the Age of Immunotherapy," *American Society of Clinical Oncology Educational Book*, 2019, *39*, 3–10.

Paik, Soonmyung, Gong Tang, Steven Shak, Chungyeul Kim, Joffre Baker, Wanseop Kim, Maureen Cronin, Frederick L Baehner, Drew Watson, and John Bryant, "Gene Expression and Benefit of Chemotherapy in Women with Node-Negative, Estrogen Receptor–Positive Breast Cancer," *Journal of Clinical Oncology*, 2006, *24* (23), 3726–3734.

Peethambaram, Prema P, Tanya L Hoskin, Courtney N Day, Matthew P Goetz, Elizabeth B Habermann, and Judy C Boughey, "Use of 21-Gene Recurrence Score Assay to Individualize Adjuvant Chemotherapy Recommendations in ER+/HER2 Node Positive Breast Cancer—A National Cancer Database study," *NPJ Breast Cancer*, 2017, *3* (1), 41.

Petkov, Valentina I, Dave P Miller, Nadia Howlader, Nathan Gliner, Will Howe, Nicola Schussler, Kathleen Cronin, Frederick L Baehner, Rosemary Cress, and Dennis Deapen, "Breast-cancer-specific mortality in patients treated based on the 21-gene assay: a SEER population-based study," *NPJ breast cancer*, 2016, *2* (1), 1–9.

Quan, Hude, Vijaya Sundararajan, Patricia Halfon, Andrew Fong, Bernard Burnand, Jean-Christophe Luthi, L Duncan Saunders, Cynthia A Beck, Thomas E Feasby, and William A Ghali, "Coding Algorithms for Defining Comorbidities in ICD-9-CM and ICD-10 Administrative Data," *Medical Care*, 2005, *43* (11), 1130–1139.

Raymond, Lindsey, "The Market Effects of Algorithms," Report, Working Paper 2023.

Roberts, Megan C, Dave P Miller, Steven Shak, and Valentina I Petkov, "Breast cancer-specific survival in patients with lymph node-positive hormone receptor-positive invasive breast cancer and Oncotype DX Recurrence Score results in the SEER database," *Breast cancer research and treatment*, 2017, *163* (2), 303–310.

_ , Morris Weinberger, Stacie B Dusetzina, Michaela A Dinan, Katherine E Reeder-Hayes, Lisa A Carey, Melissa A Troester, and Stephanie B Wheeler, "Racial Variation in the Uptake of Oncotype DX Testing for Early-Stage Breast Cancer," *Journal of Clinical Oncology*,

2016, *34* (2), 130.

Romano, Patrick S, Leslie L Roos, and James G Jollis, "Presentation Adapting a Clinical Comorbidity Index for Use with ICD-9-CM Administrative Data: Differing Perspectives," *Journal of Clinical Epidemiology*, 1993, *46* (10), 1075–1079.

Schaafsma, Evelien, Baoyi Zhang, Merit Schaafsma, Chun-Yip Tong, Lanjing Zhang, and Chao Cheng, "Impact of Oncotype DX Testing on ER+ Breast Cancer Treatment and Survival in the First Decade of Use," *Breast Cancer Research*, 2021, *23* (1), 1–11.

Smith-Bindman, Rebecca, Diana L Miglioretti, Nicole Lurie, Linn Abraham, Rachel Ballard Barbash, Jodi Strzelczyk, Mark Dignan, William E Barlow, Cherry M Beasley, and Karla Kerlikowske, "Does Utilization of Screening Mammography Explain Racial and Ethnic Differences in Breast Cancer?," *Annals of Internal Medicine*, 2006, *144* (8), 541–553.

Sparano, Joseph A, Robert J Gray, Della F Makower, Kathleen I Pritchard, Kathy S Albain, Daniel F Hayes, Charles E Geyer Jr, Elizabeth C Dees, Edith A Perez, and John A Olson Jr, "Prospective Validation of a 21-Gene Expression Assay in Breast cancer," *New England Journal of Medicine*, 2015, *373* (21), 2005–2014.

– , – , – , – , – , – , – , – , – , Matthew P Goetz, and John A Olson Jr, "Adjuvant Chemotherapy Guided by a 21-gene Expression Assay in Breast Cancer," *New England Journal of Medicine*, 2018, *379* (2), 111–121.

Su, Kevin W, Jane Hall, Pamela R Soulos, Maysa M Abu-Khalaf, Suzanne B Evans, Sarah S Mougalian, Charles E Rutter, Amy J Davidoff, and Cary P Gross, "Association of 21-Gene Recurrence Score Assay and Adjuvant Chemotherapy Use in the Medicare Population, 2008–2011," *Journal of Geriatric Oncology*, 2016, *7* (1), 15–23.

Sun, Liyang and Sarah Abraham, "Estimating Dynamic Treatment Effects in Event Studies with Heterogeneous Treatment Effects," *Journal of Econometrics*, 2021, *225* (2), 175–199.

Unger, Joseph M, Elise Cook, Eric Tai, and Archie Bleyer, "The Role of Clinical Trial Participation in Cancer Research: Barriers, Evidence, and Strategies," *American Society of Clinical Oncology Educational Book*, 2016, *36*, 185–198.

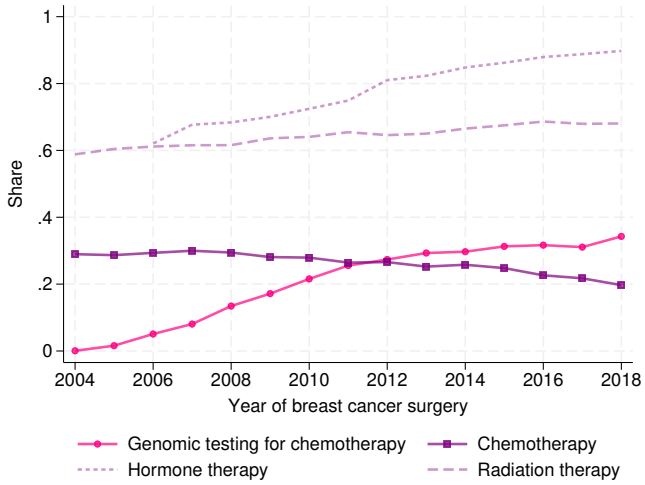
Zhang, Lu, Mei-Chin Hsieh, Valentina Petkov, Qingzhao Yu, Yu wen Chiu, and Xiao-Cheng Wu, "Trend and Survival Benefit of Oncotype DX Use Among Female Hormone Receptor-Positive Breast Cancer Patients in 17 SEER Registries, 2004–2015," *Breast Cancer Research and*

Treatment, 2020, *180*, 491–501.

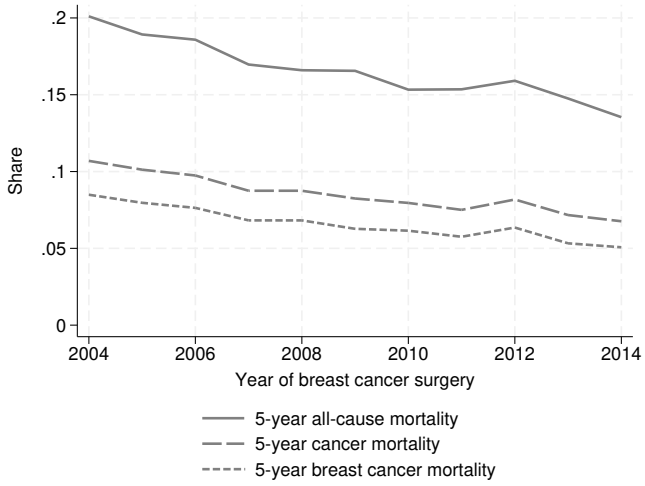
8 List of Exhibits

Figure 1: Technology Diffusion and Trends in Breast Cancer Treatments and Mortality

(a) Trends in Genomic Testing and Targeted and Non-Targeted Treatments



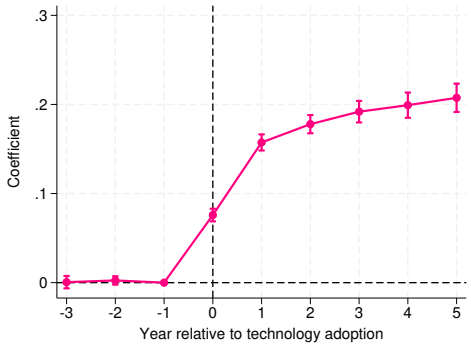
(b) Trends in 5-year Mortality



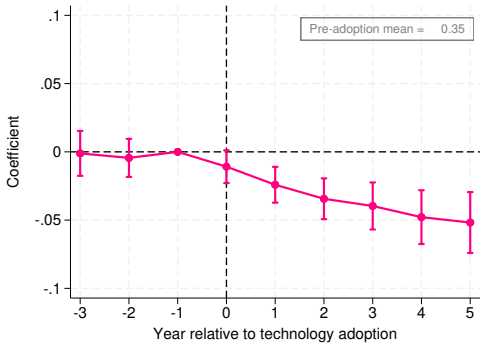
Note. The figure shows technology diffusion and trends in treatment (panel a) as well as the five-year mortality rate (panel b) in the sample of breast cancer surgery patients. See column 1 of Table 1 for characteristics of this sample. Panel (a) shows the share of patients that were tested around the time of their breast cancer surgery (pink solid line), received chemotherapy within one year (purple solid line), received hormone therapy within one year (short dash light purple line), and/or received radiation therapy within one year (long dash light purple line). The time series for hormone therapy starts in 2006, when Part D coverage was available. Panel (b) shows the share of patients who died within five years of their breast cancer surgery due to any cause (solid gray line), due to any cancer (long dashed gray line), or due to breast cancer (short dashed gray line). Panel (b) is restricted surgeries between 2004-2014 to allow for a five-year follow-up period.

Figure 2: Effect of Technology Adoption on Testing and Targeted Treatments

(a) Genomic testing within 3 months



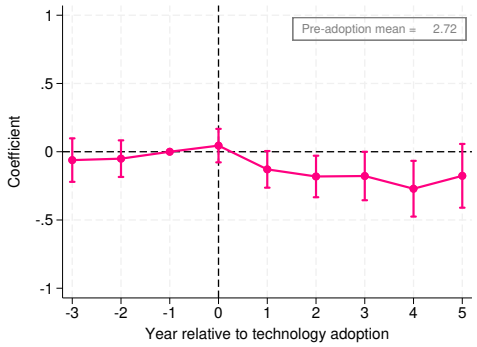
(b) Chemotherapy within 1 year



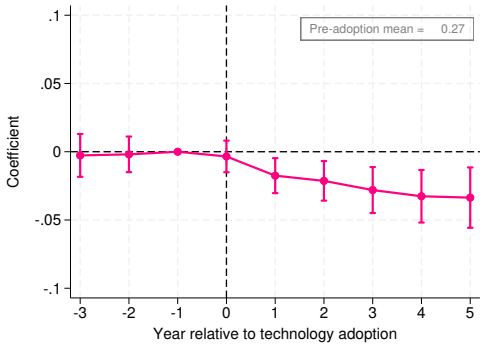
Note. The figure shows the impact of provider technology adoption on the likelihood that a patient is tested around the time point of their surgery (panel a) and treated with chemotherapy within one year (panel b) of the breast cancer surgery, respectively. The figure depicts the β coefficients from estimating the event study model specified in Equation (1) in the event study sample (see column (2) in Table 1 for summary statistics of this sample). The pre-adoption average of the outcome variable is calculated in the sample of adopting providers and displayed in the upper right corner of the graph. The event study model includes controls for a patient's disability status, an indicator for living in a metropolitan area, NCI comorbidity tertiles, health service area fixed effects, patient age (binned in 5-year intervals), and various measures of cancer characteristics, including indicators for the tumor being well-differentiated, neoadjuvant chemotherapy use, localized cancer, HER2 positive treatment use, and the cancer being classified as metastatic within 90 days of surgery. Standard errors are clustered at the provider level.

Figure 3: Effect of Technology Adoption on the Treatment and Prevention of Side Effects

(a) Number of days with anti-emetic prescription within one year



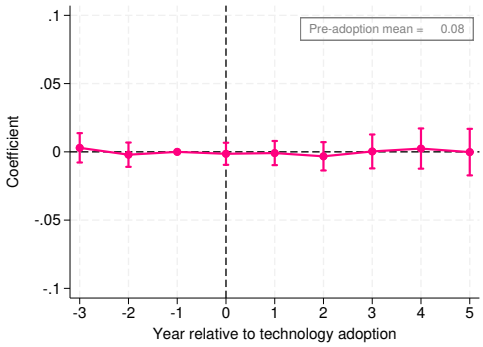
(b) Any immunosupportive prescription within one year



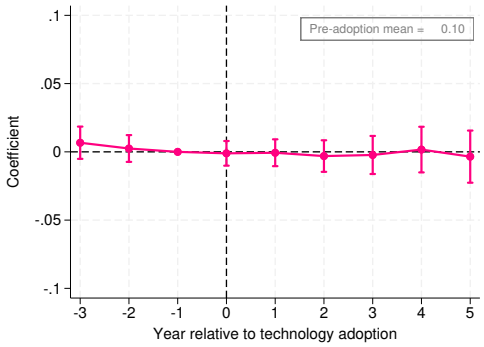
Note. The figure shows the impact of provider technology adoption on the number of days that a patient has anti-emetic prescriptions within one year of the breast cancer surgery (panel a) as well as the likelihood of having any immunosupportive prescription within one year of the surgery (panel b). The figure depicts the β coefficients from estimating the event study model specified in Equation (1) in the event study sample (see column (3) in Table 1 for summary statistics of this sample). The pre-adoption average of the outcome variable is calculated in the sample of adopting providers and displayed in the upper right corner of the graph. The event study model includes controls for a patient's disability status, an indicator for living in a metropolitan area, NCI comorbidity tertiles, health service area fixed effects, patient age (binned in 5-year intervals), and various measures of cancer characteristics, including indicators for the tumor being well-differentiated, neoadjuvant chemotherapy use, localized cancer, HER2 positive treatment use, and the cancer being classified as metastatic within 90 days of surgery. Standard errors are clustered at the provider level.

Figure 4: Effect of Technology Adoption on Cancer Mortality

(a) Died within five years due to breast cancer



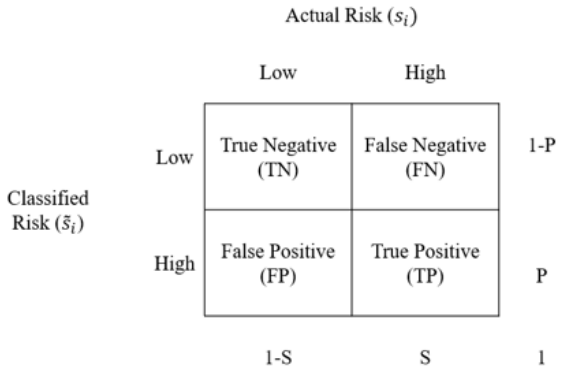
(b) Died within five years due to any cancer



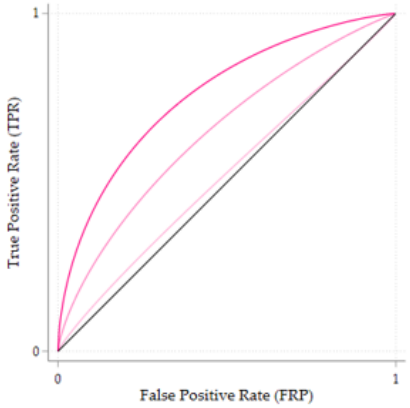
Note. The figure shows the impact of provider technology adoption on the likelihood that a patient dies within one year (panel a) and five years (panel b) of the breast cancer surgery, respectively. The figure depicts the β coefficients from estimating the event study model specified in Equation (1) in the event study sample, restricted to surgeries before 2014 to allow for a five-year follow-up period. The pre-adoption average of the outcome variable is calculated in the sample of adopting providers and displayed in the upper right corner of the graph. The event study model includes controls for a patient's disability status, an indicator for living in a metropolitan area, NCI comorbidity tertiles, health service area fixed effects, patient age (binned in 5-year intervals), and various measures of cancer characteristics, including indicators for the tumor being well-differentiated, neoadjuvant chemotherapy use, localized cancer, HER2 positive treatment use, and the cancer being classified as metastatic within 90 days of surgery. Standard errors are clustered at the provider level.

Figure 5: Technology Adoption and Productivity

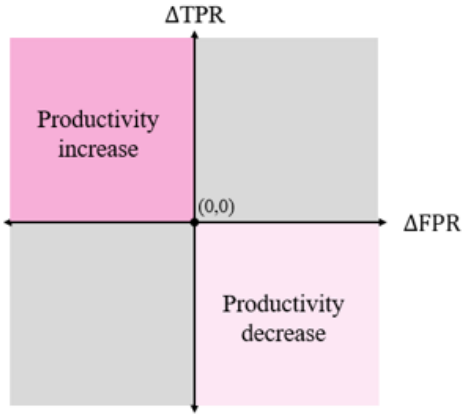
(a) Classification matrix



(b) ROC curve



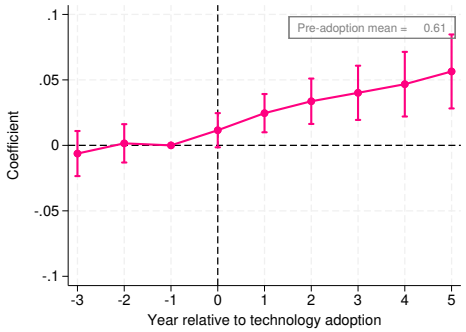
(c) Change in productivity



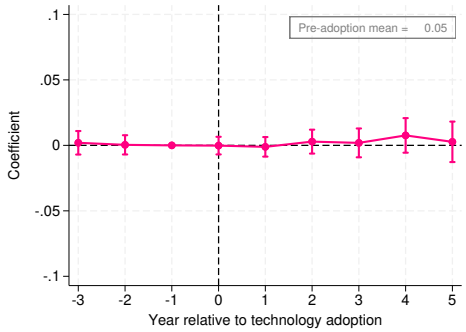
Note. The figure illustrates the framework for evaluating the productivity effects of technology adoption. Panel (a) displays a classification matrix, panel (b) displays the Receiver Operating Characteristic (ROC) curve, and panel (c) translates what changes in the True Positive Rate (TPR) and changes in the False Positive Rate (FPR) imply for productivity.

Figure 6: Effect of Technology Adoption on Medical Errors

(a) Proxy for true negatives

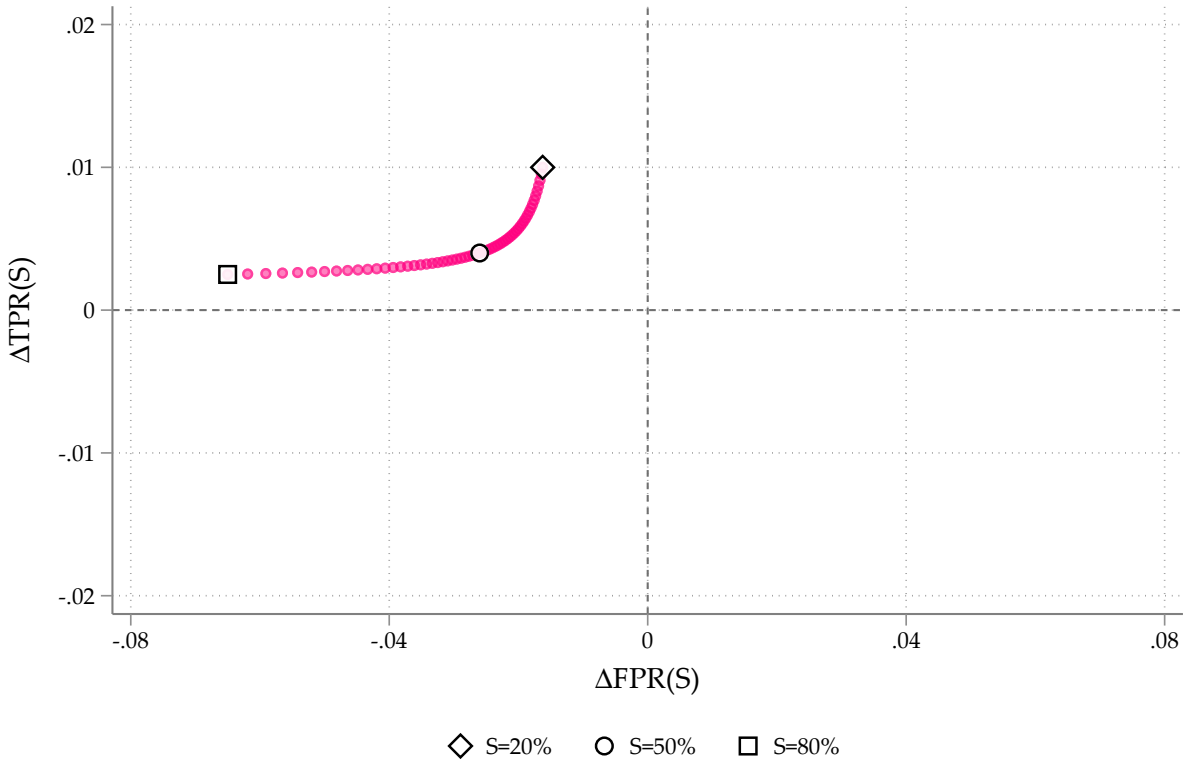


(b) Proxy for false negatives



Note. The figure shows the impact of provider technology adoption on the likelihood of a true negative (panel a) and a false negative (panel b), respectively. The figure depicts the β coefficients from estimating the event study model specified in Equation (1) in the event study sample, restricted to surgeries before 2014 to allow for a five-year follow-up period. The pre-adoption average of the outcome variable is calculated in the sample of adopting providers and displayed in the upper right corner of the graph. The event study model includes controls for a patient's disability status, an indicator for living in a metropolitan area, NCI comorbidity tertiles, health service area fixed effects, patient age (binned in 5-year intervals), and various measures of cancer characteristics, including indicators for the tumor being well-differentiated, neoadjuvant chemotherapy use, localized cancer, HER2 positive treatment use, and the cancer being classified as metastatic within 90 days of surgery. Standard errors are clustered at the provider level.

Figure 7: Effect of Technology Adoption on Productivity



Note. The figure displays the quadrant of Figure 5, panel (c), in which the impacts of genomic testing adoption are estimated to lie. More specifically, this figure shows estimates from equation (4) using as inputs the coefficients estimated in (2) (also displayed in Table 2) for varying levels of the share, S, of patients that are truly high-risk, ranging from 20 percent to 80 percent. $S = 50$ is displayed as well - it represents the empirically estimated share of patients who are likely too high-risk to be tested.

Table 1: Summary statistics

	(1)	(2)
	At-Risk Sample	Event Study Sample
Patient demographics		
Age	70.97 (6.44)	70.92 (6.53)
White (%)	85.94 (34.76)	86.65 (34.01)
Black (%)	8.56 (27.98)	8.31 (27.6)
Medicaid eligible (%)	16.88 (37.46)	16.46 (37.08)
Enrollment SSDI (%)	9.41 (29.19)	9.63 (29.5)
Metropolitan county (%)	85.02 (35.68)	85.17 (35.54)
Patient comorbidities		
NCI comorbidity score	0.26 (0.41)	0.26 (0.41)
Charlson comorbidity score	0.79 (1.25)	0.77 (1.23)
Breast cancer characteristics		
Localized (%)	70.31 (45.69)	68.92 (46.28)
Well differentiated (%)	23.1 (42.15)	22.43 (41.71)
Metastatis (%)	10.73 (30.95)	11.5 (31.9)
Breast cancer treatments		
Lumpectomy (%)	67.4 (46.87)	66.62 (47.16)
HER2 treatment (%)	7.86 (26.91)	8.4 (27.74)
Neoadjuvant chemotherapy (%)	7.29 (25.99)	7.21 (25.87)
Predicted probability of chemotherapy	20.18 (18.16)	20.77 (18.7)
Immunohistochemistry testing (%)	96.47 (18.45)	96.89 (17.36)
Genomic testing (%)	20.93 (40.68)	17.78 (38.24)
Adjuvant chemotherapy (%)	26.23 (43.99)	28.66 (45.22)
Hormone therapy (%)	73.24 (44.27)	73.01 (44.39)
Radiation therapy (%)	64.45 (47.87)	65.41 (47.57)
Patients (N)	163,613	90,389
Providers (N)		3,481

Note. The table shows summary statistics for different samples used in the analysis. Column (1) displays statistics for the sample of patients who had a breast cancer surgery between January 1 2004 to Dec 31 2018. Column (2) displays statistics for a subset of the individuals in column (1). This sample is the sample used for the event-study analysis.

Table 2: Effect of Technology Adoption on Inputs to the Classification Matrix

	(1) Tested	(2) Chemotherapy	(3) True negative	(4) False negative
Adopted	0.104***	-0.011**	0.010*	0.001
	(0.004)	(0.005)	(0.006)	(0.003)
Observations	97,164	97,164	81,526	81,526

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note. The table shows the results from estimating Equation (2) in the event study sample in columns 1-2, and restricted to surgeries before 2014 in columns 3-4 to allow for a five-year follow-up period. The model includes controls for a patient's disability status, an indicator for living in a metropolitan area, NCI comorbidity tertiles, medicaid status, health service area fixed effects, patient age (binned in 5-year intervals), and various measures of cancer characteristics, including indicators for the tumor being well-differentiated, neoadjuvant chemotherapy use, localized cancer, HER2 positive treatment use, and the cancer being classified as metastatic within 90 days of surgery. Standard errors are clustered at the provider level. *, **, ***, indicate significance at the 10%, 5% and 1% level.

Table 3: Effect of Technology Adoption on Inputs to the Classification Matrix for Black Relative to White Patients

	(1) Tested	(2) Chemotherapy	(3) True negative	(4) False negative
Adopted	0.108***	-0.010*	0.010*	-0.002
	(0.004)	(0.007)	(0.005)	(0.006)
Black	-0.001	0.044***	-0.056***	-0.017***
	(0.003)	(0.010)	(0.001)	(0.007)
Adopted ×Black	-0.055***	0.003	-0.001	-0.011
	(0.006)	(0.011)	(0.012)	(0.007)
Observations	97,164	97,164	81,526	81,526

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note. The table shows the results from estimating equation 2 interacted with an indicator for being Black, in the event study sample in columns 1-2, and restricted to surgeries before 2014 in columns 3-4 to allow for a five-year follow-up period. The model includes controls for a patient's disability status, an indicator for living in a metropolitan area, NCI comorbidity tertiles, medicaid status, health service area fixed effects, patient age (binned in 5-year intervals), and various measures of cancer characteristics, including indicators for the tumor being well-differentiated, neoadjuvant chemotherapy use, localized cancer, HER2 positive treatment use, and the cancer being classified as metastatic within 90 days of surgery. Standard errors are clustered at the provider level. *, **, ***, indicate significance at the 10%, 5% and 1% level.

Table 4: Decomposing the Within-Provider Racial Gap in Testing

Panel A. Baseline	**Gap**	**SE**	**Change in Gap**
Unadjusted	0.092	0.010	
Age-adjusted	0.090	0.010	baseline
Panel B: Medical Appropriateness			
Early stage (I-IIIA)	0.079	0.010	13%
Hormone positive (ER/PR)	0.058	0.009	36%
HER2 negative	0.084	0.010	8%
Node negative	0.078	0.001	14%
NCI Comorbidity Tertile	0.088	0.0010	4%
All medical appropriateness variables	0.046	0.010	49%
Panel C: Other			
Dual eligible	0.082	0.010	10%
Health Service Area	0.093	0.010	-2%
SEER registry	0.093	0.010	-2%
All other variables	0.084	0.010	7%
All variables	0.040	0.01	56%

Note. The table shows the results from estimating Equation (5) in a subset of the event-study sample; surgeries between 2010-2018 in a subset of SEER registries with detailed information on cancer characteristics and only including providers who saw at least one black patient and had already adopted the technology by the time they say their patients. The average share of patients tested during this period was 0.26. The column "Gap" displays β from equation 5, "SE" displays the associated standard error, and "Change in Gap", displays the change in the estimated gap when controlling for the characteristics displayed in Panels B and C, relative to the baseline. The "baseline" corresponds to the racial gap when controlling for 5-year patient age indicators measured at the time point of surgery. In Panel B, each row presents the results of estimating Equation (5) with a particular medical appropriateness variable included as a control in the regression (one at a time and not jointly), and similarly for Panel C. Standard errors are clustered at the provider-level.

Table 5: Decomposing the Overall Racial Gap in Testing

	White	Black	Gap	Log	Share of gap
Conditional probability of testing	0.20	0.14	0.67	-0.40	100%
Access	0.97	0.95	0.98	-0.02	5%
Medical Appropriateness	0.50	0.40	0.80	-0.22	56%
Testing	0.42	0.36	0.85	-0.16	39%

Note. The table shows empirical estimates of access ($A_g \equiv P(A_g = 1)$), medical appropriateness ($I_g \equiv P(I_g = 1 \mid A_g = 1)$), and testing ($T_g \equiv P(T_g = 1 \mid I_g = 1, A_g = 1)$) as defined in Appendix A.6. The empirical estimates are obtained by estimating regressions of the form $Y_i = \beta \text{Black}_i + \delta X_i' + \epsilon_i$, where Y_i is 1) having access to an adopting provider, 2) being medically appropriate for the technology, and 3) being tested. Each regression includes fixed effects for health service area, patient age dummies (binned in 5-year intervals), and surgery year fixed effects. The sample for the first regression (Access) is a subset of the event-study sample; surgeries between 2010-2018 in a subset of SEER registries with detailed information on cancer characteristics and only including providers who saw at least one black patient (N=19,675 patients, N=1,284 providers). The sample for the second regression (Medical appropriateness), is limited to patients who saw an adopting provider. The sample for the third regression (Testing), is further limited to patients who were the most medically appropriate for the test. The second and third regression also include provider fixed effects, with standard errors clustered at the provider level. β is used to compute the predicted value of the outcome for white and Black patients separately, presented in the first two columns. The third column (Gap), presents the ratio of the two. The fourth column presents the log-transformed value of the gap. The top row (Conditional probability of testing) is obtained by multiplying consecutive rows within a column. The final column (Share of gap), assigns the relative contribution of each row (Access, Medical Appropriateness, Testing) to the total racial gap in testing.

A Appendix

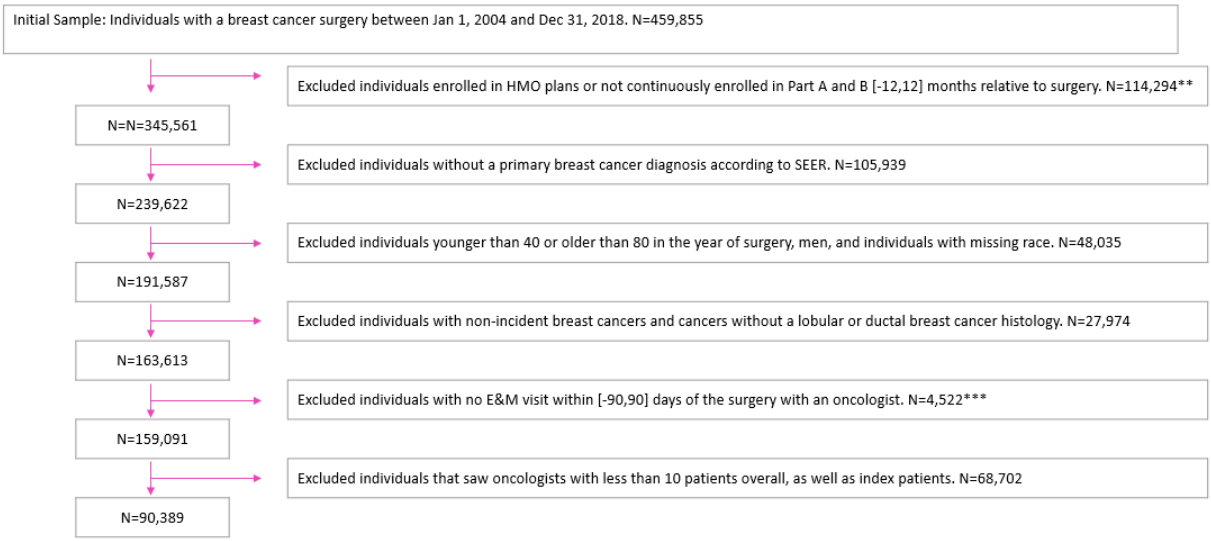
A.1 Measurement and Sample Construction

Table A.1: Measurement

Main treatments	**Codes**
Breast cancer surgery	HCPCS codes: 19300-19307, 19140, 19160, 19162, 19180, 19182, 19200, 19220, 19240
Chemotherapy	HCPCS and NDC codes from CanMED
Hormone therapy	HCPCS and NDC codes from CanMED
Radiation therapy	HCPCS codes: 77261-77799
Antiemetic drugs	HCPCS and NDC codes from CanMED
Immunostimulant drugs	HCPCS and NDC codes from CanMED

Note. The table lists the codes used to identify the main treatments studied. CanMED refers to the Cancer Medications Enquiry Database. Source: Cancer Medications Enquiry Database (CanMED). Surveillance Research Program SEER website tool. Division of Cancer Control and Population Sciences, National Cancer Institute. Available at https://seer.cancer.gov/oncologytoolbox/canmed/. Accessed on January 13, 2023.

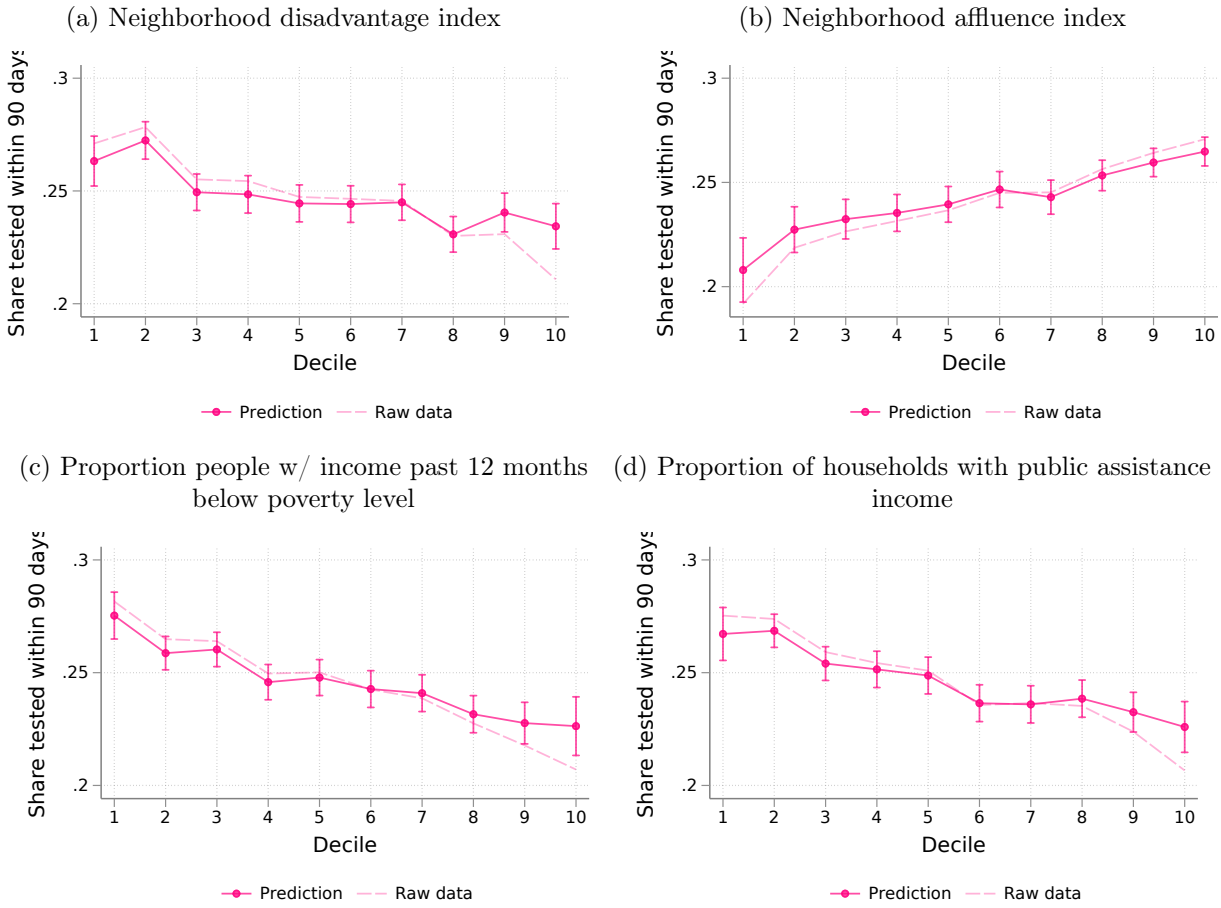
Figure A.1: Sample Construction



Note. The figure displays a flow chart of the sample construction process.

A.2 Descriptives on Technology Diffusion

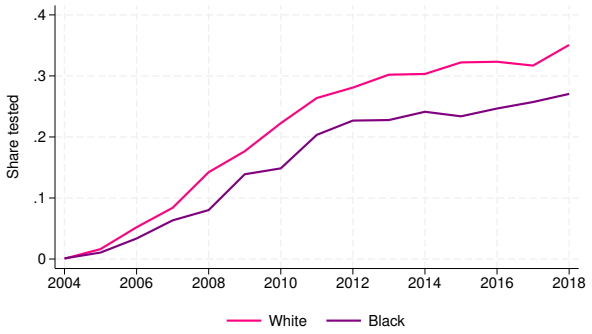
Figure A.2: Socioeconomic Gradients in Genomic Testing - 2007-2019



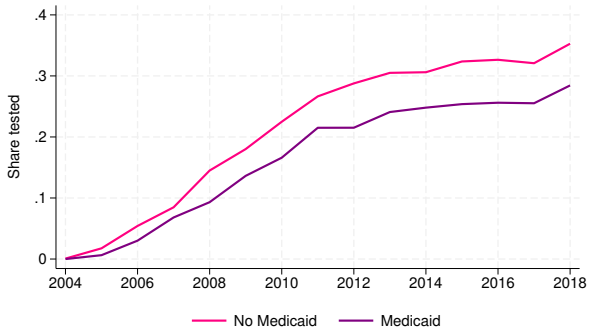
Note. The figure shows the relationship between testing within 90 days of surgery and four different measures of area-level socioeconomic (dis)advantage in the alternative surgery sample, using a 20 percent sample of unlinked Medicare patients. The raw data shows the average share of patients tested within 90 days of surgery in each decile of a given measure of socioeconomic (dis)advantage (e.g., the neighborhood disadvantage index). "Prediction" refers to the prediction from a regression that adjusts for age at surgery, comorbidity history, and HER2 status.

Figure A.3: Diffusion of Genomic Testing by Race and SES

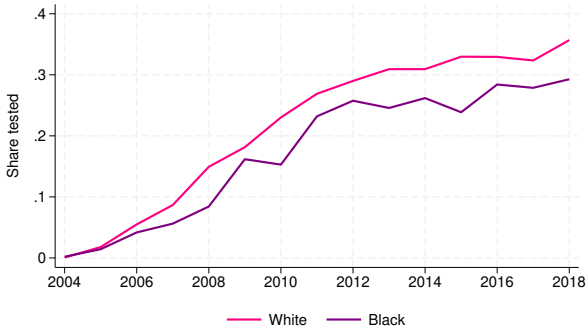
(a) Black vs white



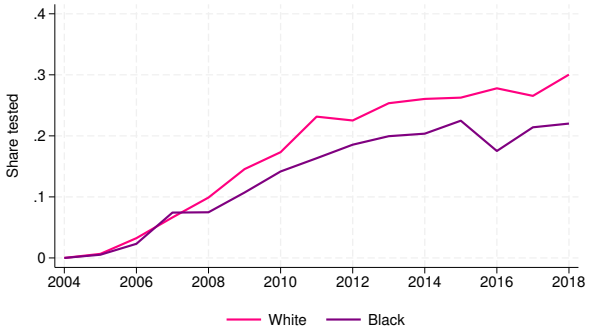
(b) Duals vs non-duals



(c) Black vs white among non-duals



(d) Black vs white among duals



Note. The figure shows the share of patients tested within 90 days of their surgery over time in the surgery sample (see column 1 of Table 1 for characteristics of this sample), for different patient populations. Panel (a) shows the share of Black and white patients tested over time. Panel (b) shows the testing share for individuals who are enrolled in Medicare only (non-duals) or both Medicare and Medicaid (duals). Panels (c) and (d) show the testing share for Black and white patients separately for non-dual eligible patients (panel c), and dual-eligible (panel d).

Figure A.4: Diffusion of Genomic Testing by Other Patient Characteristics

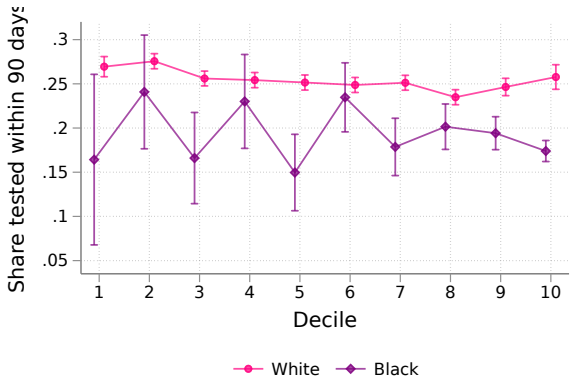
(a) Originally enrolled through old-age vs disability

(b) Metropolitan vs non-metropolitan county of residence

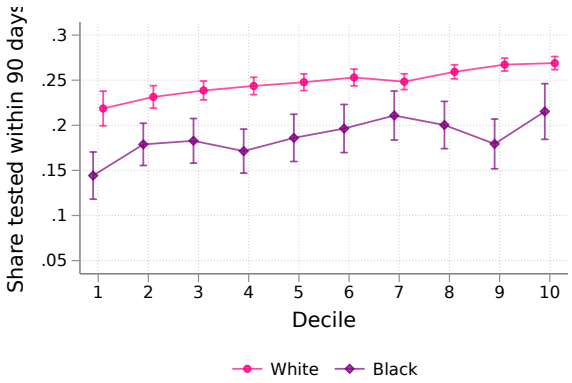
Note. The figure shows the share of patients tested within 90 days of their surgery over time in the surgery sample (see column 1 of Table 1 for characteristics of this sample), for different patient populations. Panel (a) shows the testing share for individuals who originally enrolled in Medicare by virtue of reaching the age of 65 (OASI) or through a disability (SSDI). Panel (b) shows the testing share for individuals living in metropolitan counties and non-metropolitan counties.

Figure A.5: Socioeconomic Gradients in Genomic Testing by Race - 2007-2019

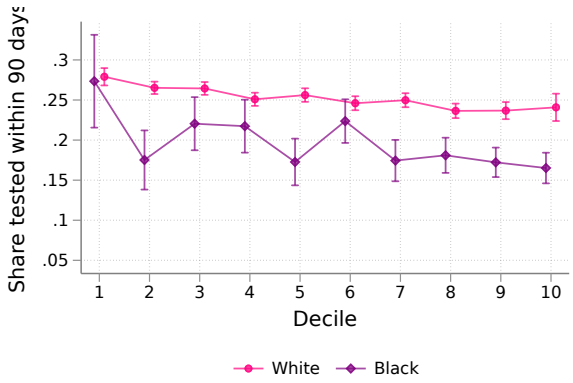
(a) Neighborhood disadvantage index



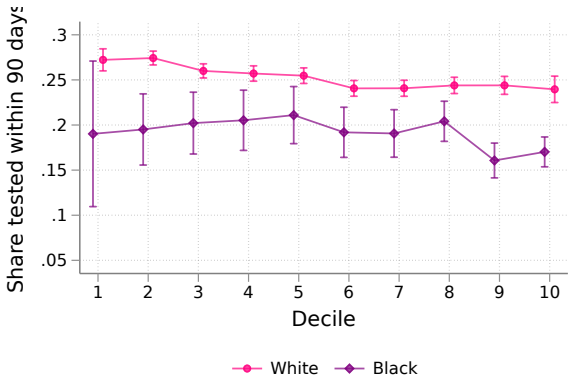
(b) Neighborhood affluence index



(c) Proportion people w/ income past 12 months below poverty level

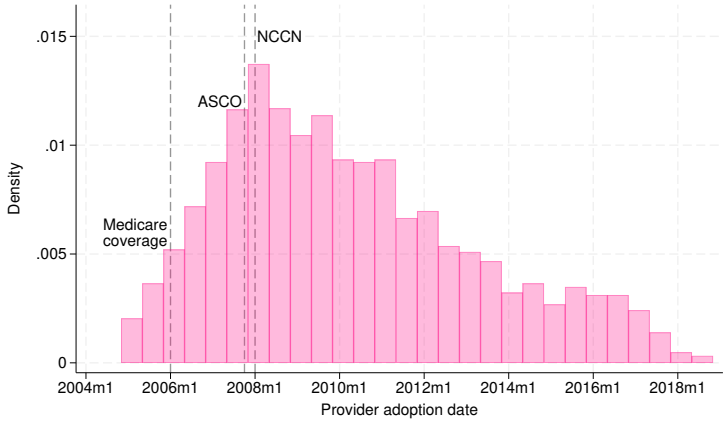


(d) Proportion of households with public assistance income



Note. The figure shows the relationship between testing within 90 days of surgery and four different measures of area-level socioeconomic (dis)advantage in the alternative surgery sample, separately for Black and white patients, using a 20 percent sample of Medicare patients. The raw data shows the average share of patients tested within 90 days of surgery in each decile of a given measure of socioeconomic (dis)advantage (e.g., the neighborhood disadvantage index). "Prediction" refers to the prediction from a regression that adjusts for age at surgery, comorbidity history, and HER2 status.

Figure A.6: Provider Variation in the Timing of Technology Adoption

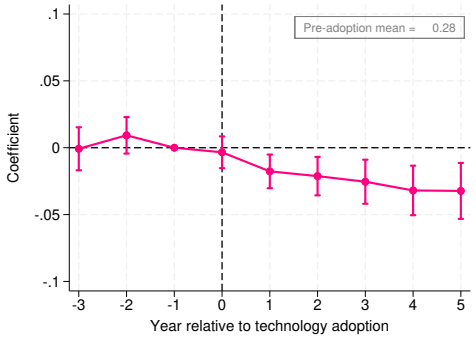


Note. The figure shows the distribution of the time of adoption of oncologists in the analysis sample (N=3,430 adopting oncologists). The gray dashed vertical lines indicate three distinct events; Medicare's decision to cover the test in 2006, the American Society of Clinical Oncology (ASCO) 2007 Update of Recommendation for the Use of Tumor Markers in Breast Cancer (published on-line in November 2007), and the inclusion of the test in the National Comprehensive Cancer Network (NCCN) Clinical Practice Guidelines in Oncology, Breast Cancer (version 1 published in January 2008).

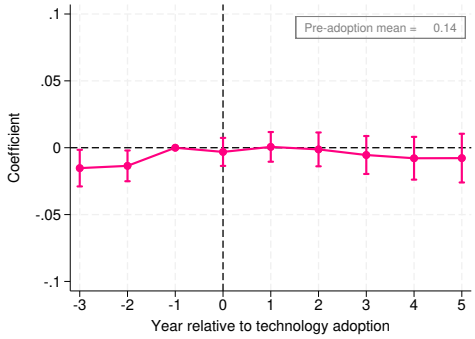
A.3 Additional event studies

Figure A.7: Effect of Technology Adoption on the Use of Specific Chemotherapy Drugs

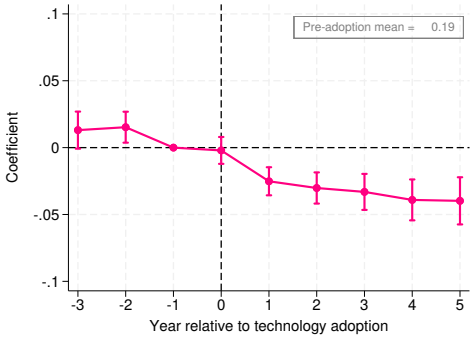
(a) Cyclophosphamide



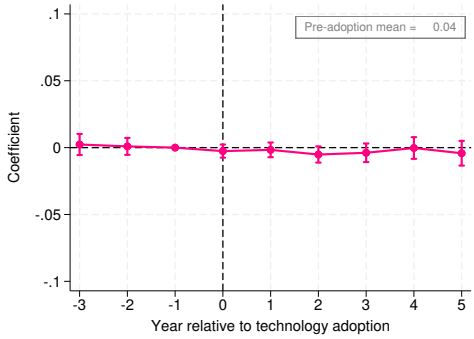
(b) Docetaxel



(c) Doxorubicin



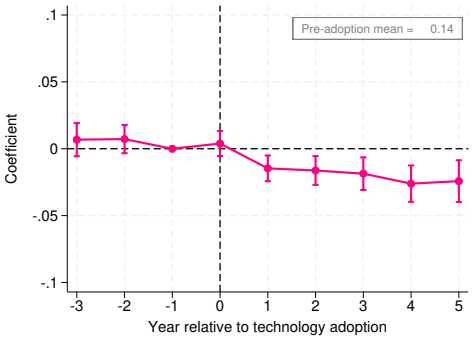
(d) Methotrexate



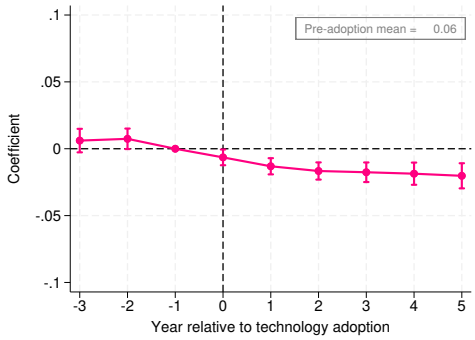
Note. The figure shows the impact of provider technology adoption on the likelihood that a patient is treated with a specific chemotherapy drug. The figure depicts the β coefficients from estimating the event study model specified in Equation (1) in the event study sample (see column (2) in Table 1 for summary statistics of this sample). The pre-adoption average of the outcome variable is calculated in the sample of adopting providers and displayed in the upper right corner of the graph. The event study model includes controls for a patient's disability status, an indicator for living in a metropolitan area, NCI comorbidity tertiles, health service area fixed effects, patient age (binned in 5-year intervals), and various measures of cancer characteristics, including indicators for the tumor being well-differentiated, neoadjuvant chemotherapy use, localized cancer, HER2 positive treatment use, and the cancer being classified as metastatic within 90 days of surgery. Standard errors are clustered at the provider level.

Figure A.8: Effect of Technology Adoption on the Use of Specific Chemotherapy Regimes and Classes of Drugs

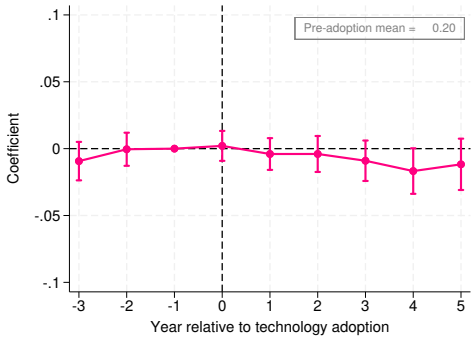
(a) ACT regimes



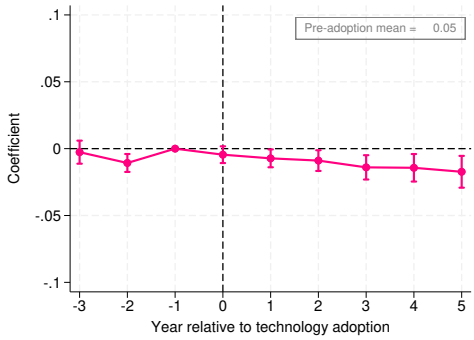
(b) AC regimen without a taxane



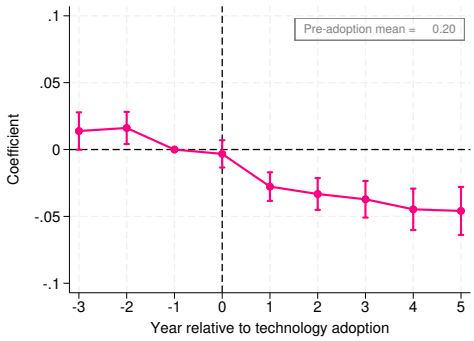
(c) TC regimes



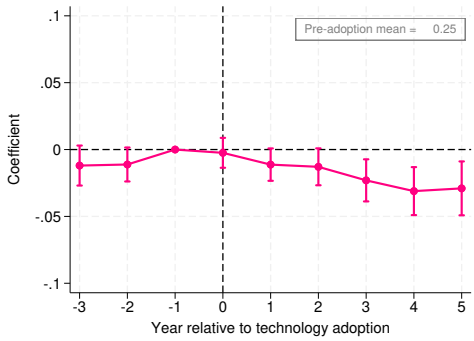
(d) Taxane without cyclophosphamide



(e) Any anthracycline



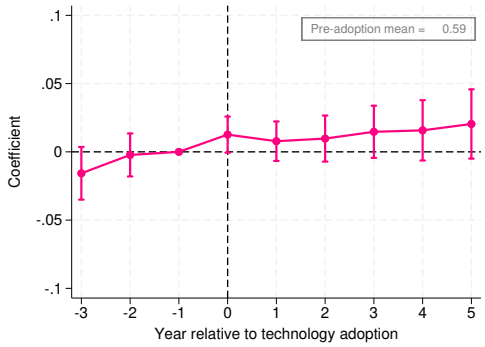
(f) Any taxane



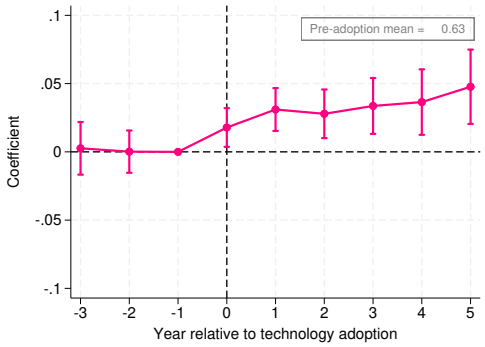
Note. The figure shows the impact of provider technology adoption on the likelihood that a patient is treated with a particular chemotherapy drug class or regime. The figure depicts the β coefficients from estimating the event study model specified in Equation (1) in the event study sample (see column (2) in Table 1 for summary statistics of this sample). The pre-adoption average of the outcome variable is calculated in the sample of adopting providers and displayed in the upper right corner of the graph. The event study model includes controls for a patient's disability status, an indicator for living in a metropolitan area, NCI comorbidity tertiles, health service area fixed effects, patient age (binned in 5-year intervals), and various measures of cancer characteristics, including indicators for the tumor being well-differentiated, neoadjuvant chemotherapy use, localized cancer, HER2 positive treatment use, and the cancer being classified as metastatic within 90 days of surgery. Standard errors are clustered at the provider level.

Figure A.9: Effect of Technology Adoption on Non-Targeted Treatments

(a) Hormone therapy within one year



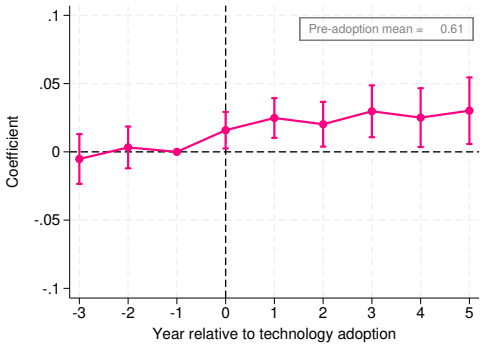
(b) Radiation therapy within one year



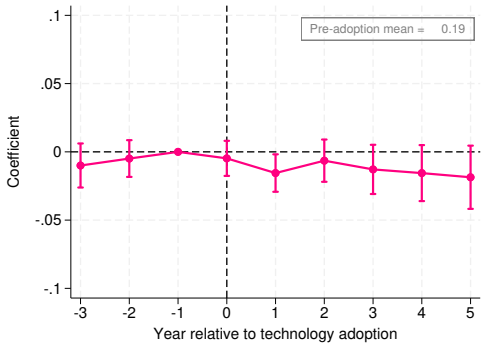
Note. The figure shows the impact of provider technology adoption on the likelihood that a patient is treated with hormone therapy (panel a) and radiotherapy (panel b) within one year of the breast cancer surgery, respectively. The figure depicts the β coefficients from estimating the event study model specified in Equation (1) in the event study sample (see column (3) in Table 1 for summary statistics of this sample). The pre-adoption average of the outcome variable is calculated in the sample of adopting providers and displayed in the upper right corner of the graph. The event study model includes controls for a patient's disability status, an indicator for living in a metropolitan area, NCI comorbidity tertiles, health service area fixed effects, patient age (binned in 5-year intervals), and various measures of cancer characteristics, including indicators for the tumor being well-differentiated, neoadjuvant chemotherapy use, localized cancer, HER2 positive treatment use, and the cancer being classified as metastatic within 90 days of surgery. Standard errors are clustered at the provider level.

Figure A.10: Effect of Technology Adoption on Surgical Choice

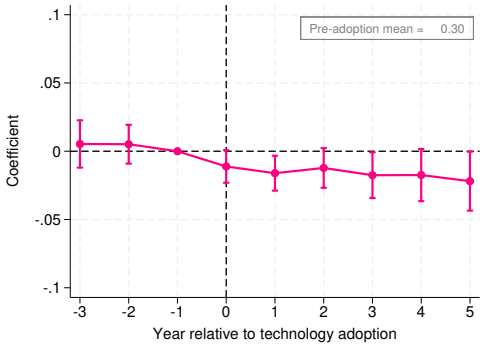
(a) Lumpectomies (least Intensive)



(b) Complete mastectomy (more intensive)



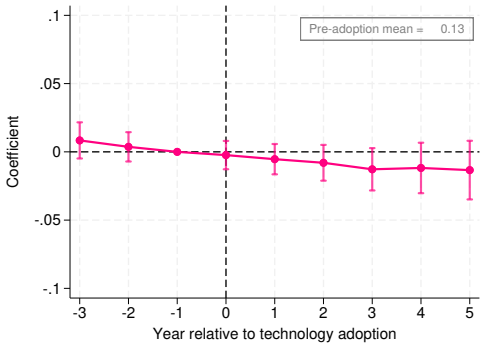
(c) Radical mastectomy (most intensive)



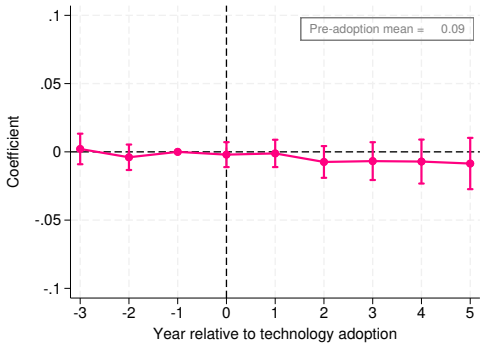
Note. The figure shows the impact of provider technology adoption on the likelihood that a patient undergoes a lumpectomy (panel a), a complete mastectomy (panel b), or a radical mastectomy (panel c). The figure depicts the β coefficients from estimating the event study model specified in Equation (1) in the event study sample (see column (2) in Table 1 for summary statistics of this sample). The pre-adoption average of the outcome variable is calculated in the sample of adopting providers and displayed in the upper right corner of the graph. The event study model includes controls for a patient's disability status, an indicator for living in a metropolitan area, NCI comorbidity tertiles, health service area fixed effects, patient age (binned in 5-year intervals), and various measures of cancer characteristics, including indicators for the tumor being well-differentiated, neoadjuvant chemotherapy use, localized cancer, HER2 positive treatment use, and the cancer being classified as metastatic within 90 days of surgery. Standard errors are clustered at the provider level.

Figure A.11: Effect of Technology Adoption on Proxies for Cancer Recurrence

(a) Metastasis within 2-5 years



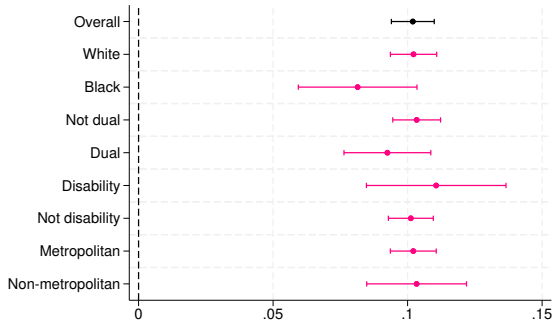
(b) Chemotherapy within 2-5 years



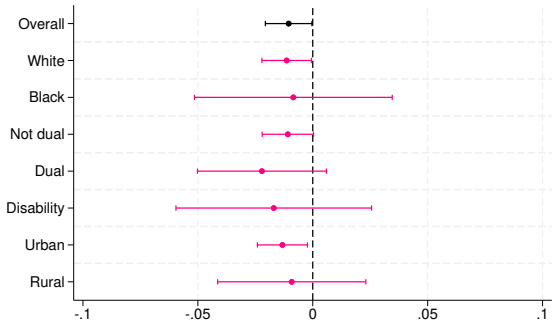
Note. The figure shows the impact of provider technology adoption on the likelihood that a patient experiences a metastasis (panel a) and is treated with chemotherapy within 2-5 year (panel b) of the breast cancer surgery, respectively. The figure depicts the β coefficients from estimating the event study model specified in Equation (1) in the event study sample, restricted to surgeries before 2014 to allow for a five-year follow-up period. The pre-adoption average of the outcome variable is calculated in the sample of adopting providers and displayed in the upper right corner of the graph. The model includes controls for a patient's disability status, an indicator for living in a metropolitan area, NCI comorbidity tertiles, health service area fixed effects, patient age (binned in 5-year intervals), and various measures of cancer characteristics, including indicators for the tumor being well-differentiated, neoadjuvant chemotherapy use, localized cancer, HER2 positive treatment use, and the cancer being classified as metastatic within 90 days of surgery. Standard errors are clustered at the provider level.

Figure A.12: Treatment Effect Heterogeneity by Patient Characteristics

(a) Genomic testing within 90 days



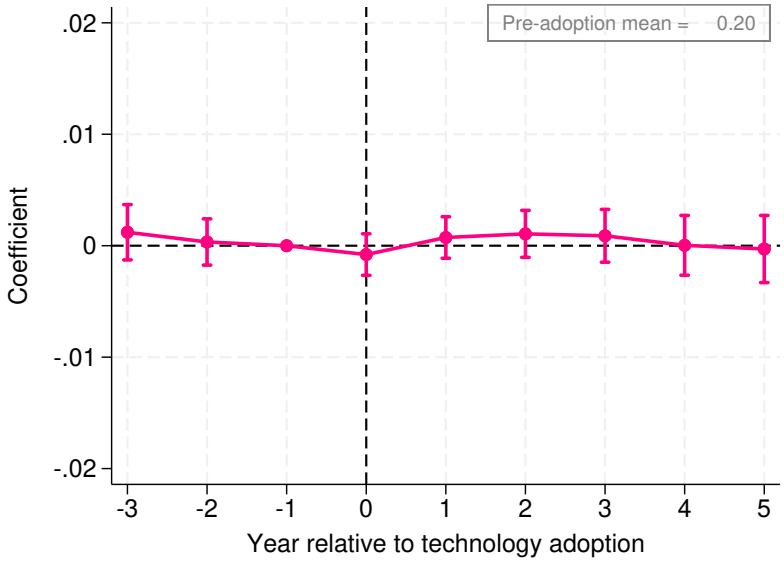
(b) Chemotherapy within one year



Note. Each panel displays the difference-in-difference coefficient for the outcome variable (e.g., testing), in the entire sample (denoted "overall" - see column 1 of Table 1 for characteristics of this sample) as well as in subsamples (e.g. white). The model includes controls for a patient comorbidities (as measured by NCI comorbidity tertiles), health service area fixed effects, patient age (binned in 5-year intervals), and various measures of cancer characteristics, including indicators for the tumor being well-differentiated, neoadjuvant chemotherapy use, localized cancer, HER2 positive treatment use, and the cancer being classified as metastatic within 90 days of surgery. Standard errors are clustered at the provider level.

A.4 Robustness

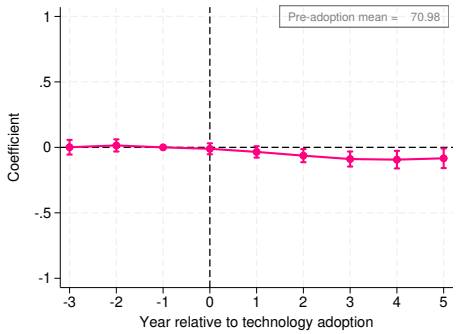
Figure A.13: Technology Adoption and Predicted Probability of Chemotherapy Treatment



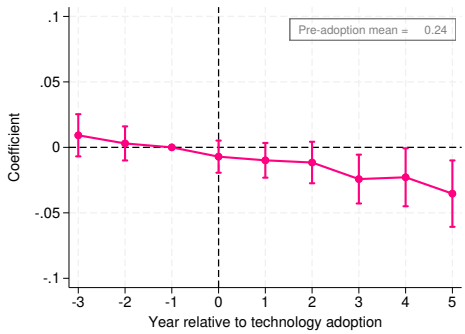
Note. The figure shows the impact of provider technology adoption on the predicted probability of chemotherapy treatment within one year of surgery, estimated in a regression that includes a patient's year of surgery, age, NCI comorbidity tertile, a patient's disability status, an indicator for living in a metropolitan area, patient age, and various measures of cancer characteristics, including indicators for the tumor being well-differentiated, neoadjuvant chemotherapy use, localized cancer, HER2 positive treatment use, and the cancer being classified as metastatic within 90 days of surgery. The figure depicts the β coefficients from estimating the event study model specified in Equation (1) in the event study sample (see column (2) in Table 1 for summary statistics of this sample). The pre-adoption average of the outcome variable is calculated in the sample of adopting providers and displayed in the upper right corner of the graph. Standard errors are clustered at the provider level.

Figure A.14: Technology Adoption and Patient Characteristics

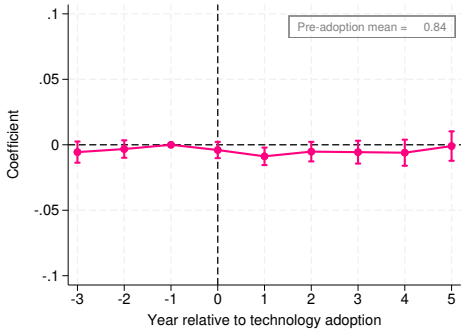
(a) Patient age at surgery



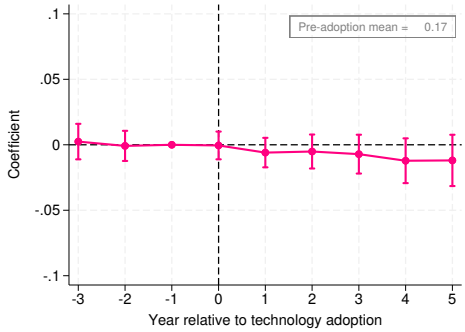
(b) NCI comorbidity score



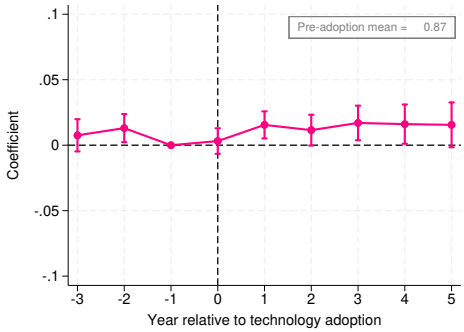
(c) Patient residing in urban county



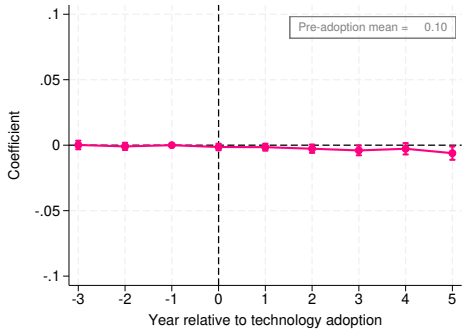
(d) Dual-eligible



(e) White patient



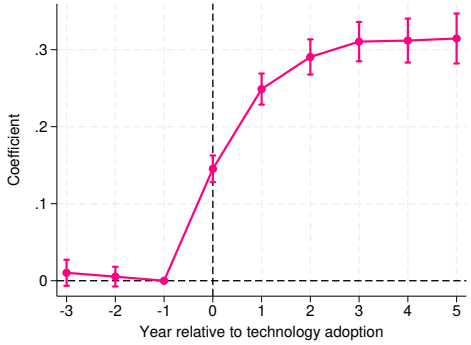
(f) Disability



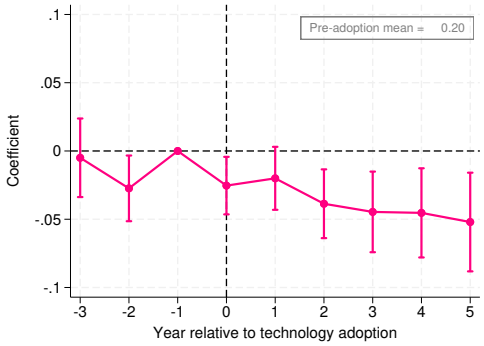
Note. The figure shows the impact of provider technology adoption on various patient characteristics. The figure depicts the β coefficients from estimating the event study model specified in Equation (1) in the event study sample. The pre-adoption average of the outcome variable is calculated in the sample of adopting providers and displayed in the upper right corner of the graph. Standard errors are clustered at the provider level.

Figure A.15: Effect of Technology Adoption on Testing and Targeted Treatments - Subsample of Early-Stage, Hormone Positive, HER2 Negative, Lymph Node Negative Cancers

(a) Genomic testing within 3 months



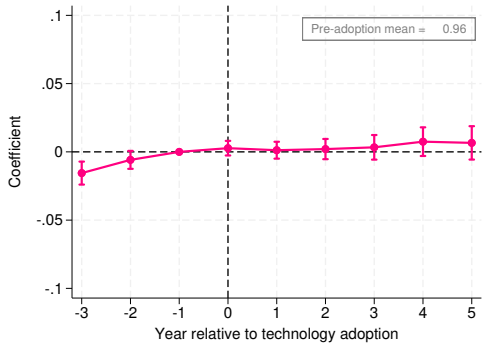
(b) Chemotherapy within 1 year



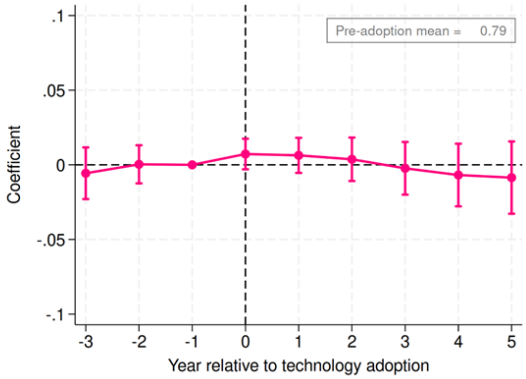
Note. The figure shows the impact of provider technology adoption on the likelihood that a patient is tested around the time point of their surgery (panel a) and treated with chemotherapy within one year (panel b) of the breast cancer surgery, respectively, in the subset of patients commonly recommended for testing by clinical guidelines. The figure depicts the β coefficients from estimating the event study model specified in Equation (1). The pre-adoption average of the outcome variable is calculated in the sample of adopting providers and displayed in the upper right corner of the graph. The event study model includes controls for a patient's disability status, an indicator for living in a metropolitan area, NCI comorbidity tertiles, health service area fixed effects, patient age (binned in 5-year intervals), and various measures of cancer characteristics, including indicators for the tumor being well-differentiated, neoadjuvant chemotherapy use, localized cancer, HER2 positive treatment use, and the cancer being classified as metastatic within 90 days of surgery. Standard errors are clustered at the provider level.

Figure A.16: Technology Adoption and the Use of Other Technologies

(a) Immunohistochemistry testing within [-5,5] days of surgery



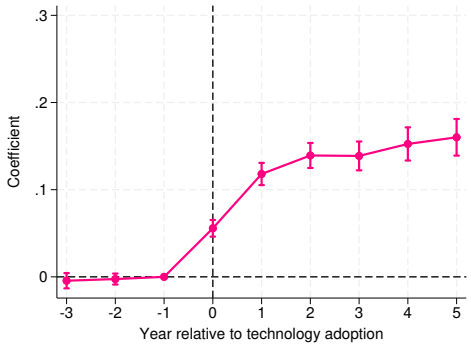
(b) Sentinel lymph node biopsy within [-90,90] days of surgery



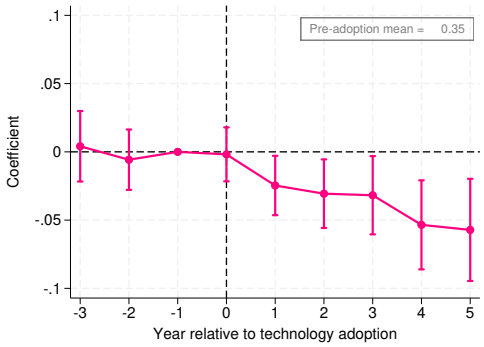
*Note.*The figure shows the impact of provider technology adoption on the likelihood that a patient receives immunohistochemistry testing (panel a) and a sentinel lymph node biopsy (panel b). The figure depicts the β coefficients from estimating the event study model specified in Equation (1) in the event study sample (see column (2) in Table 1 for summary statistics of this sample). The pre-adoption average of the outcome variable is calculated in the sample of adopting providers and displayed in the upper right corner of the graph. The event study model includes controls for a patient's disability status, an indicator for living in a metropolitan area, NCI comorbidity tertiles, health service area fixed effects, patient age (binned in 5-year intervals), and various measures of cancer characteristics, including indicators for the tumor being well-differentiated, neoadjuvant chemotherapy use, localized cancer, HER2 positive treatment use, and the cancer being classified as metastatic within 90 days of surgery. Standard errors are clustered at the provider level.

Figure A.17: Technology Adoption and Treatment Decisions - Control vs Treated Diseases

(a) Genomic testing within 90 days



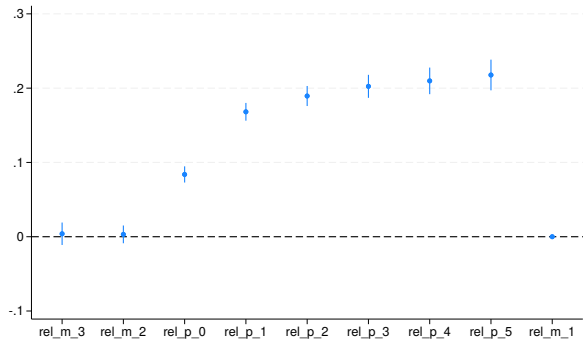
(b) Any chemotherapy within 1 year



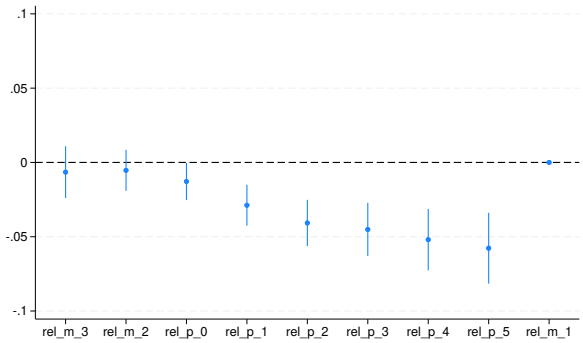
Note. The figure shows the impact of provider technology adoption on the likelihood that a patient is tested around the time point of their surgery and treated with chemotherapy within one year of the breast cancer surgery, respectively, for patients with early-stage cancers ("treated") relative to patients with late-stage cancers ("control"). The figure depicts the β coefficients from estimating the event study model specified in Equation (3) in the event-study sample. The pre-adoption average of the outcome variable is calculated in the sample of adopting providers and displayed in the upper right corner of the graph. The event study model includes controls for a patient's disability status, an indicator for living in a metropolitan area, NCI comorbidity tertiles, health service area fixed effects, patient age (binned in 5-year intervals), and various measures of cancer characteristics, including indicators for the tumor being well-differentiated, neoadjuvant chemotherapy use, HER2 positive treatment use. Standard errors are clustered at the provider level.

Figure A.18: Effect of Technology Adoption on Testing and Targeted Treatments using the Sun and Abraham Estimator

(a) Genomic testing within 90 days



(b) Any chemotherapy within 1 year



Note. The figure shows the impact of provider technology adoption on the likelihood that a patient is tested around the time point of their surgery and treated with chemotherapy within one year of the breast cancer surgery, respectively. The regression model includes controls for a patient's disability status, an indicator for living in a metropolitan area, NCI comorbidity tertiles, health service area fixed effects, patient age (binned in 5-year intervals), and various measures of cancer characteristics, including indicators for the tumor being well-differentiated, neoadjuvant chemotherapy use, HER2 positive treatment use. Standard errors are clustered at the provider level.

A.5 Deriving the *change* in the True Positive and False Positive Rate

First, note that the true positive rate and the false positive rate are defined as follows:

$$\text{TPR} = \frac{\text{TP}}{\text{TP} + \text{FN}}$$

$$\text{FPR} = \frac{\text{FP}}{\text{FP} + \text{TN}}$$

Moreover, the following set of relationships hold for the classification matrix

$$\text{TP} = \text{S} - \text{FN} \tag{7}$$

$$\text{FP} = \text{P} - \text{TP} \tag{8}$$

$$\text{TN} = 1 - \text{FN} - \text{TP} - \text{FP} \tag{9}$$

Assume there are two time periods, $t \in \{0, 1\}$ and assume that the share of the population that is actually high-risk does not change over time. Then, it follows that the *change* in the TPR can be written as follows

$$\begin{aligned}
\Delta\text{TPR} &= \frac{\text{TP}_1}{\text{TP}_1 + \text{FN}_1} - \frac{\text{TP}_0}{\text{TP}_0 + \text{FN}_0} \\
&= \frac{\text{S}_1 - \text{FN}_1}{\text{S}_1 - \text{FN}_1 + \text{FN}_1} - \frac{\text{S}_0 - \text{FN}_0}{\text{S}_0 - \text{FN}_0 + \text{FN}_0} \\
&= \frac{\text{S}_1 - \text{FN}_1}{\text{S}_1} - \frac{\text{S}_0 - \text{FN}_0}{\text{S}_0} \\
&= \frac{-\text{FN}_1 + \text{FN}_0}{S} \\
&= \frac{-\Delta\text{FN}}{\text{S}}
\end{aligned}$$

where the second line follows from Equation (7) and the last line follows from the assumption that $\text{S}_0 = \text{S}_1$.

The *change* in the FPR can be written as follows

$$
\begin{aligned}
\Delta\text{FPR} &= \frac{\text{FP}_1}{\text{FP}_1 + \text{TN}_1} - \frac{\text{FP}_0}{\text{FP}_0 + \text{TN}_0} \\
&= \frac{\text{FP}_1}{1 - \text{FN}_1 - \text{TP}_1 - \text{TN}_1 + \text{TN}_1} - \frac{\text{FP}_0}{1 - \text{FN}_0 - \text{TP}_0 - \text{TN}_0 + \text{TN}_0} \\
&= \frac{\text{FP}_1}{1 - \text{FN}_1 - \text{TP}_1} - \frac{\text{FP}_0}{1 - \text{FN}_0 - \text{TP}_0} \\
&= \frac{\text{FP}_1}{1 - \text{FN}_1 - (\text{S}_1 - \text{FN}_1)} - \frac{\text{FP}_0}{1 - \text{FN}_0 - (\text{S}_0 - \text{FN}_0)} \\
&= \frac{\text{FP}_1}{1 - \text{S}_1} - \frac{\text{FP}_0}{1 - \text{S}_0} \\
&= \frac{\text{FP}_1 - \text{FP}_0}{1 - \text{S}} \\
&= \frac{(\text{P}_1 - \text{TP}_1) - (\text{P}_0 - \text{TP}_0)}{1 - \text{S}} \\
&= \frac{(\text{P}_1 - \text{S} + \text{FN}_1) - (\text{P}_0 - \text{S} + \text{FN}_0)}{1 - \text{S}} \\
&= \frac{\Delta\text{P} + \Delta\text{FN}}{1 - \text{S}}
\end{aligned}
$$

where the second line follows from Equation (9), the fourth line follows from Equation (7), the sixth line follows from the assumption that $\text{S}_0 = \text{S}_1$, the seventh line follows from Equation (8), and the eight line follows from Equation (7).

A.5.1 Benchmark to clinical trial

How does this result compare to the evidence from clinical trials? I perform a simple back of the envelope calculation. Scaling the chemotherapy reduction (approximately 15 percent decrease) by the testing increase (approximately 122 percent increase), suggest an approximately 12 percent reduction (-0.15/1.22) in chemotherapy use. Prior research suggests the distribution of Oncotype risk scores in a population of 66-80 years old was: 13.5% high-risk, 59.3% intermediate-risk, 27.2% low-risk.[31] The risk-score distribution is relatively similar for ages 36-66. Moreover, as previously noted, the TAILORx estimates suggest the rate of over and undertreatment in the high and low risk groups of, 25 and 43 percent, respectively. Before 2018, there was little guidance on how to treat patients with intermediate-level risk score. If one assumes the arrival of the technology led to a 43 percent increase in treatment in the high-risk group, a 25 percent reduction in treatment in the low-risk group, and no change in treatment in the intermediate group, this would correspond to a lower bound of a 1 percentage point reduction at the population level. Assuming instead that the arrival of the technology led to a 50 percent reduction in the chemotherapy rate in the intermediate risk-score would imply a 31 percentage point reduction in chemotherapy use at the population level.

[31] https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7663955/

A.6 Conceptual framework

Let G denote (one of the) gains(s) from using the technology. In the current setting, the gain could be a true negative, and the technology is a genomic test. Assume there are two groups $g \in [a, b]$, and assume that the likelihood that patients within a group gain from innovation is a function of the likelihood that patients within the group 1) have access to a provider who uses the technology (A_g), conditional on access, are appropriate or indicated for the technology (I_g) , 3) conditional on access and appropriateness, are tested (T_g) , and 4) conditional on all of the above, experience a good outcome (G_g) as opposed to a bad one (B_g). In the current setting, the good outcome could be a true negative, and the bad outcome could be a false negative. This can be visualized in a patient/provider decision tree (see the below), where the ultimate probability of a good outcome is a product of the probabilities of travelling down the pink branch. Denote this branch probability $\hat{G}$. In other words:

$$P(\hat{G}_g = 1) =$$

$$P(A_g = 1) \times P(I_g = 1|A_g = 1) \times P(T_g = 1|A_g = 1, I_g = 1) \times P(G_g = 1|A_g = 1, I_g = 1, T_g = 1)$$

For a group to fully realize the gains from innovation (i.e. for $P(\hat{G}_g) = 1$), it is straightforward to see that this is achieved when all subcomponents are equal to 1; i.e. all patients have access, all patient are appropriate for the technology, all patients are tested if they are appropriate, and all patients derive value from the test if they are tested, conditional on being appropriate. In practice, not all patients will have access to an adopting provider or be indicated for the technology, rendering $P(\hat{G}_g) < 1$.

It is also straightforward to see that differences in the subcomponents may results in greater gains in one group than the other. For instance, in the extreme, assume all patients in group a have access to the technology, are indicated for the technology, are tested, and experience the intended health outcome (e.g. a true negative in this setting). Then $P(\hat{G}_g = 1) = 1$ and group a fully gains from the technology. The corresponding probabilities are identical in group b, except patients in group b have an 80 percent probability of getting tested. Then, $P(\hat{G}_a) = 1$, $P(\hat{G}_b) = 0.8$ and the relative gains from innovation is lower in group b than group a. If a denotes a group of white patients, and b denotes a group of black patients, a racial disparity arises in the gains from innovation:

$$P(\hat{G}_b) < P(\hat{G}_a) \qquad \textbf{(Racial disparity in gains)}$$

Moreover, introducing two time periods, t and $t-1$, the gains from innovation may increasingly accrue to one group relative to the other, if the subcomponents grow at a faster rate in one group:

$$P(\hat{G}_b^t)/P(\hat{G}_b^{t-1}) < P(\hat{G}_a^t)/P(\hat{G}_a^{t-1}) \qquad \textbf{(Widening of a racial disparity in gains)}$$

In particular, a disparity *emerges* if there are no initial differences in gains $P(\hat{G}_b^{t-1}) = P(\hat{G}_a^{t-1})$. Since G is a the product of a set of probabilities, it is also possible to decompose the racial disparity at a given point in time into its subcomponent. More specifically, let $G_g \equiv P(G_g = 1)$, $A_g \equiv P(A_g = 1)$, $I_g \equiv P(I_g = 1|A_g = 1)$, $T_g \equiv P(T_g = 1|I_g = 1, A_g = 1)$, and $V_g \equiv P(G_g = 1|T_g = 1, I_g = 1, A_g = 1)$ and log-transforming gives $ln(\hat{G_g}) = ln(A_g) + ln(I_g) + ln(T_g) + ln(V_g)$. The difference in gains is then given by $ln(\hat{G_b}) - ln(\hat{G_a}) = \big(ln(A_b) - ln(A_a)\big) + \big(ln(I_b) - ln(I_a)\big) + \big(ln(T_b) - ln(T_a)\big) + \big(ln(V_b) - ln(V_a)\big)$. In other words, the racial disparity in gains can then be decomposed into the relative importance of 1) access, 2) appropriateness, 3) testing, and 4) value:

$$\Delta\text{gains}_t = \Delta\text{access}_t + \Delta\text{appropriateness}_t + \Delta\text{ testing}_t + \Delta\text{value}_t \qquad \textbf{(Static decomposition of the racial disparity in gains)}$$

It is also possible to decompose the rate of change in the racial disparity from time 0 to time 1 as a function of the rate of change in the subcomponents:

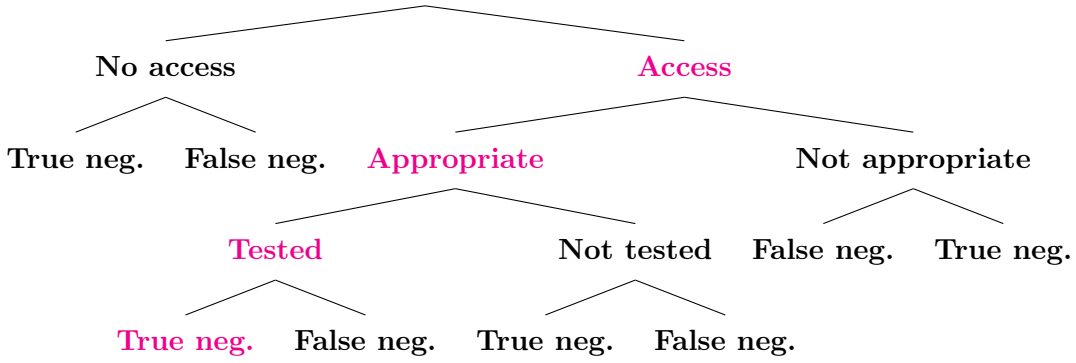
$$\Delta\text{gains}_{t-1,t} = \Delta\text{access}_{t-1,t} + \Delta\text{appropriateness}_{t-1,t} + \Delta\text{ testing}_{t-1,t} + \Delta\text{value}_{t-1,t} \qquad \textbf{(Dynamic decomposition of the racial disparity in gains)}$$

This framework illustrates two broader points. First, it highlights how the gains from medical innovation can vary across patient subpopulations as a function of factors that can to various degrees by shaped by policy. While increasing "access" and "testing" may be the most straightforward factors to influence, factors like "appropriatenss" may become increasingly important and relevant as precision medicine moves from targeting late-stage diseases to earlier stages of disease, and in the extreme, healthy populations (e.g. emerging technologies that screen for cancer). To the extent

that preventive health and screenings ensure timely diagnosis and early detection, making a patient "appropriate" for the technology, the existing case for reducing existing socioeconomic and racial disparities in preventive health becomes increasingly acute.

Second, this framework highlights how the gains from innovation in information technologies, ultimately depend on the value of information that the technology delivers, which may differ across patient groups and providers, both as a result of the inherent characteristics of the technology as well behavioral responses by patients and providers to test results (like adherence).[32] The behavioral dimension that this introduces is somewhat unique to information technologies.

Figure A.19: Clinical pathways to a true negative



[32]This is particularly relevant for the technology studies in this setting where patients/providers face a difficult decision of forgoing a treatment that has the potential to save a person from dying. Risk-aversion will play a critical role when patients decide whether or not to adhere to a recommendation of forgoing chemotherapy.