The Value of (Sub) Specialization: Evidence from Oncology

René Karadakic, David C. Chan, Nancy L. Keating, Bruce E. Landon and Michael L. Barnett*

March 20, 2025

Abstract

Specialization is a key driver of productivity in high-skill professions like medicine, yet its benefits also require geographical access to specialized expertise. Rising "subspecialization" in medicine, defined as a narrow focus within a medical specialty, has transformed health care delivery, but its impact on patient outcomes remains unclear. This paper estimates the causal effect of access to subspecialized oncologists on health care utilization, treatment decisions, and mortality. We exploit quasi-exogenous variation in Medicare fee-for-service patients' differential distance to subspecialized versus general oncologists to identify the effects of subspecialist access. Using a novel sample of 6.1 million chemotherapy episodes, we find that greater access to a subspecialized oncologist of the relevant cancer type reduces three-year mortality by 8.6% relative to the mean mortality in the population and increases disease-specific clinical trial enrollment with no effect on placebo health outcomes unrelated to cancer. Our findings demonstrate how subspecialization facilitates access to experimental treatment and improves cancer outcomes, while also raising questions about geographic disparities in specialized care. These results contribute to the broader literature on specialization, labor markets, and health care productivity by providing new causal evidence on the returns to deep expertise in complex professional services.

JEL Classification: I1, J24

Keywords: Provider Specialization, Mortality, Health Care Access

*Karadakic (corresponding): Department of Health Policy & Management, Harvard T. H. Chan School of Public Health, 677 Huntington Ave, Boston, MA 02115, rkaradakic@hsph.harvard.edu; Chan: Haas School of Business, University of California Berkeley and National Bureau of Economic Research; Keating: Department of Health Care Policy, Harvard Medical School and Department of Medicine, Brigham and Women's Hospital; Landon: Department of Health Care Policy, Harvard Medical School and Department of Medicine, Beth Israel Deaconess Medical Center; Barnett: Department of Health Policy & Management, Harvard T. H. Chan School of Public Health and Department of Medicine, Brigham and Women's Hospital. The authors gratefully acknowledge comments by Christopher Manz, Jukka-Pekka Onnela and Yuhua Zhang. This work was supported by a grant from the National Institute on Aging (R01 AG076580).

1 Introduction

Specialization is a fundamental determinant of the division of labor and the returns to skilled work, shaping how tasks are allocated and influencing productivity and economic outcomes (Smith, 2002). By allowing workers to develop expertise in narrowly defined areas, specialization often enhances efficiency and drives gains in output. However, it can also impose costs, including barriers to access, coordination inefficiencies, and a narrower focus that may neglect broader system-level objectives. These trade-offs raise the question of whether greater access to specialized services improves individual outcomes or if inefficiencies and frictions reduce its benefits.

Despite the importance of specialization in modern economies, its implications within professions remain poorly understood. This paper examines specialization in medicine, focusing on medical oncology, to evaluate how high degrees of specialization affect patient outcomes. Medicine has a long history of specialization, with distinct medical specialties emerging as early as the 19th century (Weisz, 2006). Over time, physicians organized into specialties based on procedural similarities, diagnostic approaches, and disease patterns. More recently, "subspecialization" has emerged, further dividing specialists into narrower fields based on specific disease types. Medical oncology exemplifies this trend, with "subspecialists" focusing on cancers that affect the same anatomical region (e.g. breast cancer or prostate cancer) or share similar treatment approaches (e.g., leukemia and lymphoma). In 2008, 9% of chemotherapy episodes among Medicare fee-for-service patients were managed by subspecialized oncologists; by 2020, this share had doubled to 18%, reflecting the rapid expansion of novel treatments and cancer-specific clinical guidelines (Lozinski, 2024).[1]

Subspecialization can enhance precision of care when providers have the knowledge to tailor treatments to the biological and clinical nuances of specific and even rare, cancers. However, it also presents challenges: achieving economies of scale in specialized care requires large patient volumes, leading to the uneven geographic distribution of subspecialists, with rural and underserved areas often facing limited access (Dingel et al., 2023). Additionally, many cancer patients have multiple chronic conditions and even multiple cancers, making care coordination increasingly complex. These dynamics make medical oncology an ideal setting for studying the trade-offs of specialization on patient outcomes, health care utilization, and access disparities.

In this paper, we exploit exogenous variation in Medicare fee-for-service patients' differential distance to subspecialized oncologists for their specific cancer type versus gen-

[1]See Appendix Figure B1 for an overview in the rise of subspecialization within medical oncology in traditional Medicare between 2008 and 2020.

eral oncologists to estimate the causal effect of subspecialized oncologic care on patient outcomes. The growing trend of oncologic subspecialization, combined with geographic shifts in oncologists' practice locations over time, creates quasi-exogenous variation in access to subspecialists. Using a sample of 6.1 million chemotherapy episodes among 2.2 million patients, we compare individuals within the same ZIP Code Tabulation Area (ZCTA), diagnosed with the same cancer type, whose relative proximity to the relevant subspecialized oncologist versus a general oncologist changes over time. We show that differential distance strongly predicts whether a patient has access to or receives chemotherapy management from a subspecialized oncologist of the relevant cancer type. Under the plausible assumption that patients within the same ZCTA and with the same cancer type differ in subspecialist access only due to changes in differential distance over time, our instrumental variable strategy provides causal estimates of the effects of subspecialist access on health care utilization and patient health outcomes.

Our analysis finds that patients with greater access to subspecialists experience lower mortality in the years following treatment onset and are more likely to enroll in disease-specific clinical trials. Our instrumental variable estimates indicate that access to a subspecialized oncologist for the relevant cancer type reduces three-year mortality by 1.4 percentage points, an 8.6% reduction relative to the mean. However, we find no significant short-term mortality benefits within the first year, an expected result given that the benefits of differential treatment would take months to accumulate. In contrast, ordinary least squares (OLS) estimates suggest that access to subspecialists is associated with higher mortality, consistent with subspecialists treating more severe and potentially more advanced cancers.

To validate our instrumental variable, we perform a series of tests with additional results. We show that, conditional on ZCTA-level characteristics, beneficiary demographics, and fixed effects, our instrument is uncorrelated with a broad set of over three dozen patient health and demographic characteristics that influence clinical outcomes. Our instrumental variable estimates on clinical trial enrollment further support the hypothesis of physician specialization as the main mechanism. We find that subspecialist access significantly increases enrollment in cancer trials specific to the beneficiary's diagnosis but does not affect participation in non-cancer-related trials or generic multi-cancer trials. This suggests that the effects are primarily driven by specialized oncologic expertise rather than non-specific differences in physician quality or patient selection. Additionally, we demonstrate that our instrumental variable estimates do not impact health outcomes unrelated to oncologic care, reinforcing the interpretation that physician specialization, rather than general clinical skill or patient selection, is the key mechanism behind

mortality reductions. For instance, access to a subspecialized oncologist has no effect on diagnoses of acute myocardial infarction, hip fractures, or strokes in the two years following chemotherapy initiation. Taken together, these results bolster our findings that oncologists' cancer-specific expertise is the primary driver of the observed mortality reductions.

We further investigate potential mechanisms underlying the observed mortality reductions. One mechanism could be that subspecialists are able to get state-of-the-art experimental treatment for their patients more easily. To examine this, we combine Medicare claims with clinical trial data from ClinicalTrials.gov, classifying trials by cancer category using OpenAI's GPT-4 API. Our analysis shows that access to a subspecialized oncologist of the relevant cancer type significantly increases enrollment in cancer clinical trials, particularly those specific to the patient's diagnosis, suggesting that greater access to specialized and novel treatments may contribute to improved survival outcomes. Another mechanism could be reflected in the magnitude of specialization - the most subspecialized physicians should provide greater disease-specific benefit than less subspecialized physicians, even among the most specialized professionals. To examine whether the degree of specialization of the care-coordinating oncologist influences mortality outcomes, we used detailed cancer type classification beyond the 5 large groups used in main analyses to construct Herfindahl-Hirschman Indices (HHI) of oncologists' specialization, where a higher HHI indicates a greater concentration of a physician's caseload in a narrower set of cancers. Instrumental variable estimates show that higher oncologist specialization significantly reduces mortality, with a 0.1-point increase in HHI (equivalent to moving from the 50th to the 64th percentile of the HHI distribution) lowering three-year mortality by one percentage point (a 2.6 percent reduction relative to the mean). These findings support the hypothesis that subspecialization could improve patient outcomes by facilitating access to advanced treatments and highly-specific physician expertise in a small number of cancers.

To understand how our estimated effects apply to different patient groups, we conduct a complier subgroup analysis, comparing compliers—those whose oncologist choice is influenced by differential distance—to the broader Medicare fee-for-service chemotherapy population. Compliers tend to be older, less likely to be female, and slightly more likely to be black. They also appear healthier on average, with fewer chronic condition flags for most chronic conditions. These differences help contextualize the impact of subspecialist access.

In ongoing work, we are also investigating the adoption of novel cancer treatments and cost differences between subspecialists and generalists as additional dimensions of

subspecialist care. Subspecialists may be more likely to prescribe newly approved therapies, potentially contributing to improved patient outcomes. At the same time, differences in treatment intensity, drug choices, and care coordination could lead to variation in healthcare spending between subspecialists and generalists. Understanding these factors will provide further insight into the trade-offs between specialization, treatment effectiveness, and cost in oncology care.

Our findings contribute to several strands of literature. First, we add to the broader economic literature on specialization and the division of labor (Smith, 2002; Becker and Murphy, 1992). Economic theory predicts that greater specialization enhances efficiency and expertise, yet also introduces potential trade-offs related to coordination costs and accessibility (Rosen, 1983; Baumgardner, 1988). We examine these effects in a high-skilled professional setting—medicine, and specifically medical oncology—where subspecialization has expanded in response to the rapid growth of medical knowledge and treatment options (Lozinski, 2024). Furthermore, our paper directly connects to Dingel et al. (2023) by providing empirical evidence on the benefits and constraints of higher specialization in health care markets. While their work highlights how larger markets facilitate specialization due to economies of scale, we quantify the patient-level implications of this specialization, showing that access to subspecialized oncologists improves survival and increases clinical trial enrollment.

Second, our research contributes to the broader literature on physician productivity, specialization, and its effects on treatment decisions and patient outcomes (Chan and Chen, 2022; Baicker and Chandra, 2004). While prior work has documented differences in health care practices between specialists and generalists—particularly in volume-outcome relationships among surgeons (Birkmeyer et al., 2002; Huckman and Pisano, 2006; Chandra and Staiger, 2007; Chowdhury, Dagash and Pierro, 2007; Avdic, Lundborg and Vikström, 2019)—there is limited causal evidence on the effects of specialization in settings where expertise is defined not by procedural frequency, but by the breadth and depth of knowledge required for complex decision-making, as in medical oncology. In addition, subspecialization is a relatively recent phenomenon, and the absence of granular physician classifications and detailed patient outcomes has constrained prior efforts to measure its effects (Shanafelt et al., 2012; Davidoff et al., 2020). This paper addresses these limitations by constructing a new dataset on chemotherapy episodes, allowing us to directly classify oncologists into subspecialties using micro-level data and detailed patient characteristics—a level of granularity not documented in prior research. Our findings contribute to the growing literature on the division of labor in healthcare, providing new insights into how specialization influences the adoption of advanced treatments and

impacts patient survival.

Finally, our study provides new evidence on how subspecialization influences access to clinical trials. While prior research has documented disparities in clinical trial enrollment and the diffusion of new treatments (Agha and Molitor, 2018; Alsan et al., 2022), it remains unclear whether access to subspecialists improves access to these advancements by facilitating patient participation in cutting-edge treatments. Using Medicare claims linked to clinical trial data, we show that patients with greater subspecialist access are significantly more likely to enroll in cancer trials. Future work will also address the role of subspecialists in the utilization of novel chemotherapy drugs directly measured in claims data. This evidence established that subspecialization may play a central role in access to medical innovations, contributing to broader discussions on how physician expertise affects equitable distribution of medical advancements.

The organization of this paper is as follows. Section 2 provides details on the context of medical oncology in the U.S. Section 3 describes our main data sources, the construction of our chemotherapy sample, and the definition of our instrumental variable. In Section 4, we discuss our empirical strategy and the assumptions underpinning our identification approach. Section 5 presents our main results, before concluding with Section 6.

2 Background on Medical Oncology

Cancer is the second leading cause of death in the United States, with older populations disproportionately affected. Among Medicare beneficiaries—predominantly comprised of individuals 65 and older—cancer care is a significant driver of health care utilization and costs. In 2015, cancer related health care spending accounted for 29% of overall Medicare spending, amounting to $183 billion, reflecting the high prevalence and complexity of cancer management in this population (Mariotto et al., 2020; Kaiser Family Foundation, 2025). The unique challenges posed by cancer in older adults, including comorbidities, frailty, and socioeconomic factors, necessitate specialized and coordinated approaches to care.

Medical oncology plays a central role in cancer treatment. Medical oncologists, trained in internal medicine and oncology, focus on systemic therapies, including chemotherapy, immunotherapy, hormone therapy, and targeted treatments, while coordinating care across multidisciplinary teams with surgical oncologists, radiation oncologists, and other health care professionals. Advances in treatment modalities over the last few decades—such as the development of targeted therapies addressing specific genetic mu-

tations and immunotherapy leveraging the body's immune system—have transformed the landscape of cancer care (Sharma and Allison, 2015; Carroll et al., 2023). These innovations have improved survival rates for many cancers, including those in advanced stages (Emens et al., 2017). However, the increasing complexity of treatment regimens has also driven a trend toward subspecialization within oncology since oncologists face the challenge of an expanding knowledge base (Lozinski, 2024). In Figure 1 we present the rapid increase in medical guidelines for five distinct cancer categories. Guideline page counts between 2002 and 2020 have increased by over 300 percent for all cancer categories with leukemia and lymphoma experiencing a guideline page increase of 731 percent.

Resulting from this increase in complexity and volume of specific treatment guidelines, oncologists often focus on specific cancer types or broader cancer categories, such as breast, gastrointestinal, or thoracic cancers, allowing them to develop expertise in managing the nuances of a set of specific malignancies. Subspecialization often centers around an anatomical region of the body (e.g. the breast, prostate or the gastrointestinal tract) or similar types of treatments (e.g. for leukemia and lymphoma). This makes it particularly relevant for older patients, whose treatment plans require careful balancing of efficacy against potential risks associated with aging and comorbidities. However, subspecialization also introduces challenges, particularly regarding equitable access to specialized care. Among Medicare beneficiaries, we find that the share of chemotherapy episodes treated by subspecialized oncologists in 2020 was more than three times higher in high income areas compared to low income areas.[2] This geographic disparity is especially consequential for Medicare beneficiaries, many of whom face barriers to travel or rely on local providers for care.

3 Data, Sample and Differential Distance Instrument

To define specialization of oncologists and assess the implications of access to subspecialized oncologists we draw upon a variety of data sources. The cornerstone of our analysis is a dataset containing chemotherapy episodes for the period 2008 to 2020, which is constructed utilizing 100% Medicare claims data accessed through the Center for Medicare and Medicaid Services' (CMS) Virtual Research Data Center (VRDC). We use data from 2007 to 2021 from Medicare Parts A, B and D. In addition we supplement our main sam-

[2]See Appendix Figure B2 for differences in the share of chemotherapy episodes treated by highly subspecialized oncologists across ventiles of the U.S. population ordered from lowest income to highest income.

ple with data on socioeconomic characteristics of ZCTAs obtained from the US Census Bureau. Furthermore, we are able to link information on clinical trials from clinicaltrials.gov to claims data using National Clinical Trial (NCT) numbers available in Medicare claims.

3.1 Chemotherapy Episode Data

The foundation of our analysis is the construction of a dataset containing cancer care episodes as defined by the Oncology Care Model (OCM) (RTI International and Actuarial Research Corporation, 2020). The OCM is a value-based payment and care delivery model introduced by CMS that aims at improving the quality and coordination of care for Medicare beneficiaries undergoing chemotherapy while reducing overall health care costs.[3] OCM encompasses all cancer treatment from 193 oncology practices in the US, representing 5 percent of oncology clinics in the US (Keating et al., 2023). Following guidelines from the OCM enables us to leverage the clinical and institutional experience of a large federal government initiative to capture an "industry standard" approach to measure cancer care. Using OCM definitions, we can define non-overlapping six month chemotherapy episodes for beneficiaries, assign episodes to a single cancer type, and assign a care coordinating principal medical oncologist based on the plurality of office visits during a chemotherapy episode.

In order to construct the final episode level data we include fee-for-service Medicare beneficiaries with cancer who received chemotherapy (including oral, targeted, cytotoxic, hormonal and immunotherapy). Individuals in the episode sample are enrolled in Medicare Part A and B and do not receive the Medicare Endstage Renal Disease Benefit (ESRD). Beneficiaries with cancer are identified by following the OCM Perfomance-Based Payment Methodology (CMS, 2020; Keating et al., 2021).

Chemotherapy episodes are constructed using Medicare claims data, specifically physician-administered chemotherapy claims from Part B and Outpatient files (linked to a cancer diagnosis on the claim) and prescription fills for chemotherapy agents from the Part D event file.[4] To ensure alignment between Part D claims and active treatment, prescription fills are included only if a corresponding Part B claim with a cancer diagnosis occurred within the prior 59 days. Using this approach, we define 180-day chemotherapy episodes starting from the initial chemotherapy claim. Each episode is then assigned to the med-

[3]Since the OCM went into operation in July 2016, parts of the sample period considered in this study are not part of the OCM evaluation period. Nevertheless, the guidelines on episode definition can be applied to data as far back as 2006, when Medicare Part D was introduced.

[4]A full list of all Part B drugs can be found in Appendix Table B1. National Drug Codes (NDC) are available upon request, due to the significant number of codes used for drug identification.

ical oncologist who handled the plurality of evaluation and management (E&M) office visits during the episode. The cancer type for each episode is determined based on the plurality of cancer diagnoses from E&M visits within the episode (see Appendix Table B2). To account for the look back period and episode definitions, we utilized claims data from 2007 to 2021, restricting our final analysis sample to chemotherapy episodes occurring between 2008 and 2020.

3.2 Definition of Subspecialized Oncologists

To define whether a chemotherapy episode is coordinated by a subspecialized oncologist, we classify subspecialists as oncologists who provide at least 80 percent of their chemotherapy episodes within a single cancer category in a given year. The 80 percent threshold was chosen to reflect a balance between capturing clinicians whose work is dominated by a single cancer type while allowing for the reality that many oncologists have to treat common cancers for financial and clinical reasons despite their subspecialty. In additional analyses below, we also examine a continuous approach to defining specialization. Our dataset includes chemotherapy episodes for cancers split into 9 broad categories: breast cancer, gastrointestinal (GI) cancer, gynecologic cancer, head and neck cancer, leukemia/lymphoma, prostate/genitourinary cancer, skin cancer, thoracic cancer, and other cancers. This analysis focuses on 5 of these groups (breast cancer, GI cancer, leukemia/lymphoma, prostate/genitourinary cancer, and thoracic cancer) which are typically coordinated by medical oncologists and together account for 90 percent of all identified episodes. We exclude head and neck, skin, gynecologic, and other cancers for two reasons: first, many of these cancers have care that is traditionally led by surgeons, not medical oncologists, particularly gynecologic and head/neck cancers. Second, the excluded groups have small sample sizes, which makes it difficult to define distinct subspecialties using a volume-based threshold. Under this classification, an oncologist is considered a breast cancer subspecialist if at least 80 percent of their chemotherapy episodes in a given year involve breast cancer episodes. The same 80 percent threshold applies analogously to oncologists specializing in the remaining cancer categories.

We define access to subspecialists using multiple definitions. Our main definition of access is defined as having any E&M visit with a subspecialized oncologist of the relevant cancer type during the year in which chemotherapy was initiated. A secondary measures of access includes whether a beneficiaries chemotherapy episode was managed by a subspecialized oncologist of the relevant cancer type. Both measures are highly correlated but capture contrasting dimensions of subspecialty involvement in care.

3.3 Main Sample Definition

The sample of cancer episodes includes beneficiary identifiers, the date of the initiating E&M claim for chemotherapy, the primary cancer type (see Table B2), whether the initial chemotherapy agent was administered via infusion or oral drugs, and the National Provider Identifier (NPI) of the care-coordinating oncologist.

We supplement this episode-level dataset with additional information from Medicare claims data and external sources. First, we incorporate beneficiary information including age, zip code, date of death, sex, and race—from the Master Beneficiary Summary File (MBSF), along with binary indicators for all 27 chronic conditions. Zip codes are converted into ZCTAs using publicly available crosswalks (Audirac, 2024). We also construct annual healthcare access measures at the ZCTA level using Medicare claims data. Finally, we merge in ZCTA-level population counts and annual median household income from the American Community Survey (ACS). Summary statistics of our main variables can be found in Table 1.

To examine health outcomes and healthcare utilization, we leverage 100% Medicare samples, extracting enrollment in clinical trials (NCT number) from the Carrier file and constructing acute myocardial infarction, hip fracture, and stroke indicators using diagnostic related group codes from the Inpatient file.[5] We also use Medicare claims data to construct measures of prior healthcare utilization, including hospitalizations, primary care visits, emergency room visits, and cancer screening utilization, providing insight into beneficiaries' healthcare engagement before chemotherapy initiation. Preliminary analyses on drug utilization, particularly the adoption of novel chemotherapy drugs, leverage data from the Carrier, Outpatient, and Part D files.

3.4 Differential Distance Measure

The heart of the empirical strategy is the construction of an instrumental variable based on the differential distance between a beneficiary with cancer of a specific type to a general oncologist versus the distance to the nearest subspecialist for the relevant cancer type. Since we do not observe the exact location of residence of beneficiaries, we use the centroid of the ZCTA of each beneficiary by year. For oncologists we use the centroid of the modal ZCTA of all E&M visits in the relevant year in the Part B files. Distance matrices by year are obtained from the Zip Code Distance Database (National Bureau of Economic Research, 2025).

[5] See Appendix Table B3 for details on the definition of these outcomes.

$$\mathbf{DD}_{cit} = \text{Dist. Subspecialist}_{ct} - \text{Dist. Generalist}_{t}$$

Formally we define differential distance DD for a beneficiary with cancer type c and ZCTA i in year t as the difference between the nearest subspecialist with the relevant cancer subspecialization (e.g. breast cancer subspecialist for beneficiaries with breast cancer) and the distance to the nearest general oncologist.[6] Due to increases in subspecialization among medical oncologists over time differential distances between subspecialists and generalists become smaller over time.

To construct our main sample, we restrict chemotherapy episodes to those with non-negative differential distances, meaning cases where subspecialists are located further away than general oncologists. Additionally, we exclude episodes with distances exceeding the 95th percentile of the annual distribution for either subspecialists or general oncologists.[7] This restriction avoids drawing inferences from outlier observations in geographically remote locations. After applying these criteria, our final sample retains 90% of the original chemotherapy episodes, encompassing 2.2 million unique beneficiaries treated by 17,679 distinct medical oncologists between the years 2008 to 2020.

Due to the non-linear relationship between the instrument and our measure of subspecialist access we additionally transform our measure of differential distance using the inverse-hyperbolic-sine (arcsinh) transformation. This transformation is frequently used to approximate the logarithmic transformation in regression models, while simultaneously allowing for negative and zero values of a variable. For cases where arcsinh transformations enter regression models on the right hand side of the equation, as in our case, the interpretation of the slope parameter changes slightly particularly for small values of the explanatory variable (Bellemare and Wichman, 2020). We therefore define our instrument as follows:

$$z = \sinh^{-1}(x) = \ln\left(x + \sqrt{x^2 + 1}\right)$$

This transformation ensures that we are able to capture the non-linear relationship between differential distances and access measures while also capturing the significant number of chemotherapy episodes where distances to subspecialists and general oncologists are equivalent.[8] In Figure 2 Panel 2a we present the unadjusted first-stage relation-

[6] General oncologists are defined as all oncologists who do not manage 80% of cancers within one cancer type.

[7] Only 118,527 chemotherapy episodes are assigned negative differential distances, representing approximately one percent of all episodes.

[8] Appendix Figure B3 presents the distribution of our raw differential distance and arcsinh-transformed differential distance measure for different years.

ship between our instrument and the probability of having any E&M office visit with a subspecialized oncologist of the relevant cancer type within the same year as chemotherapy initiation. Several aspects stand out, first, with increasing differential distance between subspecialists and general oncologists the share of episodes where the beneficiary has seen any subspecialized oncologist of the relevant cancer type declines. Second, the probability of having any E&M visit with a relevant subspecialist increases over time, reflecting the increasing subspecialization of medical oncologists in the US. Third, the negative relationship between differential distance and access to subspecialized oncologists is mostly increasing over time.[9] Notably, the instrument and a binary indicator of having a subspecialized oncologists of the relevant cancer type as the care coordinating oncologist is also highly correlated and extremely similar to the prior measure of access mentioned previously (see Figure 2 Panel 2b).

4 Empirical Strategy

The goal of our empirical analysis is to understand the effect of access to subspecialized oncologists on health outcomes and spending of patients. Due to differential access and selection of patients to more specialized medical oncologists it is not possible to simply compare individuals who have access to subspecialized oncologists versus those who do not. Evidence on the benefits of access to highly specialized physicians mainly comes from surgical specialties, where there is a clear relationship between increased volume of a surgical procedure and patient outcomes (Halm, Lee and Chassin, 2002; Birkmeyer et al., 2002; Avdic, Lundborg and Vikström, 2019). For medical oncologists there is only limited evidence on the effects of specialized oncologists on patient outcomes, with none of the studies significantly addressing selection effects with respect to access to subspecialized oncologists (see Shanafelt et al. (2012); Davidoff et al. (2020)). However, due to the geographic concentration of subspecialized oncologists and the resulting differences in patient populations treated by subspecialized versus general medical oncologists simple selection on observable strategies will not disentangle causal effects from adjusted associations.

Our approach to estimating the impact of access to subspecialized oncologists leverages variation in patients' exposure to subspecialized care based on their geographic location and the timing of chemotherapy initiation among cancer patients. We focus on

[9] In 2020 the negative relationship between the instrument and access to subspecialists attenuates slightly but statistically significantly towards zero compared to the year 2019. We attribute this to effects of the COVID-19 pandemic, which made physical access to cancer treatment and oncologic care more difficult (Nogueira et al., 2024).

patients undergoing chemotherapy because this group represents the "marginal" population of particular interest to policymakers due to their cost and high clinical risk. Chemotherapy is also central to the clinical implications of receiving specialized versus general oncology care.

The empirical design in this study uses a distance-based instrument which has been employed in a variety of studies in the health economics literature (McClellan, McNeil and Newhouse, 1994; Card, Fenizia and Silver, 2023; Gruber et al., 2025). One of the main concerns with distance based instruments is the endogeneity of provider location. The distance of more subspecialized oncologists might be associated with patient characteristics that both influence our measure of access and specific health outcomes. For example, patients in more affluent suburban areas might live further away from subspecialized oncologists often located in academic medical centers in downtown areas, while also generally having better health outcomes. We address this issue in two ways. First, we do not solely base our empirical strategy on distance, but instead construct differential distance, a measure that captures the relative ease of access of one oncologist over the other. Second, we augment our instrumental variable strategy by including ZCTA fixed effects, so that we compare individuals in the same ZCTA at different points in time when access to subspecialized oncologists differed. The variation in access therefore results from changes in differential distances within the same ZCTA over time, resulting from oncologists specializing, exiting and entering markets.

We use two-stage least squares to estimate the effect of subspecialist access on outcomes of beneficiaries. In the first stage we estimate the effect our differential distance instrument on access to subspecialized oncologists:

$$\text{Access}_i = \alpha + \beta \text{DD}_{t(i)z(i)} + \delta X_i + \tau_{t(i)} + \gamma_{z(i)} + \psi \text{D}_{t(i)z(i)} + \varepsilon_i \tag{1}$$

for episode i in ZCTA z in year by cancer type t, where DD captures the inverse hyperbolic sine of the differential distance between a subspecialized oncologist of the relevant cancer type and a general oncologist at the ZCTA and year level. The vector X_i includes beneficiary demographic and chronic conditions, as well as ZCTA level controls which vary over time. D is a simple distance measure capturing the distance to the nearest oncologist of any kind, $\tau_{t(i)}$ is a cancer type by year fixed effect and $\gamma_{z(i)}$ a ZCTA fixed effect. The dependent variable Access_i is a binary indicator variable equal to one if a beneficiary has had any E&M visit with a subspecialized oncologist of the relevant cancer type during the year in which chemotherapy was initiated.

Next, we estimate the effect of access to a subspecialized oncologist on mortality, clinical trial enrollment, drug use and spending. We estimate:

$$Y_i = \alpha + \beta \widehat{\text{Access}}_i + \delta X_i + \tau_{t(i)} + \gamma_{z(i)} + \psi \text{D}_{t(i)z(i)} + \varepsilon_i \tag{2}$$

where Y_i are different mortality indicators, clinical trial enrollment indicators and other outcome measures.

Using this instrumental variable design allows us to estimate a local average treatment effect (LATE) for a complier subgroup for whom the instrument, differential distance, decreases the probability to have any E&M visit with a specialist of the relevant cancer type. The validity of our instrument necessitates the standard monotonicity and exclusion assumptions. Monotonicity in our setting means that individuals for whom the distance to the subspecialist versus a general oncologist increases the probability of access to a subspecialist weakly decreases. The exclusion restriction in our specific case requires that differential distance to a subspecialist versus a general oncologist only affects spending, mortality and clinical trial enrollment through access to subspecialized oncologists and we discuss this in more detail in Section 5.3.

Table 2 presents a direct test of the relevance condition by reporting first-stage estimates for our primary measure of access to subspecialized oncologists of the relevant cancer type. The first-stage coefficient remains highly significant, even after controlling for our full set of design covariates and fixed effects. Additionally, the first-stage F-statistic is large, indicating that our instrument generates sufficiently strong variation in access to subspecialists (Lee et al., 2022).

5 Main Results

In this section, we present our main findings on the effects of access to subspecialized oncologists of the relevant cancer type on patient outcomes. We find that access to a subspecialized oncologist significantly reduces mortality in the medium term, but has no significant effect on short-term mortality (under one year). These estimates contrast with the positive mortality effects suggested by reduced form OLS estimates, indicating substantial selection effects in determining who receives care from subspecialists. Following our main mortality results, we examine the characteristics of the complier subgroup and assess the validity of the exclusion restriction. We then present results on the relationship between subspecialist access and clinical trial enrollment, followed by an analysis of the intensive margin of oncologist specialization. Finally, we provide additional evidence on health outcomes that should remain unaffected by access to subspecialized oncologists, further supporting our identification strategy.

5.1 Mortality

Our primary health outcome measure is mortality, defined as the time from chemotherapy initiation until the date of death. We construct binary indicators for mortality at various time intervals, setting the variable to one if a patient dies within a given period after starting chemotherapy. These indicators range from 90-day mortality to 1,080-day mortality, in 90-day increments. To estimate the effect of subspecialist access, we apply Equation 2 and plot the 2SLS estimates in Figure 3 Panel 3a. The results indicate no statistically significant mortality effects within the first year after chemotherapy initiation. However, beyond this point, mortality declines steadily. Specifically, access to a subspecialized oncologist for the relevant cancer type reduces 1-year (360-day) mortality by 2 percentage points. Given an average 1-year mortality rate of 17 percent, this corresponds to an 11 percent reduction relative to the mean.

Notably, the mortality reduction continues to grow, reaching 3.7 percentage points at 2 years (720 days) post-initiation before stabilizing. Although the absolute reduction plateaus, baseline mortality continues to rise over time, reducing the relative effect size. To account for this, Figure 3 Panel 3b scales the mortality estimates relative to the population's mean mortality. The results show that relative mortality reductions peak at 12.8 percent for 630-day mortality before gradually converging back to zero. This pattern suggests that while access to subspecialized oncologists mitigates medium-term mortality risk, its impact diminishes over time, likely reflecting the progression of age and underlying disease including comorbidities.

In contrast to the negative mortality effects estimated using 2SLS, the OLS estimates suggest that access to subspecialists is significantly positively correlated with mortality. For example, OLS results indicate that having access to a subspecialized oncologist is associated with a 0.6 percentage point increase in 1-year (360-day) mortality and a 1.4 percentage point increase in 3-year (1,080-day) mortality (see Table 3). This pattern suggests negative selection into treatment, where patients with more severe, complex, or advanced-stage cancers are more likely to receive care from subspecialized oncologists.[10]

5.2 Complier Characteristics

Our instrumental variable estimates identify local average treatment effects (LATEs), capturing the effect of access to a subspecialized oncologist for the subgroup of compliers—patients who are quasi-randomly assigned to a subspecialist due to differential

[10]We also present results for our main mortality measures where we allow the sample to vary depending on the mortality measure and results are similar.

distance. To characterize these compliers, we follow (Gruber et al., 2025), estimating the first-stage relationship between our instrument and subspecialist access within subgroups defined by age, race, cancer type, and comorbidities. We then compute kappa values, weighting each subgroup's sample share by its first-stage coefficient relative to the full sample, to quantify its contribution to the complier population. Finally, we compare complier characteristics to the overall sample using reweighted means.

Table 4 presents this comparison, showing that compliers tend to be slightly older, less likely to be female, and more likely to be black. On average, they appear healthier based on chronic condition indicators, being less likely to have chronic condition flags for colorectal, endometrial, lung, and breast cancer, but more likely to have a chronic condition flag for prostate cancer.

5.3 Exclusion Restriction

In Section 4, we outlined the conditions necessary for the validity of our instrumental variable approach and for our 2SLS estimates to obtain a LATE interpretation. In our setting, the exclusion restriction requires that mortality—or any other relevant outcome—is affected by our instrument only through its impact on access to a subspecialized oncologist of the relevant cancer type. A broader interpretation, extending to an endogenous variable that is not strictly binary, is that the instrument should influence mortality solely through changes in the level of oncologist specialization available to a beneficiary, we will make use of this interpretation in a later section of this paper.

To assess the validity of this assumption, we provide balancing tests demonstrating that our instrument is uncorrelated with beneficiary characteristics, conditional on controls and fixed effects. Additionally, we show that our instrument is strongly correlated with many beneficiary characteristics in the absence of controls and fixed effects, underscoring the necessity of the conditional independence assumption for identification.

Table 5 examines the relationship between our instrument and three broad categories of beneficiary characteristics: chronic conditions, prior healthcare utilization, and other characteristics. We estimate these associations using both an unadjusted model, which excludes controls and fixed effects, and an adjusted model, which incorporates our full set of controls (excluding chronic conditions and demographics) and fixed effects. The unadjusted results indicate that many beneficiary characteristics are significantly correlated with the instrument, highlighting the potential for a violation of the independence assumption and exclusion restriction. However, once we include ZCTA-level controls and fixed effects, these correlations disappear, suggesting that conditional independence

holds for the set of observable characteristics presented in Table 5.

5.4 Clinical Trial Enrollment

One channel through which subspecialist access may influence mortality is by improving access to novel and potentially life-saving treatments, such as those that may only be accessible through clinical trials. A key way to assess differences in treatment approaches between subspecialists and generalists is by examining beneficiary enrollment in clinical trials. Since 2014, Medicare has required the reporting of clinical trial identifier codes (NCTs) on claims for services provided in research studies that qualify for coverage under specific conditions (CMS, 2014). This policy allows us to identify whether chemotherapy patients in our sample from 2014 onward have participated in a clinical trial based on the presence of an NCT number on any claim during the year of chemotherapy initiation.

To supplement this administrative data, we incorporate information from ClinicalTrials.gov, focusing on non-observational trials and the primary condition each trial targets. Additionally, we use OpenAI's GPT-4 to classify trials into our broader cancer categories, enabling us to distinguish between overall cancer trial enrollment and enrollment in trials specific to a patient's cancer type. This classification helps assess whether access to subspecialists increases general clinical trial enrollment or selectively improves access to trials that are more directly relevant to a patient's diagnosis. Appendix A provides further details on the classification methodology.

Overall, subspecialized oncologists are more likely to enroll beneficiaries in clinical trials of any type, as shown in Figure 4. However, enrollment in unspecified cancer trials (i.e. trials targeted toward cancer treatment in general) and non-cancer trials remains low for both general and subspecialized oncologists (see Panels 4c and 4d). On average, general oncologists enroll approximately 1 percent of beneficiaries in cancer trials, with an even lower proportion enrolling in cancer trials specific to the patient's cancer type. In contrast, subspecialists enroll beneficiaries at roughly four times this rate (see Panels 4a and 4b).

To assess whether clinical trial enrollment is causally driven by access to subspecialists rather than selection into treatment, we estimate Equation 2, using as an outcome a binary indicator for whether a beneficiary enrolled in a clinical trial within the year of chemotherapy initiation.[11] Table 6 presents 2SLS and OLS estimates on the relationship between access to a subspecialized oncologist of the relevant cancer type and clinical trial

[11] We do not directly test whether the oncologist is responsible for enrolling the beneficiary in a clinical trial, nor do we investigate whether claims with NCT numbers are specifically linked to oncologists coordinating the beneficiary's care.

enrollment across different trial types. Panel A suggests that access to a subspecialist of the relevant cancer type increases enrollment in any cancer trial by 2.2 percentage points, a 116 percent increase relative to the sample mean. The effect is similar in magnitude for cancer-type concordant trials, with an estimated increase of 1.8 percentage points, or 113 percent relative to the mean. Notably, there is no statistically significant difference in enrollment for unspecified cancer trials or non-cancer trials, in contrast to the OLS estimates and raw unadjusted shares presented in Figure 4. While we have not yet directly tested whether trial enrollment drives mortality reductions, clinical trial participation represents a key difference in physician behavior that may contribute to the observed survival benefits associated with subspecialist access.

5.5 Degree of Oncologist Specialization

Our binary measure of access to a subspecialized oncologist provides a straightforward way to compare mortality outcomes between patients who receive care from subspecialists and those who do not. This measure captures the effect of access to specialization on mortality at the extensive margin. However, an important question remains: to what extent does the degree of specialization influence patient outcomes?

Rather than simply distinguishing between general oncologists and subspecialists—where subspecialists are defined as those treating at least 80 percent of chemotherapy episodes within a single cancer category—we investigate whether further specialization within a cancer type improves patient outcomes. The underlying hypothesis is that a narrower clinical focus may enhance expertise, leading to greater improvements in mortality. To test this hypothesis we constructed oncologist level Herfindahl-Hirschman Indices (HHI) of chemotherapy episodes of different cancer types, using a more detailed classification of cancer types capturing 46 different categories, as opposed to the 5 broad categories used above. The HHI is defined as follows:

$$\text{HHI}_i = \sum_{c=1}^{N} s_{ic}^2 \text{ where } s = \frac{\text{e}_{ic}}{\sum_{c=1}^{N} \text{e}_{ic}} \tag{3}$$

where HHI for oncologist i is calculated as the sum of the squared shares of chemotherapy episodes across cancer types. The episode share for cancer type c is defined as the number of chemotherapy episodes e_{ic} for that cancer type, divided by the total number of chemotherapy episodes the oncologist manages in a given year. The HHI ranges from 0 to 1, where higher values indicate greater specialization, meaning a larger share of an oncologist's caseload is concentrated within a single cancer type. We construct this mea-

sure separately for each year and each oncologist, using a more granular classification of cancer types than our five main cancer categories used to define subspecialists. The HHI is strongly correlated with subspecialist status. In 2020, at least 70 percent of oncologists in the top three deciles of the HHI distribution were classified as subspecialists, while the top two deciles consisted entirely of subspecialists.[12]

In Table 7, we present 2SLS estimates of Equation 2, replacing our binary access measure with the care-coordinating physician's HHI as the endogenous variable. Under the assumptions of instrument independence, monotonicity, relevance, and exclusion, this specification allows us to estimate the causal effect of increased physician specialization on patient mortality. Previously, we argued that differential distance should affect patient health outcomes only through access to a subspecialized oncologist, and that any alternative pathways would violate the exclusion restriction. In this context, one could argue that the degree of specialization is a downstream consequence of access to a subspecialist, implying that the instrument primarily affects outcomes by influencing our binary access measure, which in turn operates through the care coordinating oncologist's degree of specialization. This interpretation suggests that the degree of physician specialization is an integral mechanism through which subspecialist access impacts patient mortality.

Focusing on Panel A of Table 7, we observe patterns consistent with our main findings—mortality reductions relative to the population mean diminish over time, suggesting that specialization can only lower mortality rates to a certain extent. The causal effect of an increase in the care-coordinating oncologist's HHI is not statistically significant within the first 180 days but leads to significant mortality reductions thereafter. For instance, a 0.1 increase in HHI reduces 3-year mortality by 1 percentage point, corresponding to a 2.6 percent reduction relative to the mean 3-year mortality rate. For context, the average difference in HHI between a subspecialized and a general oncologist is 0.45 points. Scaling our point estimate using this difference suggests a 4.6 percentage point decline in 3-year mortality, which is slightly larger than the corresponding estimate in Figure 3, Panel 3a, suggesting potential non-linearities in the relationship between HHI and mortality. OLS estimates in Panel B of Table 7 suggest smaller mortality reductions for the same increase in HHI, implying that despite potential negative selection of beneficiaries, the degree of physician specialization remains an important determinant of patient survival.

[12]In Appendix Figure 5 we present the distribution of the oncologist HHI and its association with the subspecialist definition used in this paper.

5.6 Placebo Outcomes

One potential mechanism that has received little attention so far is the possibility that subspecialized oncologists are superior clinicians overall compared to their general oncologist counterparts. If this were the case, the mortality improvements we observe might be driven by overall physician skill rather than subspecialists' specific expertise in managing certain cancers. The higher enrollment of beneficiaries in clinical trials relevant to their specific cancer type already suggests that subspecialists have greater knowledge of specialized cancer treatment options. The analysis focusing on the degree of specialization of oncologists also is indicative of mortality benefits resulting from a larger degree of specialization. In this subsection, we examine an additional pathway to help distinguish whether our main results are primarily driven by specialization effects rather than general clinical skill.

If overall physician skill is driving our results, we would expect to see improvements in health outcomes beyond mortality. However, because cancer and its treatments have significant systemic effects, particularly on the immune system and overall health of older individuals, there are few outcomes that can be completely isolated from cancer's influence. Given this limitation, we focus on three outcomes that are arguably less directly related to cancer treatment: acute myocardial infarctions (AMI), hip fractures, and strokes in the two years following chemotherapy initiation. These conditions are less likely to be directly influenced by cancer care because (1) myocardial infarctions and strokes are primarily driven by cardiovascular health and pre-existing risk factors rather than oncologic treatment decisions, (2) hip fractures are mostly (but not entirely) related to musculoskeletal health, falls, and osteoporosis, which are not primary concerns in oncologic care, and (3) while some cancer treatments may have secondary effects on cardiovascular and bone health, these outcomes are generally not the focus of oncologists when managing cancer treatment regimens.

Table 8 presents 2SLS and OLS estimates for our main specification, examining the effect of access to subspecialized oncologists on binary indicators for diagnoses of acute myocardial infarction, hip fracture, and stroke. We report results separately for the first and second years following chemotherapy initiation. Panel A of Table 8 presents the 2SLS estimates, which show no statistically significant relationship between access to a subspecialized oncologist and any of these placebo outcomes. Panel B reports the OLS estimates, which, while statistically significant, yield precisely estimated zero effects, indicating that any correlation between subspecialist access and these non-cancer health outcomes is extremely small and clinically negligible.

Taken together, these findings suggest that the overall skill of subspecialists—proxied

by access to a subspecialist of the relevant cancer type—does not influence selected health outcomes across multiple organ systems unrelated to cancer. This strengthens the interpretation that mortality reductions are driven by oncologist specialization and expertise in cancer treatment, rather than general physician ability.

6 Conclusion

This paper provides novel causal evidence on the effects of the implications of specialization in high skilled professions. Focusing on medical oncology, we show that access to subspecialized oncologists improves medium- and long-term patient survival and increases enrollment in disease-specific clinical trials. Leveraging quasi-exogenous variation in differential distance to subspecialists versus general oncologists, we estimate that greater access to a subspecialist of the relevant cancer type reduces three-year mortality by 8.6 percent relative to the mean, while having no significant impact on non-cancer-related health outcomes, reinforcing the role of specialized expertise rather than general physician skill.

Our findings highlight two key mechanisms through which subspecialization improves outcomes. First, subspecialist access significantly increases clinical trial enrollment, particularly for trials targeting the patient's cancer type, suggesting that greater exposure to experimental treatments may partially contribute to better survival rates. Second, more specialized oncologists—measured using Herfindahl-Hirschman Indices (HHI)—are associated with lower mortality, indicating that a deeper focus on specific cancer types improves treatment effectiveness.

Despite these benefits, subspecialization also raises concerns about equity in access to care. Subspecialists tend to cluster in high-income areas, leaving rural and underserved populations with more limited access. This geographic disparity raises important questions about how to structure healthcare markets to balance specialization-driven quality improvements with equitable access to high-quality care.

Beyond mortality and clinical trial enrollment, we are currently investigating the role of subspecialists in adopting novel treatments and the cost implications of subspecialist versus generalist care. Understanding these trade-offs will provide a more comprehensive picture of the efficiency, accessibility, and financial sustainability of increasing subspecialization in oncology and could further inform specialization in other medical specialties. Overall, this study contributes to the broader literature on specialization, labor markets, and healthcare productivity, offering new empirical evidence on the returns to deep expertise in complex professional services.

References

Agha, Leila, and David Molitor. 2018. "The local influence of pioneer investigators on technology adoption: evidence from new cancer drugs." *Review of Economics and Statistics*, 100(1): 29–44.

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

Audirac, Michelle. 2024. "Zip2zcta master xwalk."

Avdic, Daniel, Petter Lundborg, and Johan Vikström. 2019. "Estimating returns to hospital volume: Evidence from advanced cancer surgery." *Journal of health economics*, 63: 81–99.

Baicker, Katherine, and Amitabh Chandra. 2004. "The Productivity of Physician Specialization: Evidence from the Medicare Program." *American Economic Review*, 94(2): 357–361.

Baumgardner, James R. 1988. "The division of labor, local markets, and worker organization." *Journal of Political Economy*, 96(3): 509–527.

Becker, Gary S, and Kevin M Murphy. 1992. "The division of labor, coordination costs, and knowledge." *The Quarterly journal of economics*, 107(4): 1137–1160.

Bellemare, Marc F, and Casey J Wichman. 2020. "Elasticities and the inverse hyperbolic sine transformation." *Oxford Bulletin of Economics and Statistics*, 82(1): 50–61.

Birkmeyer, John D, Andrea E Siewers, Emily VA Finlayson, Therese A Stukel, F Lee Lucas, Ida Batista, H Gilbert Welch, and David E Wennberg. 2002. "Hospital volume and surgical mortality in the United States." *New England Journal of Medicine*, 346(15): 1128–1137.

Card, David, Alessandra Fenizia, and David Silver. 2023. "The Health Impacts of Hospital Delivery Practices." *American Economic Journal: Economic Policy*, 15(2): 42–81.

Carroll, Caitlin E, Mary Beth Landrum, Alexi A Wright, and Nancy L Keating. 2023. "Adoption of innovative therapies across oncology practices—evidence from immunotherapy." *JAMA oncology*, 9(3): 324–333.

Chan, David C, and Yiqun Chen. 2022. "The Productivity of Professions: Evidence from the Emergency Department." National Bureau of Economic Research.

Chandra, Amitabh, and Douglas O Staiger. 2007. "Productivity spillovers in health care: evidence from the treatment of heart attacks." *Journal of political Economy*, 115(1): 103–140.

Chowdhury, MM, H Dagash, and A Pierro. 2007. "A systematic review of the impact of volume of surgery and specialization on patient outcome." *Journal of British Surgery*, 94(2): 145–161.

CMS. 2014. "Medicare Claims Processing Manual, Pub 100-04, Transmittal 2955." Baltimore, MD, Department of Health Human Services (DHHS), Change Request 8401: Mandatory Reporting of an 8-Digit Clinical Trial Number on Claims.

CMS. 2020. "OCM Performance-Based Payment Methodology." Accessed 2023.

Currie, Janet, and Hannes Schwandt. 2016. "Mortality inequality: the good news from a county-level approach." *Journal of Economic Perspectives*, 30(2): 29–52.

Davidoff, Amy J, Jessica B Long, Natalia Neparidze, Jan Philipp Bewersdorf, Rory M Shallis, Nikolai A Podoltsev, Rong Wang, Steven D Gore, Amer M Zeidan, and Scott F Huntington. 2020. "Oncologist sub-specialization, care setting, and multiple myeloma treatment and outcomes." *Blood*, 136: 2–3.

Dingel, Jonathan I, Joshua D Gottlieb, Maya Lozinski, and Pauline Mourot. 2023. "Market Size and Trade in Medical Services." National Bureau of Economic Research.

Emens, Leisha A, Paolo A Ascierto, Phillip K Darcy, Sandra Demaria, Alexander MM Eggermont, William L Redmond, Barbara Seliger, and Francesco M Marincola. 2017. "Cancer immunotherapy: opportunities and challenges in the rapidly evolving clinical landscape." *European journal of cancer*, 81: 116–129.

Gruber, Jonathan, David H Howard, Jetson Leder-Luis, and Theodore L Caputi. 2025. "Dying or Lying? For-Profit Hospices and End-of-Life Care." *American Economic Review*, 115(1): 263–294.

Halm, Ethan A, Clara Lee, and Mark R Chassin. 2002. "Is volume related to outcome in health care? A systematic review and methodologic critique of the literature." *Annals of internal medicine*, 137(6): 511–520.

Huckman, Robert S, and Gary P Pisano. 2006. "The firm specificity of individual performance: Evidence from cardiac surgery." *Management science*, 52(4): 473–488.

Kaiser Family Foundation. 2025. "What to Know About Medicare Spending and Financing." Accessed: 2025-02-27.

Keating, Nancy L., Mary Beth Landrum, Lauren Riedel, Michael P.-H. Liu, Joyce Lii, Matthew Trombley, Sean McClellan, Andrea Hassol, Qing Zheng, Morgan Michaels, Lauren Davis, Roberta Glass, Louisa Buatti, Carol Simon, Yujing Shen, Alex Kappes, Erin Huffstetler, Darin Ullman, Maxwell Mindock, Colin Doyle, Anna Ialynytchev, Annalise Maillet, Soumita Lahiri, Gabriel A. Brooks, Nirav S. Kapadia, Colleen Kummet, Van Doren Hsu, and Stephanie Shao. 2023. "Evaluation of the Oncology Care Model: Performance Periods 1-9 Evaluation Report." Centers for Medicare Medicaid Services.

Keating, Nancy L, Shalini Jhatakia, Gabriel A Brooks, Amanda S Tripp, Inna Cintina, Mary Beth Landrum, Qing Zheng, Thomas J Christian, Roberta Glass, Colleen M Kummet, et al. 2021. "Association of participation in the oncology care model with Medicare payments, utilization, care delivery, and quality outcomes." *Jama*, 326(18): 1829–1839.

Lee, David S, Justin McCrary, Marcelo J Moreira, and Jack Porter. 2022. "Valid t-ratio Inference for IV." *American Economic Review*, 112(10): 3260–3290.

Lozinski, Maya. 2024. "Knowledge Growth and Specialization: Evidence from Oncologists."

Mariotto, Angela B, Lindsey Enewold, Jingxuan Zhao, Christopher A Zeruto, and K Robin Yabroff. 2020. "Medical care costs associated with cancer survivorship in the United States." *Cancer epidemiology, biomarkers & prevention*, 29(7): 1304–1312.

McClellan, Mark, Barbara J McNeil, and Joseph P Newhouse. 1994. "Does more intensive treatment of acute myocardial infarction in the elderly reduce mortality?: analysis using instrumental variables." *Jama*, 272(11): 859–866.

National Bureau of Economic Research. 2025. "ZIP Code Distance Database." Accessed: 2025-02-27.

Nogueira, Leticia M, Elizabeth J Schafer, Qinjin Fan, Nikita Sandeep Wagle, Jingxuan Zhao, Kewei Sylvia Shi, Xuesong Han, Ahmedin Jemal, and K Robin Yabroff. 2024.

"Assessment of changes in cancer treatment during the first year of the COVID-19 pandemic in the US." *JAMA oncology*, 10(1): 109–114.

Rosen, Sherwin. 1983. "Specialization and human capital." *Journal of Labor Economics*, 1(1): 43–49.

RTI International and Actuarial Research Corporation. 2020. "OCM Performance-Based Payment Methodology, Version 6.0." Centers for Medicare Medicaid Services, Research Triangle Park, NC and Annandale, VA. Prepared for the Centers for Medicare Medicaid Services.

Schwandt, Hannes, Janet Currie, Marlies Bär, James Banks, Paola Bertoli, Aline Bütikofer, Sarah Cattan, Beatrice Zong-Ying Chao, Claudia Costa, Libertad González, et al. 2021. "Inequality in mortality between Black and White Americans by age, place, and cause and in comparison to Europe, 1990 to 2018." *Proceedings of the National Academy of Sciences*, 118(40).

Shanafelt, Tait D, Neil E Kay, Kari G Rabe, David J Inwards, Clive S Zent, Jose F Leis, Susan M Schwager, Carrie A Thompson, Deborah A Bowen, Thomas E Witzig, et al. 2012. "Hematologist/oncologist disease-specific expertise and survival: lessons from chronic lymphocytic leukemia (CLL)/small lymphocytic lymphoma (SLL)." *Cancer*, 118(7): 1827–1837.

Sharma, Padmanee, and James P Allison. 2015. "Immune checkpoint targeting in cancer therapy: toward combination strategies with curative potential." *Cell*, 161(2): 205–214.

Smith, Adam. 2002. "An Inquiry into the Nature and Causes of the Wealth of Nations." *Readings in economic sociology*, 6–17.

Weisz, George. 2006. *Divide and conquer: a comparative history of medical specialization.* Oxford University Press.

Figures

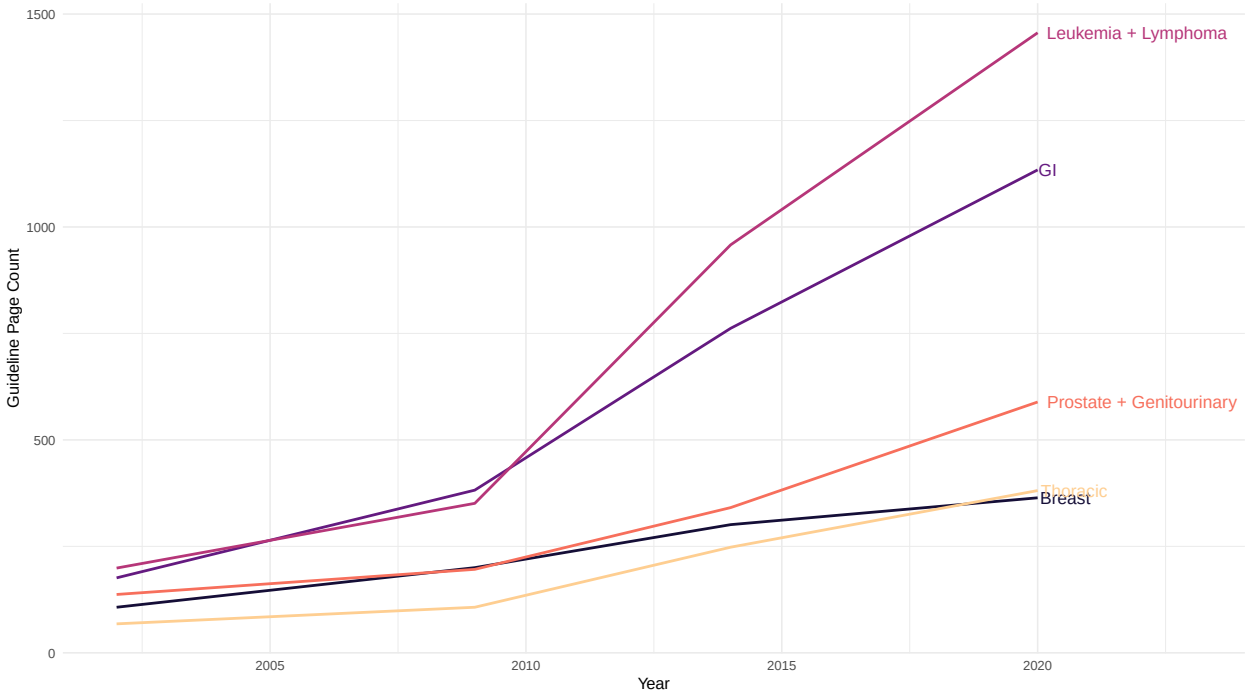
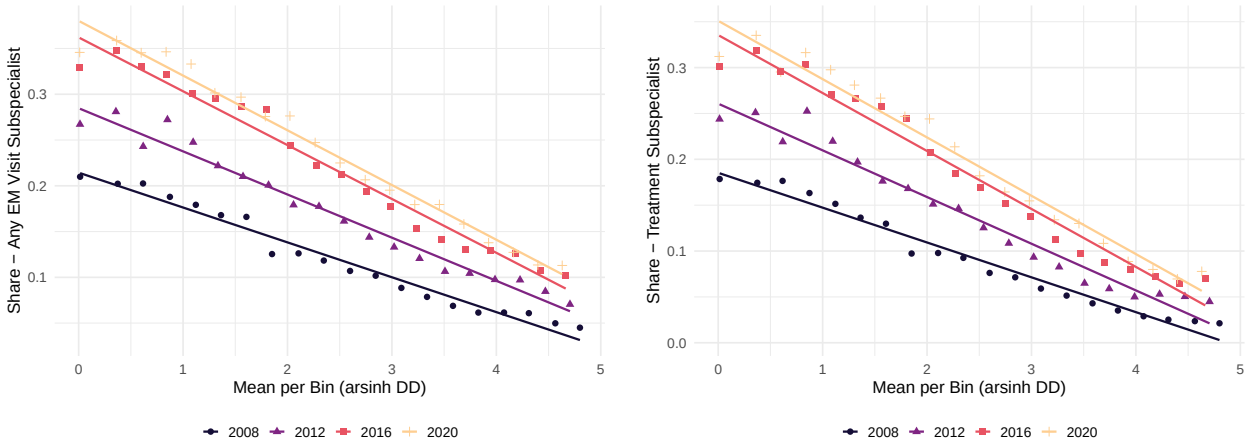


Figure 1: Guideline Page Counts by Cancer Category and Year, 2002 - 2020

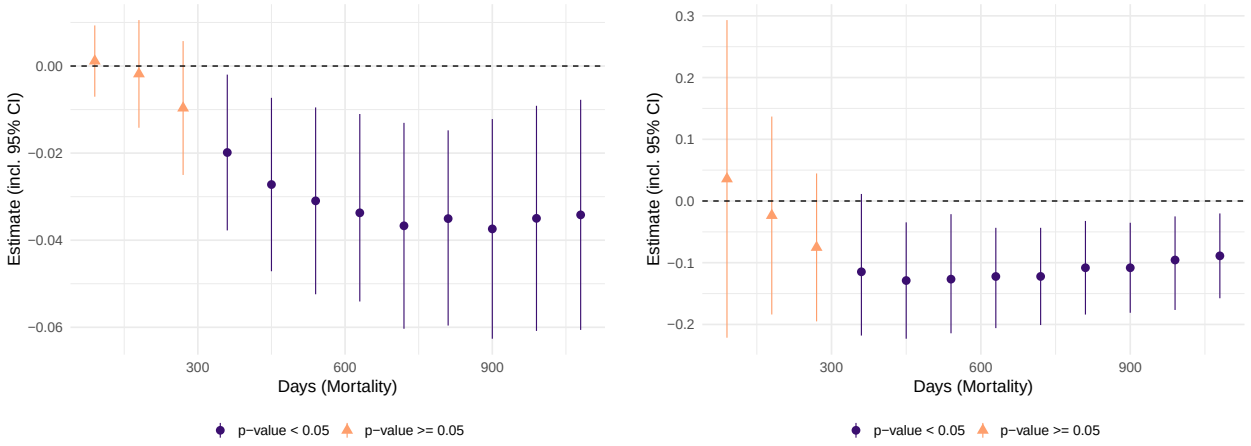
Note: This figures shows the number of pages in clinical guidelines provided by the National Comprehensive Cancer Network (NCCN). The guideline page count was separated by the main cancer categories in our sample. Guideline page counts were obtained from Lozinski (2024)



(a) Any E&M Visit Subspecialist (b) Treated by Subspecialist

Figure 2: First Stage Relationship by Year

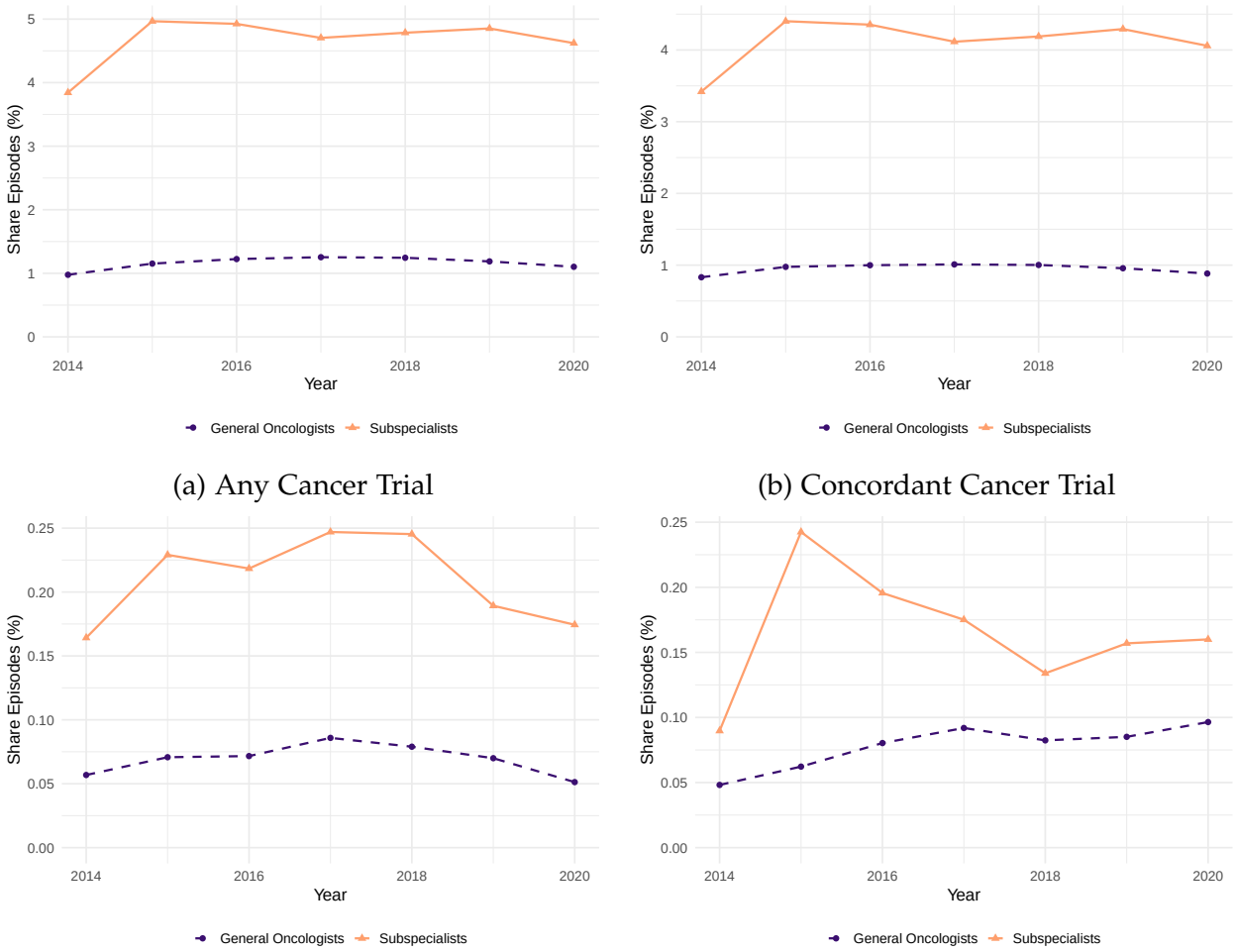
Note: The figure displays a binned scatter plot of the transformed differential distance instrument (x-axis) against measures of access to subspecialized oncologists (y-axis). Different colors represent distinct years, with linear best-fit lines plotted separately for each year. Panel A plots the share of episodes with any E&M visit with a subspecialized oncologist of the relevant cancer type per bin. Panel B plots the share of episodes where the care coordinating oncologist is a subspecialist of the relevant cancer type.



(a) Main Mortality Effect (b) Relative to Mean Mortality

Figure 3: Effect of Subspecialist Access on Mortality

Note: The figure shows the mortality effect of access to a subspecialized oncologists of the relevant cancer type estimated using 2SLS. Each point is estimated using a separate regression. The x-axis represents different mortality measures in quarterly intervals (90-days). The y-axis represents the respective 2SLS estimates including 95% confidence intervals which are constructed using standard errors clustered at the ZCTA level. We used the same sample of all chemotherapy episodes from 2008 to 2017 for estimation.



(a) Any Cancer Trial

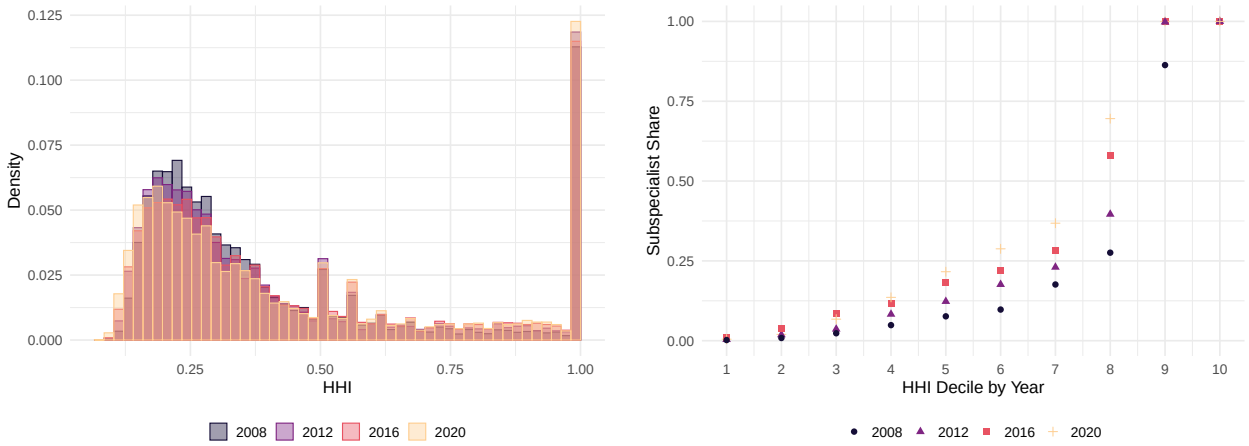
(b) Concordant Cancer Trial

(c) Unspecified Cancer Trial

(d) Non-Cancer Trial

Figure 4: Share of Episodes with Cancer Trial Enrollment by Trial Type and Oncologist Subspecialization

Note: The figure displays the share of chemotherapy episodes between 2014 and 2020 where the beneficiary had any claim with a NCT number reported in the same year as the chemotherapy was initiated. The solid orange line shows the respective share for episodes coordinated by subspecialized oncologists, while the purple dashed line shows the share of episodes for general oncologists. Panel A reports the share for any cancer trial, Panel B for cancer trials concordant to a beneficiaries cancer type, Panel C the share for cancer trials with unspecified cancers and Panel D shows the share for non-cancer trials.



(a) HHI Distribution

(b) Share of Subspecialists by HHI Decile

Figure 5: Oncologist Cancer Type Concentration and Subspecialization

Note: The figure displays the distribution of oncologist level Herfindahl-Hirschman Indices (HHI of cancer type concentration in Panel A. Panel B plots the share of subspecialist oncologist by decile of the respective HHI.

Tables

Table 1: Summary Statistics of Main Sample

Variable	Mean	SD	Min.	Max.	Missing	Unique
Panel A: Cancer Characteristics						
Breast Cancer	0.448	0.497	0.000	1.000	0.000	2.000
GI Cancer	0.102	0.303	0.000	1.000	0.000	2.000
Leukemia + Lymphoma	0.216	0.412	0.000	1.000	0.000	2.000
Prostate + Genitourinary Cancer	0.13	0.336	0.000	1.000	0.000	2.000
Thoracic Cancer	0.103	0.304	0.000	1.000	0.000	2.000
Panel B: Demographics						
Bene Age	75.977	6.545	67.000	115.000	0.000	45.000
Female Bene	0.652	0.476	0.000	1.000	0.000	2.000
Black Bene	0.080	0.271	0.000	1.000	0.000	2.000
Hispanic Bene	0.011	0.105	0.000	1.000	0.000	2.000
Asian Bene	0.015	0.121	0.000	1.000	0.000	2.000
Other non-White Race Bene	0.015	0.122	0.000	1.000	0.000	2.000
Panel C: Chronic Conditions						
Alzheimer's Disease	0.025	0.156	0.000	1.000	0.000	2.000
Alzheimer's Disease or Dementia	0.086	0.280	0.000	1.000	0.000	2.000
Acute Myocardial Infarction (Heart Attack)	0.011	0.104	0.000	1.000	0.000	2.000
Anemia	0.496	0.500	0.000	1.000	0.000	2.000
Asthma	0.062	0.241	0.000	1.000	0.000	2.000
Atrial Fibrillation	0.119	0.324	0.000	1.000	0.000	2.000
Breast Cancer	0.473	0.499	0.000	1.000	0.000	2.000
Colorectal Cancer	0.078	0.268	0.000	1.000	0.000	2.000
Endometrial Cancer	0.007	0.083	0.000	1.000	0.000	2.000
Lung Cancer	0.127	0.333	0.000	1.000	0.000	2.000
Prostate Cancer	0.141	0.348	0.000	1.000	0.000	2.000
Cataracts	0.231	0.421	0.000	1.000	0.000	2.000
Congestive Heart Failure	0.195	0.396	0.000	1.000	0.000	2.000
Chronic Kidney Disease	0.305	0.461	0.000	1.000	0.000	2.000
Chronic Obstructive Pulmonary Disease (COPD)	0.182	0.386	0.000	1.000	0.000	2.000
Depression	0.187	0.390	0.000	1.000	0.000	2.000
Diabetes	0.308	0.462	0.000	1.000	0.000	2.000
Hyperplasia	0.081	0.273	0.000	1.000	0.000	2.000
Glaucoma	0.116	0.32	0.000	1.000	0.000	2.000
Hyperlipidemia (High Cholesterol)	0.554	0.497	0.000	1.000	0.000	2.000
Hypertension (High Blood Pressure)	0.706	0.456	0.000	1.000	0.000	2.000
Hypothyroidism	0.199	0.399	0.000	1.000	0.000	2.000
Ischemic Heart Disease	0.364	0.481	0.000	1.000	0.000	2.000
Osteoporosis	0.129	0.336	0.000	1.000	0.000	2.000
Rheumatoid Arthritis/Osteoarthritis	0.391	0.488	0.000	1.000	0.000	2.000
Stroke (TIA)	0.042	0.201	0.000	1.000	0.000	2.000
Panel C: ZCTA Characteristics						
Distance Closest Oncologist	3.649	3.706	0.000	17.287	0.000	51,495.000
Mean Age	71.729	2.148	46.333	97.504	0.000	278,960.000
Share FFS	0.712	0.104	0.000	1.000	0.000	180,746.000
Share Full Dual	0.103	0.081	0.000	0.943	0.000	164,230.000
Share Disabled	0.142	0.071	0.000	0.918	0.000	174,270.000
Share Male	0.451	0.029	0.000	1.000	0.000	145,368.000
Share Black	0.097	0.167	0.000	1.000	0.000	153,926.000
Share Hispanic	0.019	0.044	0.000	0.680	0.000	100,383.000
Share Asian	0.02	0.045	0.000	1.000	0.000	97,453.000
Share Other non-White Race	0.020	0.029	0.000	1.000	0.000	109,358.000
Total Nr. Benes	4,904.062	3,477.523	1.000	43,028.000	0.000	12,386.000
Median Household Income	64,139.115	27,461.29	2,499.000	25,0001.000	0.000	80,158.000
Total Nr. Providers	109.696	198.904	0.000	7,211.000	0.000	1,660.000
Nr. Primary Care Providers	33.329	65.425	0.000	2,147.000	0.000	753.000
Nr. Mental Health Providers	10.487	20.801	0.000	338.000	0.000	268.000
Distance Closest Primary Care Provider	0.484	1.291	0.000	16.761	0.000	26,160.000

Notes: The table provides summary statistics our main sample of chemotherapy episodes.

Table 2: First Stage Estimates

	Any E&M Visit Subspecialist	**Treatement Subspecialist**
$\sinh^{-1}(\text{DD})$	-0.026***	-0.018***
	(0.001)	(0.001)
ZCTA FE	Yes	Yes
Cancer-Year FE	Yes	Yes
Observations	6,144,379	6,144,379
R^2	0.176	0.183
Mean Dep. Var.	0.206	0.163
F-test (1st stage)	22,095.726	13,890.040

Notes: The table provides estimates of the first stage relationship between the inverse hyperbolic sine of the differential distance between a subspecialized oncologist of the relevant cancer type and a general oncologist. Column 1 provides estimates for our main access measure, while column 2 provides estimates of the relationship between the instrument and having a care coordinating oncologist who is a subspecialist of the relevant cancer type. Standard errors are clustered at the ZCTA level. Signif. Codes: ***: 0.01, **: 0.05, *: 0.1.

Table 3: Mortality Effects of Access to Subspecialized Oncologist - Constant Sample

	180-day	**360-day**	**720-day**	**1080-day**
Panel A: 2SLS				
Any E&M Visit Subs.	-0.002	-0.020**	-0.037***	-0.034**
	(0.006)	(0.009)	(0.012)	(0.013)
R^2	0.118	0.218	0.292	0.317
F-test (1st stage)	13,911	13,911	13,911	13,911
Observations	4,456,173	4,456,173	4,456,173	4,456,173
Mean Dep. Var.	0.077	0.173	0.300	0.385
Panel B: OLS				
Any E&M Visit Subs.	-0.004***	0.006***	0.014***	0.014***
	(0.000)	(0.001)	(0.001)	(0.001)
R^2	0.118	0.218	0.293	0.319
Observations	4,456,173	4,456,173	4,456,173	4,456,173
Mean Dep. Var.	0.077	0.173	0.300	0.385

Notes: The table provides estimates on the effect of access to subspecialized oncologists on various outcomes of mortality for the years 2008 to 2017 of our main sample. Due to the required look forward window when constructing measures of mortality we have fixed the sample to be constant for all measures of mortality. Panel A shows two-stage least squares estimates and Panel B presents the corresponding OLS estimates. All models include demographic, ZCTA level and chronic conditions controls as well fixed effects for the beneficiaries' ZCTA and cancer type by year fixed effects. Standard errors are clustered at the ZCTA level. Signif. Codes: ***: 0.01, **: 0.05, *: 0.1.

Table 4: Differential Distance IV Complier Characteristics

Variable	Share Among All	Share Among Compliers	Total Obs.	Conditional Obs.
Panel A: Demographics				
Age 67-69	18.289	17.469	6,144,379	1,123,773
Age 70-74	29.478	28.86	6,144,379	1,811,236
Age 75-79	23.868	23.753	6,144,379	1,466,516
Age 80-84	16.316	17.186	6,144,379	1,002,531
Age 85+	12.049	12.658	6,144,379	740,323
Asian Bene	1.484	1.539	6,144,379	91,170
Female Bene	65.172	56.436	6,144,379	4,004,390
Black Bene	7.958	8.669	6,144,379	488,948
Hispanic Bene	1.118	0.824	6,144,379	68,680
Other non-White Race Bene	1.508	1.364	6,144,379	92,640
Panel B: Chronic Conditions Indicators				
Alzheimer's Disease	2.507	2.191	6,144,379	154,055
Alzheimer's Disease or Dementia	8.570	7.828	6,144,379	526,593
Acute Myocardial Infarction (Heart Attack)	1.088	0.995	6,144,379	66,880
Anemia	49.609	45.304	6,144,379	3,048,179
Asthma	6.211	6.094	6,144,379	381,642
Atrial Fibrillation	11.925	11.848	6,144,379	732,715
Breast Cancer	47.275	34.326	6,144,379	2,904,775
Colorectal Cancer	7.783	5.599	6,144,379	478,200
Endometrial Cancer	0.687	0.684	6,144,379	42,216
Lung Cancer	12.698	9.114	6,144,379	780,191
Prostate Cancer	14.110	16.651	6,144,379	866,975
Cataracts	23.095	23.935	6,144,379	1,419,052
Congestive Heart Failure	19.525	17.934	6,144,379	1,199,697
Chronic Kidney Disease	30.530	30.472	6,144,379	1,875,872
Chronic Obstructive Pulmonary Disease (COPD)	18.179	16.981	6,144,379	1,116,957
Depression	18.660	17.922	6,144,379	1,146,528
Diabetes	30.850	29.437	6,144,379	1,895,527
Glaucoma	11.567	12.783	6,144,379	710,690
Hyperlipidemia (High Cholesterol)	55.433	56.433	6,144,379	3,405,997
Hyperplasia	8.105	7.394	6,144,379	498,023
Hypertension (High Blood Pressure)	70.57	70.622	6,144,379	4,336,103
Hypothyroidism	19.866	19.334	6,144,379	1,220,625
Ischemic Heart Disease	36.418	35.488	6,144,379	2,237,663
Osteoporosis	12.949	12.559	6,144,379	795,613
Rheumatoid Arthritis/Osteoarthritis	39.051	38.632	6,144,379	2,399,464
Stroke or Transient Ischemic Attack (TIA)	4.225	4.089	6,144,379	259,615

Notes: The table presents characteristics of the differential distance compliers, in comparison to the overall sample of chemotherapy beneficiaries. For categorical variables the mean in each bin is represented.

Table 5: Balancing Test of Instrumental Variable

Variable	Mean	SD	Est. Unadj.	Std. Err. Unadj.	Est. Adj.	Std. Err. Adj.
Panel A: Chronic Conditions Indicators						
Alzheimer	0.025	0.156	-0.001	0.000	0.000	0.000
Alzheimer with Dementia	0.086	0.280	-0.001	0.000	0.000	0.000
AMI	0.011	0.104	0.001	0.000	0.000	0.000
Anemia	0.496	0.500	0.026	0.000	0.001	0.000
Asthma	0.062	0.241	-0.002	0.000	0.000	0.000
Atrial Fibrillation	0.119	0.324	0.005	0.000	0.000	0.000
Breast Cancer	0.473	0.499	-0.090	0.000	0.000	0.000
Colorectal Cancer	0.078	0.268	0.018	0.000	0.000	0.000
Endometrial Cancer	0.007	0.083	-0.001	0.000	0.000	0.000
Lung Cancer	0.127	0.333	0.035	0.000	0.000	0.000
Prostate Cancer	0.141	0.348	0.027	0.000	0.000	0.000
Cataract	0.231	0.421	-0.006	0.000	0.000	0.000
Congestive Heart Failure	0.195	0.396	0.008	0.000	-0.001	0.000
Chronic Kidney Disease	0.305	0.461	0.014	0.000	0.001	0.000
Chronic Obstructive Pulmonary Disease	0.182	0.386	0.025	0.000	-0.001	0.000
Depression	0.187	0.390	-0.002	0.000	0.000	0.000
Diabetes	0.308	0.462	0.003	0.000	0.000	0.000
Glaucoma	0.116	0.320	-0.008	0.000	0.000	0.000
Hip Fracture	0.011	0.104	0.000	0.000	0.000	0.000
Hyperlipidemia	0.554	0.497	-0.002	0.000	0.000	0.000
Hyperplasia	0.081	0.273	0.011	0.000	0.000	0.000
Hypertension	0.706	0.456	0.007	0.000	0.000	0.000
Hypothyroidism	0.199	0.399	-0.005	0.000	0.000	0.000
Ischemic Heart Disease	0.364	0.481	0.019	0.000	0.000	0.000
Osteoporosis	0.129	0.336	-0.014	0.000	0.000	0.000
Rheumatoid Arthritis	0.391	0.488	-0.009	0.000	0.000	0.000
Stroke	0.042	0.201	0.001	0.000	0.000	0.000
Panel B: Prior Healthcare Use and Diagnosis						
Any Cervical Screening	0.068	0.252	-0.010	0.000	-0.001	0.000
Any Colorectal Screening	0.022	0.147	0.000	0.000	0.000	0.000
Any Lung Screening	0.002	0.050	0.000	0.000	0.000	0.000
Any Mammogram	0.193	0.394	-0.017	0.000	0.001	0.000
Any Prostate Screening	0.057	0.232	0.014	0.000	0.000	0.000
Any Primary Care Visit	0.853	0.354	-0.002	0.000	0.000	0.000
Any ER Visit	0.348	0.476	0.013	0.000	0.000	0.000
Any Hospital Visit	0.253	0.435	0.015	0.000	0.000	0.000
Any Lymph Node Involvement	0.087	0.282	0.001	0.000	0.000	0.000
Any Metastatic Cancer Diagnosis	0.208	0.406	0.017	0.000	0.000	0.000
Panel C: Other Measures						
Any LIS	0.15	0.357	-0.006	0.000	0.000	0.000
Full Dual	0.080	0.271	-0.009	0.000	0.000	0.000
Disabled	0.000	0.002	0.000	0.000	0.000	0.000
Number of Chronic Conditions	4.859	2.932	0.098	0.001	0.001	0.002
Predicted Mortality	0.078	0.107	0.011	0.000	0.000	0.000

Notes: The table provides summary statistics for different beneficiary characteristics and healthcare indicators prior to chemotherapy. Column 1 indicates the variable name, column 2 the mean within the overall sample, column 3 the standard deviation, column 4 the estimate of an unadjusted regression of our instrument on the variable in the respective row (without any controls and without fixed effects), column 5 the corresponding heteroskedasticity robust standard error, column 6 shows estimates from a regression of our instrument on the outcome in the respective row including cancer type by year, as well as ZCTA fixed effects and demographic controls as well as ZCTA level controls, finally column 7 shows the corresponding standard error clustered at the ZCTA level.

Table 6: Clinical Trial Enrollment and Access to Subspecialized Oncologists

	Any Cancer	Concordant Cancer	Unspecified Cancer	Non Cancer
Panel A: 2SLS				
Any E&M Visit Subs.	0.022***	0.018**	0.000	0.001
	(0.005)	(0.005)	(0.000)	(0.001)
R^2	0.052	0.048	0.016	0.015
F-test (1st stage)	10,984	10,984	10,984	10,984
Observations	3,737,072	3,737,072	3,737,072	3,737,072
Mean Dep. Var.	0.019	0.016	0.001	0.001
Panel B: OLS				
Any E&M Visit Subs.	0.044***	0.038***	0.002***	0.001***
	(0.000)	(0.000)	(0.000)	(0.000)
R^2	0.056	0.051	0.016	0.015
Observations	3,737,072	3,737,072	3,737,072	3,737,072
Mean Dep. Var.	0.019	0.016	0.001	0.001

Notes: The table provides estimates of the effect of acces to a subspecialized oncologist of the relevant cancer type on measures of clinical trial enrollment. Estimates for different outcomes are presented in different columns. Column 1 presents results referring to any enrollment in a cancer trial, column 2 any enrollment in a cancer trial of the concordant cancer type, column 3 any enrollment in a cancer trial where cancer types are not clearly specified and column 4 shows results for non cancer trials. Panel A presents 2SLS estimates and Panel B presents results for simple OLS estimates. All models include demographic, ZCTA level and chronic conditions controls as well fixed effects for the beneficiaries' ZCTA and cancer type by year fixed effects. Standard errors are clustered at the ZCTA level. Signif. Codes: ***: 0.01, **: 0.05, *: 0.1.

Table 7: Mortality Effects of Oncologist Specialization

	180-day	**1-Year**	**2-Year**	**3-Year**
Panel A: 2SLS				
HHI	-0.006	-0.065**	-0.114***	-0.103**
	(0.020)	(0.029)	(0.038)	(0.042)
R^2	0.118	0.219	0.293	0.319
F-test (1st stage)	4,775	4,775	4,775	4,775
Observations	4,456,173	4,456,173	4,456,173	4,456,173
Mean Dep. Var.	0.077	0.175	0.303	0.388
Panel B: OLS				
HHI	-0.008***	-0.015***	-0.029***	-0.039***
	(0.001)	(0.001)	(0.001)	(0.002)
R^2	0.118	0.220	0.294	0.319
Observations	4,456,173	4,456,173	4,456,173	4,456,173
Mean Dep. Var.	0.077	0.175	0.303	0.388

Notes: The table provides estimates on the effect of oncologist level cancer type HHIs on various outcomes of mortality for the years 2008 to 2017 of our main sample. Panel A shows two-stage least squares estimates and Panel B presents the corresponding OLS estimates. All models include demographic, ZCTA level and chronic conditions controls as well fixed effects for the beneficiaries' ZCTA and cancer type by year fixed effects. Standard errors are clustered at the ZCTA level. Signif. Codes: ***: 0.01, **: 0.05, *: 0.1.

Table 8: Access to Subspecialists and Placebo Outcomes

	Year + 1			Year + 2		
	AMI	**Hip Fracture**	**Stroke**	**AMI**	**Hip Fracture**	**Stroke**
Panel A: 2SLS						
Any E&M Visit Subs.	0.002	-0.002	0.002	0.000	-0.002	0.000
	(0.001)	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)
R^2	0.016	0.011	0.027	0.012	0.009	0.016
F-test (1st stage)	19,061	19,061	19,061	19,061	19,061	19,061
Observations	6,144,379	6,144,379	6,144,379	6,144,379	6,144,379	6,144,379
Mean Dep. Var.	0.003	0.004	0.004	0.002	0.003	0.003
Panel B: OLS						
Any E&M Visit Subs.	0.000***	0.000***	0.000***	0.000***	0.000***	0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
R^2	0.017	0.011	0.028	0.012	0.009	0.016
Observations	6,144,379	6,144,379	6,144,379	6,144,379	6,144,379	6,144,379
Mean Dep. Var.	0.003	0.004	0.004	0.002	0.003	0.003

Notes: The table provides estimates on the effect of access to subspecialized oncologists of the relevant cancer type on binary indicators for having any diagnosis for acute myocardial infarction (AMI), hip fracture and stroke in the first year (Year + 1) and second year (Year + 2) after chemotherapy initiation. Panel A shows two-stage least squares estimates and Panel B presents the corresponding OLS estimates. All models include demographic, ZCTA level and chronic conditions controls as well fixed effects for the beneficiaries' ZCTA and cancer type by year fixed effects. Standard errors are clustered at the ZCTA level. Signif. Codes: ***: 0.01, **: 0.05, *: 0.1.

7 Appendix

A Clinical Trial Classification

To classify clinical trials obtained from ClinicalTrials.gov, we utilized OpenAI's GPT-4 API. The model was prompted with the primary condition listed for each trial and tasked with assigning it to one of several predefined cancer categories. The classification process followed these steps:

First, we linked the ClinicalTrials.gov dataset to a list of clinical trial identifiers observed in Medicare claims, obtained from the Virtual Research Data Center (VRDC). This allowed us to subset the data, retaining only trials that were present in Medicare claims. We then extracted key trial characteristics, including study type, study status, conditions, interventions, and trial identifiers, for further processing.

Since some trials list multiple conditions, we focused on the primary condition recorded in the dataset to ensure consistency in classification. We then defined a set of predefined cancer categories: breast cancer, GI cancer, leukemia/lymphoma, skin cancer, head/neck cancer, prostate/genitourinary cancer, thoracic cancer, gynecologic cancer, other cancer, general cancer, and no cancer. With the exception of the last two categories, these classifications align with the definitions used in Table B2.

The general cancer category was used to classify trials that were clearly cancer-related but did not fall into a specific cancer type. For example, some trials listed broad terms such as "neoplasm" or "malignancies" as their primary condition. While these terms indicate a cancer-related trial, they do not provide enough specificity to assign the trial to a distinct cancer category. The no cancer category was used for trials that were entirely unrelated to cancer.

For each classification request, we set the model's temperature to zero to ensure deterministic outputs and limited responses to a maximum of 20 tokens. The API was prompted with the following format:

Label as one of the following: breast cancer, GI cancer, leukemia/lymphoma, skin cancer, head/neck cancer, prostate/genitourinary cancer, thoracic cancer, gynecologic cancer, other cancer, general cancer, no cancer – for medical condition: non-small cell lung cancer.

We manually reviewed a subset of the clinical trials and obtained correct classification for more than 90% of our requests. In the final step, we merged the classification results back into the clinical trial dataset, ensuring each trial was assigned a cancer category for further analysis.

B Additional Figures and Tables

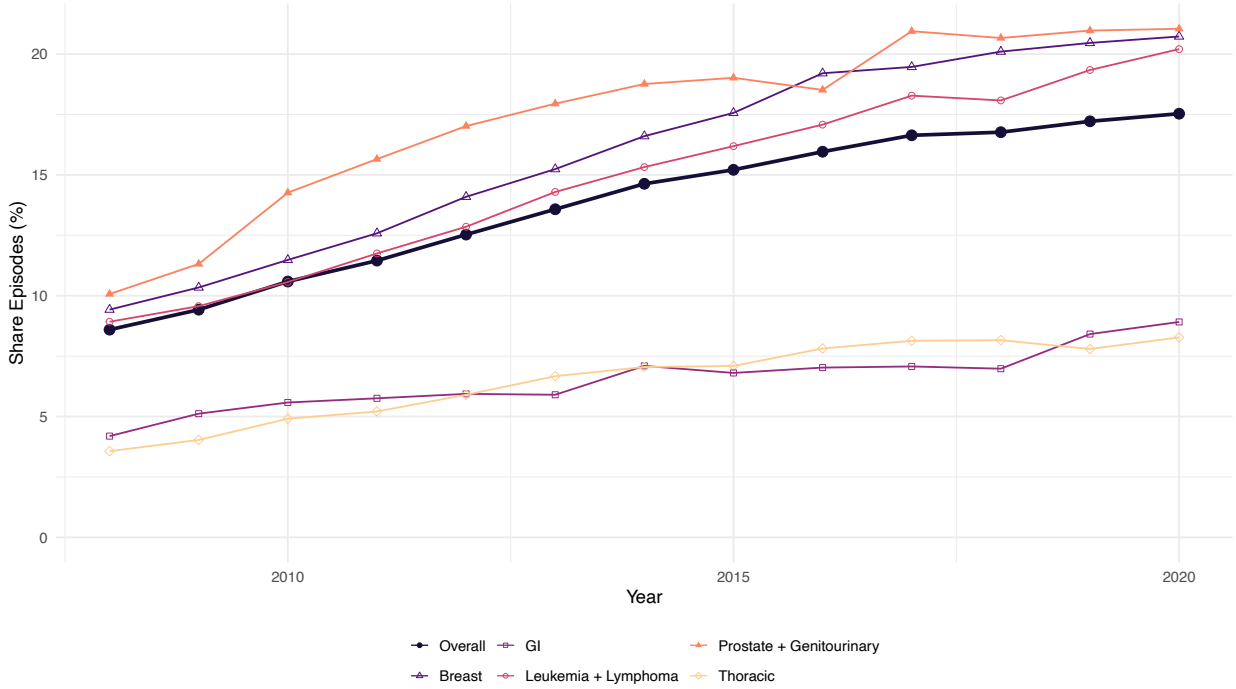


Figure B1: Trends in Subspecialization in Cancer Care 2008 - 2020

Note: This figure presents trends in subspecialization for beneficiary chemotherapy episodes of different cancer types. The figure depicts the share of chemotherapy episodes managed by highly subspecialized oncologists of the relevant cancer type separately by cancer type. The "Overall" group (black line with black dots) indicates the trend in subspecialization for all cancer types combined. Abbreviations: GI=gastrointestinal.

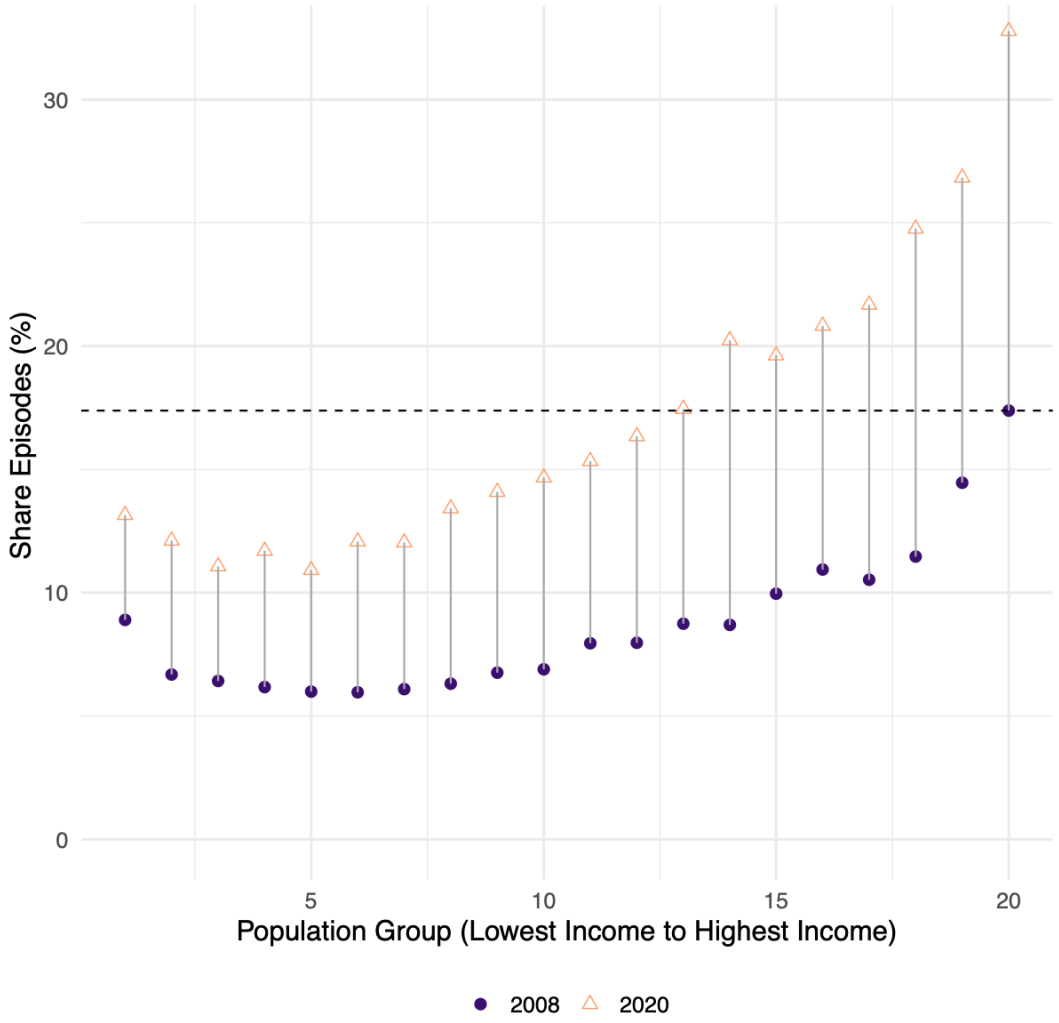
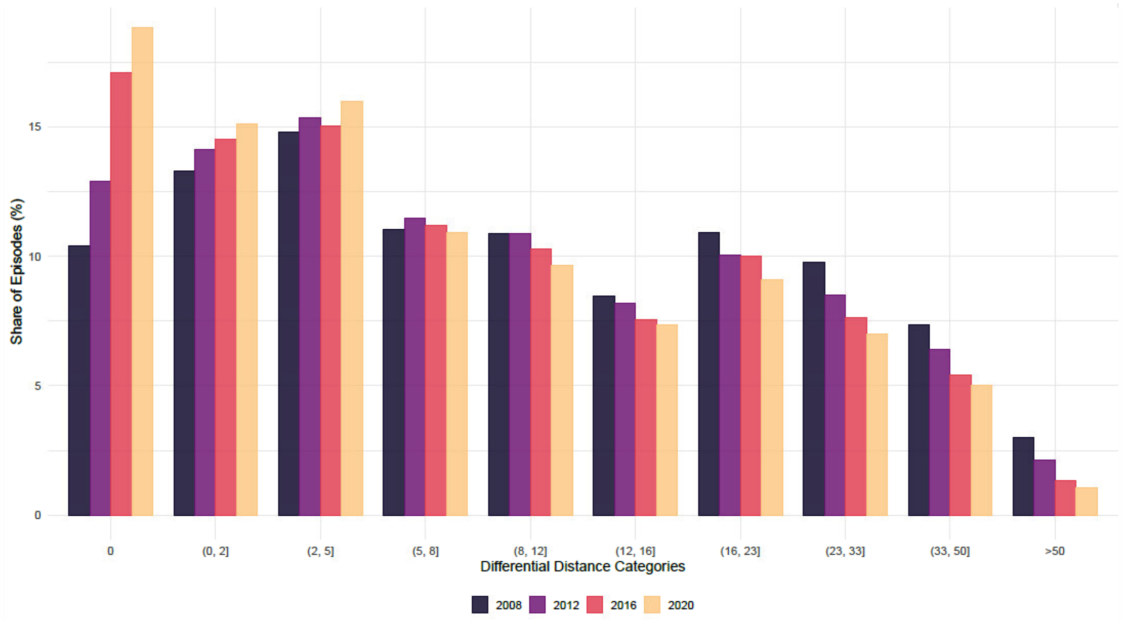
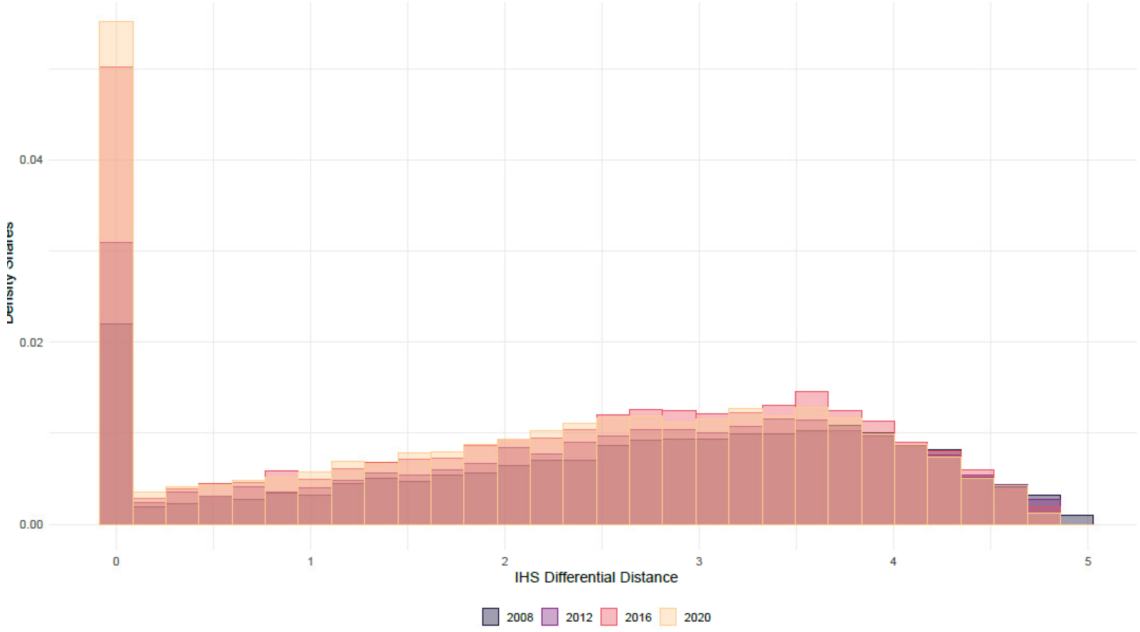


Figure B2: Socioeconomic Gradients in Access to Highly Subspecialized Cancer Care

Note: This figure illustrates the socioeconomic gradients in access to highly subspecialized cancer care, using beneficiary ZCTA and median household income data from the American Community Survey. The figure plots the share of chemotherapy episodes managed by highly subspecialized oncologists on the y-axis, with the x-axis representing population groups (ventiles), ordered from lowest income to highest income, for the years 2008 and 2020. Vertical grey lines indicate differences in access to highly subspecialized cancer care across years for each population group. The horizontal dashed line represents the level of access to highly subspecialized cancer care for the highest income population group in 2008. This figure follows prior work on mortality differences across different areas in the U.S. (Currie and Schwandt, 2016; Schwandt et al., 2021).



(a) DD Distribution (miles)



(b) arcsin (DD) Distribution

Figure B3: Distribution of Differential Distances over Time

Note: This figure presents the distribution of differential distances for four different years (2008, 2012, 2016, 2020) for our entire main sample. Panel A shows the distribution of differential distances in miles by different bins. Panel B shows the histogram of arcsinh-transformed differential distances across different years.

Table B1: Chemotherapy Drug HCPCS Codes

Generic Drug Name	HCPCS Code
LUTETIUM LU 177 DOTATATE	A9513
IBRITUMOMAB	A9543
TOSITUMOMAB	A9545
IOBENGUANE I 131	A9590
TRIPTORELIN	C9016
OBINUTUZUMAB	C9021
DAUNORUBICIN AND CYTARABINE	C9024
RAMUCIRUMAB	C9025
PEMBROLIZUMAB	C9027
INOTUZUMAB OZOGAMICIN	C9028
COPANLISIB	C9030
LUTETIUM LU 177 DOTATATE	C9031
BENDAMUSTINE	C9042
CEMIPLIMAB-RWLC	C9044
MOXETUMOMAB PASUDOTOX-TDFK	C9045
TAGRAXOFUSP-ERZS	C9049
ADO-TRASTUZUMAB EMTANSINE	C9131
BRENTUXIMAB VEDOTIN	C9287
ASPARAGINASE ERWINIA	C9289
PERTUZUMAB	C9292
CARFILZOMIB	C9295
ZIV-AFLIBERCEPT	C9296
OMACETAXINE	C9297
IOBENGUANE I 131	C9408
BELINOSTAT	C9442
BLINATUMOMAB	C9449
NIVOLUMAB	C9453
SILTUXIMAB	C9455
RITUXIMAB AND HYALURONIDASE	C9467
TALIMOGENE LAHERPAREPVEC	C9472
IRINOTECAN, LIPOSOMAL	C9474
NECITUMUMAB	C9475
DARATUMUMAB	C9476
ELOTUZUMAB	C9477
TRABECTEDIN	C9480
ATEZOLIZUMAB	C9483
OLARATUMAB	C9485
AVELUMAB	C9491
DURVALUMAB	C9492
ALEMTUZUMAB	J0202
BUSULFAN	J0594
DECITABINE	J0894
HISTRELIN	J1675
LANREOTIDE	J1930
LEUPROLIDE	J1950
OCTREOTIDE	J2353
OCTREOTIDE	J2354
SILTUXIMAB	J2860
TRIPTORELIN	J3315
TRIPTORELIN	J3316
ANTI-THYMOCYTE GLOBULIN, EQUINE	J7504

(Table B1 continued)

ANTI-THYMOCYTE GLOBULIN, RABBIT	J7511
BUSULFAN	J8510
CAPECITABINE	J8520
CAPECITABINE	J8521
CYCLOPHOSPHAMIDE	J8530
ETOPOSIDE	J8560
FLUDARABINE	J8562
GEFITINIB	J8565
MELPHALAN	J8600
TEMOZOLOMIDE	J8700
TOPOTECAN	J8705
ANTINEO, NOC	J8999
DOXORUBICIN	J9000
DOXORUBICIN, LIPOSOMAL	J9001
DOXORUBICIN, LIPOSOMAL	J9002
ALEMTUZUMAB	J9010
ALDESLEUKIN	J9015
ARSENIC TRIOXIDE	J9017
ASPARAGINASE ERWINIA	J9019
ASPARAGINASE	J9020
ATEZOLIZUMAB	J9022
AVELUMAB	J9023
AZACITIDINE	J9025
CLOFARABINE	J9027
BCG (BACILLUS CALMETTE-GUERIN)	J9030
BCG (BACILLUS CALMETTE-GUERIN)	J9031
BELINOSTAT	J9032
BENDAMUSTINE	J9033
BENDAMUSTINE	J9034
BEVACIZUMAB	J9035
BENDAMUSTINE	J9036
BLINATUMOMAB	J9039
BLEOMYCIN	J9040
BORTEZOMIB	J9041
BRENTUXIMAB VEDOTIN	J9042
CABAZITAXEL	J9043
BORTEZOMIB	J9044
CARBOPLATIN	J9045
CARFILZOMIB	J9047
CARMUSTINE	J9050
CETUXIMAB	J9055
COPANLISIB	J9057
CISPLATIN	J9060
CISPLATIN	J9062
CLADRIBINE	J9065
CYCLOPHOSPHAMIDE	J9070
CYCLOPHOSPHAMIDE	J9080
CYCLOPHOSPHAMIDE	J9090
CYCLOPHOSPHAMIDE	J9091
CYCLOPHOSPHAMIDE	J9092
CYCLOPHOSPHAMIDE	J9093
CYCLOPHOSPHAMIDE	J9094
CYCLOPHOSPHAMIDE	J9095
CYCLOPHOSPHAMIDE	J9096

(Table B1 continued)

CYCLOPHOSPHAMIDE	J9097
CYTARABINE, LIPOSOMAL	J9098
CYTARABINE	J9100
CALASPARGASE PEGOL-MKNL	J9118
CEMIPLIMAB-RWLC	J9119
DACTINOMYCIN	J9120
DACARBAZINE	J9130
DACARBAZINE	J9140
DARATUMUMAB	J9145
DAUNORUBICIN	J9150
DAUNORUBICIN, LIPOSOMAL	J9151
DAUNORUBICIN AND CYTARABINE	J9153
DEGARELIX	J9155
DENILEUKIN DIFTITOX	J9160
DOCETAXEL	J9170
DOCETAXEL	J9171
DURVALUMAB	J9173
ELOTUZUMAB	J9176
ENFORTUMAB VEDOTIN-EJFV	J9177
EPIRUBICIN	J9178
ERIBULIN	J9179
ETOPOSIDE	J9181
ETOPOSIDE	J9182
FLUDARABINE	J9185
FLUOROURACIL	J9190
GEMCITABINE	J9198
GEMCITABINE	J9199
FLOXURIDINE	J9200
GEMCITABINE	J9201
GOSERELIN	J9202
GEMTUZUMAB OZOGAMICIN	J9203
IRINOTECAN, LIPOSOMAL	J9205
IRINOTECAN	J9206
IXABEPILONE	J9207
IFOSFAMIDE	J9208
IDARUBICIN	J9211
INTERFERON, GAMMA 1-B	J9216
LEUPROLIDE	J9217
LEUPROLIDE	J9218
LEUPROLIDE	J9219
HISTRELIN	J9225
IPILIMUMAB	J9228
INOTUZUMAB OZOGAMICIN	J9229
MECHLORETHAMINE	J9230
MELPHALAN	J9245
MELPHALAN	J9246
NELARABINE	J9261
OMACETAXINE	J9262
OXALIPLATIN	J9263
PACLITAXEL, PROTEIN-BOUND	J9264
PACLITAXEL	J9265
PEGASPARGASE	J9266
PACLITAXEL	J9267
PENTOSTATIN	J9268

(Table B1 continued)

TAGRAXOFUSP-ERZS	J9269
PEMBROLIZUMAB	J9271
MITOMYCIN	J9280
OLARATUMAB	J9285
MITOMYCIN	J9290
MITOMYCIN	J9291
MITOXANTRONE	J9293
NECITUMUMAB	J9295
NIVOLUMAB	J9299
GEMTUZUMAB OZOGAMICIN	J9300
OBINUTUZUMAB	J9301
OFATUMUMAB	J9302
PANITUMUMAB	J9303
PEMETREXED	J9305
PERTUZUMAB	J9306
PRALATREXATE	J9307
RAMUCIRUMAB	J9308
POLATUZUMAB VEDOTIN-PIIQ	J9309
RITUXIMAB	J9310
RITUXIMAB AND HYALURONIDASE	J9311
RITUXIMAB	J9312
MOXETUMOMAB PASUDOTOX-TDFK	J9313
ROMIDEPSIN	J9315
STREPTOZOCIN	J9320
TALIMOGENE LAHERPAREPVEC	J9325
TEMOZOLOMIDE	J9328
TEMSIROLIMUS	J9330
THIOTEPA	J9340
TOPOTECAN	J9350
TOPOTECAN	J9351
TRABECTEDIN	J9352
ADO-TRASTUZUMAB EMTANSINE	J9354
TRASTUZUMAB	J9355
TRASTUZUMAB AND HYALURONIDASE-OYSK	J9356
VALRUBICIN	J9357
FAM-TRASTUZUMAB DERUXTECAN-NXKI	J9358
VINBLASTINE	J9360
VINCRISTINE	J9370
VINCRISTINE, LIPOSOMAL	J9371
VINCRISTINE	J9375
VINCRISTINE	J9380
VINORELBINE	J9390
FULVESTRANT	J9395
ZIV-AFLIBERCEPT	J9400
Not otherwise classified, antineoplastic drugs	J9999
TENIPOSIDE	Q2017
TISAGENLECLEUCEL	Q2040
AXICABTAGENE CILOLEUCEL	Q2041
TISAGENLECLEUCEL	Q2042
SIPULEUCEL-T	Q2043
DOXORUBICIN, LIPOSOMAL	Q2048
DOXORUBICIN, LIPOSOMAL	Q2049
DOXORUBICIN, LIPOSOMAL	Q2050
BEVACIZUMAB-AWWB	Q5107

(Table B1 continued)

TRASTUZUMAB-DTTB	Q5112
TRASTUZUMAB-PKRB	Q5113
TRASTUZUMAB-DKST	Q5114
RITUXIMAB-ABBS	Q5115
TRASTUZUMAB-QYYP	Q5116
TRASTUZUMAB-ANNS	Q5117
BEVACIZUMAB-BVZR	Q5118
RITUXIMAB-PVVR	Q5119
ALEMTUZUMAB	Q9979
TEMOZOLOMIDE	WW002
TEMOZOLOMIDE	WW003
TEMOZOLOMIDE	WW004
TEMOZOLOMIDE	WW005
TEMOZOLOMIDE	WW006
TEMOZOLOMIDE	WW007
TEMOZOLOMIDE	WW008
TEMOZOLOMIDE	WW009
BUSULFAN	WW020
ETOPOSIDE	WW030
ETOPOSIDE	WW031
ETOPOSIDE	WW032
MELPHALAN	WW080
MELPHALAN	WW081
CAPECITABINE	WW089
CAPECITABINE	WW090
CAPECITABINE	WW091
CAPECITABINE	WW093
CAPECITABINE	WW094
CAPECITABINE	WW096
TOPOTECAN	WW140

Notes: The table provides the generic drug names and HCPCS codes for chemotherapy drugs used to construct chemotherapy episodes following the Oncology Care Model.

Table B2: ICD 9 and ICD 10 Codes for Cancer Type Classification

Cancer Type Label	**Cancer Types Included**	**ICD-9/ICD-10 Codes**
Breast		174.xx, 175.xx, 233.0x, C50.xx, D05.xx
GI	Esophagus, stomach, pancreas, liver, small intestine, colon, rectum	154.2x, 154.3x, 154.8x, 230.0x, 230.1x, 230.2x, 230.3x, 230.4x, 230.5x, 230.6x, 230.7x, 230.8x, 230.9x, 150.xx, 151.xx, 155.xx, 156.0x, 156.1x, 156.2x, 156.8x, 156.9x, 195.2x, 159.xx, 157.xx, 152.xx, 153.xx, 154.0x, 154.1x, C21.xx, D00.xx, D01.xx, C15.xx, C16.xx, C22.xx, C23.xx, C24.xx, C76.2x, C25.xx, C17.xx, C18.xx, C19.xx, C20.xx
Gynecologic	Ovaries, uterus, cervix, vulva, vagina	233.1x, 179.xx, 180.xx, 182.xx, 184.0x, 184.1x, 184.2x, 184.3x, 184.4x, 183.2x, 183.3x, 183.4x, 183.5x, 183.8x, 183.9x, 184.8x, 184.9x, 181.xx, 183.0x, D06.xx, C51.xx, C52.xx, C53.xx, C54.xx, C55.xx, C57.xx, C58.xx, C56.xx
Head and Neck	Lip, tongue, salivary gland, gum, floor of mouth, oropharyngeal, nasopharyngeal, hypopharyngeal, sinuses, laryngeal, trachea, nasal cavity, middle ear, eye, adnexa, head, face, neck, palate, parotid gland	231.xx, 140.xx, 141.0x, 141.1x, 141.2x, 141.3x, 141.4x, 141.5x, 141.6x, 141.8x, 141.9x, 142.0x, 142.1x, 142.2x, 142.8x, 142.9x, 143.xx, 144.xx, 145.0x, 145.1x, 145.2x, 145.3x, 145.4x, 145.5x, 145.6x, 145.8x, 145.9x, 146.0x, 146.1x, 146.2x, 146.3x, 146.4x, 146.5x, 146.6x, 146.7x, 146.8x, 146.9x, 147.xx, 148.0x, 148.1x, 148.2x, 148.3x, 148.8x, 148.9x, 149.xx, 160.0x, 160.1x, 160.2x, 160.3x, 160.4x, 160.5x, 160.8x, 160.9x, 161.xx, 162.0x, 190.xx, 195.0x, D02.xx, C00.xx, C01.xx, C02.xx, C03.xx, C04.xx, C05.xx, C06.xx, C07.xx, C08.xx, C09.xx, C10.xx, C11.xx, C12.xx, C13.xx, C14.xx, C30.xx, C31.xx, C32.xx, C33.xx, C69.xx, C76.0x
Leukemia + Lymphoma	Acute leukemia, chronic leukemia, lymphoma, multiple myeloma	205, 205.01, 205.02, 204.0x, 205.3x, 206.0x, 207.0x, 207.2x, 208.0x, 205.2x, 204.1x, 205.1x, 208.1x, 206.1x, 238.71, 208.2x, 208.8x, 208.9x, 204.9x, 238.72, 238.73, 238.74, 238.75, 206.2x, 206.9x, 203.81, 203.0x, 203.1x, 289.83, 205.9x, 238.76, 289.89, 202.3x, 202.5x, 202.6x, 202.9x, 204.2x, 204.8x, 206.8x, 205.8x, 207.8x, 207.1, 207.11, 207.12, 238.4x, 202.8, 202.81, 202.82, 202.83, 202.84, 202.85, 202.86, 202.87, 202.88, 203.8, 203.82, 200.0x, 200.1x, 200.2x, 200.3x, 200.4x, 200.5x, 200.6x, 200.7x, 200.8x, 201.xx, 202.0x, 202.1x, 202.2x, 202.4x, 202.7x, 273.3x, C91.0x, C91.3x, C91.5x, C91.6x, C91.ax, C92.0x, C92.3x, C92.4x, C92.5x, C92.6x, C92.ax, C93.0x, C94.0x, C94.2x, C94.3x, C95.0x, C94.4x, C92.2x, C91.1x, C92.1x, C95.1x, C93.1x, D47.1x, D47.3x
Prostate + Genitourinary	Kidney, bladder, ureter, prostate, testis, penis	188.xx, 189.1x, 189.2x, 189.3x, 189.4x, 189.8x, 189.9x, 233.2x, 233.3x, 233.4x, 233.5x, 233.6x, 189.0x, 187.1x, 187.2x, 187.3x, 187.4x, 187.5x, 187.6x, 187.7x, 187.8x, 187.9x, 186.xx, 185.xx, C65.xx, C66.xx, C67.xx, C68.xx, D07.xx, C64.xx, C60.xx, C63.xx, C62.xx, C61.xx
Skin	Melanoma, non-melanoma skin cancers	232.xx, 172.xx, 209.31, 209.32, 209.33, 209.34, 209.35, 209.36, 173.xx, D04.xx, C43.xx, D03.xx, C4A.xx, C44.xx
Thoracic	Lung, pleura, mediastinum, thymus	162.2x, 162.3x, 162.4x, 162.5x, 162.8x, 162.9x, 165.xx, 163.xx, 164.1x, 164.2x, 164.3x, 164.8x, 164.9x, 195.1x, 164.0x, C34.xx, C39.xx, C45.xx, C38.xx, C76.1x, C37.xx
Other	All other cancers	170.4x, 170.5x, 170.7x, 170.8x, 170.0x, 170.1x, 170.2x, 170.3x, 170.6x, 170.9x, 209.3, 193.xx, 194.0x, 194.1x, 194.3x, 194.4x, 194.5x, 194.6x, 194.8x, 194.9x, 209.0x, 209.1x, 209.2x, 191.xx, 192.0x, 192.1x, 192.2x, 192.3x, 192.8x, 192.9x, 233.7x, 233.9x, 234.xx, 176.xx, 195.5x, 195.8x

Note: The table presents the classification of cancer types used in our main episode data. Column 1 provides the cancer type label, column 2 provides the included cancer types and column 3 the corresponding ICD-9 and ICD-10 codes used for identification of cancers in the Medicare claims.

Table B3: Diagnostic Related Group Codes for Placebo Outcomes

Condition	**DRG Code Nr.**
Acute Myocardial Infarction	280, 281, 282
Hip Fracture	480, 481, 482
Stroke	061, 062, 063, 064, 065, 066

Notes: The table provides the diagnostic related group codes used to define placebo outcomes from Medicare Inpatient files.

Table B4: Mortality Effects of Access to Subspecialized Oncologist - Varying Sample

	180-day	**1-Year**	**2-Year**	**3-Year**
Panel A: 2SLS				
Any E&M Visit Subs.	-0.001	-0.015*	-0.034***	-0.033**
	(0.005)	(0.008)	(0.011)	(0.013)
R^2	0.114	0.214	0.290	0.318
F-test (1st stage)	17,177	17,177	15,478	13,911
Observations	5,598,571	5,598,571	5,016,515	4,456,173
Mean Dep. Var.	0.074	0.168	0.298	0.388
Panel B: OLS				
Any E&M Visit Subs.	-0.004***	0.006***	0.013***	0.014***
	(0.000)	(0.000)	(0.001)	(0.001)
R^2	0.114	0.215	0.292	0.319
Observations	5,598,571	5,598,571	5,016,515	4,456,173
Mean Dep. Var.	0.074	0.168	0.298	0.388

Notes: The table provides estimates on the effect of access to subspecialized oncologists on various outcomes of mortality for the years 2008 to 2019 of our main sample. Due to the required look forward window when constructing measures of mortality the exact sample varies for the different mortality outcomes. Panel A shows two-stage least squares estimates and Panel B presents the corresponding OLS estimates. All models include demographic, ZCTA level and chronic conditions controls as well fixed effects for the beneficiaries' ZCTA and cancer type by year fixed effects. Standard errors are clustered at the ZCTA level. Signif. Codes: ***: 0.01, **: 0.05, *: 0.1.