The Value of (Sub) Specialization: Evidence from Oncology

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

February 13, 2026

CLICK HERE FOR UPDATED DRAFT

Abstract

Specialization enhances professional productivity, but its benefits depend on access to the relevant expertise. In oncology, subspecialization—the narrowing of clinical focus within cancer care—has become increasingly common, yet its effects on patient outcomes remain poorly understood. This paper examines the impact of physician subspecialization in medical oncology on patient outcomes, health care spending, and access to innovation. Using detailed US Medicare data on 2.2 million first-time chemotherapy episodes from 2008 to 2020, we exploit quasi-exogenous variation in access to subspecialized oncologists through a differential distance instrument. We find that access to a subspecialist reduces three-year mortality by up to 10% relative to the mean, without increasing total Medicare spending. Subspecialist care also increases the use of newer chemotherapy agents and more than doubles enrollment in clinical cancer trials—particularly for trials aligned with the patient's cancer type. Falsification and selection analyses support the identifying assumptions and suggest that observed gains reflect differences in treatment pathways rather than patient selection. These findings provide new evidence on the productivity effects of specialization in a non-routine, knowledge-intensive profession, and underscore the organizational trade-offs involved in delivering complex care.

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 Services, Policy & Practice, Brown University School of Public Health, Brigham and Women's Hospital. The authors gratefully acknowledge comments by Christopher Manz, Jukka-Pekka Onnela and Yuhua Zhang, as well as seminar participants at the NBER SI on Aging 2025 and the Division of Population Sciences at the Dana Farber Cancer Institute. This work was supported by a grant from the National Institute on Aging (R01 AG076580). The content of this manuscript is solely the responsibility of the authors and does not necessarily represent the official views of the National Institute on Aging or the National Institutes of Health.

1 Introduction

Despite the centrality of specialization in modern economies, its effects on service quality and outcomes within professional domains remain understudied. 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). Between 2008 and 2020, the share of chemotherapy episodes managed by subspecialized oncologists in Medicare nearly doubled from 9% to 18%, mirroring the rapid expansion of targeted treatments and cancer-specific guidelines (Karadakic et al., 2025; Lozinski, 2024).[1]

Subspecialization can improve the precision of cancer care by allowing physicians to tailor treatments to the biological and clinical features of specific—and sometimes rare—cancers. Yet it also introduces important challenges: delivering specialized care efficiently often requires large patient volumes, leading to geographic concentration and limited access in rural or underserved areas (Dingel et al., 2023; Karadakic et al., 2025). Moreover, many cancer patients present with multiple chronic conditions or even multiple cancers, complicating coordination across providers (Cebul et al., 2008). These tradeoffs make medical oncology a compelling setting to study how specialization affects patient outcomes, health care utilization, spending, 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 general 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 2.2 million first-time, six-month chemotherapy episodes—defined as episodes initiated by oral or physician-administered systemic anticancer therapy—we compare individuals diagnosed with the same detailed cancer type and residing in the same ZIP Code Tabulation Area (ZCTA), whose access to subspecialized oncologists varies over time due to changes in their relative proximity to

[1]See Appendix Figure D1 for an overview in the rise of subspecialization within medical oncology in traditional Medicare between 2008 and 2020.

subspecialized versus general oncologists. 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 assumption that changes in differential distance to subspecialized versus general oncologists are unrelated to unobserved determinants of outcomes—conditional on cancer type, ZCTA, and time fixed effects—our instrumental variable strategy yields causal estimates of the effect 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 1,080-day mortality by 4.5 percentage points, a 10% 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.

Beyond improved survival, we examine how access to subspecialized oncologists affects health care spending during six-month chemotherapy episodes. Using detailed Medicare claims, we construct comprehensive, episode-level spending measures across Parts A, B, and D. Our two-stage least squares estimates reveal that subspecialist access does not significantly alter total spending per episode. However, a closer look at spending subcategories shows meaningful reductions in Part B subcategory spending, primarily driven by lower expenditures on injectable and infused chemotherapy agents, as well as related chemotherapy administration services. These results suggest that while overall spending remains flat, subspecialist access is associated with different treatment patterns—achieving better survival outcomes without increasing costs.

To validate our instrumental variable, we conduct a series of robustness and falsification tests. First, we show that, conditional on ZCTA-level characteristics, beneficiary demographics, and fixed effects, the instrument is uncorrelated with more than three dozen patient health and demographic characteristics that influence clinical outcomes. Second, we perform falsification exercises by randomly reassigning differential distances within years and find no significant effect on 1,080-day mortality. We further reassign differential distances to subspecialists of unrelated cancer types and again observe no systematic effects on 1,080-day mortality, supporting the exclusion restriction. Finally, our instrumental variable estimates on clinical trial enrollment reinforce the specializa-

tion mechanism: access to a subspecialized oncologist significantly increases enrollment in cancer trials specific to the patient's diagnosis, but has no effect on enrollment in non-cancer-related 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.

To assess whether the observed survival gains are driven by selective treatment of healthier patients, we construct a new sample of first office visits with medical oncologists linked to a cancer diagnosis. Applying our identification strategy, we find no evidence that patients seen by subspecialized oncologists are more likely to initiate chemotherapy within 90, 180, or 360 days following the visit. Among those who do initiate chemotherapy within 360 days, patients seen by subspecialists start treatment slightly earlier; however, this modest difference in timing is unlikely to be clinically meaningful. Estimated mortality effects in the first visit sample align in direction with our main results, though confidence intervals are wider. Crucially, among patients who do not initiate chemotherapy after their first visit, we observe no significant difference in mortality depending on whether the visit was with a generalist or subspecialist. This could be cautiously interpreted as support that survival gains are not driven by selection on unobserved patient characteristics correlated with mortality risk.

We explore several mechanisms that may explain the observed reductions in mortality. First, we examine whether subspecialists facilitate greater access to cutting-edge treatments by linking Medicare claims to clinical trial data from ClinicalTrials.gov, using GPT-4 to classify trials by cancer type. We find that access to a subspecialized oncologist of the relevant cancer type significantly increases clinical trial enrollment, particularly in diagnosis-specific studies—suggesting improved access to novel therapies as a potential driver of better outcomes.

Second, we assess whether subspecialists are more likely to prescribe newer cancer therapies by analyzing the average FDA approval year of chemotherapy agents. We find that access to subspecialists increases the probability to have received a drug that was newly approved within the last two years, but do not find any conclusive evidence on

the average chemotherapy drug age mix. Third, we assess end-of-life treatment intensity and find that access to subspecialists reduces hospice use in the final 30 to 3 days of life by 35%, suggesting a shift toward slightly more aggressive care near death. This is accompanied by an average $272 reduction in hospice spending per episode, further indicating lower reliance on hospice services among patients with access to subspecialists. Finally, we examine whether subspecialist-led care alters the frequency of visits with providers during chemotherapy. Using episode-level claims data, we find no effect on the number of visits or the diversity of provider specialties, but observe a modest reduction in the number of unique providers. This suggests that concerns about increased care fragmentation with subspecialization are not borne out in this setting.

We further investigate the magnitude of specialization - the most subspecialized physicians should provide greater disease-specific benefit than less subspecialized physicians, even among the most specialized professionals. In our primary analysis, we focused on specialization in five groups of cancers: breast, gastrointestinal, hematologic, prostate/genitourinary, and thoracic. To examine whether the degree of specialization of the care-coordinating oncologist influences mortality outcomes, we used detailed cancer type classification beyond the five groups used in the 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 cancer types. 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 65th percentile of the HHI distribution) lowering 1,080-day mortality by 1.4 percentage points (a 3.1 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 assess the external validity of our findings, we conduct a complier subgroup analysis by comparing compliers—patients whose oncologist assignment is influenced by differential distance—to the broader population of Medicare fee-for-service beneficiaries receiving chemotherapy. Overall, compliers are broadly representative of the full sample. They exhibit slightly fewer chronic conditions and marginally lower predicted mortality, but otherwise closely resemble the general chemotherapy population across key demographic and clinical characteristics.

Our findings contribute to several strands of literature. First, we add to the broader economic literature on specialization and the division of labor (Smith, 1819; 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; Garicano, 2000; Cebul et al., 2008). We examine these effects in a high-skilled professional setting—medicine, and specifically medical oncology—where specialization has expanded (Karadakic et al., 2025) at a time of rapid growth of medical knowledge and treatment options (Lozinski, 2024; Cutler and McClellan, 2001). 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, reduces spending 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; Sahni et al., 2016; Avdic, Lundborg and Vikström, 2019)—other studies emphasize the role of physician expertise and information in improving decision-making and outcomes, especially in complex domains like cardiac surgery (Cutler, Huckman and Landrum, 2004). However, 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 beyond associational studies on select malignancies (Shanafelt et al., 2012; Davidoff et al., 2020; Caswell-Jin et al., 2025). More broadly, recent research has emphasized the role of physician beliefs, practice styles, and cognitive specialization in shaping treatment choices and patient outcomes—even outside surgical contexts (Cutler et al., 2019; Currie, MacLeod and Van Parys, 2016). 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 shapes access to

medical innovation. While prior research highlights disparities in clinical trial enrollment and treatment diffusion (Coleman, Katz and Menzel, 1957; Agha and Molitor, 2018; Alsan et al., 2022; Chandra and Skinner, 2012), it remains unclear whether subspecialists facilitate access to cutting-edge therapies. Linking Medicare claims to clinical trial data, we find that patients with greater access to subspecialized oncologists are significantly more likely to enroll in cancer trials. In addition, subspecialist access is associated with the use of newer chemotherapy drugs, as reflected by a lower average age of prescribed agents. These findings suggest that subspecialization enhances access to medical innovation, with implications for both treatment equity and the role of physician expertise in the diffusion of new therapies.

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 episodes (i.e., six-month episodes in which patients received systemic anti-cancer therapy), 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, Section 6 provides details on potential mechanisms, before concluding with Section 7.

2 Background on Medical Oncology

Cancer is the second leading cause of death in the United States, with older populations disproportionately affected.[2] 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 was equivalent to 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.[3]

Medical oncology is a cornerstone of cancer treatment. Medical oncologists—physicians trained in both internal medicine and oncology—primarily manage

[2]The median age at cancer diagnosis in the United States is 67 years, meaning that half of all cancer cases occur among individuals aged 67 and older, even though this group represents less than 17% of the U.S. population (National Cancer Institute, 2025*a*).

[3]The importance of subspecialization in ensuring up-to-date knowledge and optimal treatment is explicitly recognized by some cancer clinics and providers, who may emphasize it as part of their core mission statements (see e.g. Yale Cancer Clinic (2025)).

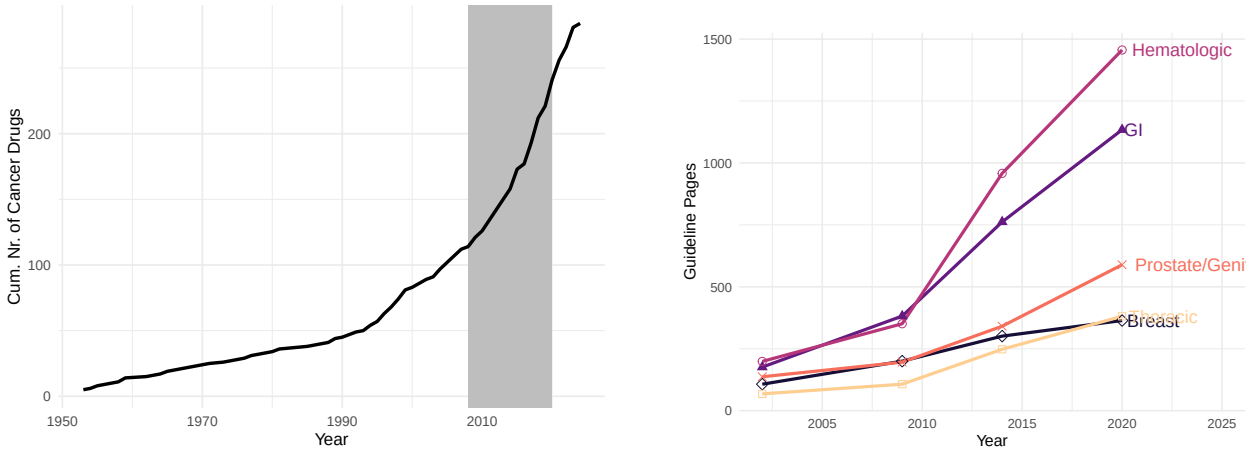
systemic therapies (i.e., cytotoxic chemotherapy, immunotherapy, targeted therapy, and hormone therapy). For simplicity, we refer to all of these as chemotherapy throughout the paper. Chemotherapy is typically delivered in one of two ways: infused or injected therapy, which is administered under the supervision of a healthcare professional, and oral therapy, which involves prescription medications taken by the patient in pill form.

In addition to prescribing and administering systemic anti-cancer and supportive care treatments, medical oncologists play a central role in 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 mutations 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, not all novel therapies provide meaningful clinical benefit, and recent approvals often show only modest survival gains despite high costs (Mailankody and Prasad, 2015). As the volume of treatment options expands, clinical judgment becomes increasingly important—not only to adopt new therapies but also to withhold or deviate from guideline-recommended treatments when evidence is weak or risks outweigh benefits.[4] The growing complexity of cancer treatment has fueled a trend toward subspecialization in oncology, as oncologists must stay current with an expanding knowledge base and the rapid introduction of new therapies (Lozinski, 2024). Two developments underscore this increasing complexity and are illustrated in Figure 1. Panel 1a shows that the cumulative number of cancer drugs—defined by 7-digit ATC codes—more than doubled from 114 in 2008 to 241 in 2020. At the same time, clinical guidelines from the National Comprehensive Cancer Network have grown substantially. As shown in Panel 1b, guideline page counts for five major cancer types increased by over 300% between 2002 and 2020, with hematologic cancer guidelines expanding by more than 700%.

The increasing complexity of oncologic care presents opportunities for oncologists to 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 cancer type (hematologic cancers). Subspecialized care may be particularly relevant for older patients, whose treatment plans require careful balancing of efficacy against potential risks

[4]The guidelines of the National Comprehensive Cancer Network (NCCN) include information on the quality of evidence guidelines are based on, acknowledging uncertainty and the difficulty in providing clear recommendations in specific clinical circumstances.

associated with aging and comorbidities. However, subspecialization also may introduce challenges, particularly regarding equitable access to specialized care and potential fragmentation of 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 (Karadakic et al., 2025).[5] This geographic disparity is especially consequential for Medicare beneficiaries, many of whom face barriers to travel or rely on local providers for care.



(a) Cumulative Nr. of Cancer Drugs

(b) Medical Guideline Page Count

Figure 1: Cancer Drug and Medical Guideline Expansion across Time

Note: This figure presents two proxies for rising complexity in the treatment of cancer. Panel A displays the cumulative number of unique cancer drugs over time, defined by 7-digit Anatomical Therapeutic Chemical (ATC) classification codes, based on data from Pantziarka et al. (2021). Each unique ATC code corresponds to a distinct chemical substance. The shaded gray area highlights the time period of our main chemotherapy sample (2008–2020). Panel B 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 (hematologic, gastrointestinal, prostate/genitourinary, breast and thoracic cancer). Guideline page counts were obtained from Lozinski (2024). Abbreviations: GI=gastrointestinal, Prostate/Geni=prostate/genitourinary.

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

[5]See Appendix Figure D2 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.

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 sample with data on socioeconomic characteristics of ZCTAs obtained through 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. Information on the Food and Drug Administration (FDA) approval year of novel cancer drugs was obtained from the National Cancer Institutes's (NCI) Surveillance, Epidemiology, and End Results (SEER) Program which provides FDA approval year through the Cancer Medication Enquiry Database (CanMED) and we link those to the relevant HCPCS and NDC codes in our data (National Cancer Institute, 2025*b*).

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) (CMS, 2025). OCM was a value-based payment and care delivery model introduced by CMS that aimed to improve the quality and coordination of care for Medicare beneficiaries undergoing chemotherapy while reducing overall health care costs.[6] Following OCM methodology 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 with cancer, 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.[7]

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

Chemotherapy episodes are constructed using 100% Medicare claims data, specif-

[6]OCM was in operation from July 2016-December 2022, we applied the episode identification methodology throughout our study period.

[7]We only focus on office visits with medical oncologists defined using CMS specialty codes 82 (hematology), 83 (hematology/oncology), 90 (medical oncology) and 98 (gynecological/oncology). While there are other provider specialties providing oncologic care (e.g. urology, internal medicine) or pathways into chemotherapy we abstract from those and focus entirely on medical oncologists as defined by these specialty codes.

ically 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.[8] 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 medical 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 office visits within the episode (see Appendix Table E2). 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 initiated between 2008 and 2020.

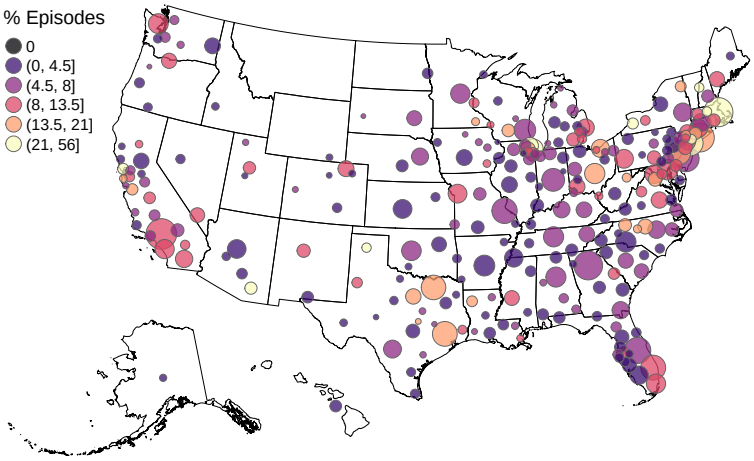
3.2 Definition of Subspecialized Oncologists

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 or set of related cancer types while allowing for the reality that many oncologists have to treat common cancers for financial and clinical reasons despite their subspecialty (Karadakic et al., 2025). In additional analyses below, we also examine a continuous approach to defining specialization. Our episode level dataset includes chemotherapy episodes for cancers split into 9 broad categories: breast cancer, gastrointestinal (GI) cancers, gynecologic cancers, head and neck cancer, hematologic cancers, prostate/genitourinary cancer, melanoma, thoracic cancer, and other cancers. 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.

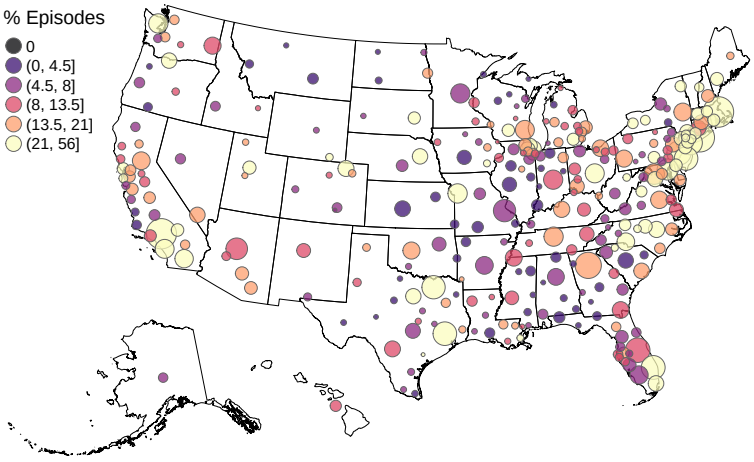
As noted earlier, the share of chemotherapy episodes managed by subspecialized oncologists nearly doubled between 2008 and 2020, rising from approximately 9% to 17.5%. This increase was especially pronounced in hematologic, breast, and prostate/genitourinary cancers, though subspecialization rose across all cancer types to varying degrees (see Appendix Figure D1).

[8] A full list of all Part B drugs can be found in Appendix Table E1. National Drug Codes (NDC) are available upon request, due to the significant number of codes used for drug identification.

To illustrate the geographic distribution of subspecialist utilization, Figure 2 plots the share of chemotherapy episodes managed by subspecialized oncologists across Hospital Referral Regions (HRRs) in 2008 and 2020. Subspecialist-managed care is disproportionately concentrated in large metropolitan areas, particularly in the Northeast, Southwest, and other major urban centers. The figure highlights both the spatial clustering of subspecialist care and how its prevalence has expanded over time.



(a) 2008



(b) 2020

Figure 2: Utilization of Subspecialized Oncologists by Hospital Referral Region

Note: This figure shows the geographic distribution of chemotherapy episodes by Hospital Referral Region (HRR) for the years 2008 and 2020 based on data from Karadakic et al. (2025). Each bubble is positioned at the centroid of the largest polygon within the HRR (based on the beneficiaries location). Bubble size reflects the total number of chemotherapy episodes in the HRR, while bubble color indicates the share of episodes managed by subspecialized oncologists of the relevant cancer type. State borders are included for geographic reference. Due to data output restrictions, we have omitted HRRs with less than eleven chemotherapy episodes managed by subspecialized oncologists in a each year.

Our primary measure of access to a subspecialized oncologist is an indicator for whether a beneficiary had any office visit with a subspecialist of the relevant cancer type during the calendar year in which chemotherapy was initiated. As a secondary measure, we also consider whether the chemotherapy episode was managed by a subspecialist. Because these two measures are highly correlated, we focus on the office visit–based definition in our empirical analyses.

3.3 Main Sample and Variable Definitions

The sample of chemotherapy episodes includes beneficiary identifiers, the date of the initiating E&M claim for chemotherapy, the primary cancer type (see Table E2), whether the initial chemotherapy agent was administered via infusion or oral drugs, and the National Provider Identifier (NPI) of the care-coordinating oncologist.

Our main analysis focuses on five of nine cancer type groups—breast, GI, hematologic, prostate/genitourinary, and thoracic cancers—where systemic therapy is commonly used and typically managed by medical oncologists. We exclude head and neck, skin, gynecologic, and other cancers because (i) care is often led by surgeons rather than oncologists, and (ii) small sample sizes limit the ability to define subspecialties based on volume. We restrict the sample to each beneficiary's first chemotherapy episode to capture patients at a similar point in their treatment trajectory and to minimize confounding from prior care. This approach also ensures that mortality effects are estimated on unique individuals rather than repeat observations. Finally, we limit the sample to beneficiaries aged 67 and older to observe at least two full years of Medicare claims history prior to treatment initiation.

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, along with binary indicators for all 27 chronic conditions. Furthermore, we construct a measure of predicted mortality using linear probability models trained on beneficiaries not included in our main sample and applied to our analytic sample.[9] Zip codes are converted into ZCTAs using publicly available crosswalks (Audirac, 2024). We also

[9]To construct predicted mortality, we use Medicare claims and enrollment data from 2006–2021 to build a detailed beneficiary-level dataset with demographics, coverage indicators, chronic condition flags, and prior-year utilization measures (e.g., provider visits, ER use, hospitalizations). After addressing missingness through additional binary indicators equal to one if a variable is missing, we estimate a linear probability model of death in the current year using beneficiaries not included in our main sample. We then apply this model to all beneficiaries included in our main sample to generate individual-specific predicted mortality scores, capturing underlying health risk without reflecting treatment choices.

construct annual health care 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. Summary statistics of our main variables can be found in Appendix Table E3.

To measure spending associated with chemotherapy, we aggregate episode-level expenditures beginning on the date of chemotherapy initiation and continuing through 180 days post-initiation, or until the beneficiary's date of death if death occurs within that window. Our measure of total spending includes payments made by Medicare, out-of-pocket spending by beneficiaries, and payments from other primary non-Medicare payers. This approach captures a more comprehensive view of the financial burden associated with care, beyond government expenditures alone. We disaggregate spending by Medicare Parts (A, B, and D), by claims source (e.g., Carrier, Outpatient, Inpatient), and further by Restructured-Berenson-Eggers Type of Service Code subcategories to shed light on the underlying components and drivers of spending patterns (CMS, 2024). We provide a detailed overview of our spending definitions in Appendix B.

To examine other health outcomes and health care 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.[10] 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.

3.4 Differential Distance Measure

The core of the empirical strategy relies on an instrumental variable constructed from the differential distance between a beneficiary's ZCTA and the nearest general oncologist versus the nearest subspecialized oncologist for the relevant cancer type. Because exact residential addresses are not available, we proxy beneficiary location using the centroid of their ZCTA in each year. For oncologists, we use the centroid of their modal ZCTA based on office visits recorded in Medicare Part B claims. Year-specific distance matrices are drawn from the NBER ZIP Code Distance Database (National Bureau of Economic Research, 2025).

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

[10] See Appendix Table E17 for details on the definition of these outcomes.

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.[11] Due to increases in subspecialization among medical oncologists over time differential distances between subspecialists and generalists become smaller over time (see Figure 3).

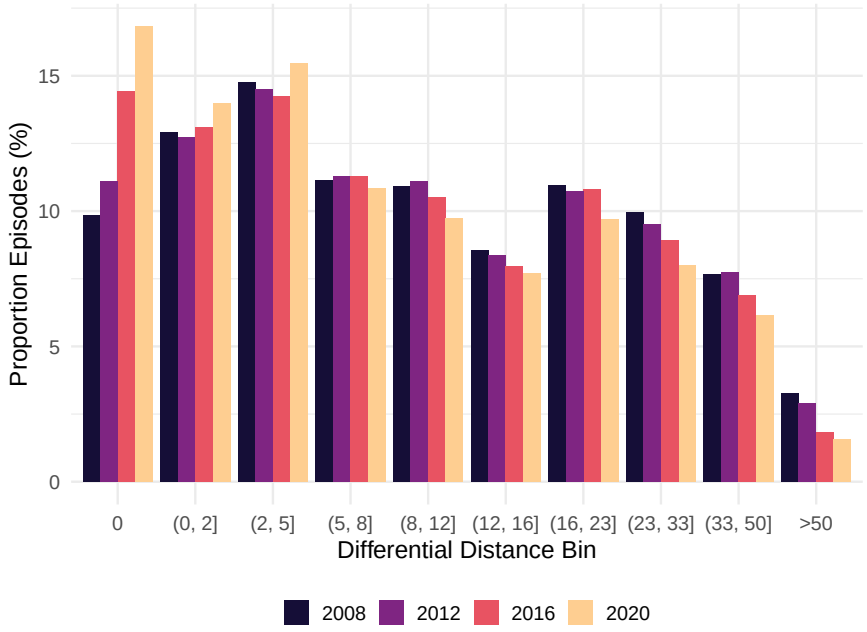


Figure 3: Distribution of Differential Distances over Time

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

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.[12] This restriction avoids drawing inferences from outlier observations in geographically remote locations. Applying these criteria yields a final sample of 2.2 million first chemotherapy episodes, treated by 17,325 distinct medical oncologists across five major cancer groups and 45 detailed cancer types from 2008 to 2020.[13]

[11] General oncologists are defined as all oncologists who do not manage 80% of cancers within one cancer type or set of related cancer types.

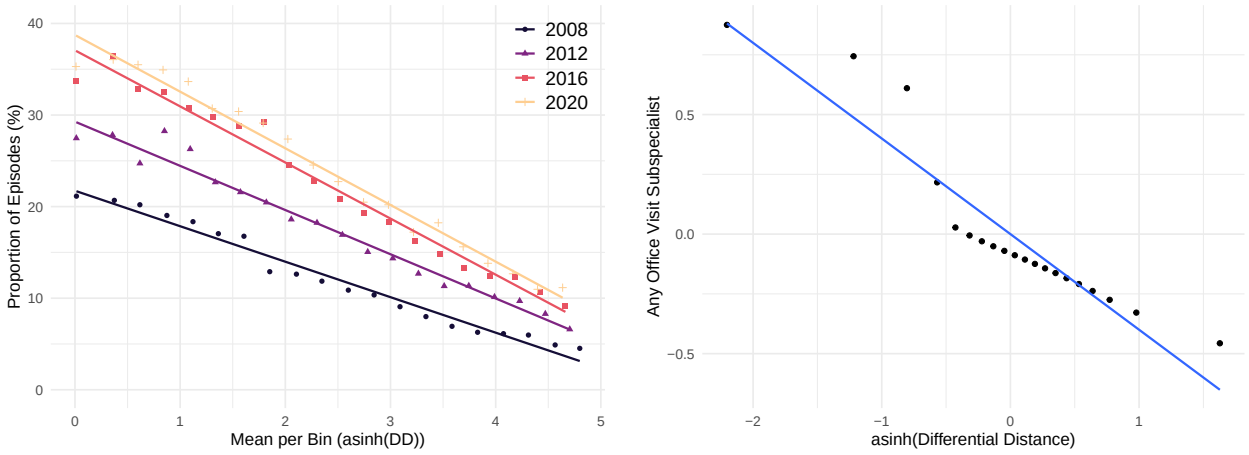
[12] Only small number of chemotherapy episodes are assigned negative differential distances.

[13] A table with detailed cancer types and corresponding ICD codes can be found in Appendix Table E2.

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 (IHS) 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 the IHS 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:

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

In Figure 4 Panel 4a we present the unadjusted first-stage relationship between our instrument and the probability of having any office visit with a subspecialized oncologist of the relevant cancer type within the same year as chemotherapy initiation.



(a) Unadjusted First Stage by Year

(b) Residualized First Stage

Figure 4: First Stage Relationship

Note: This figure shows binned scatter plots of the relationship between the differential distance instrument (x-axis) and access to subspecialized oncologists (y-axis). Panel A displays the unadjusted first-stage relationship, with each color representing a different year and separate linear best-fit lines plotted by year. Panel B presents the residualized relationship after adjusting for a rich set of covariates and fixed effects, as specified in Equation 1.

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 office 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.[14] Panel 4b presents a binned scatterplot of the residualized first-stage relationship between our access measure and the instrument. The negative association between greater differential distance and access to subspecialists remains evident, even after residualizing both variables using a rich set of controls and fixed effects.

4 Empirical Strategy

The goal of our empirical analysis is to understand the effect of access to subspecialized oncologists on patient health outcomes and spending. 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; Sahni et al., 2016; 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); Caswell-Jin et al. (2025)). 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 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 oncologic 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

[14] In 2020 the negative relationship between the instrument and access to subspecialists attenuates slightly 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).

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 cancer type-by-year 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.[15] The dependent variable Access_i is a binary indicator variable equal to one if a beneficiary has had any office 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, spending, clinical trial enrollment and drug use. 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

[15] We define fixed effects using 45 detailed cancer types, whereas subspecialist classification is based on five broader cancer categories. Table E2 provides an overview of the ICD codes corresponding to each category.

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 office visit with a subspecialist 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.

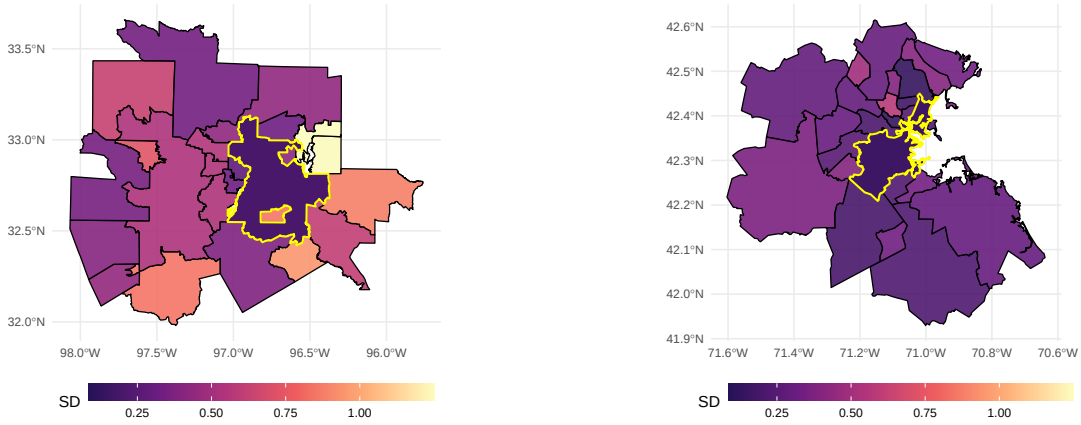
To illustrate the variation leveraged in our empirical strategy, Figure 5 presents the standard deviation of the residualized differential distance measure across selected metropolitan areas, aggregated at the Hospital Service Area (HSA) level.[16] We construct the residualized differential distance by regressing our instrument on ZCTA and cancer type-by-year fixed effects. We then standardize the resulting residuals to have mean zero and unit variance, average them by HSA and year, and finally calculate the standard deviation of these yearly HSA-level averages over time. This measure captures the within-HSA temporal variation in access to subspecialized oncologists that underlies our identification strategy.

Figure 5 provides graphical intuition for our identifying variation. Panel 5a (Dallas, TX) highlights a metropolitan area with substantial within-HSA variation in residualized differential distance over time, while Panel 5b (Boston, MA) depicts a metropolitan area with relatively little temporal within-HSA variation. These maps provide visual intuition for the source of identifying variation in our design: we exploit temporal changes in access to subspecialized oncologists within the same geographic area—driven by physician entry, exit, or changes in specialization—rather than broader, potentially confounded cross-sectional differences, such as those between rural upstate New York and Manhattan.[17] This is further supported by Appendix Table E4, which presents a variance decomposition of the instrument. The majority of raw variation is explained by cross-sectional differences across ZCTAs (Adj. $R^2 = 0.527$) and cancer types (Adj. $R^2 = 0.102$). After including ZCTA and cancer type-by-year fixed effects (Adj. $R^2 = 0.633$), the remaining

[16] While our empirical strategy exploits variation at the ZCTA level, data use restrictions prevent us from displaying beneficiary-level ZCTA data.

[17] Appendix Figure D3 shows the distribution of our instrument before and after residualizing for ZCTA and cancer type-by-year fixed effects, highlighting how our empirical design restricts identification to within-ZCTA, over-time variation.

identifying variation primarily reflects over-time changes in differential distance within ZCTAs—precisely the variation our empirical strategy leverages, and one that is plausibly exogenous.



(a) Dallas, TX

(b) Boston, MA

Figure 5: Variation in Residualized Differential Distance across Hospital Service Areas

Note: The figure displays the within-HSA standard deviation of our residualized instrumental variable. To construct this measure, we first regress the instrument on cancer type-by-year and ZCTA fixed effects. We then compute the average residual by HSA and year, and calculate the standard deviation of these HSA-level means over time. We have to rely on HSAs for this instead of ZCTAs due to file output restrictions of the data environment. The maps illustrate geographic variation in this measure across selected metropolitan areas. Panels A depicts the greater Dallas metropolitan areas and highlights regions with high variation in residualized differential distance over time. Panel B depicts the greater Boston metropolitan areas and highlights regions with comparably low variation in residualized differential distance over time. HSAs outlined in yellow represent the central HSA of the respective metropolitan area (e.g. downtown Boston for the Boston metropolitan area).

To isolate the sources of variation in our instrumental variable, we construct three alternative versions of the differential-distance measure. First, we hold the stock of oncologists fixed at its 2008 composition, preventing physicians from entering or leaving the workforce. This allows us to assess how much of the identifying variation reflects physician entry and exit. Second, we hold each oncologist's practice location (ZCTA) fixed at the first year in which they appear in our data, thereby removing all variation arising from changes in where oncologists practice over time. Finally, we construct a version of the instrument that holds each oncologist's subspecialization status fixed at the year they first enter our sample. This isolates how much of the baseline instrument's variation is driven by changes in oncologists' subspecialization patterns rather than changes in supply or location.

Appendix Table E5 shows that most of the identifying variation in our differential–distance instrument comes from changes in oncologist subspecialization rather than

workforce churn or physician relocation. When holding the oncologist stock or their practice ZCTA fixed, the resulting instruments remain closely aligned with the baseline measure, have a almost identical first stage and retain substantial explanatory power in the within-specification (within-adjusted-R^2 of 0.31 and 0.26). In contrast, fixing each oncologist's subspecialization status reduces both the correlation with the baseline instrument and the within-adjusted-R^2 to just 0.13, and the first-stage coefficient falls by more than half. This pattern indicates that time-varying subspecialization of oncologists is the primary driver of the within-ZCTA identifying variation used in our empirical strategy.[18]

In Appendix Table E6 we present 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). A visual representation of our first stage relationship is also presented in Figure 4.

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 structural form OLS estimates, indicating potential selection effects in determining who receives care from subspecialists. The second part of our main results is comprised of a detailed look at health care spending per chemotherapy episode, where we find no significant overall spending differences. Following our main results on mortality and spending, we assess the validity of the exclusion restriction and examine the characteristics of the complier subgroup.

[18]This pattern is intuitive. Entry and exit of oncologists primarily occur in markets that already have substantial oncology capacity and oncologists are geographically highly concentrated (Milligan et al., 2024; Karadakic et al., 2025), so fixing the workforce or practice ZCTAs leaves most of the differential-distance variation intact. By contrast, subspecialization has changed meaningfully over time: many general oncologists have shifted into subspecialized practice. These changes directly alter relative distances to subspecialists versus generalists, which explains why subspecialization accounts for the bulk of the identifying variation in the instrument.

5.1 Mortality

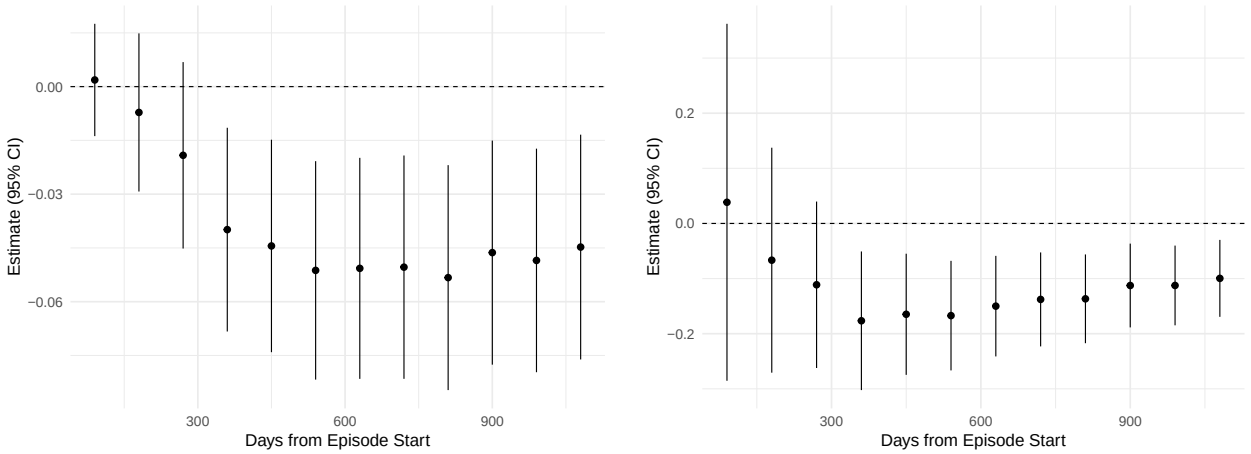
Our primary health outcome measure is mortality, defined as the time from chemotherapy initiation until the date of death.[19] We construct binary indicators for mortality at various time intervals, setting the variable to one if a patient dies within a given period after starting chemotherapy. These 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 6 Panel 6a. The results indicate no statistically significant mortality effects before 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 360-day mortality by 4 percentage points. Given an average 1-year mortality rate of 22.6 percent, this corresponds to an 17.7 percent reduction relative to the mean.[20]

Notably, the mortality reduction continues to grow, reaching 5.1 percentage points at 540 days post chemotherapy initiation before marginally declining thereafter. In addition to the marginal decrease in the point estimate, mortality rates continue to rise the further we move away from chemotherapy initiation, leading to a reduction in the overall effect size expressed in relation to mean. To account for this, Figure 6 Panel 6b scales the mortality estimates relative to the population's mean mortality. The results show that relative mortality reductions peak at 17.7 percent for 360-day mortality before gradually declining again. 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 of the structural form suggest that access to subspecialists is significantly positively correlated with mortality after 540 days. For example, OLS results indicate that having access to a subspecialized oncologist is associated with a 0.3 percentage point increase in 720-day mortality and a 0.5 percentage point increase in 1,080-day mortality (see Appendix Table E7). This pattern suggests negative selection into treatment, where patients with more severe, complex, or advanced-stage cancers may be more likely to receive care from subspecialized oncologists.

[19] Chemotherapy initiation is defined following the OCM and corresponds to the date of the first qualifying chemotherapy claim. Specifically, the trigger date is the line first expense date on the chemotherapy drug line in Carrier claims, the revenue center date in Outpatient claims, and the fill date in Part D Event (PDE) data.

[20] In Appendix Figure D4 we provide binned scatterplots of the structural form and reduced form relationship.



(a) Mortality Effect

(b) Mortality Effect Relative to Mean

Figure 6: Effect of Subspecialist Access on Mortality

Note: The figure presents the effect of access to a subspecialized oncologist of the relevant cancer type on mortality, estimated using two-stage least squares (2SLS) in the sample of first chemotherapy episodes from 2008 to 2017. Each point reflects a separate regression estimated with the full set of controls and fixed effects specified in Equation 2. The x-axis denotes mortality measured in quarterly intervals (every 90 days), and the y-axis displays the corresponding 2SLS point estimates with 95% confidence intervals, based on standard errors clustered at the ZCTA level. Panel A plots the raw mortality estimates. Panel B rescales the estimates relative to the mean mortality rate at each time interval.

We recognize that tumor characteristics such as cancer stage, grade, histology, and tumor markers are important drivers of cancer outcomes that we cannot measure in claims data. Thus, we conduct a series of robustness checks to ensure that our mortality results are not driven by selection into treatment or by the volume of cases treated by the coordinating oncologist within specific cancer types. In Appendix Table E8 we therefore compare our main mortality estimates to estimates obtained from different specifications and samples. First, we estimate models using samples that vary by the time window over which mortality is measured (Panel B of Appendix Table E8), and find consistent results across specifications. Second, we control for the care-coordinating oncologist's episode volume across the five main cancer types by including inverse hyperbolic sine–transformed measures of episode volume for each category, and again find that results remain similar (Panel C of Appendix Table E8). Finally, we run a version of Equation 2 where we additional allow for unique slope parameters for every ZCTA to account for linear trends and find that although marginally smaller our results remain robust to our main specification (Panel D of Appendix Table E8).

To address concerns that observed mortality improvements could reflect selective entry into chemotherapy rather than differences in care quality, we conduct a complementary analysis using a new sample of over 4.5 million first oncology consultations (Ap-

pendix A). We find no evidence that subspecialists are more likely to initiate chemotherapy within 90, 180, or 360 days following the first visit, suggesting no differential selection on the extensive margin. Among those who do initiate treatment, subspecialist visits are associated with a modest 3-day reduction in time to chemotherapy start—a difference that is marginally significant but unlikely to be clinically meaningful. Most importantly, we find no evidence of mortality effects following subspecialist consultations that are not followed by chemotherapy, and we replicate our main survival effects when restricting the chemotherapy sample to patients with observed consultations. Together, these findings reinforce the interpretation that the survival benefits of subspecialized care are not driven by selective treatment of healthier patients, but by differences in how chemotherapy is delivered and differences in how subspecialists treat patients in comparison to general oncologists.

5.2 Episode Spending

Motivated by the observed mortality benefits of subspecialist access, we next examine its impact on chemotherapy episode spending. We construct comprehensive episode-level spending measures that include Medicare payments, beneficiary out-of-pocket costs, and payments from non-Medicare primary payers, capturing all spending from chemotherapy initiation through 180 days or until death, if earlier. This approach reflects the total cost of care rather than Medicare spending alone. Appendix B provides details on measure construction and descriptive statistics. Briefly, average episode spending has risen over time, driven largely by increases in Part B and Part D expenditures. The share of Part D spending grew from 8% in 2008 to 18% in 2020, highlighting a shift toward oral chemotherapy agents.

Table 1 reports the estimated effect of access to a subspecialized oncologist on various spending measures. We find no statistically significant impact on total spending at the episode level. When disaggregating by spending source, access to subspecialists is not significantly associated with changes in Part A, Part B, or Part D spending. However, while the 2SLS point estimates for Part B and Part D spending are negative and imprecisely estimated, they contrast with the positive and statistically significant associations observed in OLS models. This divergence suggests that subspecialists may be more likely to treat patients with more complex or costly cancers, and that the spending reductions in the IV models likely reflect differences in treatment approach rather than differences in patient severity.

Table 1: Access to Subspecialized Oncologist and Spending

	Total	**Part A**	**Part B**	**Part D**
Panel A: First Stage				
$\sinh^{-1}(\text{DD})$	-0.024***	-0.024***	-0.024***	-0.024***
	(0.001)	(0.001)	(0.001)	(0.001)
Panel B: Reduced Form				
$\sinh^{-1}(\text{DD})$	8.47	-19.71	16.98	11.21
	(27.36)	(12.42)	(22.18)	(11.20)
Panel C: Structural Form				
Any Office Visit Subs.	4,813.83***	1,642.52***	2,726.58***	444.74***
	(83.89)	(36.00)	(68.68)	(32.71)
Panel D: 2SLS				
Any Office Visit Subs.	-357.95	833.76	-717.66	-474.06
	(1,156.93)	(524.94)	(937.57)	(474.08)
Adj. R^2	0.155	0.176	0.190	0.242
Observations	2,165,024	2,165,024	2,165,024	2,165,024
Mean Dep. Var.	41,279.05	7,769.20	28,461.13	5,048.72
F-Stat (1st Stage)	1,823	1,823	1,823	1,823

Notes: The table provides estimates on the effect of access to subspecialized oncologists on different measures of spending for chemotherapy episodes in our main sample. Column 1 provides estimates for total spending, column 2 Part A spending, column 3 Part B spending and column 4 Part D spending. Panel A presents first stage estimates, Panel B the respective reduced form estimates, Panel C the structural form estimates and Panel D provides 2SLS 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. First-stage strength is reported using the Kleibergen-Paap F-statistic. Standard errors are clustered at the ZCTA level. Signif. Codes: ***: 0.01, **: 0.05, *: 0.1.

We further decompose Part B spending and find that the negative point estimate associated with subspecialist access is primarily driven by lower spending in the Carrier file, while spending recorded in the Outpatient file increases (Appendix Table E9).[21] To isolate chemotherapy-related spending, we use HCPCS codes from the Oncology Care Model (OCM) to construct a measure of chemotherapy drug spending, and apply the 2024 Restructured BETOS Classification System (RBCS) to identify injection and infusion services. These breakdowns reveal that the observed cost savings are concentrated in these two components: OCM-defined chemotherapy drug spending falls by approx-

[21]This increase likely reflects that many subspecialists practice in hospital-based settings, including NCI-designated cancer centers, where chemotherapy-related costs are billed through outpatient hospital claims.

imately $2,568, and injection/infusion services (as defined by RBCS) decline by about $700 (Appendix Table E1).

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.

Figure 7 examines the relationship between our instrument and variables falling into 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 (solid purple dots), and an adjusted model (hollow orange dots), which incorporates our full set of controls (excluding chronic conditions and demographics) and fixed effects. The unadjusted results indicate that the instrument is significantly associated with many beneficiary characteristics, highlighting the potential for a violation of the independence assumption and exclusion restriction. However, once we include ZCTA-level controls and fixed effects, these associations effectively disappear, suggesting that conditional independence holds for the set of observable characteristics presented in Figure 7.[22]

[22]We also provide the balancing results as Appendix Table E11.

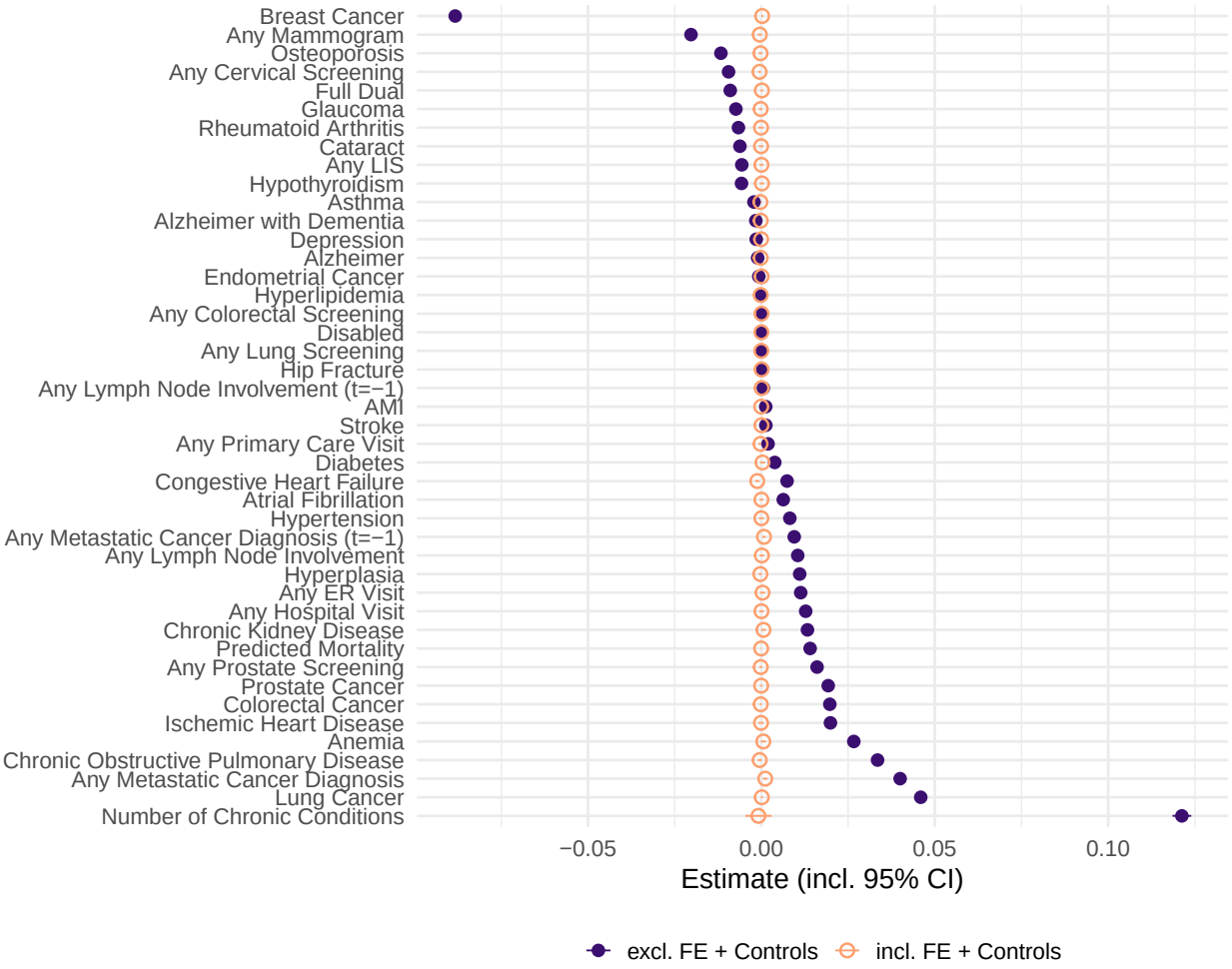


Figure 7: Instrument Balance across Beneficiary Characteristics

Note: This figure displays coefficients from separate regressions of the instrument on the beneficiary characteristics listed on the y-axis. Each point represents a distinct regression. Solid purple dots show unadjusted associations without controls or fixed effects. Hollow orange dots include our full set of ZIP Code Tabulation Area (ZCTA) controls, ZCTA fixed effects, and cancer type-by-year fixed effects. Confidence intervals are based on heteroskedasticity-robust standard errors in unadjusted regressions, and ZCTA-clustered standard errors in the adjusted specifications.

To further strengthen the validity of our exclusion restriction, we conduct two falsification tests. First, for each chemotherapy episode, we randomly reassign differential distances within years and re-estimate our main specification from Equation 2 using 1,080-day mortality as the outcome. We repeat this randomization 100 times and find that none of the 2SLS estimates are statistically significant at conventional levels (see Appendix Figure D5). The null results arise because randomizing differential distances breaks the first-stage relationship—differential distance becomes uncorrelated with access to a subspecialized oncologist of the relevant cancer type—supporting the idea that

our instrument relies on meaningful geographic variation in access.

As a second falsification exercise, we reassign differential distances based on the nearest subspecialist for an unrelated cancer type. For example, we assign breast cancer patients the differential distance to the nearest gastrointestinal cancer subspecialist. We conduct this reassignment across multiple cancer-type pairings and re-estimate Equation 2 using 1,080-day mortality as the outcome (see Appendix Figure D6). The logic of this test is straightforward: if cancer-type-specific expertise is the key mechanism, then proximity to subspecialists of unrelated cancer types should not systematically affect mortality. Moreover, if mortality effects are instead driven by features of the broader practice environment—such as being treated at a high-quality hospital—we would expect to see similar effects regardless of the oncologist's cancer specialization. Consistent with the specificity of expertise as the main mechanism, we find no evidence of systematic mortality effects when using differential distance to subspecialists in unrelated cancer types.

5.4 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 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 cancer type, demographic characteristics, comorbidities, and other health characteristics. 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.

Appendix Table E12 summarizes characteristics of compliers, treated individuals, and the full chemotherapy population. Compliers appear broadly representative of the overall sample. Across key domains—such as cancer type, age, sex, race/ethnicity, comorbidity burden, and predicted mortality—the differences in shares between compliers and the full sample are modest. For example, the share of patients aged 67–70 is 0.302 among compliers versus 0.309 in the full population. Comorbidity profiles and mortality risk also align closely, with slightly lower values among compliers but no major deviations. These patterns suggest that the local average treatment effects (LATE) identified by our instrumental variable strategy are likely generalizable to a wide and policy-relevant

group, rather than reflecting idiosyncratic effects for a narrow subset.

6 Mechanisms

In this section, we examine potential mechanisms underlying our main mortality findings in greater detail. We separate these into three categories. First, we analyze differences in the utilization of health care during chemotherapy that could potentially be linked to differences in mortality. We particularly focus on the enrollment in clinical trials, the age of chemotherapy drugs used for chemotherapy, the role of end of life care and the importance of the health care provider mix for potential care fragmentation. Second, we analyze the effects of the intensive margin of oncologists' specialization on patient outcomes. Finally, we present additional evidence on health outcomes that should be unaffected by access to subspecialized oncologists, providing further support for the validity of our identification strategy.

6.1 Health Care Utilization during Chemotherapy

6.1.1 Clinical Trial Enrollment

One potential mechanism through which subspecialist access improves survival is by facilitating access to novel and potentially life-saving treatments. While clinical trial enrollment alone is unlikely to account for the full mortality effect, it highlights meaningful differences in care utilization between patients with and without access to subspecialists. To assess this, we examine whether beneficiaries enrolled in clinical trials, using Medicare claims data from 2014 onward—when reporting of clinical trial identifier codes (NCTs) became mandatory for covered research services (CMS, 2014). We define clinical trial participation based on the presence of an NCT number on any claim during the year of chemotherapy initiation.

To supplement these 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 C provides further details on the classification methodology.

Figure 8 shows the proportion of first chemotherapy episodes linked to a clinical cancer trial in the same year of chemotherapy initiation separately for episodes managed by subspecialized and general oncologists. Overall, subspecialized oncologists are more likely to manage episodes where the beneficiary is enrolled in any clinical cancer trial. Importantly, 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 Appendix Figure D7 Panels D7c and D7d). On average, episodes managed by general oncologists are linked to a cancer trial in one percent of cases in our sample, with an even lower proportion enrolling in cancer trials specific to the patient's cancer type. In contrast, subspecialists managed episodes are linked to a clinical cancer trial at roughly five to six times this rate.

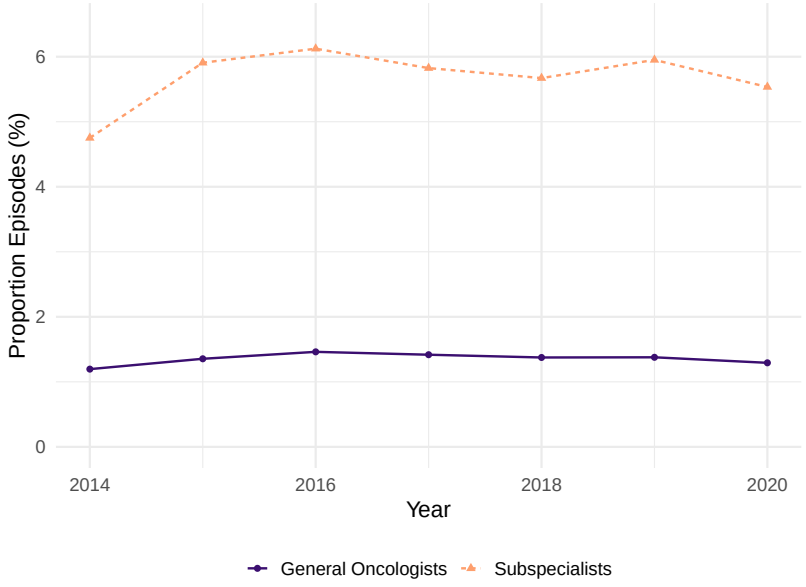


Figure 8: Proportion of Chemotherapy Episodes with Clinical Cancer Trial Enrollment

Note: This figure shows the share of first chemotherapy episodes from 2014 to 2020 in which the beneficiary had any claim containing an NCT number in the same year as chemotherapy initiation. NCT numbers were linked to cancer clinical trials listed in ClinicalTrials.gov. The solid orange line represents episodes coordinated by subspecialized oncologists, while the dashed purple line represents episodes managed by general oncologists.

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.[23] Table 2 presents 2SLS and OLS estimates on the relationship

[23]We do not directly test whether the oncologist is responsible for enrolling the beneficiary in a clin-

between access to a subspecialized oncologist of the relevant cancer type and clinical trial enrollment across different trial types.

Table 2: Clinical Trial Enrollment and Access to Subspecialized Oncologists

	Cancer Trials				
	Any	**Concordant**	**Discordant**	**Unspecified**	**Non-Cancer**
Panel A: First Stage					
$\sinh^{-1}$(DD)	-0.025***	-0.025***	-0.025***	-0.025***	-0.025***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Panel B: Reduced Form					
$\sinh^{-1}$(DD)	-0.001***	0.000***	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Panel B: Structural Form					
Any Office Visit Subs.	0.046***	0.041***	0.006***	0.002***	0.001***
	(0.001)	(0.001)	(0.000)	(0.000)	(0.000)
Panel D: 2SLS					
Any Office Visit Subs.	0.023***	0.020**	0.003	-0.001	0.001
	(0.007)	(0.006)	(0.002)	(0.002)	(0.001)
Adj R^2	0.025	0.022	0.006	0.001	0.000
Observations	1,122,816	1,122,816	1,122,816	1,122,816	1,122,816
Mean Dep. Var.	0.021	0.018	0.003	0.001	0.001
F-Stat (1st Stage)	1,167	1,167	1,167	1,167	1,167

Notes: The table provides estimates of the effect of access 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 concordant to the patients cancer, column 3 any enrollment in a cancer trial discordant to the patients cancer, column 4 provides results for enrollment in in unspecified or multi-cancer trials and column 5 provides for enrollment in non cancer trials. Panel A presents first stage estimates, Panel B the respective reduced form estimates, Panel C the structural form estimates and Panel D provides 2SLS estimates. All models include demographic, ZCTA level and chronic conditions controls as well fixed effects for the beneficiaries' ZCTA and cancer type by year. Standard errors are clustered at the ZCTA level. First-stage strength is reported using the Kleibergen-Paap F-statistic. Signif. Codes: ***: 0.01, **: 0.05, *: 0.1.

Panel A suggests that access to a subspecialist of the relevant cancer type increases enrollment in any cancer trial by 2.3 percentage points, a 110 percent increase relative to the sample mean. The effect is similar in magnitude for cancer-type concordant trials, with an estimated increase of 2 percentage points, or 111 percent relative to the mean.

ical trial, nor do we investigate whether claims with NCT numbers are specifically linked to oncologists coordinating the beneficiary's care.

Notably, there is no statistically significant difference in enrollment into trials discordant to the beneficiaries cancer type, unspecified cancer trials or non-cancer trials, in contrast to the structural form estimates and raw unadjusted shares presented in Appendix Figure D7. While we do not directly test whether trial enrollment drives mortality reductions and few Medicare beneficiaries enroll in clinical trials, these findings reflect a key difference in healthcare utilization between individuals with access to subspecialized oncologists that may contribute modestly to the observed survival benefits of subspecialist care. Importantly it also suggests clear differences in practice patterns of subspecialized oncologists.

6.1.2 Age of Cancer Drugs

To assess the age of cancer drugs administered during treatment, we linked FDA approval years—sourced from the NCI SEER Program and mapped to corresponding HCPCS and NDC codes—to Medicare claims data. We conducted two complementary analyses. First, we examined whether access to subspecialized oncologists increases the likelihood that a patient receives a chemotherapy drug approved within the two years prior to treatment initiation. Second, we calculated the average approval age of all chemotherapy drugs administered in the year of treatment initiation. For Part B drugs, we used a weighted average based on the number of claim lines per HCPCS code; for Part D drugs, we weighted by the number of prescription fills per NDC code. Due to limitations in linking approval dates to cancer-specific indications and the absence of detailed clinical information in Medicare claims, we focus on the approval year of the generic substance rather than indication-specific approvals.[24]

Appendix Table E13, Panel D, presents 2SLS estimates for the effect of subspecialist access on the probability of receiving a newly approved cancer drug (within two years of treatment initiation). Column 1 shows a marginally significant 2 percentage point increase in the probability of receiving a newly approved chemotherapy drug—corresponding to a 6.9% increase relative to the mean. While no significant effect is observed for newly approved Part B drugs (Column 2), we do detect a marginally significant increase in the use of recently approved Part D drugs (Column 3). Together, these results provide suggestive evidence that subspecialist access may modestly increase the likelihood of receiving newly approved therapies.

Appendix Table E14 reports 2SLS estimates for the effect of subspecialist access on the average approval age of chemotherapy drugs. Panel D shows no statistically signif-

[24]This approach does not account for dosage, intensity, or drug quantity, and therefore reflects average drug age based on frequency of use rather than treatment intensity or volume.

icant effects on the weighted average age of drugs overall (Column 1), for Part B drugs (Column 2), or for Part D drugs (Column 3). However, all point estimates are negative, indicating a consistent—though imprecise—trend toward use of slightly newer drugs among subspecialists. These findings should be interpreted with caution, as the weighting methodology does not capture differences in drug dosage or treatment intensity, which likely limits the accuracy of this average age measure.

Taken together, the results suggest that while access to subspecialized oncologists does not substantially shift the overall age profile of chemotherapy drugs, it may slightly increase the likelihood of receiving more recently approved treatments, particularly oral agents covered under Part D.

6.1.3 End of Life Care

Chemotherapy patients in our main sample—who are aged 67 and older and often have multiple chronic conditions—experience high mortality. Within one year of treatment initiation, 22.6% of beneficiaries die; this rises to 36.5% within two years and 44.9% within three.[25] Given this context, the quality and intensity of care delivered near the end of life becomes a key dimension of cancer treatment, with implications for both patient experience and health system costs (Zeltzer et al., 2023). Access to subspecialized oncologists may influence not only treatment outcomes but also end-of-life decision-making.

To assess this, we construct four binary indicators of care intensity in the last 30 days of life, defined for patients who die during or within 30 days after their chemotherapy episode: (i) any emergency room (ER) visit, (ii) any intensive care unit (ICU) stay, (iii) hospice enrollment between 30 and 3 days before death, and (iv) hospice initiation in the final 3 days of life. ER and ICU use are typically viewed as indicators of aggressive, high-intensity care (Jang et al., 2015), while earlier hospice use is associated with more comfort-oriented, less intensive care. In contrast, very late hospice initiation often reflects poor transitions at the end of life.

To avoid conditioning on survival, we define these outcomes across the full sample of chemotherapy patients, interacting each measure with an indicator for short-term mortality (i.e., death during or shortly within 30-days of chemotherapy termination). This approach allows us to estimate the effect of subspecialist access on both the likelihood of short-term death and, conditional on death, the intensity of care received.

Appendix Table E15 presents the results. We find no statistically significant effect of

[25] In comparison, the age-adjusted death rate in the general U.S. population in 2015 was 733 per 100,000—less than 1% (Xu et al., 2016).

subspecialist access on ER or ICU use at the end of life, nor on late hospice enrollment. However, access to a subspecialized oncologist significantly reduces the likelihood of earlier hospice use (between 30 and 3 days before death). The point estimate corresponds to a 35% relative decline compared to the sample mean, suggesting a lower likelihood of comfort-focused care. This finding is reinforced by a reduction in hospice-related Medicare spending: as shown in Appendix Table E9, access to subspecialists reduces hospice spending by $272 per episode.

Taken together, these results suggest that subspecialist access may be associated with a modest shift toward more intensive treatment patterns at the end of life, potentially reflecting a preference for continued active treatment over early palliative care transitions.

6.1.4 Provider Mix and Fragmentation

Next, we assess whether access to subspecialized oncologists affects the composition and diversity of providers involved during chemotherapy. Using the Carrier and Outpatient files, we extract all office visits occurring within each beneficiary's chemotherapy episode and construct three episode-level measures: (i) the total number of office visits, (ii) the number of unique providers, and (iii) the number of unique provider specialties. We estimate the effect of subspecialist access on each outcome using our main instrumental variable strategy (Equation 2).

Results in Panel D of Appendix Table E16 show that subspecialist access has no discernible effect on the number of office visits or the diversity of unique provider specialties for which patients have office visits during the chemotherapy episode. However, it is associated with a modest but statistically significant 4.3% reduction in the number of unique providers relative to the mean.

These findings suggest that subspecialists deliver care through a stable and coordinated provider team, without increasing provider diversity or visit frequency. While specialization is often associated with greater care fragmentation, we find no evidence of such an effect. Instead, the modest decline in provider count points to more streamlined care delivery as a result of access to subspecialists.

6.2 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 45 different categories, as opposed to the 5 broad categories used above (see Appendix Table E2 for an overview of those detailed cancer types). 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 measure 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.[26]

In Panel D of Table 3, 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 oncologist specialization on patient mortality. Previously, we argued that differential distance should affect patient health outcomes only through access to a subspecialized oncologist, and

[26] Appendix Figure D8 illustrates the distribution of oncologists' Herfindahl-Hirschman Index (HHI) values and their association with the subspecialist definition used in this study. As expected, higher HHI percentiles—indicating greater concentration in treating a specific cancer type—are associated with a higher proportion of oncologists classified as subspecialists. Notably, some subspecialists also appear in the lower HHI deciles. This is consistent with our definition, which is based on broader cancer categories: an oncologist may qualify as a subspecialist even while treating a mix of detailed cancer types within a broader category.

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.

Table 3: Mortality Effects of Oncologist Specialization

	180-Day	**360-Day**	**720-Day**	**1080-Day**
Panel A: First Stage				
$\sinh^{-1}(\text{DD})$	-0.008***	-0.008***	-0.008***	-0.008***
	(0.000)	(0.000)	(0.000)	(0.000)
Panel B: Reduced Form				
$\sinh^{-1}(\text{DD})$	-0.000	0.001***	0.001***	0.001***
	(0.000)	(0.000)	(0.000)	(0.000)
Panel C: Structural Form				
HHI	-0.014***	-0.024***	-0.039***	-0.048***
	(0.001)	(0.001)	(0.002)	(0.002)
Panel D: 2SLS				
HHI	-0.023	-0.126***	-0.159***	-0.141**
	(0.036)	(0.046)	(0.050)	(0.051)
Adj R^2	0.120	0.224	0.302	0.325
Observations	1,681,119	1,681,119	1,681,119	1,681,119
Mean Dep. Var.	0.108	0.226	0.365	0.449
F-Stat (1st Stage)	715	715	715	715

Notes: This table reports estimates of the effect of the degree of cancer type subspecialization of the chemotherapy care coordinating oncologist on mortality for all first chemotherapy episodes between 2008 and 2017. Panel A shows first stage estimates. Panel B shows the reduced form estimates. Panel C shows the structural form estimates, and Panel D provides the 2SLS estimates. All models control for demographics, ZCTA characteristics, comorbidities, and include ZCTA and cancer type-by-year fixed effects. Standard errors are clustered at the ZCTA level. First-stage strength is reported using the Kleibergen-Paap F-statistic. Significance levels: *** p<0.01, ** p<0.05, * p<0.1.

Focusing on Panel D of Table 3, 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 HHI point increase reduces 1,080-day mortality by 1.4 percentage points, corresponding to a 3.1 percent reduction relative to the respective mean mortality rate. For context, the average difference in HHI between a subspecialized and a general oncologist is 0.4 HHI points. Scaling our point estimate using this difference suggests a 5.6 percentage point decline in 1,080-day mortality, which is larger than the corresponding estimate in Figure 6, Panel 6a, suggesting potential non-linearities in the relationship between HHI and mortality. OLS estimates of the structural form in Panel C of Table 3 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.

6.3 Placebo Outcomes

One potential concern is that patients with access to subspecialized oncologists may also have better access to higher-quality clinicians more broadly, which could independently contribute to improved survival. If true, the observed mortality benefits might reflect the overall quality of a patient's care team rather than the specific expertise of subspecialized oncologists. However, prior evidence—such as higher enrollment in cancer-specific clinical trials and improved outcomes among more narrowly focused oncologists—suggests that subspecialist access itself plays a meaningful role. To further test whether our results are driven by cancer-specific expertise rather than broader differences in care quality, we examine placebo outcomes unrelated to oncology care, such as mortality following hip fracture, stroke, and acute myocardial infarction.

If broader care quality 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 (i) myocardial infarctions and strokes are primarily driven by cardiovascular health and pre-existing risk factors rather than oncologic treatment decisions, (ii) hip fractures are mostly (but not entirely) related to musculoskeletal health, falls, and osteoporosis, which are not primary concerns in oncologic care, and (iii) while some cancer treatments may have secondary effects on car-

diovascular and bone health, these outcomes are generally not the focus of oncologists when managing cancer treatment regimens.

Appendix Table E18 reports results from our main specification, assessing the effect of access to subspecialized oncologists on placebo outcomes—specifically, inpatient diagnoses of acute myocardial infarction, hip fracture, and stroke. These binary outcomes are measured separately for the first and second years following chemotherapy initiation. Panel D shows that the 2SLS estimates reveal no statistically significant associations between subspecialist access and any of the placebo outcomes at the conventional 5% level. While we observe a marginally significant effect (at the 10% level) for stroke in the year following chemotherapy initiation, the point estimate is positive, indicating a higher likelihood of stroke associated with subspecialist access. If improved overall care quality were the primary mechanism behind our observed mortality reductions, this finding would not align with that interpretation. Panel C reports OLS estimates of the structural form equation, which—despite some statistical significance—indicate precisely estimated effects close to zero, suggesting no meaningful association. These results support the interpretation that subspecialist access affects cancer-specific outcomes without influencing unrelated health events.[27]

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.

7 Conclusion

This paper provides novel causal evidence on the implications of specialization in high-skill professions. Focusing on medical oncology, we demonstrate that access to subspecialized oncologists improves medium- and long-term survival and reduces episode-level spending. Using quasi-exogenous variation in differential distance to subspecialists versus general oncologists, we estimate that access to a cancer-type–matched subspecialist reduces 1,080-day mortality by 10% relative to the mean, without affecting non-cancer-related health outcomes—highlighting the value of specialized clinical expertise beyond general physician skill. We also find that subspecialist access lowers spending on Part B chemotherapy drugs per episode, but does not have overall effects on total episode

[27] Using alternative definitions based on chronic condition indicators for AMI, hip fracture, and stroke yields similarly null results.

spending within Medicare.

Our findings are robust to a range of model specifications and control variables, and we provide additional support for the validity of our instrument through falsification tests. While subspecialists initiate chemotherapy slightly earlier, we find no evidence that they selectively treat patients with lower baseline mortality risk, reinforcing the exogeneity of our identification strategy. Mechanistically, we document that patients with subspecialist access are significantly more likely to enroll in clinical trials—especially those relevant to their cancer type—and receive chemotherapy regimens composed of newer drugs. We also find suggestive evidence of more intensive end-of-life care, reflected in lower hospice use. Importantly, we do not observe increased fragmentation of care, as measured by provider diversity or visit patterns.

In additional analyses, we show that oncologists with a deeper focus on specific cancer types—as measured by Herfindahl-Hirschman Indices (HHI)—achieve better mortality outcomes, reinforcing the returns to deep clinical focus. At the same time, we find no meaningful impact of subspecialist access on unrelated health outcomes such as hip fractures, strokes, or myocardial infarctions, further underscoring the domain-specific nature of their effects.

Despite the benefits of subspecialization, our findings raise important concerns about equitable access subspecialist oncologists. Subspecialists are disproportionately located in higher-income and urban areas, limiting access for rural and underserved populations. This geographic concentration highlights the need to consider how healthcare systems can structure access to specialized expertise while ensuring equity. Overall, our study contributes to broader discussions in economics and health policy on specialization, labor markets, and 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.

Caswell-Jin, Jennifer L, Marissa B Reitsma, Hao Tang, James C Dickerson, Shannon Phillips, Esther M John, Allison W Kurian, Becky Staiger, and Jeremy D Goldhaber-Fiebert. 2025. "Physician Specialization and Receipt of Updated Breast Cancer Care in the United States: A SEER-Medicare Analysis." *JCO Oncology Practice*, OP–25.

Cebul, Randall D., James B. Rebitzer, Lowell J. Taylor, and Mark E. Votruba. 2008. "Organizational Fragmentation and Care Quality in the U.S. Healthcare System." *Journal of Economic Perspectives*, 22(4): 93–113.

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.

Chandra, Amitabh, and Jonathan Skinner. 2012. "Technology growth and expenditure growth in health care." *Journal of Economic Literature*, 50(3): 645–680.

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." https://www.cms.gov/priorities/innovation/files/x/ocm-cancercodelists.pdf, Accessed 2023.

CMS. 2024. "Restructured BETOS Classification System." *https://data.cms.gov/provider-summary-by-type-of-service/provider-service-classifications/restructured-betos-classification-system*, Accessed June 2025.

CMS. 2025. "Oncology Care Model." *https://www.cms.gov/priorities/innovation/innovation-models/oncology-care*, Accessed July 2025.

Coleman, James, Elihu Katz, and Herbert Menzel. 1957. "The diffusion of an innovation among physicians." *Sociometry*, 20(4): 253–270.

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

Currie, Janet, W Bentley MacLeod, and Jessica Van Parys. 2016. "Provider practice style and patient health outcomes: The case of heart attacks." *Journal of health economics*, 47: 64–80.

Cutler, David, Jonathan S Skinner, Ariel Dora Stern, and David Wennberg. 2019. "Physician beliefs and patient preferences: a new look at regional variation in health care spending." *American Economic Journal: Economic Policy*, 11(1): 192–221.

Cutler, David M, and Mark McClellan. 2001. "Is technological change in medicine worth it?" *Health affairs*, 20(5): 11–29.

Cutler, David M, Robert S Huckman, and Mary Beth Landrum. 2004. "The role of information in medical markets: an analysis of publicly reported outcomes in cardiac surgery." *American Economic Review*, 94(2): 342–346.

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.

Garicano, Luis. 2000. "Hierarchies and the Organization of Knowledge in Production." *Journal of political economy*, 108(5): 874–904.

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.

Jang, Raymond W, Monika K Krzyzanowska, Camilla Zimmermann, Nathan Taback, and Shabbir MH Alibhai. 2015. "Palliative care and the aggressiveness of end-of-life care in patients with advanced pancreatic cancer." *Journal of the National Cancer Institute*, 107(3): dju424.

Kaiser Family Foundation. 2025. "What to Know About Medicare Spending and Financing." https://www.kff.org/medicare/issue-brief/what-to-know-about-medicare-spending-and-financing/, Accessed: 2025-02-27.

Karadakic, René, Christopher Manz, Arno Cai, David C Chan, Bruce E Landon, Jukka-Pekka Onnela, Nancy L Keating, and Michael L Barnett. 2025. "Geographic Variation in the Utilization of Cancer Care From Subspecialized Medical Oncologists in the United States, 2008 to 2020." *Annals of internal medicine*.

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." https://dx.doi.org/10.2139/ssrn.4960603.

Mailankody, Sham, and Vinay Prasad. 2015. "Five years of cancer drug approvals: innovation, efficacy, and costs." *JAMA oncology*, 1(4): 539–540.

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.

Milligan, Michael, Parsa Erfani, E John Orav, Stephen Schleicher, Gabriel A Brooks, and Miranda B Lam. 2024. "Practice consolidation among US medical oncologists, 2015-2022." *JCO oncology practice*, 20(6): 827–834.

National Bureau of Economic Research. 2025. "ZIP Code Distance Database." https://www.nber.org/research/data/zip-code-distance-database, Accessed: 2025-02-27.

National Cancer Institute. 2025*a*. "Age and Cancer Risk." U.S. Department of Health and Human Services, National Institutes of Health.

National Cancer Institute. 2025*b*. "Observational Research in Oncology Toolbox: Cancer Medications Enquiry Database." https://seer.cancer.gov/oncologytoolbox/, Accessed 2025.

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.

Pantziarka, Pan, Rica Capistrano I, Arno De Potter, Liese Vandeborne, and Gauthier Bouche. 2021. "An open access database of licensed cancer drugs." *Frontiers in pharmacology*, 12: 627574.

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

Sahni, Nikhil R, Maurice Dalton, David M Cutler, John D Birkmeyer, and Amitabh Chandra. 2016. "Surgeon specialization and operative mortality in United States: retrospective analysis." *Bmj*, 354.

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. 1819. *An inquiry into the nature and causes of the wealth of nations.* London:Printed for William Allason, and J. Maynard, and W. Blair.

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

Xu, Jiaquan, Sherry L. Murphy, Kenneth D. Kochanek, and Elizabeth Arias. 2016. "Mortality in the United States, 2015." National Center for Health Statistics NCHS Data Brief 267, Hyattsville, MD.

Yale Cancer Clinic. 2025. "Delivering Cancer Care in our Communities." https://medicine.yale.edu/news-article/delivering-cancer-care-in-our-communities/, Accessed: 2025-07-11.

Zeltzer, Dan, Liran Einav, Amy Finkelstein, Tzvi Shir, Salomon M Stemmer, and Ran D Balicer. 2023. "Why is end-of-life spending so high? Evidence from cancer patients." *Review of Economics and Statistics*, 105(3): 511–527.

A Selection into Chemotherapy

Our main analysis focuses on beneficiaries who initiate chemotherapy, as defined by the OCM framework. However, a potential concern is that improved mortality outcomes associated with subspecialist access could reflect selection into chemotherapy rather than differences in treatment quality. For instance, subspecialists may be more selective, initiating treatment only for healthier patients, thereby improving observed survival among those who receive chemotherapy.

To examine this, we constructed a new sample designed to capture the first point of contact between patients and medical oncologists. Specifically, we identify, for each beneficiary, the first office visit with a medical oncologist between 2008 and 2020, where the claim includes a cancer diagnosis corresponding to one of the major cancer types in our main analysis (breast, gastrointestinal, hematologic, prostate/genitourinary, or thoracic cancers). For each beneficiary, we retain only the earliest such visit within this period.

The purpose of this "first visit" sample is twofold. First, it aims to isolate the initial clinical encounter at which an oncologist and patient discuss and decide on potential systemic therapy—representing the key decision point that precedes treatment initiation. Second, by focusing on the first oncologist visit linked to a relevant cancer diagnosis, we reduce the likelihood of capturing routine follow-up or surveillance visits that do not reflect a true treatment decision.

In the next step, we link this first-visit information to our chemotherapy episode data, restricting the latter to first chemotherapy episodes for each beneficiary. This linkage allows us to determine, for each patient's initial oncologist visit, whether it subsequently resulted in the initiation of chemotherapy. We do not require the cancer diagnosis recorded at the first visit to match the primary cancer diagnosis in the chemotherapy sample. Although the vast majority of first visits that lead to chemotherapy share the same cancer type, differences in diagnosis codes may arise due to coding errors or changes in diagnostic classification.

We then enrich this dataset by incorporating information on each oncologist's subspecialization status and retain only visits involving oncologists for whom subspecialization information is available.[28] We further merge in beneficiary characteristics, including demographics, chronic conditions, and data on the characteristics of the ZCTA of the beneficiary, as well as measures of geographic access—specifically, the distance to the nearest generalist and subspecialized oncologist for the relevant cancer type, following

[28]Subspecialization measures are derived from chemotherapy episode data available from 2008 to 2020.

the same procedure as in our main analysis.

To reduce the influence of extreme values, we exclude first visits for which the distances to both the nearest general and subspecialized oncologist exceed the 95th percentile of the respective distance distributions. Finally, we exclude cases where chemotherapy was initiated prior to the identified first office visit, as these typically represent episodes in which the initial point of contact occurred with a provider outside medical oncology—most commonly general surgeons, radiation oncologists, urologists, or internal medicine physicians.[29]

This results in a sample of 4,554,039 first office visits, where 33.4% have initiated chemotherapy within 360 days of the first visit, and 32.2% have initiated chemotherapy within 360 days for the same cancer type as the initial visit. We follow a similar identification strategy as in our main analysis, but specifically we estimate the following equation:

$$\text{Visit}_i = \alpha + \beta \cdot \text{DD}_{t(i)z(i)} + \delta X_i + \tau_{t(i)} + \gamma_{z(i)} + \psi \text{D}_{t(i)z(i)} + \varepsilon_i \tag{4}$$

for visit 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 Visit_i is a binary indicator variable equal to one if a beneficiary had the first office visit with a subspecialized oncologist of the relevant cancer type and zero otherwise.

In a second step, we estimate the effect of initial subspecialist consultation on the probability to initiate chemotherapy. We estimate:

$$Y_i = \alpha + \beta \cdot \widehat{\text{Visit}}_i + \delta X_i + \tau_{t(i)} + \gamma_{z(i)} + \psi \text{D}_{t(i)z(i)} + \varepsilon_i \tag{5}$$

where Y_i is the outcome variable for example a binary indicator equal to one if chemotherapy was initiated within a specific time windows (e.g. within 90, 180, or 360 days after the initial consultation).

The first-stage relationship between the instrument and the probability of having a first office visit with a subspecialized oncologist is strong, exhibiting a statistically sig-

[29] These specialties frequently treat cancer patients but are not included in our chemotherapy sample as medical oncologists, and subspecialty classifications are not available for them.

nificant coefficient and an F-statistic well above conventional thresholds for instrument relevance (Lee et al., 2022). In Table A1 we present the first stage relationship for our first visit sample. Notably, the estimated coefficient is slightly smaller in magnitude (-0.018) compared to the corresponding first-stage estimate in our chemotherapy sample (-0.024). This suggests that differential distance between subspecialized and general oncologists exerts a somewhat weaker influence on the likelihood of seeing a subspecialist at the initial consultation stage than at the point of chemotherapy initiation. This pattern is intuitive: many initial consultations occur before a definitive cancer diagnosis or treatment decision, when patients may not yet perceive the situation as requiring the same level of specialized expertise as at the time of active chemotherapy.

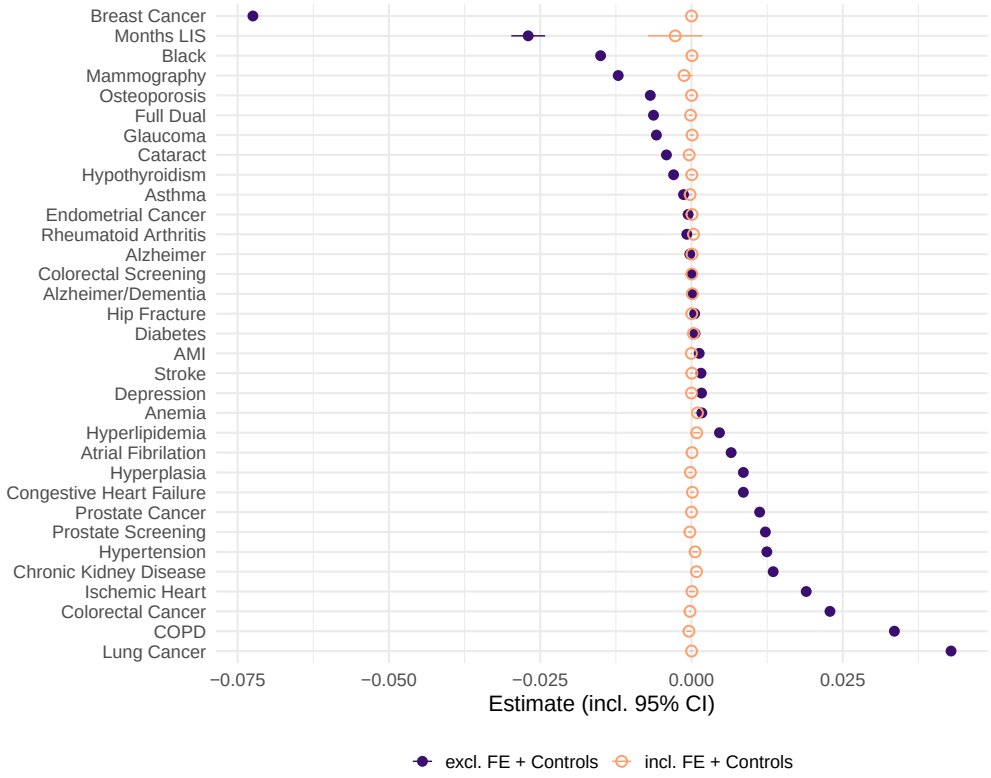


Figure A1: Instrument Balance across Beneficiary Characteristics

Note: This figure displays coefficients from separate regressions of the instrument on the beneficiary characteristics listed on the y-axis for the full sample of first office visits. Each point represents a distinct regression. Solid purple dots show unadjusted associations without controls or fixed effects. Hollow orange dots include our full set of ZIP Code Tabulation Area (ZCTA) controls, ZCTA fixed effects, and cancer type-by-year fixed effects. Confidence intervals are based on heteroskedasticity-robust standard errors in unadjusted regressions, and ZCTA-clustered standard errors in the adjusted specifications.

We provide evidence that our instrument satisfies both the relevance and independence assumptions in this new office visit sample. Table A1 shows a strong first-stage

relationship between the instrument and the likelihood of having a first new office visit with a subspecialized oncologist, with the association remaining robust after controlling for covariates and fixed effects; the large F-statistic confirms sufficient instrument strength (Lee et al., 2022). In support of the independence assumption, Figure A1 demonstrates that the instrument is uncorrelated with observable beneficiary characteristics once we condition on ZCTA and cancer type-by-year fixed effects.

Table A1: First Office Visits: First Stage

	First Office Visit Subspecialist
$\sinh^{-1}(\text{DD})$	-0.018***
	(0.000)
ZCTA FE	Yes
Cancer-Year FE	Yes
Adj. R^2	0.120
Observations	4,554,039
Mean Dep. Var.	0.110
F-Stat (1st stage)	1,767

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 for the sample of first office visits. Standard errors are clustered at the ZCTA level. Reported first-stage F-statistics are Kleibergen-Paap statistics. Signif. Codes: ***: 0.01, **: 0.05, *: 0.1.

We estimate Equation 5 using binary indicators for chemotherapy initiation within 90, 180, and 360 days following a patient's first office visit with an oncologist. As shown in Table A2, we find no statistically significant effect of having this first visit with a subspecialist on the likelihood of initiating chemotherapy within these time frames. This result holds for both any chemotherapy initiation and initiation of chemotherapy that corresponds to the main cancer diagnosis of the first visit. These findings suggest that, at least on the extensive margin, subspecialists do not initiate chemotherapy at higher rates than generalists. While this analysis cannot fully rule out differential pathways into the initial office visit based on prior care, we find no evidence of systematic differences in observable beneficiary characteristics correlated with our instrument—supporting the validity of the identification strategy in this subsample.

Table A2: First Office Visits and Probability of Chemotherapy Initiation

	≤ 90	≤ 180	≤ 360
Panel A: Any Chemotherapy			
First Office Visit Subs.	0.022	0.015	0.012
	(0.014)	(0.014)	(0.015)
Adj R^2	0.091	0.095	0.093
Observations	4,554,039	4,554,039	4,554,039
Mean Dep. Var.	0.275	0.310	0.334
F-Stat (1st Stage)	1,767	1,767	1,767
Panel B: Cancer Type Concordant Chemotherapy			
First Office Visit Subs.	0.020	0.013	0.012
	(0.014)	(0.014)	(0.014)
Adj R^2	0.093	0.097	0.0.095
Observations	4,554,039	4,554,039	4,554,039
Mean Dep. Var.	0.267	0.300	0.322
F-Stat (1st Stage)	1,767	1,767	1,767

Notes: This table reports 2SLS estimates of the effect of having a first office visit with a subspecialized oncologist of the relevant cancer type on indicators for chemotherapy initiation. Each column represents a separate regression. Column 1 reports effects on initiation within 90 days, Column 2 within 180 days, and Column 3 within 360 days of the office visit. Panel A estimates effects on initiating any chemotherapy, while Panel B restricts the outcome to chemotherapy corresponding to the main diagnosis of the first visit. All models include demographic controls, ZCTA-level characteristics, indicators for chronic conditions, and fixed effects for ZCTA and cancer type-by-year. Standard errors are clustered at the ZCTA level. First-stage strength is measured using the Kleibergen–Paap F-statistic. Significance levels: *** p<0.01, ** p<0.05, * p<0.1.

In the next step, we examine whether the timing of chemotherapy initiation differs depending on whether the first oncology visit was with a subspecialist. For the subset of patients who initiated chemotherapy within 360 days of their first visit, we construct an outcome measure capturing the number of days between the office visit and chemotherapy initiation. Figure A2 presents the unadjusted distribution of the log-transformed time to chemotherapy initiation, comparing patients seen by general oncologists and those seen by subspecialized oncologists.

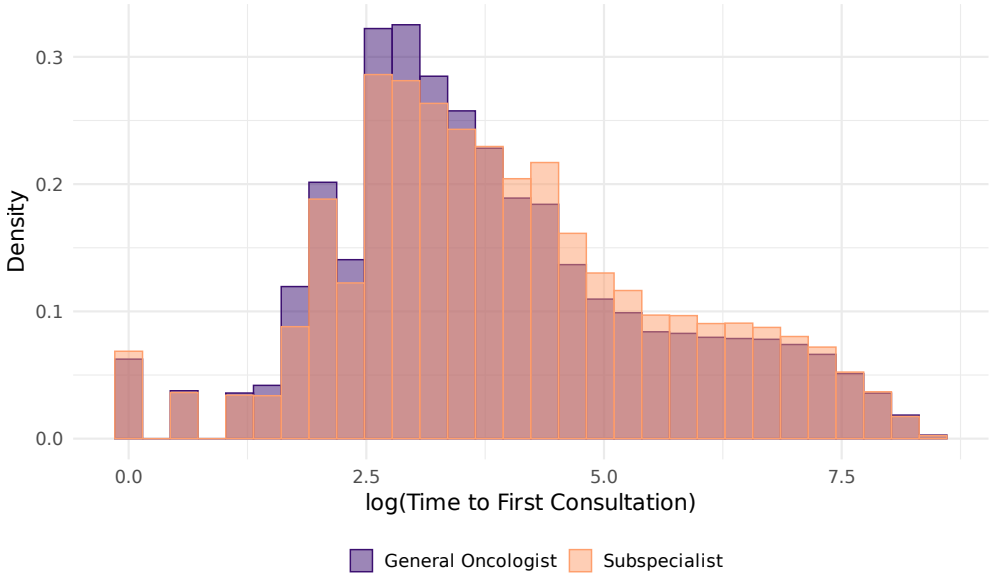


Figure A2: Distribution of (Log) Time to Chemotherapy Start

Note: This figure displays the distribution of time (in days) between a beneficiary's first oncology office visit and the initiation of chemotherapy, restricted to episodes where chemotherapy was initiated within 360 days. The histogram is shown separately for visits with general oncologists (purple) and subspecialized oncologists of the relevant cancer type (orange).

We next estimate Equation 5 using the log of time to chemotherapy initiation as the outcome and report the results in Table A3. While we previously found no difference in initiation rates, the results suggest that a first office visit with a subspecialist is associated with an average reduction in time to chemotherapy initiation of approximately 11.3%. This estimate is marginally statistically significant at the 10% level. Given an average initiation time of 27 days, this corresponds to a reduction of roughly 3.1 days. Although timely diagnosis and treatment are critical for some cancers, such a modest difference is unlikely to be clinically meaningful. A delay of three days could easily reflect scheduling logistics—such as the difference between starting treatment before or after a long weekend—rather than a substantive difference in care quality.

Table A3: First Office Visit with as Subspecialist and Time to Chemotherapy Initiation

	Log(Time to Chemotherapy Initiation)
First Office Visit Subs.	-0.113*
	(0.061)
Adj. R^2	0.061
Observations	1,518,945
Mean Dep. Var.	3.306
F-Stat (1st stage)	1,356

Notes: The table provides the 2SLS estimate for effect of a first office visit with a subspecialist on the natural logarithm of time to chemotherapy initiation. All models include demographic, ZCTA level and chronic conditions controls as well as fixed effects for the beneficiaries' ZCTA and cancer type by year fixed effects. Reported first-stage F-statistics are Kleibergen-Paap statistics. Signif. Codes: ***: 0.01, **: 0.05, *: 0.1.

The final analysis using the first office visit sample examines the effect of a first office visit with a subspecialist on patient mortality. We construct binary indicators for mortality within 180, 360, 720, and 1080 days following a patient's first office visit with a medical oncologist. Figure A3 presents the corresponding 2SLS estimates from Equation 5 for four samples: (i) all first office visits from 2008 to 2017 (FV – Full), (ii) first visits followed by chemotherapy initiation (FV – With Chemo), (iii) first visits not followed by chemotherapy (FV – No Chemo), and (iv) our main sample of first chemotherapy episodes (Main Chemo), where estimates are obtained from Equation 2. Three key patterns emerge. First, there is no evidence of short-term mortality effects in any subsample. Second, the largest mortality reductions are consistently observed in the main chemotherapy sample. Third, while the estimates differ slightly in magnitude across subsamples, the overall direction of effects is consistent, lending support to the idea that subspecialized oncologic care may improve longer-term outcomes, particularly for patients actively receiving chemotherapy. Importantly, we find no statistically significant mortality effects among patients whose first visit does not result in chemotherapy initiation, supporting the notion that subspecialists are not selecting patients based on unobserved mortality risk. This lends further credibility to our interpretation that differences in outcomes are driven by treatment pathways rather than selection into subspecialist care.

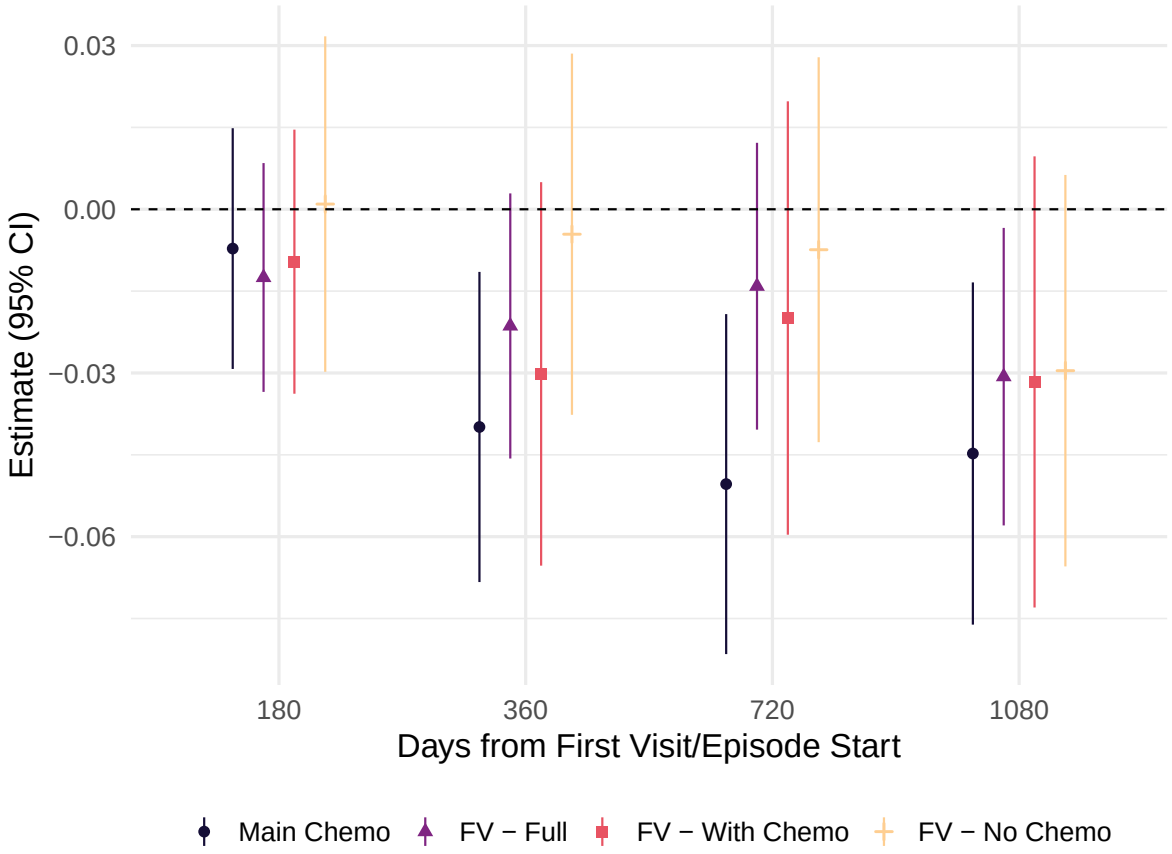


Figure A3: First Office Visit with a Subspecialist and Mortality

Note: This figure presents 2SLS estimates of the effect of a first office visit with a subspecialized oncologist on mortality, based on regressions of Equation 5. Each point reflects a separate regression for mortality measured at 180, 360, 720, and 1080 days after the initial visit. Black circles represent estimates from our main analysis sample of first chemotherapy episodes ("Main Chemo"). Purple triangles depict estimates for the full sample of first office visits ("FV – Full"). Red squares correspond to the subsample of first visits followed by chemotherapy initiation ("FV – With Chemo"), and yellow crosses show estimates for first visits not followed by chemotherapy ("FV – No Chemo"). All models include our standard set of controls: beneficiary demographics, ZCTA characteristics, chronic conditions, as well as ZCTA and cancer type-by-year fixed effects. 95% confidence intervals are based on standard errors clustered at the beneficiary ZCTA level.

B Spending Definitions

We construct detailed measures of healthcare spending using individual claims from Medicare Parts A, B, and D. This includes inpatient, hospice, home health, and skilled nursing facility (SNF) claims from Part A; carrier, outpatient, and durable medical equipment (DME) claims from Part B; and prescription drug claims from Part D. For each claims file, we identify relevant records and associated payment variables, classifying payments into three categories: (1) Medicare payments, (2) beneficiary out-of-pocket payments, and (3) payments by non-Medicare primary payers. These spending amounts are aggregated by date at the beneficiary level to allow for analysis over defined time periods.

Episode-level spending is defined as the total spending from the date of chemotherapy initiation through 180 days post-initiation, or until the date of death if the beneficiary dies within that period. For each claims file, we select relevant variables and exclude records (at the claim or line level) that do not contribute to healthcare spending. Spending is calculated separately for each file and disaggregated by payer type (Medicare, beneficiary, or other primary payer). A full overview of variable definitions and selection criteria is provided in Table B2. In order to better understand drivers of spending we also constructed episode level spending of different RBCS subcategories (CMS, 2024).

Table B1: Spending Measures Comparison with Prior Literature

	Keating et al. (2021)	**Main Sample**
Total	30,946	39,456
Part A	5,966	7,664
Part B	18,503	27,183
Part D	7,794	4,609

Notes: The table provides the average spending for chemotherapy episodes as defined by the Oncology Care Model (OCM) in USD. Column 1 provides spending measures obtained from Table 2 in Keating et al. (2021), where we averaged the OCM intervention group baseline and intervention columns. In column 2 we provide spending per chemotherapy episode from our main sample for the year 2014 corresponding to first year of observation in the comparison article.

We benchmark our spending measures against those reported in Keating et al. (2021), who evaluate spending in Oncology Care Model (OCM) participating practices compared to propensity-matched non-OCM practices. While their analysis covers a shorter time period (2014–2019), their definition of chemotherapy episodes closely aligns with ours. Due to our focus on the largest cancer types and the application of additional sample selection criteria, exact comparability is not expected. However, our spending

estimates closely track those reported in Keating et al. (2021), supporting the consistency of our measures. In Table B1, we compare average episode-level spending in our sample for the year 2014 to OCM benchmarks presented in Table 2 of Keating et al. (2021). We observe slightly higher spending in Medicare Part B and Part A and lower average spending for Part D services, which we attribute to the inclusion of both beneficiary out-of-pocket payments and non-Medicare primary payer spending in our estimates, as well as differences in the beneficiary populations and focus on a more narrow set of cancers.

Next, we provide an overview of episode level spending for all episodes in our main sample. In Figure B1 we show average spending per chemotherapy episode by Medicare Parts A, B and D from 2008 to 2020 in 1,000 USD. Over time spending per episode significantly increases, which is particularly driven by higher Part B and Part D spending. Part A spending remains almost unchanged over the entire time period.

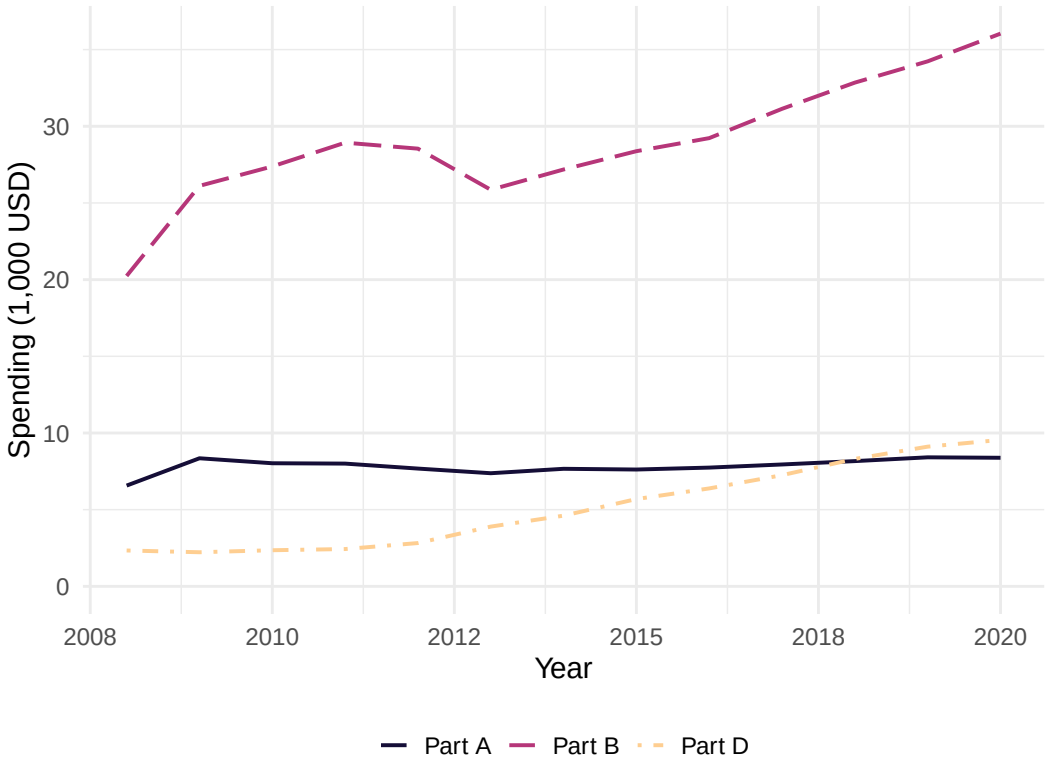


Figure B1: Spending per Episode by Medicare Part and Year

Note: The figure presents average spending per chemotherapy episode separately for Medicare Part A, B and D for the years 2008 to 2020.

Figure B2 shows the evolving share of total spending accounted for by each part of Medicare over time. In 2008—just two years after the introduction of Medicare Part D—Part D spending comprised only 8% of total episode-level spending, compared to

69% for Part B and 23% for Part A. By 2020, the share of Part D spending had increased substantially to 18%, while the share of Part B spending declined to 67%, and Part A spending fell to 16%. This shift reflects the growing importance of oral and self-administered drugs in cancer treatment and the evolving structure of Medicare-financed oncology care.

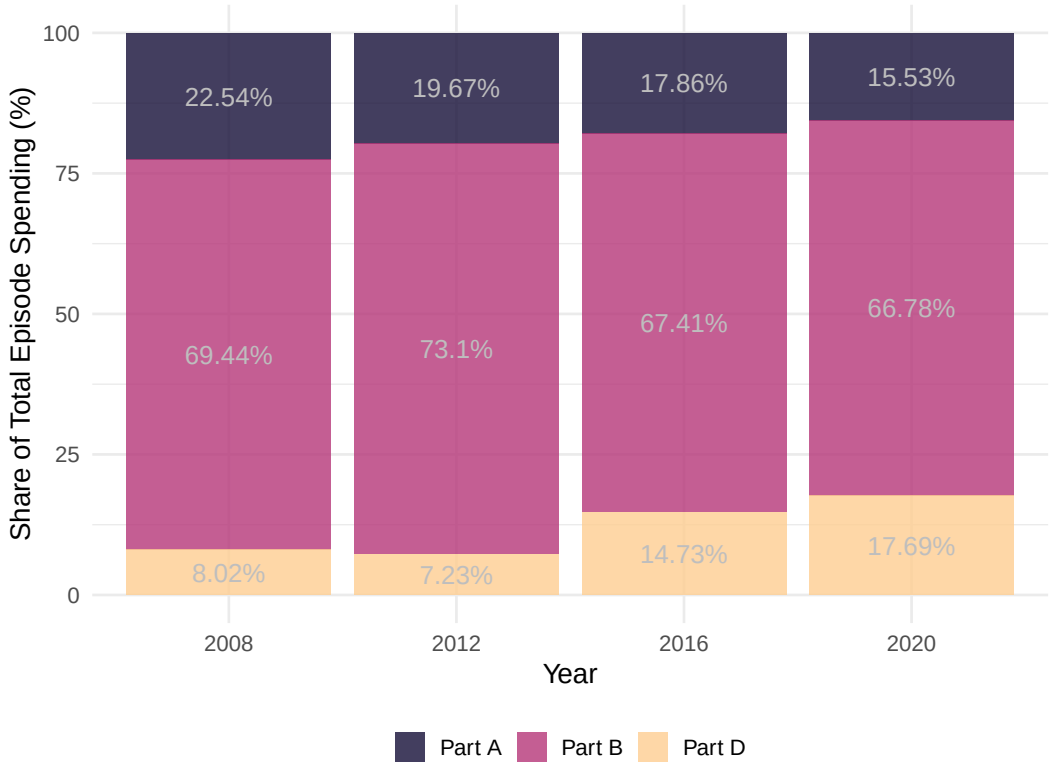


Figure B2: Share of Total Spending by Medicare Part over Time

Note: The figure presents the share of total spending by year by Medicare Part A, B and D for the years 2008, 2012, 2016 and 2020 and includes first chemotherapy episodes from our main sample.

Table B2: Variable List and Selection Criteria for Spending Definitions

Label	Variable	Payer Type	Exclude
Carrier			
Carrier Claim Payment Denial Code	CARR_CLM_PMT_DNL_CD		0 or D through Y
Line Processing Indicator Code	LINE_PRCSG_IND_CD		NOT A, R or S
Line NCH Medicare Payment Amt.	LINE_NCH_PMT_AMT	Medicare	
Line Bene. Part B Deductible Amt.	LINE_BENE_PTB_DDCTBL_AMT	Bene.	
Line Bene. Coinsurance Amt.	LINE_COINSRNC_AMT	Bene.	
Line Bene. Part B Deductible Amt.	LINE_BENE_PTB_DDCTBL_AMT	Bene.	
Line Primary Payer (if not Medicare) Paid Amt.	LINE_BENE_PRMRY_PYR_PD_AMT	Primary Payer	
Line Last Expense Date	LINE_LAST_EXPNS_DT		
Durable Medical Equipment (DME)			
Carrier Claim Payment Denial Code	CARR_CLM_PMT_DNL_CD		0 or D through Y
Line Processing Indicator Code	LINE_PRCSG_IND_CD		NOT A, R or S
Line NCH Medicare Payment Amt.	LINE_NCH_PMT_AMT	Medicare	
Line Bene. Part B Deductible Amt.	LINE_BENE_PTB_DDCTBL_AMT	Bene.	
Line Bene. Coinsurance Amt.	LINE_COINSRNC_AMT	Bene.	
Line Primary Payer (if not Medicare) Paid Amt.	LINE_BENE_PRMRY_PYR_PD_AMT	Primary Payer	
Line Last Expense Date	LINE_LAST_EXPNS_DT		
Home Health			
Claim Medicare Non-Payment Reason Code	CLM_MDCR_NON_PMT_RSN_CD		non-blank
Claim Facility Type Code	CLM_FAC_TYPE_CD		4, 5
Claim (Medicare) Payment Amt.	CLM_PMT_AMT	Medicare	
Revenue Center Non-Covered Charge Amt.	REV_CNTR_NCVRD_CHRG_AMT	Bene.	
NCH Primary Payer (if not Medicare) Claim Paid Amt.	NCH_PRMRY_PYR_CLM_PD_AMT	Primary Payer	
Claim Through Date	CLM_THRU_DT		
Hospice			
Claim Medicare Non-Payment Reason Code	CLM_MDCR_NON_PMT_RSN_CD		non-blank
Claim (Medicare) Payment Amt.	CLM_PMT_AMT	Medicare	
Revenue Center Non-Covered Charge Amt.	REV_CNTR_NCVRD_CHRG_AMT	Bene.	
NCH Primary Payer (if not Medicare) Claim Paid Amt.	NCH_PRMRY_PYR_CLM_PD_AMT	Primary Payer	
Claim Through Date	CLM_THRU_DT		
Inpatient			
Claim Medicare Non-Payment Reason Code	CLM_MDCR_NON_PMT_RSN_CD		non-blank
Claim (Medicare) Payment Amt.	CLM_PMT_AMT	Medicare	
Claim PPS Capital Disproportionate Share Amt.	CLM_PPS_CPTL_DSPRPRTNT_SHR_AMT	Medicare	
Claim PPS Capital Indirect Medical Education (IME) Amt.	CLM_PPS_CPTL_IME_AMT	Medicare	
Operating Indirect Medical Education (IME) Amt.	IME_OP_CLM_VAL_AMT	Medicare	
Operating Disproportionate Share (DSH) Amt.	DSH_OP_CLM_VAL_AMT	Medicare	
Revenue Center Non-Covered Charge Amt.	REV_CNTR_NCVRD_CHRG_AMT	Bene.	
NCH Bene. Inpatient (or other Part A) Deductible Amt.	NCH_BENE_IP_DDCTBL_AMT	Bene.	
NCH Primary Payer (if not Medicare) Claim Paid Amt.	NCH_PRMRY_PYR_CLM_PD_AMT	Primary Payer	
Claim Through Date	CLM_THRU_DT		
Outpatient			
Claim Medicare Non-Payment Reason Code	CLM_MDCR_NON_PMT_RSN_CD		non-blank
Claim Facility Type Code	CLM_FAC_TYPE_CD		4, 5
Claim (Medicare) Payment Amt.	CLM_PMT_AMT	Medicare	
NCH Bene. Part B Deductible Amt.	NCH_BENE_PTB_DDCTBL_AMT	Bene.	
NCH Bene. Part B Coinsurance Amt.	NCH_BENE_PTB_COINSRNC_AMT	Bene.	
Revenue Center Non-Covered Charge Amt.	REV_CNTR_NCVRD_CHRG_AMT	Bene.	
NCH Primary Payer (if not Medicare) Claim Paid Amt.	NCH_PRMRY_PYR_CLM_PD_AMT	Primary Payer	
Claim Through Date	CLM_THRU_DT		
Skilled Nursing Facility (SNF)			
Claim Medicare Non-Payment Reason Code	CLM_MDCR_NON_PMT_RSN_CD		non-blank
Claim (Medicare) Payment Amt.	CLM_PMT_AMT	Medicare	
Revenue Center Non-Covered Charge Amt.	REV_CNTR_NCVRD_CHRG_AMT	Bene.	
NCH Primary Payer (if not Medicare) Claim Paid Amt.	NCH_PRMRY_PYR_CLM_PD_AMT	Primary Payer	
Claim Through Date	CLM_THRU_DT		
Part D			
Amt. paid for by Part D low income subsidy	LICS_AMT	Medicare	
Amt. Paid by Patient	PTNT_PAY_AMT	Bene.	
Other True Out-of-Pocket (TrOOP) Amt.	OTHR_TROOP_AMT	Primary Payer	
Reduction in patient liability (PLRO)	PLRO_AMT	Primary Payer	
Amt. paid by Part D plan for the PDE	CVRD_D_PLAN_PD_AMT	Primary Payer	
RX Service Date	SRVC_DT		

Notes: The table provides an overview of the variables and criteria used to construct spending measures for our main analysis from individual claims. The first column describes the variable, the second column provides the variable label as outlined in the Chronic Condition Warehouse (CCW) codebooks for fee-for-service claims and Part D events. The third column indicates which payer is attributed the respective payment variable and column four indicates exclusion criteria for claim and lines for relevant variables.

C 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 E2.

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.

The final share of classified trials which have a corresponding clinical trial number both among beneficiaries in our chemotherapy sample and in clinical trials is presented

in Figure C1. Of the 11,049 classified trials relevant for our sample 24 percent are related to leukemia and lymphoma, 15 percent are not related to cancer, 12 percent are related to other cancers (cancers not covered by our broader categories).

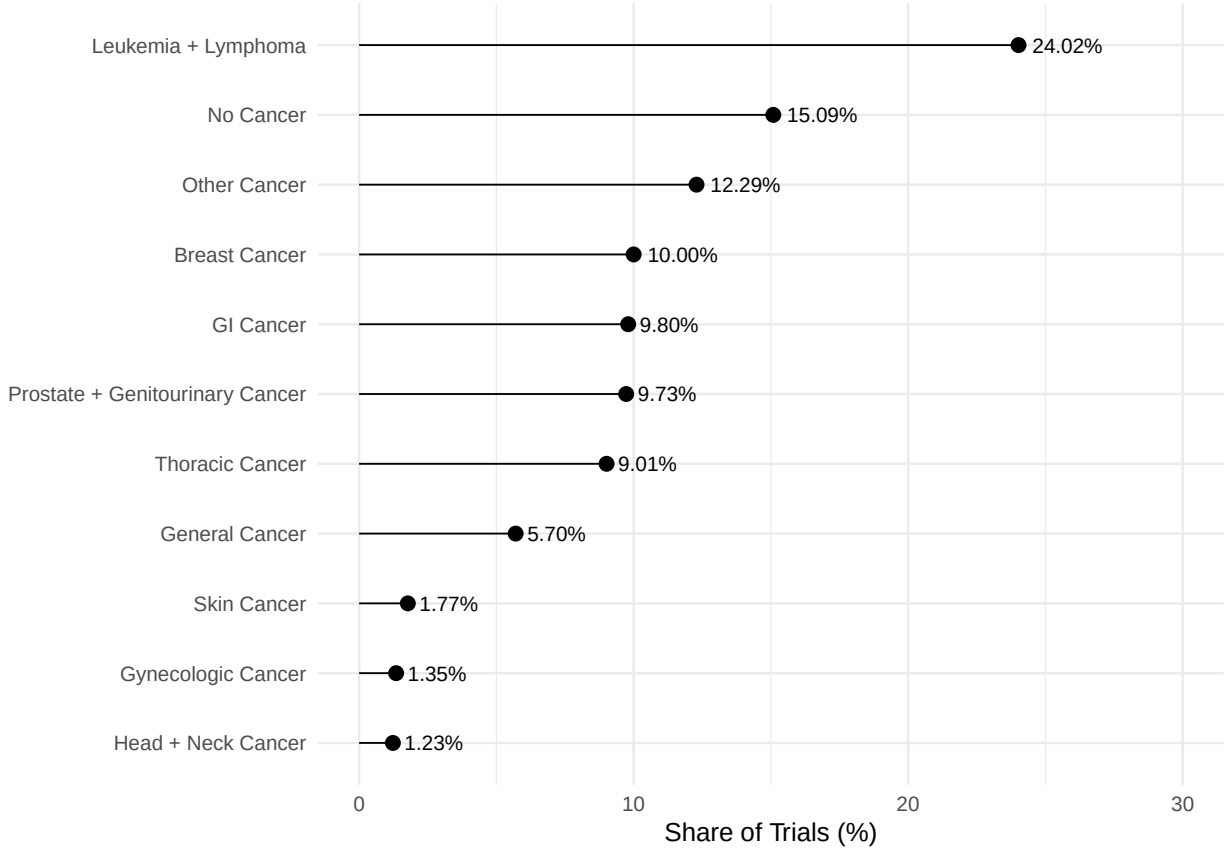


Figure C1: Share of Clinical Trials by Cancer Classification

Note: This figure presents the share of cancer trials within different categories as classified using GTP-4 in combination with data from clinicaltrials.gov. In total there are 11,049 classified trials, which are both linked to a beneficiary in our sample during the year their chemotherapy was initiated and where information is available via clinicaltrials.gov

D Additional Figures

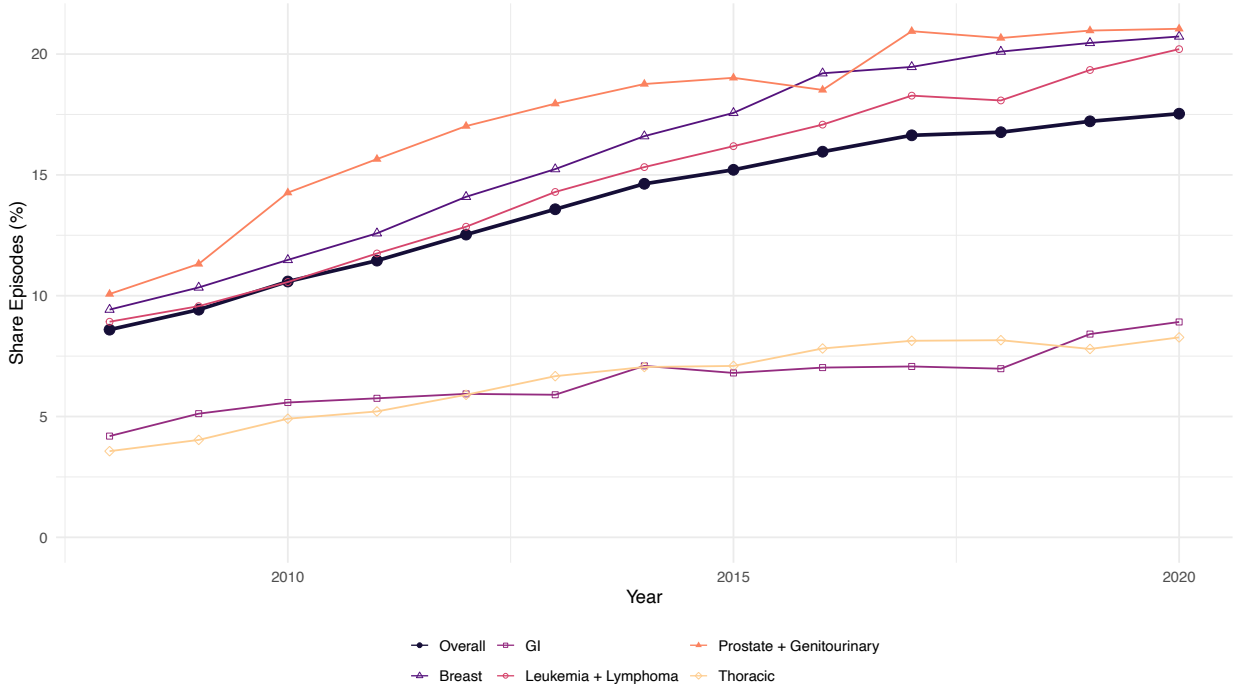


Figure D1: Trends in Subspecialization in Cancer Care 2008 - 2020

Note: This figure presents trends in subspecialization for beneficiary chemotherapy episodes of different cancer types based on results from Karadakic et al. (2025). 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.

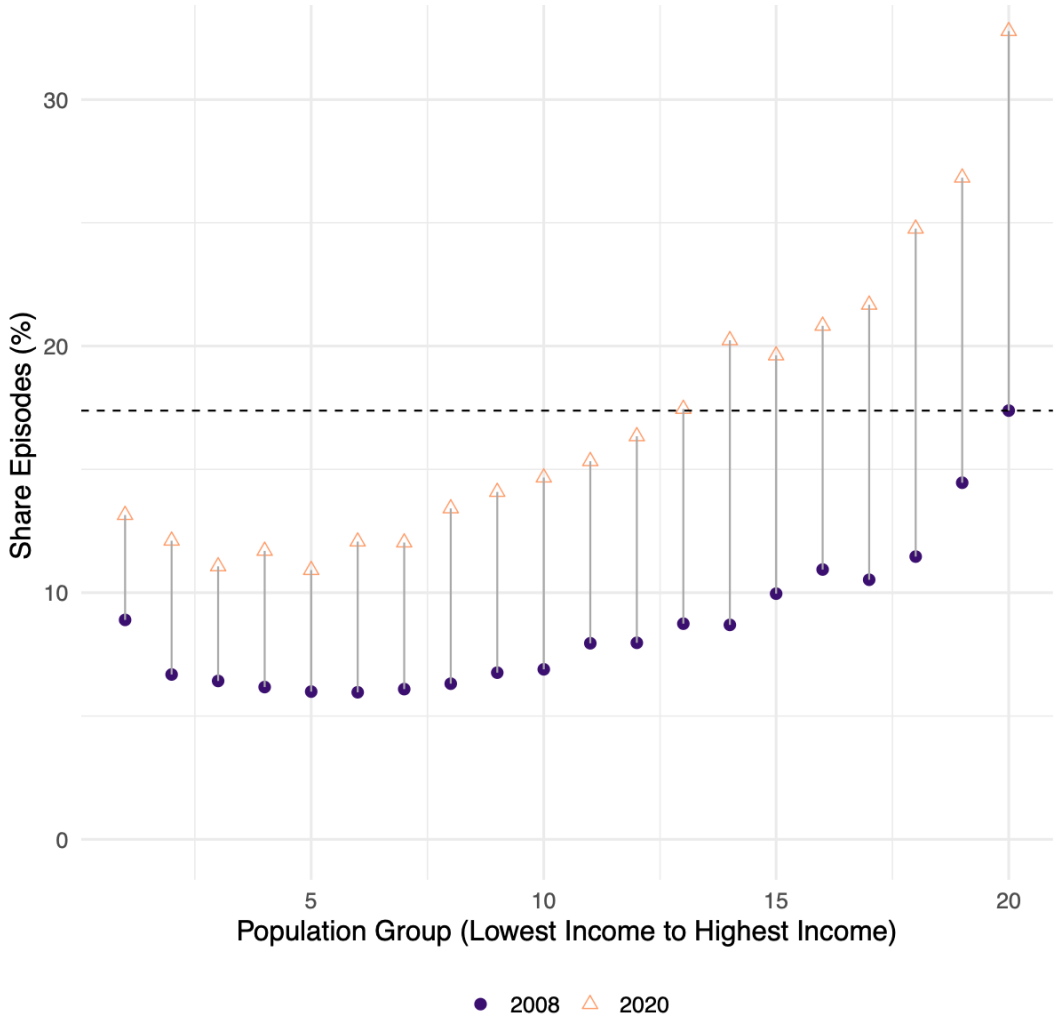
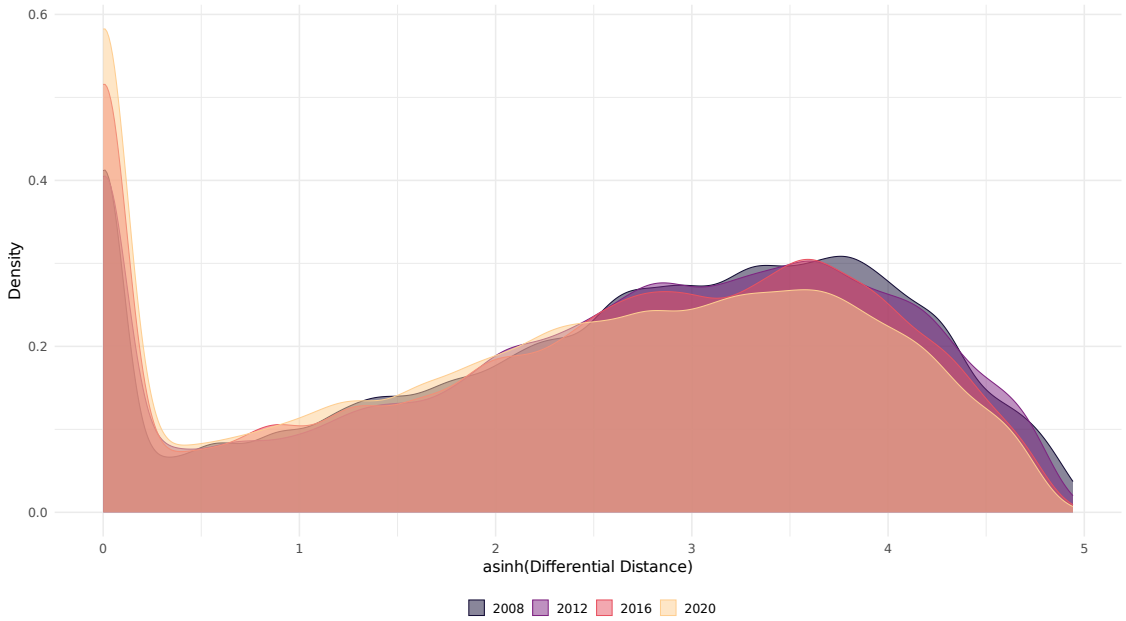
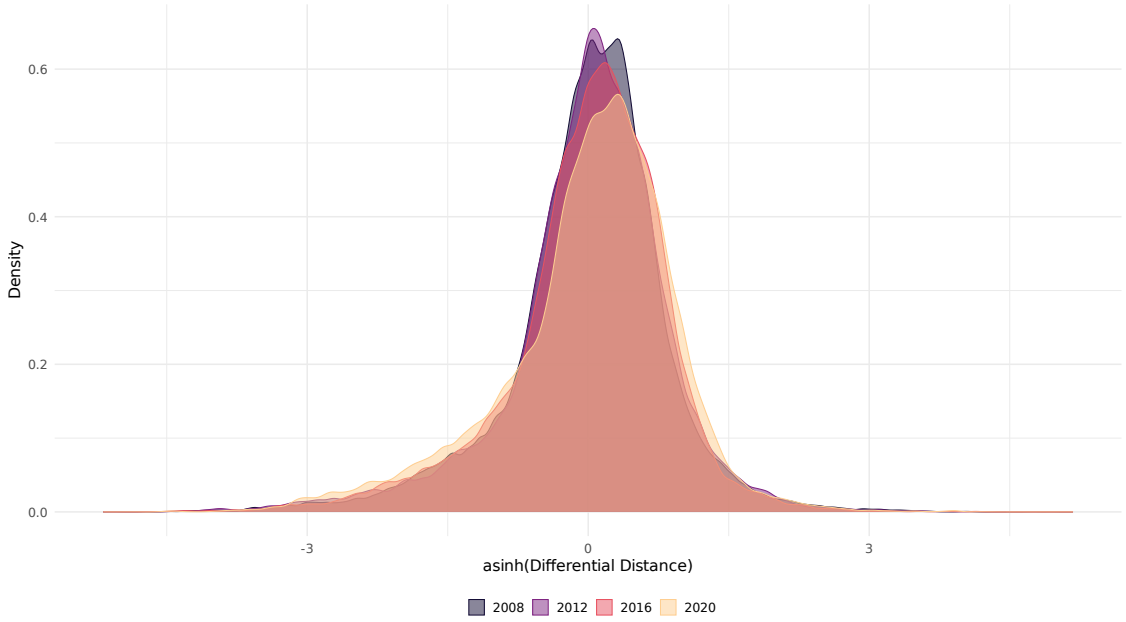


Figure D2: Socioeconomic Gradients in Access to Highly Subspecialized Cancer Care

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).



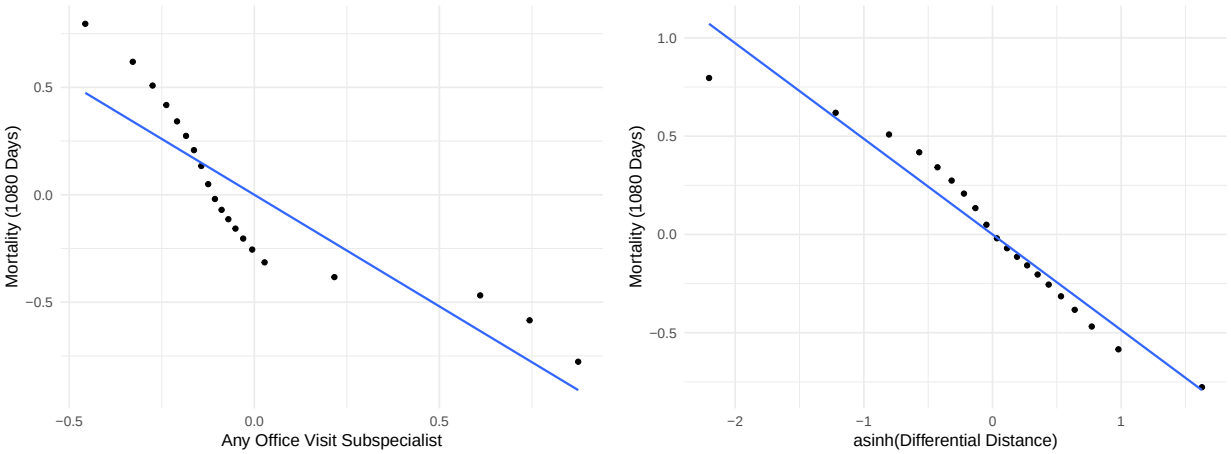
(a) asinh(Differential Distance)



(b) Residualized asinh(Differential Distance)

Figure D3: Distribution of Instrumental Variable

Note: This figure displays the distribution of the instrumental variable—the inverse hyperbolic sine (IHS) of differential distance—for the main sample of first chemotherapy episodes. All values are standardized to have a mean of 0 and a standard deviation of 1. Panel A shows kernel density plots of the raw IHS-transformed differential distance for the years 2008, 2012, 2016, and 2020. Panel B shows the corresponding residualized values, obtained by regressing the instrument on cancer type-by-year and ZCTA fixed effects.



(a) Structural Form (b) Reduced Form

Figure D4: Structural and Reduced Form Relationship (1,080-day Mortality)

Note: This figure presents binned scatter plots illustrating the structural and reduced-form relationships for 1,080-day mortality. Panel A plots the relationship between the residualized endogenous variable—subspecialist access—and the residualized mortality outcome. Panel B plots the relationship between the residualized instrument—differential distance—and the same residualized outcome. All variables have been residualized by regressing them on the full set of covariates and fixed effects included in the second-stage regression (Equation 2).

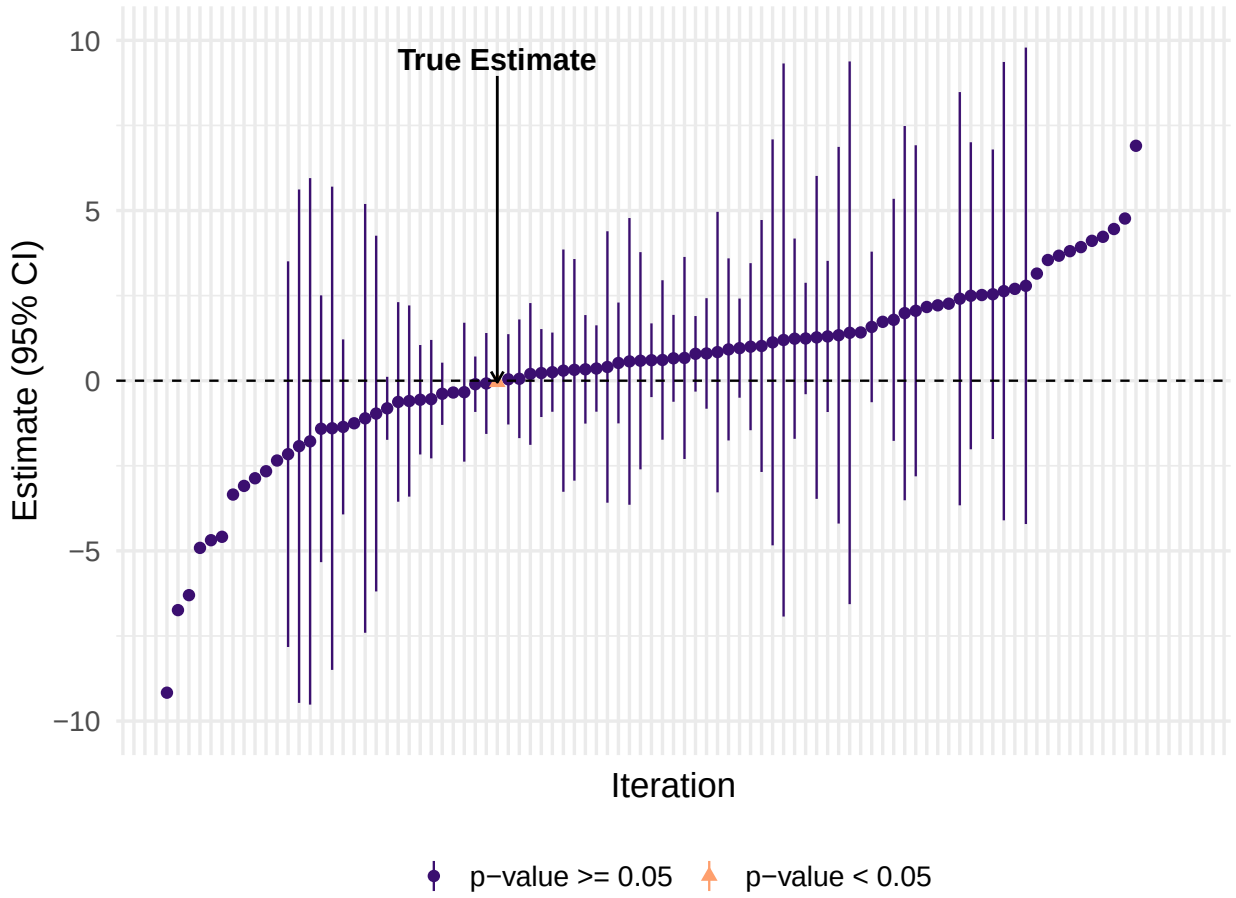


Figure D5: Falsification Test - Randomly Reassign Differential Distances

Note: This figure presents 2SLS estimates of the effect of access to a subspecialized oncologist of the relevant cancer type on 1,080 day mortality. Each dot represents the result from a separate regression using our main specification, where differential distances were randomly reassigned to chemotherapy episodes within year. This randomization was repeated 100 times. For reference, the figure also includes our main (non-randomized) estimate. All models control for beneficiary demographics, ZCTA-level characteristics, and chronic conditions, and include ZCTA fixed effects and cancer type-by-year fixed effects. Standard errors are clustered at the ZCTA level.

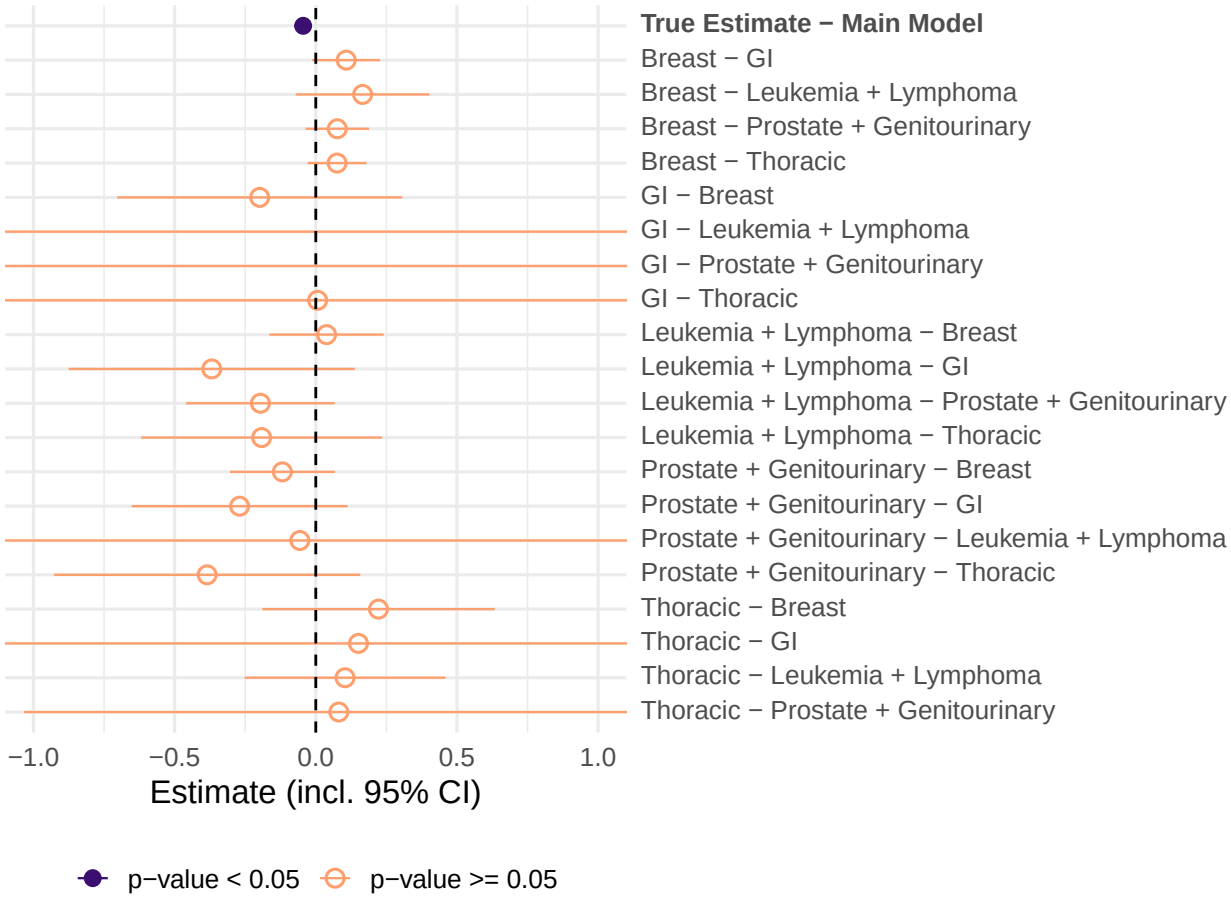
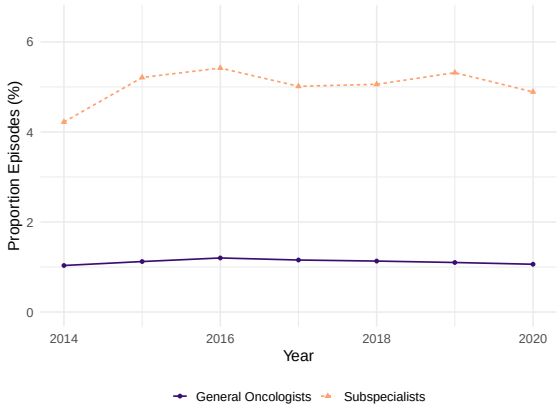
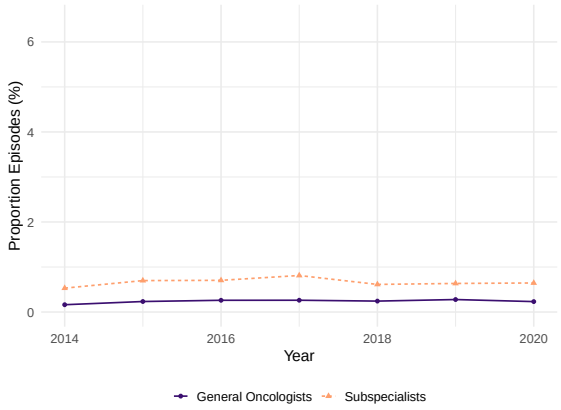


Figure D6: Falsification Test - Assignment of Distance to Unrelated Subspecialist

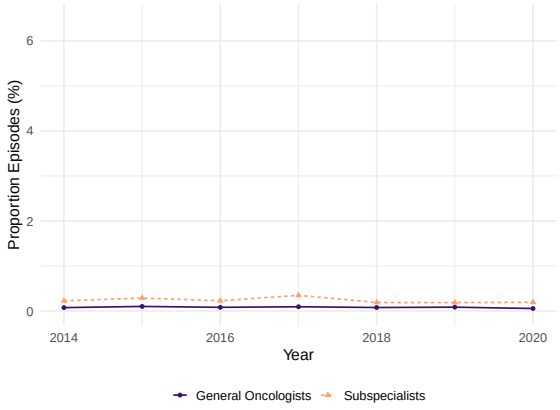
Note: This figure presents falsification tests in which beneficiaries are assigned distances to subspecialists for unrelated cancer types (e.g., the distance to the nearest breast cancer subspecialist for a beneficiary with thoracic cancer). Each dot is estimated from a separate regression. Each row on the y-axis represents a specific reassignment of distances; for example, "Breast – GI" means all beneficiaries were assigned the distance to the nearest breast cancer subspecialist, except those with breast cancer, who were assigned the distance to the nearest GI subspecialist. The top row shows the main estimates using the true distance to the nearest subspecialist of the relevant cancer type. Solid purple circles indicate statistically significant estimates ($p < 0.05$), while hollow orange circles indicate statistically insignificant results ($p \geq 0.05$). The analysis sample includes first chemotherapy episodes from 2008 to 2017. All models control for demographics, ZCTA-level characteristics, and chronic conditions, and include ZCTA fixed effects and cancer type-by-year fixed effects. Standard errors are clustered at the ZCTA level.



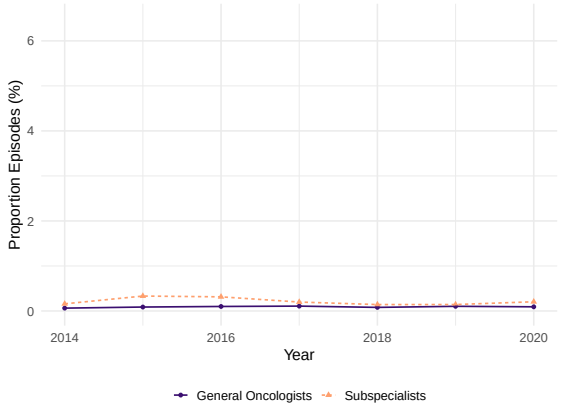
(a) Concordant Cancer Trial



(b) Discordant Cancer Trial



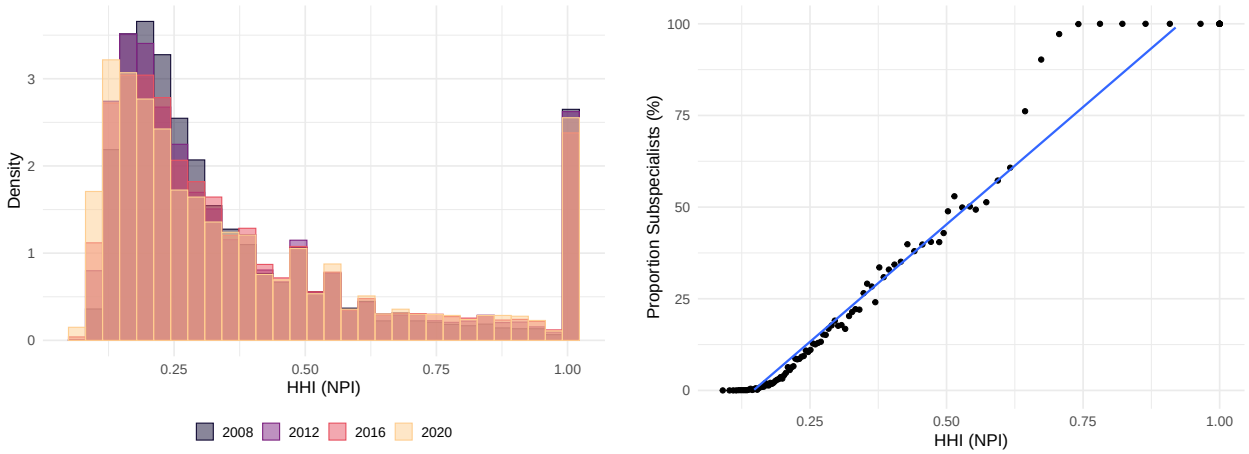
(c) Unspecified Cancer Trial



(d) Non-Cancer Trial

Figure D7: Proportion of Episodes with Cancer Trial Enrollment by Trial Type and Oncologist Subspecialization

Note: The figure shows the share of chemotherapy episodes from 2014 to 2020 in which the beneficiary had any claim containing an NCT number in the same year as chemotherapy initiation. The solid orange line represents episodes coordinated by subspecialized oncologists, while the dashed purple line represents episodes managed by general oncologists. Panel A reports the share of claims linked to cancer trials concordant with the beneficiary's cancer type; Panel B shows discordant cancer trials; Panel C depicts trials for unspecified cancers; and Panel D presents the share associated with non-cancer trials.



(a) Distribution of HHI

(b) HHI and Subspecialist Share

Figure D8: Descriptive Measures of Herfindahl–Hirschman Index

Note: This figure presents descriptive statistics of the cancer-type Herfindahl-Hirschman Index (HHI), a measure of oncologist specialization within a specific cancer type. Panel A plots the distribution of HHI values for all providers in our first chemotherapy sample for the years 2008, 2012, 2016, and 2020. Panel B displays the relationship between HHI and the share of subspecialized oncologists by plotting the average HHI and corresponding average subspecialist share across percentiles of the HHI distribution, constructed separately for each year.

E Additional Tables

Table E1: 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

(Table E1 continued)

TRIPTORELIN	J3315
TRIPTORELIN	J3316
ANTI-THYMOCYTE GLOBULIN, EQUINE	J7504
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

(Table E1 continued)

CYCLOPHOSPHAMIDE	J9094
CYCLOPHOSPHAMIDE	J9095
CYCLOPHOSPHAMIDE	J9096
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

(Table E1 continued)

PEGASPARGASE	J9266
PACLITAXEL	J9267
PENTOSTATIN	J9268
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

(Table E1 continued)

DOXORUBICIN, LIPOSOMAL	Q2049
DOXORUBICIN, LIPOSOMAL	Q2050
BEVACIZUMAB-AWWB	Q5107
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 E3: Summary Statistics of Main Sample

Variable	Mean	SD	Min.	Max.
Panel A: Cancer Characteristics				
Breast Cancer	0.355	0.479	0	1
GI Cancer	0.158	0.365	0	1
Hematologic Cancer	0.222	0.416	0	1
Prostate + Genitourinary Cancer	0.127	0.333	0	1
Thoracic Cancer	0.163	0.369	0	1
Panel B: Demographics				
Bene Age	75.175	6.626	67	114
Female Bene	0.585	0.493	0	1
Black Bene	0.080	0.271	0	1
Hispanic Bene	0.011	0.104	0	1
Asian Bene	0.014	0.119	0	1
Other Non-White Race Bene	0.015	0.121	0	1
Panel C: Chronic Conditions				
Acute Myocardial Infarction	0.014	0.118	0	1
Alzheimer's Disease	0.023	0.150	0	1
Alzheimer's or Dementia	0.082	0.274	0	1
Anemia	0.559	0.496	0	1
Asthma	0.069	0.253	0	1
Atrial Fibrillation	0.129	0.335	0	1
Cataracts	0.206	0.405	0	1
Chronic Kidney Disease	0.314	0.464	0	1
COPD	0.231	0.422	0	1
Colorectal Cancer	0.102	0.303	0	1
Congestive Heart Failure	0.209	0.407	0	1
Depression	0.196	0.397	0	1
Diabetes	0.315	0.465	0	1
Endometrial Cancer	0.008	0.087	0	1
Glaucoma	0.104	0.305	0	1
Hyperlipidemia	0.570	0.495	0	1
Hyperplasia	0.112	0.315	0	1
Hypertension	0.726	0.446	0	1
Hypothyroidism	0.188	0.391	0	1
Ischemic Heart Disease	0.395	0.489	0	1
Lung Cancer	0.188	0.391	0	1
Osteoporosis	0.107	0.309	0	1
Prostate Cancer	0.135	0.342	0	1
Rheum. Arthritis / Osteoarthritis	0.372	0.483	0	1
Stroke (TIA)	0.047	0.213	0	1
Panel D: ZCTA Characteristics				
Distance to Closest Oncologist	3.703	3.756	0	17.287
Distance to Closest PCP	0.492	1.307	0	16.761
Mean Age	71.643	2.160	46.987	91.902
Median Household Income	64443.452	27127.046	2499	250001
Nr. of Mental Health Providers	9.985	20.118	0	338
Nr. of PCPs	31.271	61.304	0	2147
Share Asian	0.019	0.044	0	1
Share Black	0.098	0.168	0	1
Share Disabled	0.146	0.071	0	0.885
Share FFS	0.713	0.103	0.024	1
Share Full Dual	0.103	0.080	0	0.943
Share Hispanic	0.018	0.043	0	0.680
Share Male	0.451	0.029	0	1
Share Other Race	0.019	0.029	0	1
Total Beneficiaries	4734.646	3383.978	1	43028
Total Providers	103.964	189.181	0	7211

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

Table E4: Explained Instrument Variance under Alternative Fixed Effect Combinations

Fixed effects	Adjusted R^2
ZCTA	0.527
Cancer Type	0.102
Year	0.009
ZCTA + Year	0.535
ZCTA + Cancer Type	0.625
Cancer Type + Year	0.110
ZCTA + Cancer Type + Year	0.633
ZCTA + Cancer Type $\times$ Year	0.634
Baseline (full specification)	0.642

Notes: Entries report the Adjusted R^2 from regressions of the differential distance instrument on alternative sets of fixed effects. Higher Adusted R^2 values indicate that the corresponding set of fixed effects explains a larger share of the variation in the instrument.

Table E5: Drivers of Variation in Instrumental Variable

	Baseline	**Fixed Stock**	**Fixed ZCTA**	**Fixed Specialization**
Correlation with Baseline Z	1.000	0.737	0.711	0.645
First Stage β	-0.024^{***}	-0.023^{***}	-0.022^{***}	-0.010^{***}
First Stage F	1,823	3,036	2,613	684
Simple Adj. R^2	—	0.528	0.501	0.416
Full Adj. R^2	—	0.746	0.730	0.682
Within Adj. R^2	—	0.307	0.263	0.133

Notes: The table reports how much of the variation in the baseline differential–distance instrument (Z) is reproduced when holding different components of the oncology workforce fixed. Column 1 (Baseline) shows results for the original instrument. Column 2 (Fixed Stock) holds the set of oncologists constant at the 2008 workforce. Column 3 (Fixed ZCTA) holds each oncologist's practice ZIP code (ZCTA) fixed at the first year they appear in our chemotherapy episode data. Column 4 (Fixed Specialization) holds only the oncologist's subspecialization status fixed at first appearance. The correlation column reports pairwise correlations with the baseline instrument. The first-stage coefficient (β) and Kleibergen–Paap F-statistic are taken from regressions of access to a subspecialist on each instrument using the main specification. The "Simple Adj. R^2" is the adjusted R^2 from a univariate regression of the baseline instrument on each alternative instrument. The "Full Adj. R^2" is from the full first-stage model including demographic controls, chronic conditions, ZCTA controls, ZCTA fixed effects, and cancer-type-by-year fixed effects. The "Within Adj. R^2" reports the adjusted R^2 after absorbing fixed effects. Significance levels: *** p<0.01, ** p<0.05, * p<0.1.

Table E6: First Stage Estimates

	Any Office Visit Subspecialist	**Treatment Subspecialist**
$\sinh^{-1}(\text{DD})$	-0.024***	-0.018***
	(0.001)	(0.001)
ZCTA FE	Yes	Yes
Cancer-Year FE	Yes	Yes
Adj. R^2	0.174	0.161
Observations	2,165,024	2,165,024
Mean Dep. Var.	0.186	0.141
F-Stat (1st stage)	1,823	1,176

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 for our main sample of first chemotherapy episodes. 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. Reported first-stage F-statistics are Kleibergen-Paap statistics. Signif. Codes: ***: 0.01, **: 0.05, *: 0.1.

Table E7: Mortality Effects of Access to Subspecialized Oncologist

	180-Day	**360-Day**	**720-Day**	**1080-Day**
Panel A: First Stage				
$\sinh^{-1}(\text{DD})$	-0.024***	-0.024***	-0.024***	-0.024***
	(0.001)	(0.001)	(0.001)	(0.001)
Panel B: Reduced Form				
$\sinh^{-1}(\text{DD})$	-0.000	0.001***	0.001***	0.001***
	(0.000)	(0.000)	(0.000)	(0.000)
Panel C: Structural Form				
Any Office Visit Subs.	-0.014***	-0.005***	0.003***	0.005***
	(0.001)	(0.001)	(0.001)	(0.001)
Panel D: 2SLS				
Any Office Visit Subs.	-0.007	-0.040***	-0.050***	-0.045**
	(0.011)	(0.014)	(0.016)	(0.016)
Adj R^2	0.120	0.224	0.302	0.325
Observations	1,681,119	1,681,119	1,681,119	1,681,119
Mean Dep. Var.	0.108	0.226	0.365	0.449
F-Stat (1st Stage)	1,649	1,649	1,649	1,649

Notes: This table reports estimates of the effect of access to subspecialized oncologists on mortality for all first chemotherapy episodes between 2008 and 2017. Panel A shows first stage estimates. Panel B shows the reduced form estimates. Panel C shows the structural form estimates, and Panel D provides the 2SLS estimates. All models control for demographics, ZCTA characteristics, comorbidities, and include ZCTA and cancer type-by-year fixed effects. Standard errors are clustered at the ZCTA level. First-stage strength is reported using the Kleibergen-Paap F-statistic. Significance levels: *** p<0.01, ** p<0.05, * p<0.1.

Table E8: Robustness of Mortality Outcomes by Samples and Specifications

	180-Day	360-Day	720-Day	1080-Day
Panel A: Fixed Sample (2008-2017)				
Any Office Visit Subs.	-0.007	-0.040***	-0.050***	-0.045**
	(0.011)	(0.014)	(0.016)	(0.016)
Adj. R^2	0.134	0.224	0.302	0.335
Observations	1,681,119	1,681,119	1,681,119	1,681,119
Mean Dep. Var.	0.108	0.226	0.365	0.449
F-Stat (1st Stage)	1,649	1,649	1,649	1,649
Panel B: Varying Sample				
Any Office Visit Subs.	-0.001	-0.024*	-0.041***	-0.045***
	(0.010)	(0.031)	(0.015)	(0.016)
Adj. R^2	0.119	0.223	0.302	0.325
Observations	2,014,625	2,014,625	1,845,326	1,681,119
Mean Dep. Var.	0.105	0.220	0.361	0.449
F-Stat (1st Stage)	1,767	1,767	1,731	1,649
Panel C: Volume Controls				
Any Office Visit Subs.	-0.005	-0.046**	-0.058***	-0.050***
	(0.014)	(0.018)	(0.020)	(0.020)
Adj. R^2	0.119	0.225	0.302	0.325
Observations	1,681,119	1,681,119	1,681,119	1,681,119
Mean Dep. Var.	0.108	0.226	0.365	0.449
F-Stat (1st Stage)	1,540	1,540	1,540	1,540
Panel D: ZCTA Slopes				
Any Office Visit Subs.	0.003	-0.032**	-0.048***	-0.039***
	(0.012)	(0.015)	(0.016)	(0.016)
Adj. R^2	0.120	0.224	0.302	0.325
Observations	1,681,119	1,681,119	1,681,119	1,681,119
Mean Dep. Var.	0.108	0.226	0.365	0.449
F-Stat (1st Stage)	1,605	1,605	1,605	1,605

Notes: This table reports 2SLS estimates of the effect of access to subspecialized oncologists on mortality. Panel A shows main results for first chemotherapy episodes from 2008–2017. Panel B varies the sample based on available follow-up for each mortality horizon. Panel C adds controls for the treating oncologist's episode volume by cancer group, and Panel D includes ZCTA-specific slopes. All models control for demographics, ZCTA characteristics, comorbidities, and include ZCTA and cancer type-by-year fixed effects. Standard errors are clustered at the ZCTA level. First-stage strength is reported using the Kleibergen-Paap F-statistic. Significance levels: *** p<0.01, ** p<0.05, * p<0.1.

Table E2: ICD 9 and ICD 10 Codes for Cancer Type Classification

Cancer Type Label	**Cancer Types Included**	**ICD-9/ICD-10 Codes**
Breast	Breast Cancer; Carcinoma in situ of breast	174.xx, 175.xx, 233.0x, C50.xx, D05.xx
GI	Anal Cancer; Carcinoma in situ of oral cavity, esophagus and stomach; Carcinoma in situ of other and unspecified digestive organs; Gastro/Esophageal Cancer; Liver Cancer; Malignant neoplasm of abdomen; Malignant neoplasm of other and ill-defined digestive organs; Pancreatic Cancer; Small Intestine / Colorectal Cancer	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	Carcinoma in situ of cervix uteri; Female GU Cancer other than Ovary; Malignant neoplasm of other and unspecified female genital organs; Malignant neoplasm of placenta; Ovarian Cancer	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	Carcinoma in situ of middle ear and respiratory system; Head and Neck Cancer	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; Acute panmyelosis with myelofibrosis; Atypical chronic myeloid leukemia, BCR/ABL negative; Chronic Leukemia; Chronic leukemia of unspecified cell type; Chronic myelomonocytic leukemia; Chronic myeloproliferative disease; Essential (hemorrhagic) thrombocythemia; Juvenile myelomonocytic leukemia; Leukemia, unspecified; Lymphoid Leukemia, unspecified; Lymphoma; MDS; Monocytic Leukemia, unspecified; Multiple Myeloma; Myelofibrosis; Myeloid leukemia, unspecified; Osteomyelofibrosis; Other and unspecified malignant neoplasms of lymphoid, hematopoietic and related tissue; Other lymphoid leukemia; Other monocytic leukemia; Other myeloid leukemia; Other specified leukemias; Polycythemia vera; Secondary and unspecified malignant neoplasm of lymph nodes	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	Bladder Cancer; Carcinoma in situ of other and unspecified genital organs; Kidney Cancer; Malignant neoplasm of penis, other, and unspecific male organs; Malignant neoplasm of testis; Prostate Cancer	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	Carcinoma in situ of skin; Malignant Melanoma; Melanoma in situ; Merkel cell carcinoma; Other and unspecified malignant neoplasm of skin	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 Cancer; Malignant neoplasm of heart, mediastinum and pleura; Malignant neoplasm of thorax; Malignant neoplasm of 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	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 used for subspecialist classification, column 2 provides the more detailed cancer types which are used for statistical modeling purposes (e.g. inclusion of cancer type fixed effects) and column 3 the corresponding ICD-9 and ICD-10 codes used for identification of cancers in the Medicare claims.

Table E9: Access to Subspecialized Oncologist and Spending by Subcategory

	Part A				Part B			
	Inpatient	**HHA**	**Hospice**	**SNF**	**Carrier**	**Outpatient**	**DME**	**Part D**
Panel A: First Stage								
$\sinh^{-1}$(DD)	-0.024***	-0.024***	-0.024***	-0.024***	-0.024***	-0.024***	-0.024***	-0.024***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Panel B: Reduced Form								
$\sinh^{-1}$(DD)	-16.66	-0.93	6.44*	-8.57***	123.42***	-105.64***	-0.80	11.21
	(10.52)	(1.50)	(3.68)	(2.87)	(20.57)	(21.19)	(1.35)	(11.20)
Panel C: Structural Form								
Any Office Visit Subs.	1,932.95***	-29.99***	-69.27***	-191.17***	-4,207.06***	6,900.43***	33.21***	444.74***
	(33.258)	(3.62)	(5.58)	(6.86)	(57.33)	(74.64)	(3.53)	(32.71)
Panel D: 2SLS								
Any Office Visit Subs.	704.42	39.30	-272.21*	362.27***	-5,217.32***	4,465.74***	33.92	-474.06
	(444.37)	(63.55)	(155.94)	(121.50)	(874.14)	(899.49)	(57.14)	(474.08)
Adj. R^2	0.146	0.136	0.024	0.082	0.197	0.114	0.068	0.242
Observations	2,165,024	2,165,024	2,165,024	2,165,024	2,165,024	2,165,024	2,165,024	2,165,024
Mean Dep. Var.	5,699.26	778.14	527.70	764.10	15,370.09	12,670.91	420.14	5,048.72
F-Stat (1st Stage)	1,823	1,823	1,823	1,823	1,823	1,823	1,823	1,823

Notes: The table provides estimates on the effect of access to subspecialized oncologists on different measures of spending for chemotherapy episodes in our main sample. Column 1 provides estimates for total spending, column 2 Part A spending, column 3 Part B spending and column 4 Part D spending. Panel A presents first stage estimates, Panel B the respective reduced form estimates, Panel C the structural form estimates and Panel D provides 2SLS 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. First-stage strength is reported using the Kleibergen-Paap F-statistic. Standard errors are clustered at the ZCTA level. Signif. Codes: ***: 0.01, **: 0.05, *: 0.1.

Table E10: Access to Subspecialized Oncologist and Spending on Chemotherapy Drugs and Services

	Chemotherapy Drugs	**Injections & Infusions (RI)**
Panel A: First Stage		
$\sinh^{-1}(\text{DD})$	-0.024***	-0.024***
	(0.001)	(0.001)
Panel B: Reduced Form		
$\sinh^{-1}(\text{DD})$	60.74***	15.51***
	(13.88)	(5.835)
Panel C: Structural Form		
Any Office Visit Subs.	-1,023.72***	-458.82***
	(37.01)	(14.24)
Panel D: 2SLS		
Any Office Visit Subs.	-2,567.73***	-655.78***
	(587.17)	(246.70)
Adj R^2	0.167	0.104
Observations	2,165,024	2,165,024
Mean Dep. Var.	9,411.18	3,332.22
F-Stat (1st Stage)	1,822	1,822

Notes: This table reports estimates of the effect of access to subspecialized oncologists on different categories of spending for the full sample of first chemotherapy episodes. Column 1 focuses on Part B spending for chemotherapy-related HCPCS codes, defined according to the Oncology Care Model. Column 2 presents Part B spending estimates for HCPCS codes related to the infusion and injection of drugs, based on the 2024 Restructured BETOS Classification System (RBCS). Panel A shows first-stage results; Panel B reports reduced-form estimates; Panel C provides structural-form estimates; and Panel D presents the 2SLS results. All specifications control for beneficiary demographics, ZIP code–level characteristics, and comorbidities, and include fixed effects for ZIP code and cancer type-by-year. Standard errors are clustered at the ZIP code (ZCTA) level. First-stage strength is summarized using the Kleibergen-Paap F-statistic. Significance levels: *** p<0.01, ** p<0.05, * p<0.1.

Table E11: Balancing Test of Instrumental Variable

Variable	Mean	SD	Est. Unadj.	Std. Err. Unadj.	Est. Adj.	Std. Err. Adj.
Panel A: Chronic Conditions Indicators						
Alzheimer	0.023	0.150	-0.001	0	0	0
Alzheimer with Dementia	0.082	0.274	-0.002	0	0	0
AMI	0.014	0.118	0.001	0	0	0
Anemia	0.559	0.496	0.027	0	0.001	0
Asthma	0.069	0.253	-0.002	0	0	0
Atrial Fibrillation	0.129	0.335	0.006	0	0	0
Breast Cancer	0.355	0.479	-0.088	0	0	0
Colorectal Cancer	0.102	0.303	0.020	0	0	0
Endometrial Cancer	0.008	0.087	-0.001	0	0	0
Lung Cancer	0.188	0.391	0.046	0	0	0
Prostate Cancer	0.135	0.342	0.019	0	0	0
Cataract	0.206	0.405	-0.006	0	0	0
Congestive Heart Failure	0.209	0.407	0.007	0	-0.001	0
Chronic Kidney Disease	0.314	0.464	0.013	0	0.001	0
Chronic Obstructive Pulmonary Disease	0.231	0.422	0.034	0	0	0
Depression	0.196	0.397	-0.001	0	0	0
Diabetes	0.315	0.465	0.004	0	0	0
Glaucoma	0.104	0.305	-0.007	0	0	0
Hip Fracture	0.011	0.107	0	0	0	0
Hyperlipidemia	0.570	0.495	0	0	0	0
Hyperplasia	0.112	0.315	0.011	0	0	0
Hypertension	0.726	0.446	0.008	0	0	0
Hypothyroidism	0.188	0.391	-0.006	0	0	0
Ischemic Heart Disease	0.395	0.489	0.020	0	0	0
Osteoporosis	0.107	0.309	-0.012	0	0	0
Rheumatoid Arthritis	0.372	0.483	-0.007	0	0	0
Stroke	0.047	0.213	0.001	0	0	0
Panel B: Prior Healthcare Use and Diagnosis						
Any Cervical Screening	0.061	0.239	-0.009	0	-0.001	0
Any Colorectal Screening	0.021	0.142	0	0	0	0
Any ER Visit	0.315	0.464	0.011	0	0	0
Any Hospital Visit	0.235	0.424	0.013	0	0	0
Any Lung Screening	0.003	0.055	0	0	0	0
Any Lymph Node Involvement (t=-1)	0.056	0.229	0.001	0	0	0
Any Lymph Node Involvement	0.187	0.390	0.010	0	0	0
Any Mammogram	0.182	0.386	-0.020	0	0	0
Any Metastatic Cancer Diagnosis (t=-1)	0.118	0.323	0.009	0	0.001	0
Any Metastatic Cancer Diagnosis	0.356	0.479	0.040	0	0.001	0
Any Primary Care Visit	0.810	0.393	0.002	0	0	0
Any Prostate Screening	0.078	0.268	0.016	0	0	0
Panel C: Other Measures						
Any LIS	0.140	0.347	-0.006	0	0	0
Disabled	0.000	0.002	0	0	0	0
Full Dual	0.073	0.260	-0.009	0	0	0
Number of Chronic Conditions	5.095	2.930	0.121	0.001	-0.001	0.002
Predicted Mortality	0.098	0.119	0.014	0	0	0

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 E12: Complier Characteristics Overview

Variable	Share	Share Among Compliers	Share Among Treated	N (Conditional)	N (Unconditional)
Panel A: Chemotherapy Episode Cancer Type					
Breast Cancer	0.330	0.241	0.347	713,977	2,166,050
GI Cancer	0.158	0.104	0.103	342,279	2,166,050
Hematologic Cancer	0.222	0.094	0.309	481,932	2,166,050
Prostate/Genito. Cancer	0.127	0.155	0.154	274,981	2,166,050
Thoracic Cancer	0.163	0.111	0.096	352,881	2,166,050
Panel B: Demographic Information					
Bene. Female	0.585	0.548	0.577	1,267,710	2,166,050
Bene. Black	0.080	0.090	0.081	172,722	2,166,050
Bene. Hispanic	0.011	0.008	0.011	23,713	2,166,050
Bene. Asian	0.014	0.011	0.019	31,227	2,166,050
Bene. Other/Non-White	0.015	0.013	0.020	32,070	2,166,050
Bene. Age 67–70	0.309	0.302	0.375	668,247	2,166,050
Bene. Age 70–74	0.274	0.269	0.299	593,753	2,166,050
Bene. Age 75–79	0.220	0.224	0.204	475,614	2,166,050
Bene. Age 80–84	0.151	0.156	0.118	327,810	2,166,050
Bene. Age 85+	0.105	0.105	0.071	228,195	2,166,050
Bene. Full Dual	0.073	0.072	0.063	157,284	2,166,050
Panel C: Chronic Conditions Indicators					
Alzheimer	0.023	0.021	0.016	49,912	2,166,050
Alz. with Dementia	0.082	0.076	0.065	177,334	2,166,050
AMI	0.014	0.013	0.012	30,411	2,166,050
Anemia	0.559	0.501	0.526	1,210,975	2,166,050
Asthma	0.069	0.069	0.071	149,164	2,166,050
Atrial Fibrillation	0.129	0.126	0.116	278,790	2,166,050
Colorectal Cancer	0.102	0.066	0.052	221,423	2,166,050
Endometrial Cancer	0.008	0.010	0.009	16,684	2,166,050
Lung Cancer	0.188	0.136	0.116	407,495	2,166,050
Prostate Cancer	0.135	0.153	0.156	292,052	2,166,050
Cataract	0.206	0.217	0.223	447,094	2,166,050
CHF	0.209	0.192	0.177	453,140	2,166,050
CKD	0.314	0.304	0.311	679,220	2,166,050
COPD	0.231	0.211	0.156	501,019	2,166,050
Depression	0.196	0.193	0.206	425,230	2,166,050
Diabetes	0.315	0.302	0.284	682,425	2,166,050
Hyperplasia	0.112	0.107	0.121	242,540	2,166,050
Glaucoma	0.104	0.119	0.121	225,445	2,166,050
Hip Fracture	0.011	0.008	0.010	24,866	2,166,050
Hyperlipidemia	0.570	0.573	0.569	1,233,660	2,166,050
Hypertension	0.726	0.718	0.689	1,572,380	2,166,050
Hypothyroidism	0.190	0.186	0.191	406,991	2,166,050
Ischemic Heart Disease	0.395	0.381	0.364	855,986	2,166,050
Osteoporosis	0.107	0.107	0.115	232,089	2,166,050
Rheumatoid Arthritis	0.372	0.372	0.380	806,324	2,166,050
Stroke	0.047	0.044	0.042	102,774	2,166,050
Panel D: Other Characteristics					
Chronic Cond. (0–3]	0.299	0.304	0.348	646,574	2,166,050
Chronic Cond. (3–4]	0.129	0.137	0.135	279,510	2,166,050
Chronic Cond. (4–6]	0.260	0.266	0.253	563,048	2,166,050
Chronic Cond. (6–8]	0.184	0.176	0.163	399,571	2,166,050
Chronic Cond. >8	0.128	0.122	0.108	277,347	2,166,050
Pred. Mortality Q1	0.200	0.205	0.260	433,210	2,166,050
Pred. Mortality Q2	0.200	0.193	0.243	433,210	2,166,050
Pred. Mortality Q3	0.200	0.198	0.211	433,210	2,166,050
Pred. Mortality Q4	0.200	0.192	0.173	433,210	2,166,050
Pred. Mortality Q5	0.200	0.161	0.118	433,210	2,166,050

Notes: The table presents characteristics of the differential distance compliers, in comparison to the overall sample of chemotherapy beneficiaries.

Table E13: Subspecialist Access and Use of Newly FDA Approved Drugs

	Any	**Part B**	**Part D**
Panel A: First Stage			
$\sinh^{-1}(\text{DD})$	-0.024***	-0.024***	-0.024***
	(0.001)	(0.001)	(0.001)
Panel B: Reduced Form			
$\sinh^{-1}(\text{DD})$	0.000*	0.000	0.000
	(0.000)	(0.000)	(0.000)
Panel C: Structural Form			
Any Office Visit Subs.	0.000***	0.005***	0.005***
	(0.000)	(0.000)	(0.000)
Panel D: 2SLS			
Any Office Visit Subs.	0.002*	0.007	0.007*
	(0.001)	(0.005)	(0.524)
Adj R^2	0.016	0.114	0.064
Observations	2,165,024	2,165,024	2,165,024
Mean Dep. Var.	0.029	0.026	0.017
F-Stat (1st Stage)	1,823	1,823	1,823

Notes: The table provides estimates on the effect of access to subspecialized oncologists on the probability to utilize a cancer drug where the generic substance has received FDA approval within the last two years. Column 1 indicates the effect on the probability to receive any such drug, either through infusion (Part B) or orally (Part D). Column 2 provides estimates of the effect for Part B drugs only and column 3 provides estimates for Part D drugs. Panel A presents first stage estimates, Panel B the respective reduced form estimates, Panel C the structural form estimates and Panel D provides 2SLS estimates. All models include demographic, ZCTA level and chronic conditions controls as well fixed effects for the beneficiaries' ZCTA and cancer type by year. Standard errors are clustered at the ZCTA level. First-stage strength is reported using the Kleibergen-Paap F-statistic. Signif. Codes: ***: 0.01, **: 0.05, *: 0.1.

Table E14: Subspecialist Access and Average Age of Cancer Drugs

	All	Part B	Part D
Panel A: First Stage			
$\sinh^{-1}(DD)$	-0.024***	-0.024***	-0.024***
	(0.001)	(0.001)	(0.001)
Panel B: Reduced Form			
$\sinh^{-1}(DD)$	0.012	0.000	0.018
	(0.009)	(0.013)	(0.012)
Panel C: Structural Form			
Any Office Visit Subs.	-0.871***	-1.217***	-0.377***
	(0.023)	(0.031)	(0.029)
Panel D: 2SLS			
Any Office Visit Subs.	-0.502	-0.011	-0.797
	(0.386)	(0.552)	(0.524)
Adj R^2	0.240	0.239	0.353
Observations	2,159,340	1,726,189	1,052,397
Mean Dep. Var.	27.46	31.68	25.65
F-Stat (1st Stage)	1,827	1,690	1,218

Notes: The table provides estimates on the effect of access to subspecialized oncologists on the average age of cancer drugs used during the year of chemotherapy initiation. Column 1 indicates the effect on the average age of Part B and Part D drugs combined, column 2 provides estimates of the effect on the average age of Part B drugs and column 3 provides estimates for the effect on the average age of Part D drugs. anel A presents first stage estimates, Panel B the respective reduced form estimates, Panel C the structural form estimates and Panel D provides 2SLS estimates. All models include demographic, ZCTA level and chronic conditions controls as well fixed effects for the beneficiaries' ZCTA and cancer type by year. Standard errors are clustered at the ZCTA level. First-stage strength is reported using the Kleibergen-Paap F-statistic. Signif. Codes: ***: 0.01, **: 0.05, *: 0.1.

Table E15: Subspecialist Access and End of Life Care (30 Days before Death)

	Any ER	**Any ICU**	**Any Hospice (30-3)**	**Any Hospice (3-0)**
Panel A: First Stage				
$\sinh^{-1}(\text{DD})$	-0.024***	-0.024***	-0.024***	-0.024***
	(0.001)	(0.001)	(0.001)	(0.001)
Panel B: Reduced Form				
$\sinh^{-1}(\text{DD})$	-0.000	0.000	0.000***	0.000
	(0.000)	(0.000)	(0.000)	(0.000)
Panel C: Structural Form				
Any Office Visit Subs.	-0.010***	-0.004***	-0.003***	-0.002***
	(0.000)	(0.000)	(0.000)	(0.000)
Panel D: 2SLS				
Any Office Visit Subs.	0.009	-0.006	-0.015***	0.000
	(0.008)	(0.006)	(0.007)	(0.005)
Adj R^2	0.086	0.057	0.049	0.023
Observations	2,165,024	2,165,024	2,165,024	2,165,024
Mean Dep. Var.	0.072	0.038	0.043	0.020
F-Stat (1st Stage)	1,823	1,823	1,823	1,823

Notes: The table provides estimates on the effect of access to subspecialized oncologists on different measures of end of life care within the last 30 days of a beneficiaries life on the first chemotherapy episode sample. Column 1 shows the effect on the probability of emergency room admission, column 2 on the effect of intensive care unit admission, column 3 shows the effect on whether a beneficiary has had a hospice claim within the last 30 to 3 days before death (30-3) and column 4 whether a beneficiary had any claim within the last 3 days of life (3-0). All outcomes additionally include the condition that a person died during an ongoing chemotherapy episode or within 30 days after. Panel A shows first stage estimates. Panel B shows the reduced form estimates. Panel C shows the structural form estimates, and Panel D provides the 2SLS estimates. All models control for demographics, ZCTA characteristics, comorbidities, and include ZCTA and cancer type-by-year fixed effects. Standard errors are clustered at the ZCTA level. First-stage strength is reported using the Kleibergen-Paap F-statistic. Significance levels: *** p<0.01, ** p<0.05, * p<0.1.

Table E16: Access to Subspecialized Oncologist and Provider Mix

	Nr. Visits	Unique Providers	Unique Specialties
Panel A: First Stage			
$\sinh^{-1}(\text{DD})$	-0.024***	-0.024***	-0.024***
	(0.001)	(0.001)	(0.001)
Panel B: Reduced Form			
$\sinh^{-1}(\text{DD})$	-0.002	0.005***	0.000
	(0.006)	(0.002)	(0.002)
Panel C: Structural Form			
Any Office Visit Subs.	0.789***	0.666***	0.417***
	(0.018)	(0.006)	(0.004)
Panel D: 2SLS			
Any Office Visit Subs.	0.094	-0.194***	-0.003
	(0.271)	(0.083)	(0.064)
Adj R^2	0.221	0.175	0.169
Observations	2,165,024	2,165,024	2,165,024
Mean Dep. Var.	12.08	4.56	4.02
F-Stat (1st Stage)	1,822	1,822	1,822

Notes: The table provides estimates on the effect of access to subspecialized oncologists on different measures relevant for the beneficiary provider mix for our full sample of first chemotherapy episodes. Column 1 provides estimates for the effect on the number of E&M office visits, column 2 the number of unique provider NPIs, column 3 the number of unique specialties. Panel A shows first stage estimates. Panel B shows the reduced form estimates. Panel C shows the structural form estimates, and Panel D provides the 2SLS estimates. All models control for demographics, ZCTA characteristics, comorbidities, and include ZCTA and cancer type-by-year fixed effects. Standard errors are clustered at the ZCTA level. First-stage strength is reported using the Kleibergen-Paap F-statistic. Significance levels: *** p<0.01, ** p<0.05, * p<0.1.

Table E17: 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 E18: Access to Subspecialists and Placebo Outcomes

	Year + 1			Year + 2		
	AMI	**Hip Fracture**	**Stroke**	**AMI**	**Hip Fracture**	**Stroke**
Panel A: First Stage						
$\sinh^{-1}(\text{DD})$	-0.024***	-0.024***	-0.024***	-0.024***	-0.024***	-0.024***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Panel B: Reduced Form						
$\sinh^{-1}(\text{DD})$	0.000	0.000	0.000*	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Panel C: Structural Form						
Any Office 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)
Panel D: 2SLS						
Any Office Visit Subs.	-0.001	-0.001	0.005*	0.000	-0.001	0.001
	(0.002)	(0.002)	(0.003)	(0.002)	(0.002)	(0.003)
Adj R^2	0.006	0.002	0.002	0.003	0.002	0.001
Observations	2,014,625	2,014,625	2,014,625	1,845,326	1,845,326	1,845,326
Mean Dep. Var.	0.005	0.005	0.008	0.003	0.004	0.006
F-Stat (1st Stage)	1,767	1,767	1,767	1,731	1,731	1,731

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 in the Part A Inpatient file. Results are estimated on the sample of first chemotherapy episode. Panel A shows first stage estimates. Panel B shows the reduced form estimates. Panel C shows the structural form estimates, and Panel D provides the 2SLS estimates. All models control for demographics, ZCTA characteristics, comorbidities, and include ZCTA and cancer type-by-year fixed effects. Standard errors are clustered at the ZCTA level. First-stage strength is reported using the Kleibergen-Paap F-statistic. Significance levels: *** p<0.01, ** p<0.05, * p<0.1.