Does Early Pain Management Pay Off? Evidence on Utilization, Costs, and Surgeries

Buyun Li
Kelley School of Business, Indiana University, Bloomington, IN 47405.
libu@iu.edu

Mohammad Zhalechian • Christopher Chen
Kelley School of Business, Indiana University, Bloomington, IN 47405,
mzhale@iu.edu •cch3@iu.edu

Chronic pain is an increasing public health concern that affects an estimated 50 million adults in the United States each year and contributes to substantial healthcare costs and lost productivity. We examine whether early intervention by pain management specialists can reduce surgeries and costs. Leveraging a 2019 pilot program by insurers that eliminated copays and deductibles for three physical therapy sessions for low back pain patients across five states, we implement a novel two-stage model that employs a difference-in-differences (DiD) estimate as an instrumental variable (IV). The first stage estimates the exogenous change in the likelihood of patients receiving early pain management intervention in a DiD framework. The second stage uses the change as an IV for early pain management intervention. We find that early intervention significantly reduces healthcare utilization, costs, surgeries, and opioid reliance for patients with new-onset chronic pain. We validated our findings using insurance claims from a chronic-pain specialist network in Southern California. Our work provides large-scale causal evidence of the value of strategically shifting PM consultation in the care trajectory and underscores the critical need for providers, payers, and policymakers to restructure referral pathways and benefit designs to prioritize timely pain management specialist access.

Key words: Early intervention, pain management, chronic pain, healthcare operations, empirical healthcare
August 1, 2025

1. Introduction

Chronic pain represents a major public health and economic challenge in the United States, imposing substantial burdens on patients' quality of life and driving significant healthcare expenditure (Institute of Medicine 2011, U.S. HHS 2019). Estimates suggest that the total cost ($560-$635 billion) for treating these patients exceeds those of heart disease and cancer combined, and it is further exacerbated by the downstream consequences of suboptimal treatment, such as the opioid crisis (Gaskin and Richard 2012, Murphy et al. 2020). Treatment pathways for chronic pain vary considerably, ranging from conservative care (e.g., physical therapy, medication) to pain management (PM) techniques (e.g., nerve blocks, epidural injections, neuromodulation) and invasive surgical procedures (e.g., spinal fusion).

While clinical guidelines suggest conservative, non-operative treatments early for most non-emergent pain conditions before resorting to surgery. However, current practice patterns do not reflect this progression, as the number of surgeries for conditions like degenerative back pain has risen dramatically in recent decades (Martin et al. 2019). Surgical procedures for chronic pain are costly and often have uncertain long-term efficacy, with a notable proportion of patients experiencing persistent issues post-operatively (Martin et al. 2019, Mirza and Deyo 2007). Moreover, there is significant evidence linking major surgery and the subsequent management of post-surgical pain to risks of opioid dependence or addiction (Alam et al. 2012, Chou et al. 2016, Brummett et al. 2017). Despite these well-documented concerns and the substantial resource utilization associated with surgical pathways (Martin et al. 2011), surgery remains a longstanding treatment option, particularly when initial conservative measures are deemed insufficient.

PM as a specialty aims to diagnose and treat pain sources, often employing a more conservative care pathway that utilizes a range of diagnostic and therapeutic options, including minimally invasive procedures where appropriate. A key objective of PM is to provide effective relief and functional improvement in order to avoid more costly and risky surgeries and reduce long-term reliance on problematic medications like opioids (Staats et al. 2022, Manchikanti et al. 2009). However, the optimal timing and role of PM specialists within the broader chronic pain care pathway remain unclear. This uncertainty reflects a fundamental operational trade-off. While early engagement with PM specialists (hereafter, "early PM") may avoid invasive procedures and lead to more efficient care, it is also conceivable that it could, for some patients, delay necessary surgical interventions. This delay could potentially result in increased costs and worse outcomes, with early PM merely adding an unnecessary layer of services without fundamentally altering downstream trajectories. This ambiguity around the role of early PM also raises key operational questions about how early PM access influences treatment trajectories, care coordination, and downstream resource use.

Existing research, such as the observational studies by Staats et al. (2022, 2023), provides valuable evidence on pain management (PM). Using Medicare and commercial payer data, their work suggests that patients initially treated by PM specialists had lower risk-adjusted costs. This reduction in cost, driven by lower inpatient, outpatient, and post-acute care costs, occurred despite the PM cohort appearing sicker at baseline when compared to those initially treated by surgical specialists. However, these observational studies are inherently limited due to potential selection bias. Patients who opt for early PM versus those who seek surgical consultations may differ systematically in observed and unobserved ways. This may include differences in pain severity tolerance, functional goals, health literacy, primary care physician referral patterns, and access barriers. All of these factors can independently influence long-term outcomes and costs, meaning that the observed correlations might not reflect the true causal effect of the timing and type of initial specialist care.

This paper aims to address this gap by estimating the *causal* effect of early consultation and treatment by PM specialists on subsequent healthcare utilization and total cost of care for patients with new-onset chronic pain. We leverage a unique natural experiment embedded within a large-scale database of administrative health claims (Optum's de-identified Clinformatics® Data Mart Database (Optum 2022)). Specifically, in 2019, Optum implemented a pilot program in select US states that removed patient cost-sharing (co-pays) for the first three physical therapy and chiropractic (PT/C) visits. This policy change provides an exogenous shock to the relative cost and accessibility of these initial conservative care services. We argue that this increased access to PT/C care impacted treatment trajectories by enhancing patient screening and thus plausibly increased the likelihood of early referral to PM specialists for patients in the policy-affected divisions.

We employ a novel instrumental variable (IV) strategy, specifically using a difference-in-differences (DiD) framework to construct the instrument. The instrument, derived from a policy that does not directly target PM specialist visits but rather access to conservative care upstream, offers a novel source of exogenous variation to identify the causal effects of early PM engagement. The DiD estimate leverages the interaction between residing in a policy-affected division and the post-policy period. This instrument isolates exogenous variation in the propensity to receive early PM care, allowing us to overcome selection bias and estimate the Local Average Treatment Effect (LATE) of early PM on total cost and utilization. Our primary contribution is providing, to our knowledge, the first large-scale causal evidence of the downstream operational and economic impact of early access to PM specialists. By doing so, we offer empirical insights into the design of high-value care pathways. We quantify the potential cost savings and utilization shifts, offering valuable insights for healthcare operations management (OM), value-based care design, and insurance benefit structuring.

Our analysis reveals that early access to PM specialists leads to significant reductions in healthcare utilization and opioid exposure among chronic pain patients. Specifically, patients with early PM access incur, on average, \$6,173 (12.2%) lower total healthcare costs over a 24-month period and experience 2.96 (23.4%) fewer healthcare visits. The rate of undergoing surgery also declines by 10%, suggesting that early non-surgical interventions may serve as effective substitutes for invasive procedures. Among patients who eventually receive surgery, early PM is associated with a \$4,254 (8.8%) reduction in post-surgical spending and 11.6% fewer follow-up visits. These results imply that early PM may enhance recovery and reduce downstream utilization. Moreover, early PM access leads to more favorable opioid-related outcomes. Patients exposed to early PM receive 147 fewer morphine milligram equivalents (MME), a 5.2% reduction, and have 2.8% lower odds of being prescribed high-dose opioids (defined as >50 MME per day). These effects are robust across multiple model specifications and hold across subgroups. We further validated our findings using insurance claims data from a Southern California chronic pain specialist network, which includes over 300 provider

groups and 2,000 clinicians. The reductions in costs, visits, and surgeries remain in line with the main estimates, underscoring the robustness and generalizability of the early PM engagement benefits.

Delving further into the observed benefits of early PM, we find that the gains appear to stem from several operational mechanisms. First, we find evidence of a direct substitution effect, where early PM reduces the subsequent need for major spinal surgeries and advanced imaging, possibly through targeted non-operative treatments and more precise diagnostics. Second, early PM shifts treatment towards cost-effective and minimally invasive pain-management procedures. We observe that early PM adoption increases low-cost PM procedures (e.g., epidural injections or facet-nerve blocks) and lowers high-cost PM procedures. Third, we observe that early PM improves care coordination, which reduces the overall intensity of healthcare interactions. This is evidenced by patients who visit a pain management specialist within their first six encounters subsequently interacting with fewer distinct medical specialties over the follow-up horizon. From an OM perspective, these mechanisms demonstrate how early PM adoption optimizes the care pathway. Our findings suggest that it achieves this by promoting more efficient coordination and a more judicious mix of interventions, substituting high-cost services with targeted, lower-cost procedures. This shift not only generates significant cost savings but also has broader implications for system-level resource management. By averting avoidable surgical episodes and their associated follow-up care, early PM can alleviate strain on high-demand resources such as operating rooms and surgical specialists, potentially improving access for other patients and creating positive spillover effects across the health system.

Our findings also have substantial managerial and policy implications for the design and delivery of chronic pain care. For healthcare payers and integrated delivery systems, our results highlight the potential return on investment from payer (public or private) policies that lower barriers for early access to PM specialists, through strategic offerings of benefits that can guide patients towards more efficient care trajectories. For providers, our research highlights the pivotal role of PM specialists in pathway management and effective gatekeeping, informing referral decisions from primary care. More broadly, this causal evidence contributes to the ongoing policy discussion around value-based care, offering a quantifiable argument for prioritizing specialized, non-operative treatment strategies in the early stages of chronic pain that can achieve both better economic outcomes and potentially improve patient care by avoiding unnecessary invasive procedures.

2. Literature Review and Hypothesis

2.1. Literature Review

Our study is related to the healthcare operations and health economics literature on efficient care delivery, as well as the research on clinical pain management strategies.

2.1.1. Healthcare Operations. The management of chronic pain presents a significant operational challenge for healthcare systems due to its inherent patient heterogeneity and diverse, often prolonged, treatment pathways. The healthcare operations literature extensively explores care coordination, referral systems, and the role of gatekeeping in managing access to specialized resources. Studies on referral patterns and team formation, such as Agha et al. (2022), suggest that established relationships and consistent referral practices can enhance care efficiency. Analytical and empirical works have examined the incentives and optimal strategies for primary care physicians acting as gatekeepers when referring patients to specialists. This involves balancing the costs of specialist care against the benefits of their expertise (Shumsky and Pinker 2003, Freeman et al. 2021, Adida and Bravo 2023). While much of this work focuses on the decision *whether* to refer or *to whom*, we study *when* a PM specialist might enter the care pathway, which is different than the gatekeeping or triage function studied in prior works.

Furthermore, the principle of care continuity, shown to improve outcomes in primary care settings (Ahuja et al. 2020), is highly relevant to chronic pain management. PMs engage in coordinating multi-modal care and having earlier PM specialist involvement affects continuity and coordination for a complex chronic condition. Our work seeks to provide empirical evidence on these operational hypotheses by linking an exogenous policy shock influencing referral pathways to tangible system-level outcomes like costs and utilization (Chen and Savva 2018).

2.1.2. Chronic Pain Management. There is a large literature on the economic burden of chronic pain, which covers direct medical expenditures, lost productivity, and diminished quality of life (Gaskin and Richard 2012, Institute of Medicine 2011). Most of the health economics and clinical literature compares the effectiveness and cost-effectiveness of various treatment modalities for chronic pain, including conservative care, pharmacotherapy, interventional procedures, and surgery (Turk 2002, Turk et al. 2011). While numerous randomized controlled trials (RCTs) and comparative effectiveness studies evaluate specific interventions (e.g., spinal fusion vs. intensive rehabilitation (Mannion et al. 2013), epidural injections vs. placebo (Chou et al. 2015)), these often focus on narrowly defined patient subgroups or specific procedures. This approach tends to overlook the broader strategic impact of when a patient accesses specialized pain management care. Consequently, large-scale causal evidence on the system-level economic impact of early access to PM specialists that includes the full episode of care is limited. Our study seeks to fill this void by quantifying these system-level effects.

Clinical research has explored optimal referral timing using longitudinal analysis for various chronic conditions, often finding benefits to earlier specialist intervention for conditions like kidney disease (Smart et al. 2014, Heatley 2009), cystic fibrosis (Lebecque et al. 2009), and rheumatoid arthritis

(Emery et al. 2002). Chronic pain, however, presents a different challenge where evidence suggests low-intensity frontline treatment may be preferable. For instance, in the context of non-specific lower back pain (LBP), studies suggest that specialist visits may lead to earlier, potentially unnecessary imaging services and greater healthcare costs without health benefit (Flynn et al. 2011, Srinivas et al. 2012, Webster et al. 2013, Frogner et al. 2018). Fritz et al. (2015), using propensity score matching, found that initial management by a PT for LBP leads to less surgery, fewer injections, and reduced specialist and Emergency Department visits, and lower health care charges. Harwood et al. (2022) corroborates these findings through an IV analysis, finding that care beginning with a specialist visit led to the highest total 1-year costs and a greater frequency of early opioid prescriptions compared to care beginning with an acupuncturist or chiropractor.

Our work differs from and extends the LBP-focused stream of literature in a few ways. First, while these studies often examine which provider is seen first, our research question examines the impact of policies that facilitate earlier referral to PM specialists as a strategic point in the care pathway. Second, we consider a broader population of patients with new-onset chronic pain, encompassing conditions potentially more severe or complex than the typical self-limited LBP, where specialist PM expertise might be more critical earlier.

Studies examining broader chronic pain care pathways using large administrative datasets offer valuable descriptive insights but often face challenges in establishing causality. Observational studies, such as Staats et al. (2022, 2023), have highlighted correlations between the specialty of the initial treating provider and subsequent healthcare costs and utilization patterns. For instance, their work suggested lower risk-adjusted costs for Medicare Fee-for-Service (FFS) beneficiaries with new-onset pain who were initially managed by interventional pain specialists compared to surgical specialists. Similarly, other research investigating treatment sequences and predictors of specific interventions, like lumbar fusion (Deyo et al. 2011), often points to significant regional variation and provider-level influences (Weinstein et al. 2006). However, a fundamental limitation of these studies is the potential for *selection bias*. That is, patients who are referred to or choose early PM may systematically differ in both observed and unobserved ways, such as pain severity, functional goals, or health beliefs from those who follow other pathways. These differences can independently influence outcomes, making it difficult to disentangle the effect of the early PM intervention itself from underlying patient heterogeneity. Our study advances this line of research by employing a quasi-experimental design specifically aimed at mitigating such endogeneity concerns.

2.2. Hypothesis

Building on the existing literature concerning chronic pain management, we hypothesize that timely engagement with PM specialists can reduce overall costs and utilization, optimize patient care, and impact post-surgery outcomes.

2.2.1. Overall Costs and Utilization. The net effect of early PM engagement on total healthcare costs over a 24-month period is unclear. On one hand, early specialist intervention inherently incurs immediate costs and adds initial visits. If early PM primarily functions as an additional layer of services without substantially altering the downstream care trajectory, or if it leads to more intensive diagnostic tests or multi-modal treatment plans involving frequent follow-ups, then both total costs and total service visits could conceivably *increase*. Furthermore, if early PM inappropriately delays inevitable surgery that subsequently becomes more complex or emergent, both metrics could be adversely affected.

Conversely, early PM could decrease both cost and utilization. This would occur if expert diagnosis and the application of targeted, often minimally invasive, treatments occur sooner, leading to more rapid symptom control. This, in turn, would reduce expenditures on prolonged or potentially less effective care, minimize reliance on long-term medications, and crucially, avert or delay expensive surgical procedures and their associated hospitalizations (Manchikanti et al. 2009, Staats et al. 2022). From an OM perspective, this pathway signifies improved process efficiency by more accurately matching patient needs to appropriate service intensity earlier in the care trajectory and avoiding resource misallocation towards less effective or higher-risk interventions. If early PM effectively resolves pain, improves function, or facilitates superior care coordination (e.g., mitigating fragmented care or excessive "doctor shopping" (Adelman et al. 2013)), the initial PM visits might be more than offset by a substantial reduction in subsequent visits. Given prior observational evidence suggesting favorable economic outcomes (Staats et al. 2022, 2023), we hypothesize:

Hypothesis 1. Early PM engagement will lead to lower total healthcare costs over a 24-month follow-up period for patients with new-onset chronic pain.

Hypothesis 2. Early PM engagement will lead to a net reduction in the total number of outpatient and professional service visits over a 24-month follow-up period for patients with new-onset chronic pain.

2.2.2. Care Pathways. Two key mechanisms through which early PM may achieve overall cost and utilization efficiencies involve influencing the likelihood of major surgical interventions and modifying long-term opioid use.

A pivotal outcome in chronic pain care is the progression to major surgical intervention. Early PM could theoretically influence surgery rates in either direction. On one hand, a PM visit might identify surgical candidates more precisely or sooner. Some PM interventions are also a necessary prerequisite for surgery, serving to either confirm the source of pain or optimize a patient's condition to ensure a successful surgical outcome. This would lead to no change or even an increase in surgery rates for appropriately selected individuals. On the other hand, by employing advanced diagnostic techniques

and non-operative treatments, early PM may resolve or adequately manage pain for a substantial subset of patients who might otherwise have proceeded to surgical intervention (Manchikanti et al. 2012). This could, in turn, reduce the surgery rates for these patients.

The role of early PM in influencing long-term opioid therapy is critical given the ongoing public health concerns. It is conceivable that for some patients with severe, acute-on-chronic pain, a PM specialist might initiate or optimize opioid therapy as part of a comprehensive plan where benefits are deemed to outweigh risks, at least in the short term. However, one key principle of pain medicine is opioid stewardship and the promotion of multimodal, non-opioid strategies. PM specialists are trained to identify underlying pain generators for targeted treatment, utilize a range of non-opioid pharmacotherapies and interventional techniques as alternatives or adjuncts to opioids, and apply principles of careful patient selection, risk assessment, monitoring, and tapering for any opioid prescriptions (Dowell 2022). Furthermore, by potentially reducing the likelihood of major surgery (H2), which itself is a significant trigger for new persistent opioid use (Brummett et al. 2017, Chou et al. 2016), early PM may indirectly contribute to lower overall opioid exposure in the patient population. We therefore hypothesize:

Hypothesis 3. Early PM engagement will reduce the number of surgeries undergone by patients with new-onset chronic pain during a 24-month follow-up period after major surgery.

Hypothesis 4. Early PM engagement will decrease opioid therapy reliance among patients with new-onset chronic pain over a 24-month follow-up period.

2.2.3. Post-Surgery Outcomes. For the subset of patients who proceed to major surgery, despite any early interventions, the involvement of PM specialists prior to the surgical event may still influence post-operative trajectories. Early PM might lead to better patient selection or pre-operative optimization, potentially resulting in *smoother recoveries and lower subsequent resource utilization*. This could manifest as reduced post-surgery costs due to fewer complications, shorter initial hospital stays (especially if PM enhances pre-operative planning), or less need for intensive post-acute care and readmissions (Wanderer et al. 2018, Hollenbeck et al. 2011). Similarly, a less complicated post-operative treatment trajectory would likely translate into fewer follow-up visits.

Conversely, patients receiving early PM and still requiring surgery might inherently represent a more complex cohort whose conditions proved resistant to non-operative treatments. For this group, their post-surgery needs might be higher regardless of pre-operative PM. Given the complexity of surgical outcomes, we present two-sided hypotheses:

Hypothesis 5a. Among patients with new-onset chronic pain who ultimately undergo major surgery, prior early PM will be associated with lower total post-surgery healthcare costs and visits within the 24-month follow-up period.

HYPOTHESIS 5B. Among patients with new-onset chronic pain who ultimately undergo major surgery, prior early PM will be associated with higher total post-surgery healthcare costs and visits within the 24-month follow-up period.

3. Empirical Setting and Data

In this section, we begin by providing background on chronic pain and the clinical specialty of PM. We then describe the data sources used in our analysis, followed by detailed definitions of the outcome and control variables included in the empirical models.

3.1. Context

Chronic pain affects over 60 million U.S. adults, about one in four, creating significant public health and economic burdens (National Center for Health Statistics 2024, Zajacova and Grol-Prokopczyk 2024). Based on the International Classification of Diseases (ICD), seven types of chronic pain are identified, including primary, cancer, postsurgical/posttraumatic, neuropathic, headache/orofacial, visceral, and musculoskeletal (Treede et al. 2015). These involve nociceptive, neuropathic, or nociplastic mechanisms, which influence treatment and prognosis (Cohen et al. 2021). Diagnosis of chronic pain is challenging due to often absent pathology. More than 90% of back pain and 40% of overall pain are nonspecific, forcing physicians to treat pain without a clear diagnosis. Objective tests are limited and are based on clinical judgment and exclusion (see Appendix Table EC.1 for ICD-10 codes). While targeting mechanisms is ideal (Woolf et al. 1998), multimodal approaches, such as physical therapy, psychotherapy, medications, interventional procedures, and surgery, are also common (Cohen et al. 2021). Pain specialists coordinate these, especially for complex cases that require advanced expertise.

Pain management is a multidisciplinary specialty that focuses on the treatment of chronic pain through pharmacologic, interventional, and rehabilitative methods. In the US, pain specialists typically complete residency training in anesthesiology, physical medicine and rehabilitation (PM&R), neurology, or internal medicine, followed by dedicated pain fellowships. Certification by the American Board of Pain Medicine (ABPM) requires advanced training in pain treatments and a validated examination (American Board of Pain Medicine 2023). The diverse training background within pain medicine enables providers to address the complex and multifactorial nature of chronic pain. Anesthesiology-trained specialists contribute expertise in interventional pain management procedures such as nerve blocks, epidural injections, and spinal cord stimulation. PM&R-trained physicians emphasize functional restoration through physical therapy and neuromuscular rehabilitation. Providers from internal medicine or neurology often focus on medication, behavioral therapy, and psychosocial support. Interventional pain management specialists, often with anesthesiology or PM&R backgrounds, add further proficiency in minimally invasive techniques such as radiofrequency ablation and facet joint injections, which target specific areas of pain and offer effective relief when

conservative treatments are not sufficient. Together, these specialties form a holistic model of care, combining different strategies to provide well-coordinated personalized pain management.

3.2. Data Source and Sample Construction

Our empirical analysis draws on administrative claims from the Optum's de-identified Clinformatics® Data Mart Database (Optum® CDM) spanning 2007 to 2022 (Optum 2022). Optum® CDM is a comprehensive, longitudinal dataset of administrative health claims for members of large commercial and Medicare Advantage health plans across all U.S. census divisions. It includes rich de-identified information on healthcare utilization, diagnoses, procedures, prescription drug fills, and expenditures.

For our analysis, we draw on three main sources of patient data, including (i) patient-level demographics and characteristics, (ii) medical claims, and (iii) pharmacy claims. The patient-level demographics and characteristics include age, sex, and census division of residence, as well as health plan enrollment start and end dates. We use these records to define continuous insurance eligibility windows and to construct baseline covariates for adjustment. The medical claims include all reimbursed services delivered in inpatient, outpatient (including emergency departments and ambulatory surgical centers), and professional settings. Each claim record provides service dates, provider specialty and facility identifiers, diagnostic information (ICD-10-CM), procedure codes (CPT and HCPCS), and standardized cost fields such as allowed amounts, plan-paid amounts, and patient cost-sharing. The pharmacy claims capture dispensed outpatient prescriptions. These include drug identifiers (NDC codes), fill dates, quantity dispensed, supply days, and associated cost components. We use the pharmacy claims data to construct patient-level measures of opioid use and dosing, including morphine milligram equivalents (MME).

We construct the analytic cohort in four sequential stages. First, using the medical claims from 2016–2021, we identify the earliest occurrence of a qualifying chronic pain diagnosis (e.g., ICD-10 codes M54, M47, G89, M54, M96; see Table EC.1 for the full list) and designate that encounter as the patient's index date. We use the introduction of a new zero-copay policy in 2019 as an instrument. Consequently, the observations between 2016 and 2019 are defined as the pre-policy period, whereas encounters from 2020 to 2021 fall into the post-policy period. Data from the year 2019 is removed as it represents the policy roll-out window.

To ensure a clean analysis, we next impose a 12-month new-onset requirement. Patients must be continuously enrolled for the full year prior to the index date and free of any qualifying pain diagnoses during that baseline interval. We then require continuous enrollment for 24 months after the index date so that healthcare utilization and cost outcomes can be observed without censoring. Finally, we exclude patients with cancer diagnoses (ICD-10 codes beginning with "C" or in the D00–D49 range) to avoid pain trajectories driven by malignancy. This sampling construction yields a pain cohort with

an uninterrupted 36-month observation window, 12 months before the index diagnosis and 24 months for outcome assessment, thus aligning the data structure with our quasi-experimental design. Figure EC.1 shows the details of the sample construction, together with the relevant exclusion criteria and the impact on the sample size.

3.3. Variables

3.3.1. Dependent Variables. To comprehensively assess the impact of early PM, we examine a set of dependent variables that capture both the economic and operational dimensions of patient care pathways over a 24-month period after index pain diagnosis.

Total Healthcare Costs, Visits, and Surgeries. We consider three outcomes related to total costs and utilization.

Total Costs: The primary economic outcome is total all-cause healthcare cost per patient, defined as the sum of the standardized allowed amounts in all inpatient, outpatient, professional, and pharmacy claims. Allowed amounts represent the full negotiated payment, inclusive of insurer contributions and patient cost-sharing, thereby capturing total resource utilization from a system-wide perspective. To ensure temporal comparability, all costs are standardized to 2019 Resource-Based Relative Value Scale (RBRVS) using a fixed national rate schedule (Centers for Medicare & Medicaid Services 2018).

Total Visits: The total number of medical service encounters, aggregated at the patient level. Because a single visit may generate multiple claims, we define visits by collapsing claims on the same service date and provider. This outcome captures overall healthcare utilization and reflects the intensity of patient interaction with the healthcare system.

Surgeries: We define surgery as the count of specific pain-related surgical operations received by a patient, e.g. spinal fusion and laminectomy. Surgeries are identified using Current Procedural Terminology (CPT) codes corresponding to major pain-related surgical operations (see Table EC.2 for details[1]). Thereafter in the paper, we refer to these major pain-related surgical operations as 'surgeries'.

Post-Surgery Costs and Visits. For patients who undergo major spinal surgery, we examine two main measures.

Post-Surgery Costs: We measure the total healthcare costs incurred after the date of surgery, calculated over the remaining duration of the 24-month follow-up period.

Post-Surgery Visits: Similarly, we measure the total number of medical visits incurred after the date of surgery, calculated over the remaining duration of the 24-month follow-up period.

These outcomes capture the intensity of downstream healthcare utilization following surgical intervention. Conceptually, if surgery is effective in resolving the underlying condition, we would expect

[1]The definitions follow the Centers for Disease Control and Prevention (CDC) Operative Procedure Code document: https://www.cdc.gov/nhsn/psc/ssi/.

lower post-surgery healthcare utilization, reflected in fewer follow-up visits and reduced spending. Conversely, higher utilization may signal complications, unmet care needs, or unclear surgical benefits.

Opioid-Related Outcomes. We examine four patient-level outcomes related to opioid use, derived from pharmacy claims.

Total MME (MME): This outcome refers to the cumulative milligram equivalent dose (MME) of morphine administered to a patient over the follow-up period. This measure aggregates across all opioid prescriptions and captures the overall extent of opioid exposure, standardizing between different drug formulations and strengths.

Any Opioid Prescription (Any Rx): This is a binary outcome that equals 1 if the patient filled at least one opioid prescription during the follow-up period, and 0 otherwise. This outcome captures the extensive margin of access to opioids.

MME Conditional on Prescription (MME | Rx): We measure the total MME received among the subset of patients who filled at least one opioid prescription. This conditional measure is crucial for understanding the intensity of opioid use, specifically among those who have already used opioids.

High-Dose Opioid Use (High-Dose): Our other binary outcome equals 1 if a patient filled any opioid prescription with a daily dosage exceeding 50 MME. This threshold follows the CDC prescribing guide, which classifies doses above 50 MME/day as high risk because they are approximately two to three times higher in the likelihood of opioid overdose or long-term dependence (Dowell et al. 2016).

3.3.2. Independent Variable. We define the independent variable, $\texttt{EarlyPM}_i$, as a binary indicator equal to 1 if patient i receives PM care within the first *six* visits following their index pain diagnosis, and 0 otherwise. The $\texttt{EarlyPM} = 0$ group therefore includes both patients who receive later PM care and those who never receive PM. In Section 5.2, we examine alternative definitions of early PM, including further discussion on different compositions of the late- and no-PM subgroups.

3.3.3. Control Variables. We control for patient-, time-, and division-level heterogeneity by including a rich set of covariates. At the patient level, we adjust for clinical status via ICD-10 diagnosis codes and indicators for specialties consulted in early treatment, demographic characteristics such as age, gender, race, and household composition (number of children/adults), and socioeconomic factors including income, education level, and occupation type. To absorb temporal shocks, we add year-of-index diagnosis dummies. Also, to account for geographical variation, we include division fixed effects together with interactions between division and each demographic covariate, allowing demographic trends to differ across divisions. A detailed description of the control variables is included in Table EC.3 of the Appendix.

4. Econometric Strategy

To answer our research questions, we need to estimate the causal impact of early PM on healthcare utilization, costs, and surgeries. The fundamental challenge centers on patient selection. Patients who receive early PM intervention may systematically differ from those who do not receive it in unobservable ways, and these differences can correlate with future health outcomes. To address this endogeneity, we employ an instrumental variable (IV) strategy by leveraging a natural experiment arising from a health plan policy change. Below, we describe our empirical approach, starting with a baseline model and then outlining our IV identification strategy.

4.1. Baseline Model

As an initial benchmark, we estimate the association between early PM and our outcomes of interest using an ordinary least squares (OLS) regression. For a given patient i, our baseline model is:

$$\begin{aligned} Y_i = \alpha_0 + \beta_{\text{OLS}} \cdot \texttt{EarlyPM}_i + \Gamma_0 \cdot \texttt{Controls}_i \\ + \psi_0 \cdot \texttt{Div\_FE}_i + \pi_0 \cdot \texttt{Year\_FE}_i + \kappa_0 \cdot (\texttt{Div\_FE}_i \times \texttt{Demographics}_i) + \varepsilon_{0i}, \end{aligned} \tag{1}$$

where Y_i represents an outcome variable (e.g., total costs, total visits, surgeries) for patient i measured over the 24-month follow-up period. $\texttt{EarlyPM}_i$ is the independent variable. $\texttt{Controls}_i$ is a vector of baseline patient demographic and clinical characteristics detailed in Section 3.3.3. $\texttt{Div\_FE}_i$ are fixed effects for the CDC division of patient residence, and $\texttt{Year\_FE}_i$ are fixed effects for the calendar year of the patient's index pain diagnosis. The term $\texttt{Div\_FE}_i \times \texttt{Demographics}_i$ allows for differential demographic trends across divisions.

4.2. Endogeneity Concerns

The decision for a patient to receive early PM is likely non-random. Patients choosing early PM often differ from those who do not along unobservable dimensions that might be correlated with outcomes. For example, patients with higher unmeasured pain severity, greater functional impairment not captured by claims, or specific psychosocial profiles (e.g., lower pain catastrophizing, higher self-efficacy) might be more likely to be referred to or seek out PM specialists. Similarly, patient preferences for minimally invasive treatments versus surgery, or variations in primary care physician referral habits and local PM specialist availability, represent unobserved factors that can drive both early PM uptake and subsequent healthcare trajectories. These unobserved confounders create a correlation between $\texttt{EarlyPM}_i$ and the error term ε_{0i}, leading to biased and inconsistent OLS estimates. While we control for observable clinical status using baseline ICD codes and demographics, residual unobserved heterogeneity remains a significant concern.

4.3. Difference-in-Differences Approach within an Instrumental Variable Framework

4.3.1. Exogenous Shock. To address the endogeneity concerns, especially those driven by time-varying unobservable factors, we employ a DiD approach within an IV framework, which we term *DiD-IV*. This strategy exploits a natural experiment introduced by Optum. We consider the "zero-copay" policy, beginning in 2019, that removed payments for the first three physical therapy and chiropractic (PT/C) visits for beneficiaries in select U.S. divisions. This policy provides an exogenous shock to the accessibility of these initial conservative care services, which we argue influences the likelihood of subsequent early PM engagement.

The policy was implemented in 6 of the 8 states within the South Atlantic (SA) CDC division[2], covering approximately 93% of that division's population. Our control group comprises patients from CDC Census divisions where this specific "zero-copay" policy was not implemented during our study period, including the Pacific, Mountain, West North Central, and East North Central divisions. This setup necessitates a more nuanced identification strategy. Unlike the classic method of using a policy shock as an IV, where assignment is truly random at the individual level, the "zero-copay" policy was assigned to an entire geographic division. This non-random assignment means we cannot assume the treatment (SA division) and control divisions are identical from the outset as they may have had pre-existing differences and been subject to different temporal trends. Therefore, an IV approach that only compares the post-policy groups may not be valid, as it can not distinguish the policy's true effect. The DiD component is thus essential in our setting.

More specifically, our approach has several advantages and is integral to recovering a robust effect. Instead of a policy shock as the instrument, this approach uses a *differenced* policy shock that leverages both cross-sectional (treatment vs. control divisions) and temporal (pre vs. post policy) variation. It explicitly controls for (i) any pre-existing, time-invariant differences between the South Atlantic and control divisions that might affect the propensity for early PM or the dependent variables through division fixed effects, and (ii) secular trends over time, such as increases in healthcare costs and national changes in pain management guidelines, using time fixed effects. Furthermore, under the parallel trends assumption, it cleanly isolates the effect of the instrument, strengthening the argument that the exogenous variation captured by the instrument is due to the policy shock rather than any diverging trends.

4.3.2. Instrument Validity. The validity of the zero-copay policy as an instrument for early PM intervention depends on its relevance and satisfying the exclusion restriction.

[2]Divisions are defined by the CDC; see https://www.cdc.gov/nchs/hus/sources-definitions/geographic-region.htm.

Relevance. The instrument must be significantly correlated with the $EarlyPM_i$. The "zero-copay" policy exogenously reduced financial barriers to accessing initial PT/C visits for patients with new-onset pain in the treatment divisions after 2019. We hypothesize that this increased access to PT/C, in turn, exogenously increased the probability of these patients subsequently receiving early PM through two plausible mechanisms.

The first possible mechanism is *enhancing screening and appropriate triage*. Increased utilization of PT services facilitates a more systematic patient assessment. PTs often employ screening tools, such as the STarT Back questionnaire, which risk-stratifies patients with back pain. Patients identified as medium or high risk (e.g., STarT Back Score ≥ 4 out of 9) are often recommended for, or directly escalated to, more specialized care, including PM, potentially bypassing extended, ineffective courses of conservative care alone (Hill et al. 2011, Foster et al. 2015, Soar et al. 2022). This auto-escalation or co-management protocol can expedite access to PM for appropriate candidates.

The second possible mechanism is *accelerating identification of non-responders*. For patients whose PT/C interventions is initiated but proves insufficient, the "zero-copay" policy allows this determination to be made more quickly. Health plan protocols often require a trial of conservative care (e.g., 4-6 weeks of PT) before authorizing referrals to specialists like PM physicians (Hutchins et al. 2021). By removing initial financial barriers, patients can complete this prerequisite trial sooner if indicated, thus accelerating the timeline to qualifying for and receiving early PM if they are non-responders to initial PT/C interventions (Fritz et al. 2012). We formally test the strength of this relationship in Section 4.3.3.

Exclusion Restriction. The exclusion restriction is valid if our IV affects patient outcomes of interest (e.g., total healthcare costs or major surgery rates) *only* through its effect on the probability of receiving early PM, conditioning on all included covariates and fixed effects. This implies that the policy change itself (i.e., eliminating co-pays for the initial PT/C visits in specific divisions) should not have a direct causal pathway to long-term pain-related healthcare outcomes, other than by shifting patients towards or away from early PM intervention.

Our primary justification for this assumption depends on the policy targeting a specific, relatively low-cost set of initial conservative care services. While PT/C services can have therapeutic value and are recommended as first-line treatments for many chronic pain conditions (O'Connell et al. 2016), we contend that for a population presenting with new-onset chronic pain of a severity or complexity that leads to consideration for specialized treatment (including PM or surgery within a 24-month horizon), the therapeutic impact of a limited number of initial PT/C visits on major, long-term outcomes is likely to be modest compared to the impact of subsequent specialized interventions.

The literature does suggest that there exists short-term benefits for pain and function from PT for conditions like chronic low back pain (Hayden et al. 2005); however, there is limited evidence for

definitive, long-term resolution of complex or persistent chronic pain from very short course PT/C (e.g., 1-4 visits) preventing progression to more intensive care. Many effective programs, such as the GLA:D Back initiative, involve extensive patient education and exercise sessions over several weeks (Kjaer et al. 2018). Furthermore, a notable proportion of patients with chronic pain do not achieve satisfactory long-term relief from initial conservative measures (Foster et al. 2018). This aligns with stepped care models in pain management, where initial, low-intensity interventions are followed by reassessment and potential intensification of treatment if needed (Von Korff and Moore 2001). Thus, we posit that the policy primarily influences major downstream outcomes by reconfiguring the *early* treatment pathway, specifically by facilitating more timely or appropriate access to PM specialists, rather than by providing a long-term solution for the marginal patient affected by the policy.

Despite the literature, the potential for the direct therapeutic effects of the subsidized PT/C visits to independently influence outcomes, leading to a violation of the exclusion restriction, represents the most significant challenge to our identification. We address this critical concern through the following robustness checks. We examine the stability of our estimates to different controls and considering potential confounding factors (see Section 5.2). By controlling for division-level fixed effects and interactions between division and time-varying demographics, we mitigate concerns about time-varying unobserved heterogeneity across divisions that could otherwise violate this restriction. Next, we conduct placebo tests using dermatological procedures which should not theoretically be affected by early PM or PT/C visits. Additionally, we isolate the subset of patients for whom the therapeutic effect is unlikely to play a role. Specifically, we repeat our analysis on the subset of patients with four or fewer PT/C visits. The intuition is that conditions that result in only a limited number of PT/C visits are either minor and easily resolved or major enough to require a referral. This analysis helps to isolate the effect of the policy change on early PM engagement, rather than on the possible direct therapeutic benefits of initial PT/C visits. It also tests the sensitivity of our findings to patients with varying levels of PT/C to demonstrate the consistency of the observed effects across different levels of initial conservative care intensity.

However, we acknowledge that unobservable direct therapeutic benefits of the incentivized PT/C care, or subtle concurrent regional factors not captured by our extensive controls and fixed effects, could remain. Therefore, our local average treatment effect (LATE) estimates should be interpreted as the effect of early PM for patients whose access was influenced by the policy, under the assumption that the policy's primary pathway to affecting major downstream outcomes is primarily mediated by its effect on engagement with specialized pain management.

Finally, it is important to clarify the interpretation of the estimated treatment effect. In our primary specification, the control group ($\texttt{EarlyPM} = 0$) consists of a mixture of patients who receive PM later

in their care trajectory ("late PM") and those who receive no specialized PM at all within the 24-month follow-up period ("never PM"). Consequently, our main LATE estimate reflects the causal effect of early PM relative to this mixed counterfactual. While this accurately captures the impact of a policy that shifts patients into the early PM group from the combined pool of late and never PM, it does not perfectly isolate the effect of timing alone (early vs. late) from the effect of receiving PM at all (early vs. never). To disentangle these effects and provide a cleaner estimate against a "never PM" counterfactual, we conduct a targeted subgroup analysis as a key component of our robustness checks (see Section 5.2.4).

4.3.3. First-Stage Model: Predicting Early PM Likelihood. The first stage models the probability that patient i receives early PM ($\texttt{EarlyPM}_i$) as a function of the instrumental variable ($\texttt{TreatmentPolicy}_i = \texttt{SA}_i \times \texttt{Post}_i$) and the set of control variables. We estimate the following logistic regression model:

$$\begin{aligned}\text{Prob}(\texttt{EarlyPM}_i = 1) = \Lambda\Big(&\alpha_1 + \beta_1 \cdot (\texttt{SA}_i \times \texttt{Post}_i) + \Gamma_1 \cdot \texttt{Controls}_i + \psi_1 \cdot \texttt{Div\_FE}_i \\ &+ \pi_1 \cdot \texttt{Year\_FE}_i + \kappa_1 \cdot (\texttt{Div\_FE}_i \times \texttt{Demographics}_{\boldsymbol{i}}) + \varepsilon_{1i}\Big),\end{aligned} \tag{2}$$

where $\Lambda(\cdot)$ is the logit function, $\texttt{Controls}_{\boldsymbol{i}}$ is a vector of baseline patient demographic, socioeconomic, and clinical characteristics, $\texttt{Div\_FE}_i$ are fixed effects for the U.S. Census division of patient residence, and $\texttt{Year\_FE}_i$ are fixed effects for the calendar year of the patient's index pain diagnosis. The term $\texttt{Div\_FE}_i \times \texttt{Demographics}_{\boldsymbol{i}}$ allows for differential demographic trends across divisions. The coefficient of interest is β_1, which captures the effect of the policy on the likelihood of early PM. From this stage, we obtain the predicted probability of early PM, $\hat{\Pr}(\texttt{EarlyPM}_i)$.

An essential component of validating the instrument is the parallel trends assumption for the DiD component. This requires that, conditional on covariates, the trends in the propensity for early PM are parallel between the treatment (SA) and control divisions *before* the policy implementation in 2019. In Figure 1, we visually inspect these pre-policy trends. The figure shows similar quarterly trends in early PM adoption probability for new-onset chronic pain patients in both the treatment and control divisions during the pre-policy period (2016–2018). This similarity supports the use of the control divisions as a valid counterfactual for the SA division. The figure also illustrates the relevance of our instrument: following the policy implementation in 2019, a clear and sustained divergence emerges, with the probability of early PM adoption increasing in the SA division relative to the control divisions.

Model-free descriptive statistics comparing patient characteristics and unadjusted outcomes between those receiving early PM and those not are presented in Table EC.4. Formal statistical tests for differences in pre-period trends are discussed in Appendix Section EC.2.2. Table EC.4 reports

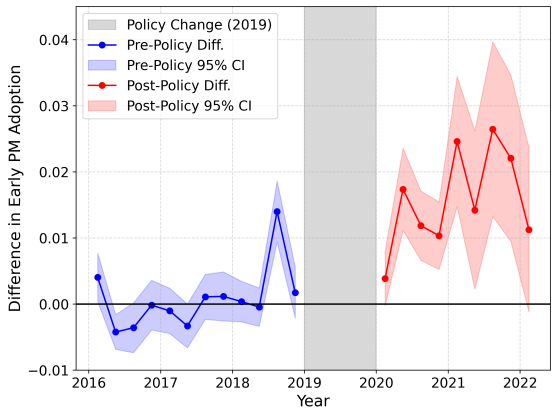

Figure 1 Differences in Pain Management Specialist Adoption Probability: SA Division vs. Control Divisions

mean *pre-policy* outcomes for the South Atlantic (treated) and the pooled control divisions. Early PM adoption is virtually identical in the two groups (1.2% vs. 1.1%), and they are statistically indistinguishable on every other utilization and cost metric. Mean total spending is \$50,715 versus \$48,770 ($p = 0.91$); average outpatient encounters are 9.02 versus 8.77 ($p = 0.10$); and major surgery operations is 2.6 versus 2.5 per 100 patients ($p = 0.47$). The prevalence of any opioid prescription (≈21%) and of high-dose use (≈10%) is likewise balanced. These similarities bolster the plausibility of the parallel-trends assumption underlying our DiD design.

The first-stage regression results, including the estimate for β_1 and the relevant log-likelihood ratio p-value for instrument strength, are reported in Table 1. We find that β_1 is significant, and the log-likelihood ratio p-value shows that the instrument is strong.

Table 1 First-Stage Logistic Regression Predicting `EarlyPM` Engagement

SA × Post	0.18^{***}
	(0.002)
Division FE	Yes
Year FE	Yes
Div × Demographic FE	Yes
Controls	Yes
Pseudo R^2	0.24
Observations	72,160
Log-Likelihood Ratio p-value	0.00

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Coefficient estimated via logistic regression. Standard error, clustered at the division level and bootstrapped ($M = 1{,}000$), is reported in parentheses.

4.3.4. Second-Stage Model: Causal Effect Estimation. In the second stage, we estimate the effect of (instrumented) early PM on our various outcome variables (Y_i). We adopt a two-stage residual inclusion (2SRI) approach as specified by Terza et al. (2008). Because the first stage uses

logistic regression rather than OLS, the 2SRI method yields consistent estimates by adding the latent propensity-shock residuals to the outcome equation. In particular, the outcome is regressed on both the early PM indicator, $\texttt{EarlyPM}_i$, and the residual obtained from the first stage, $(\texttt{EarlyPM}_i - \hat{\text{Prob}}(\texttt{EarlyPM}_i))$, controlling for the same set of covariates and fixed effects:

$$\begin{aligned} Y_i = \alpha_2 + \beta_{IV} \cdot \texttt{EarlyPM}_i + \xi \cdot \texttt{Residual}_i + \Gamma_2 \cdot \texttt{Controls}_{\boldsymbol{i}} + \psi_2 \cdot \texttt{Div\_FE}_i + \pi_2 \cdot \texttt{Year\_FE}_i \\ + \kappa_2 \cdot (\texttt{Div\_FE}_i \times \texttt{Demographics}_{\boldsymbol{i}}) + \varepsilon_{2i}, \end{aligned} \tag{3}$$

where $\texttt{Residual}_i = (\texttt{EarlyPM}_i - \hat{\text{Prob}}(\texttt{EarlyPM}_i))$ is the residual from the first stage.[3]

The coefficient β_{IV} is our estimate of the LATE, which represents the average causal effect of early PM on the outcome Y_i specifically for the subpopulation of "compliers", i.e., those patients whose decision to receive early PM was influenced by the "zero-copay" policy for PT/C visits. A critical element of this estimation strategy is the need to reconcile the group-level nature of the instrument with the individual-level analysis. While the policy shock is applied to entire geographic divisions, we use it to instrument for an individual-level choice ($\texttt{EarlyPM}_i$). Standard errors are estimated using a bootstrap procedure ($M = 1000$) with clustering at the census division level to adjust for (i) the intra-group correlation of errors that arises because all individuals within a given division are subject to the same policy shock and other shared regional factors and (ii) the additional variability introduced by the two-stage estimation process. Given the discrete nature of some outcome variables (e.g., surgery as binary, visits as counts), we estimate different parameterizations of the model. Specifically, we estimate second-stage linear probability models for binary outcomes and Poisson models for counts.

5. Results

In this section, we present evidence that early PM access leads to significant reductions in healthcare costs, service utilization, and surgery rates. We also examine the effects of early PM on post-surgical outcomes and opioid utilization patterns. Next, we validate the robustness of our empirical findings. Finally, we assess the external generalizability of our findings using a distinct dataset from a multi-site pain-management specialist network in Southern California.

5.1. Effect of Early Pain Management Engagement

Our primary results indicate that early PM engagement generates substantial operational and economic efficiencies (see Table 2). As established in our first-stage analysis (see Table 1), the zero-copay policy is a strong and statistically significant predictor of early PM engagement, providing a valid foundation for a causal inference. While both the DiD-IV and the naive OLS estimates show a significant benefit, the DiD-IV estimates, by addressing potential endogeneity, generally yield attenuated

[3] We use raw residuals specified by Terza et al. (2008). For robustness (Section 5.2.3), we estimate the second-stage using Pearson and generalized residuals, finding consistent results.

estimates for our outcomes. The DiD-IV specification shows that early PM leads to a causal decline in total healthcare spending of \$6,173 per patient over 24 months (SE = \$2,065; $p < 0.05$), corresponding to a 12% reduction relative to the pre-policy mean. This provides direct support for Hypothesis 1. There is also a significant reduction in utilization with early PM leading to 2.96 fewer medical service encounters (SE = 1.13; $p < 0.01$), which is a 23% decrease from the baseline average. This finding supports Hypothesis 2. Finally, consistent with these two results, there is also a significant reduction in the incidents of surgery, supporting Hypothesis 3. Specifically, we find that early PM access is associated with a 10.5% lower expected surgery counts per patient (Poisson coefficient = -0.11, p < 0.05).[4]

Table 2 Comparison of OLS and DiD-IV Estimates for Healthcare Costs, Utilization, and Surgeries

	Baseline			**DiD-IV**		
	Total Costs	Total Visits	Surgeries	Total Costs	Total Visits	Surgeries
Early PM	-7759.49^{**}	-4.52^{**}	-0.11^{**}	-6173.26^{**}	-2.96^{***}	-0.11^{**}
	(3417.30)	(1.72)	(0.06)	(2065.14)	(1.13)	(0.04)
Division FE	Yes					
Year FE	Yes					
Div × Demographic FE	Yes					
Controls	Yes					
Model	OLS	OLS	Poisson	OLS	OLS	Poisson
Adj./Pseudo R^2	0.16	0.15	0.23	0.31	0.26	0.22
Observations	72,160					

Notes: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Adjusted R^2 is reported for OLS models, and Pseudo R^2 for Poisson, both averaged across replications. Standard errors, bootstrapped with clustering (M=1,000), are reported in parentheses.

Among patients who undergo surgery, early access to PM significantly lowers the use of post-surgery resources (see Table 3), which supports our Hypothesis 5a. Even for those patients that have undergone surgery, those who had an early PM visit experienced reduced post-surgery costs of \$4,254.46 ($p < 0.05$), about 8%, and about 10% fewer follow-up visits ($p < 0.05$). Combined with the results on the number of surgeries in Table 2, these findings indicate that early PM intervention not only prevents surgeries, but also ensures that post-surgical costs are lower. Possible explanations include better patient preparation for surgery or improved targeting of appropriate surgical candidates. Importantly, this result suggests that early PM consultation is beneficial even if the patient requires surgery later on. This alleviates concerns that early consultations may delay surgery, leading to more expensive or intensive procedures later in the treatment pathway.

[4]With the log link, a coefficient in Poisson regression is a log incidence-rate ratio. Exponentiating $\beta = -0.11$ gives $\exp(-0.11) = 0.895$, which means that the expected surgery count is multiplied by 0.895, i.e., a reduction of $1 - 0.895 = 0.105$ (10.5%)

Table 3 Comparison of OLS and DiD-IV Estimates for Post-Surgery Costs and Visits

	Baseline		DiD-IV	
	Costs	Visits	Costs	Visits
Early PM	-5320.11^{**}	-0.09^{*}	-4254.46^{**}	-0.11^{**}
	(2115.42)	(0.048)	(1642.37)	(0.046)
Division FE		Yes		
Year FE		Yes		
Div × Demographic FE		Yes		
Controls		Yes		
Model	OLS	Poisson	OLS	Poisson
Adj./Pseudo R^2	0.22	0.23	0.21	0.23
Observations		1,119		

Notes: * $p<0.1$, ** $p<0.05$, *** $p<0.01$. Adjusted R^2 is reported for OLS models, and Pseudo R^2 for Poisson, both averaged across replications. Standard errors, bootstrapped with clustering (M=1,000), are reported in parentheses.

Finally, we consider the impact of early PM visits on opioid use. Table 4 presents a nuanced trade-off in opioid use. Early PM engagement increases the likelihood that a patient receives any opioid prescription (*the extensive margin*), yet among those treated, it lowers average dosage and the probability of high-dose therapy (*the intensive margin*) that is the cause of overdose and addiction. Specifically, we find that early PM leads to an 11.8% increase in the odds of any opioid use and an average dosage increase of 79.2 MME in the follow-up period. Among patients who do receive opioids, total dosage falls by roughly 150 MME and the likelihood of a high-dose (>50 MME/day) regimen drops by 2.8 points. Similar patterns have been documented when policy nudges broaden access but blunt escalation risk (Meara et al. 2016, Buchmueller and Carey 2018). Consequently, our findings provide mixed support for Hypothesis 4. This result also corresponds to our earlier result that early PM lowers surgery rates. Fewer operations potentially result in less postoperative pain, a major driver of high-dose opioid use, thus reinforcing the observed attenuation in intensive opioid exposure. In practice, PM specialists are willing to initiate low-dose opioids within a multimodal regimen; however, by substituting non-operative interventions (e.g., nerve blocks, image-guided injections, and behavioral therapies), they may avoid the progression to risky and high-dose opioids.

We note that the OLS coefficients exceed the DiD-IV estimations in magnitude for all main outcomes, implying that simple regression models may overstate the benefits of early PM access. This observation is consistent with *positive selection* into early PM adoption. It implies that patients who receive early PM, absent our instrument, are systematically different from those who do not in ways that correlate with better outcomes. There are several mechanisms that may drive this selection. First, patients who proactively seek specialist PM consultation may be more "healthcare savvy," i.e., more informed, motivated, and engaged in their own care. These attributes often translate into better adherence to complementary self-management strategies (e.g., PT, mindfulness) that independently

Table 4 Comparison of Models on Opioid Outcomes, by Estimation Method

	Baseline				DiD-IV			
	Total MME	Any Rx	MME | Rx	High-Dose†	Total MME	Any Rx	MME | Rx	High-Dose†
Early PM	93.6*	0.122***	−174.6***	−0.037**	79.2**	0.118***	−147.3***	−0.028**
	(54.0)	(0.009)	(35.4)	(0.016)	(34.3)	(0.007)	(30.6)	(0.013)
Division FE	Yes							
Year FE	Yes							
Div × Demo FE	Yes							
Controls	Yes							
Model	OLS	Logit	OLS	Logit	OLS	Logit	OLS	Logit
Adj./Pseudo R^2	0.30	0.24	0.28	0.22	0.32	0.26	0.29	0.25
Observations	72,160		14,823		72,160		14,823	

Notes: * $p<0.10$, ** $p<0.05$, *** $p<0.01$. Standard errors, bootstrapped with clustering (M=1,000), are reported in parentheses. The sample for 'MME | Rx' and 'High-Dose' outcomes is conditional on patients having at least one opioid prescription.

† High-dose is defined as prescriptions exceeding a daily dose threshold of 50 Morphine Milligram Equivalents (MME).

contribute to reduced downstream healthcare utilization. Second, structural advantages can play a significant role. Patients with better resource access, such as those residing in areas with a denser supply of specialists or those enrolled in superior insurance plans with well-coordinated care networks, may find it easier to obtain early PM referrals. These structural factors, not solely the PM specialist visit itself, are also correlated with fewer avoidable encounters and better health outcomes. Finally, provider network dynamics can also induce selection. Clinicians working within integrated or multidisciplinary pain centers might selectively refer patients they deem lower-risk, more adherent, or more likely to have a good outcome to their PM colleagues. This selective referral behavior would naturally amplify the apparent benefit of early PM in simple observational comparisons.

While it is known that IV estimators can be less precise than OLS, it is unlikely that the consistent attenuation we observe across multiple outcomes is merely a statistical artifact of increased noise. First, our IV estimates have smaller standard errors. Second, the systematic pattern of attenuation across distinct measures suggests the correction of a consistent underlying bias rather than random statistical noise. By leveraging a policy shock that exogenously shifts PM uptake for a broader and potentially less selected margin of patients, our IV approach addresses the confounding influence of this positive selection.

5.2. Robustness Checks

To validate our empirical strategy and the robustness of our empirical findings, we conduct a comprehensive set of robustness checks targeting the first stage, exclusion restriction checks on IV, second stage, control group definition, and external generalizability.

5.2.1. First-Stage Validation: Instrument Relevance and Specificity. The rigor of our instrumental variable hinges on its specific effect on early PM engagement only in the designated

treatment group and period. We validate this through a series of placebo and pre-trend analyses (see Sections EC.2.1 and EC.2.2 for details). First, we examine a timing placebo that relocates the "policy" to an earlier pre-policy (2016–2019 vs. 2012–2015) window (see Table EC.6). Second, we analyze a geographic placebo that applies the instrument to the never-treated (i.e., control) census divisions (see Table EC.7). We find no significant results in either of our placebo tests. Third, we investigate differential pre-trends in early PM adoption between treatment and control groups by interacting the South Atlantic indicator with year indicators for 2016 and 2017 (the pre-policy period), with 2018 serving as the omitted baseline year (see Table EC.9). We find no differential pre-trends in early PM adoption.

5.2.2. Exclusion Restriction: Pathway-Specific Identification. The exclusion restriction, although fundamentally untestable, is our most critical identifying assumption. We conduct two key analyses to strengthen its plausibility. First, we analyze a diagnostic placebo. We implement the same DiD-IV analysis on a cohort of patients diagnosed with dermatitis, a condition clinically unrelated to chronic pain and thus its care trajectory should not be affected by PM access. As shown in Table EC.8, the instrumented effect of early PM on all our outcome variables is statistically insignificant. This finding supports the argument that the instrument is not simply capturing spurious correlation with general healthcare-seeking behavior.

Second, we investigate the primary threat to the exclusion restriction, which is the potential direct therapeutic effect of the initial PT/C visits. We re-estimate our model using a subsample of patients who had no more than four PT/C sessions before a potential PM referral. We consider four or fewer PT/C visits as this limited exposure typically serves a diagnostic or triage function rather than a definitive, long-term therapeutic resolution for a chronic pain condition (see Section EC.2.3 for details). Therefore, this cohort is unlikely to experience direct PT/C benefits, which allows us to better isolate the PM referral pathway. We find that the instrument remains robust in this subsample, with 2SRI effects on total costs, visits, surgeries, and opioid outcomes closely aligned with those in the full population (see Table EC.10). The stability of our results in this subsample indicates that our findings are not primarily driven by patients who may have experienced complete pain resolution from the PT/C visits alone.

5.2.3. Second-Stage Robustness: Alternative Definitions and Follow-Up Windows. We confirm that our core second-stage results are robust to multiple modeling choices around our definition of early PM and follow-up period, as well as alternative parameterizations of the second-stage (see Section EC.2.4 for details). Varying the definition of early PM (e.g., 2, 4, or 6 months after diagnosis; before 4, 5, or 7 outpatient visits) consistently shows significant and negative effects on

cost, utilization, and opioid exposure. We also test different follow-up durations (12 and 18 months), and the results show that the observed benefits of early PM manifest relatively early and persist.

Furthermore, we assess robustness to our control function specification by re-estimating the second stage using both the generalized score and the Pearson residuals. In both specifications, the effects on cost, utilization, and opioid dependency remain negative and generally significant (see Tables EC.11, EC.12, and EC.13).

5.2.4. Control Group Composition As noted in our econometric strategy, our main control group ($\texttt{EarlyPM} = 0$) contains a mixture of "late PM" and "never PM" patients. To assess the robustness of our findings against a cleaner "never PM" counterfactual (i.e., to insure that our findings are no solely driven by a comparison against a potentially costly late PM pathway), we conduct a subgroup analysis focusing on clinical populations where late PM intervention is rare and its prevalence remained stable after the policy. Specifically, we focus on patients with ICD codes R10 (abdominal and pelvic pain) and M25 (joint disorders), where the prevalence of late PM intervention remains consistently low—around 0.3% and 0.4%, respectively—both before and after policy implementation.

As shown in Table EC.14, re-estimating our DiD-IV specification on this restricted sample yields robust and consistent effects. Early PM is associated with a reduction of approximately $5,910 in total costs ($p < 0.01$) and 2.92 fewer total visits ($p < 0.05$). The effect on surgery remains negative and marginally significant, despite the reduced statistical power in the smaller subsample. These results reinforce the robustness of our main findings and provide evidence that early PM access generates substantial benefits compared to a pathway with no specialized PM intervention at all.

5.3. External Validation

To assess the generalizability of our results, we replicate the core analysis with insurance-claims data from a chronic-pain specialist network in Southern California (2018–2023). This network comprises roughly 300 provider groups and more than 2,000 physicians, therapists, and other practitioners. The claims span eight key specialties, including neurology, orthopedics, rheumatology, pain management, neurosurgery, orthopedic surgery, diagnostic imaging, and anesthesiology. This allows for comparisons of diverse care pathways at the specialty level.

Due to the absence of a comparable policy shock in this dataset, we employ a coarsened-exact-match (CEM) approach to match patients on age, gender, initial pain diagnosis, insurance type, and specialty utilization patterns, i.e., the diversity of specialties visited during their first six encounters. This yields 4,432 matched patients for overall outcomes and 228 for the post-surgical subsample. Using this dataset, we are also able to include granular geographic information, as the dataset provides zipcode-level details on each patient. Additionally, the dataset records referral pathways, identifying

whether each patient first entered the network through a PCP, a physical therapist, or another specialist. We then estimate average treatment effects using OLS for continuous measures (e.g., costs) and Poisson models for count outcomes (e.g., visits and surgeries). Standard errors are clustered on the early PM indicator to allow for group-specific heteroskedasticity. Our dataset does not include information on prescriptions; therefore, outcomes on opioid prescriptions are not included in our analysis.

Table 5 External Validation: Matched Comparison of Early PM Patients (2018–2023) using CEM

Outcome	Matched Num.	Estimate	SE
Total Costs	4,432	$-2{,}503^{***}$	(687)
Total Visits	4,432	-3.21^{***}	(1.12)
Surgeries	4,432	-0.089^{***}	(0.030)
Post-Surgery Costs	228	$-3{,}978^{**}$	(1,605)
Post-Surgery Visits	228	-0.18^{*}	(0.095)

Notes: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$. Effects estimated by OLS or Poisson regression. Standard errors clustered on the early PM indicator.

The external validation results are reported in Table 5. As can be seen, the early PM patients incur about $2,500 lower total costs, attend over 3 fewer outpatient visits, and undergo around 9% fewer surgeries per patient. Among those receiving surgery, early PM access is associated with nearly $4,000 in post-operative savings and 16% (-0.18 for Poisson estimate) fewer follow-up visits.

The external validation estimations closely mirror our main findings from the Optum® CDM. In particular, the estimated effects of early PM engagement on key utilization and post-surgical metrics remain consistent with our main results (see Table 2 and Table 3). This overall consistency underscores the robustness of our core findings.

6. Mechanism Exploration

Our results show that early PM engagement leads to significant reductions in utilization, costs, and surgeries. In this section, we explore the operational and clinical pathways through which these efficiencies are realized. The guiding intuition is that a PM specialist can proactively redirect patients away from costly escalations (serial imaging, specialty handoffs, and costly procedures) and toward lower-intensity, high-value care. We test this by examining four candidate mechanisms: (1) substitution toward less invasive procedures, (2) optimization of diagnostic imaging, (3) shifts in the mix of interventional procedures, and (4) improved care coordination.

6.1. Surgical Modality

A primary driver of the observed cost savings is a clear substitution away from major surgical interventions. Table 6 reports how early PM changes the likelihood of two downstream surgical choices,

including a major operation (as defined in Table EC.2) versus a minimally invasive procedure.[5] Under a logit specification, we find that early PM intervention lowers the odds that patients undergo major surgery by 4.0% ($p < 0.01$) while increasing their odds of receiving a minimally invasive procedure by 4.6% ($p < 0.01$). Clinically, this results in a lower risk of complications and costs and a faster functional recovery for a significant share of the cohort. From an operational perspective, this represents a significant shift in the production of care, moving from high-cost, high-risk inpatient procedures toward lower-cost outpatient alternatives.

Table 6 Mechanism Analysis: Early PM Access and Surgical Modality (DiD–IV)

	Major Surgery	**Min. Invasive Procedures**
Early PM	-0.041^{***}	0.045^{***}
	(0.013)	(0.015)
Division FE	Yes	
Year FE	Yes	
Div × Demographic FE	Yes	
Controls	Yes	
Observations	72,160	
Model	Logit	

Notes: $^{*}\,p < 0.10$, $^{**}\,p < 0.05$, $^{***}\,p < 0.01$. Standard errors, bootstrapped with clustering (M=1,000), are reported in parentheses.

6.2. Diagnostic Imaging

Another mechanism through which early PM streamlines care is by optimizing the use of diagnostic imaging. As shown in Table 7, early PM significantly reduces subsequent imaging intensity both in volume and cost. Patients undergo 0.32 ($p < 0.01$) fewer imaging visits ($\approx$ 27 % relative to the counterfactual mean) and incur \$1,210 ($p < 0.01$) less in imaging charges over 24 months. A plausible clinical rationale is that for many chronic pain conditions, particularly non-specific low back pain, the results from the diagnostic imaging often do not dictate treatment choices (Maher et al. 2017). Recognizing this, PM specialists often rely on holistic treatments and targeted diagnostic injections rather than ordering additional scans that are unlikely to change the plan of care (Chou et al. 2009). This results in fewer appointments, lower radiation exposure, and reduced risk of unnecessary yet costly imaging procedures, which provides a direct patient benefit alongside financial savings (Lumbreras et al. 2010).

[5] Minimally invasive procedures, e.g., steroid injections and nerve blocks, are coded in CPT ranges 10004–10021 and 20100–29999 and do not meet the criteria for a major operation.

Table 7 Mechanism Analysis: Early PM Access and Diagnostic Imaging (DiD–IV)

	Imaging Visits	Imaging Costs
Early PM	-0.32^{***}	-1210.47^{***}
	(0.12)	(180.17)
Division FE	Yes	
Year FE	Yes	
Div × Demographic FE	Yes	
Controls	Yes	
Observations	72,160	
Model	Poisson	OLS

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors, bootstrapped with clustering (M=1,000), are reported in parentheses.

6.3. Care Coordination and Interventional Procedure Mix

Beyond simply substituting away from surgery, early PM engagement also appears to optimize the care pathway by promoting more efficient coordination and a more judicious mix of interventions.

First, we find evidence of improved care coordination. Table 8 quantifies how early PM consultation affects the breadth of care that patients subsequently navigate. Early PM intervention patients see 0.22 ($p < 0.05$) fewer specialties, with the reduction concentrated in nonsurgical specialties (-0.30, $p < 0.05$) and a smaller but significant drop in surgical specialties (–0.15, $p < 0.1$). Fewer specialties mean less duplication of in-take procedures, tests, and a lower risk of receiving conflicting advice.

Table 8 Early PM Effect on the Number of Distinct Specialties Visited After the First Six Encounters (DiD–IV)

	All	Surgical	Non-surgical
Early PM	-0.22^{**}	-0.15^{*}	-0.30^{**}
	(0.17)	(0.09)	(0.14)
Division FE	Yes		
Year FE	Yes		
Div × Demographic FE	Yes		
Controls	Yes		
Observations	53,153		
Model	OLS		

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The coefficients represent the marginal change in the *number of distinct specialties* a patient sees *after* the first six clinical encounters, conditional on early access to a pain-management specialist. Surgical specialties include orthopedic surgery, neurosurgery, spine surgery, etc.; Non-surgical specialties include neurology, physical medicine/rehab, rheumatology, psychiatry, etc. Standard errors, bootstrapped with clustering (M=1,000), are reported in parentheses.

Second, within the realm of interventional pain management itself, early PM promotes a shift toward more low-cost procedures. Table 9 examines the effect of early PM adoption on the number of interventional pain-management (IPM) procedures a patient receives, grouping those procedures

into low-, mid-, and high-cost tiers (See table EC.5 for detailed definition). Early PM intervention produces a clear substitution pattern, with patients receiving 0.57 additional low-cost procedures ($p < 0.01$), essentially no change in mid-cost procedures, and 0.13 fewer high-cost procedures ($p < 0.05$). Clinically, this means that PM specialists are adding more low-cost options, such as targeted steroid injections or radio-frequency ablation, while reducing expensive, invasive implants or device-based interventions.

Table 9 Early PM Effect on Counts of Interventional Pain-Management Procedures by Cost Tier (DiD–IV)

	Low-cost	**Mid-cost**	**High-cost**
Early PM	0.57^{***}	0.09	-0.13^{**}
	(0.18)	(0.11)	(0.05)
Division FE		Yes	
Year FE		Yes	
Div × Demographic FE		Yes	
Controls		Yes	
Observations		72,160	
Model		OLS	

Notes: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$. Standard errors, bootstrapped with clustering (M=1,000), are reported in parentheses.

Finally, a well-coordinated care plan often involves targeted rehabilitative services. Table 10 reports how early engagement of a PM specialist changes subsequent use of PT/C care, two common therapies, measured after the patient's first six clinical encounters. Early PM access leads to 1.25 additional PT sessions per patient ($p < 0.01$) but does not produce a statistically significant change in chiropractic visits (-0.48, n.s.). The increase in physical therapy aligns with clinical guidelines that prioritize active rehabilitation, supervised exercise, mobility training, and functional re-education, for nonspecific musculoskeletal pain.

7. Discussion and Conclusion

This study provides causal evidence that early PM engagement reduces healthcare utilization, costs, surgeries, and opioid reliance for patients with new-onset chronic pain. Using a novel difference-in-differences instrumental variable framework, we demonstrate that early PM is a high-value strategy for optimizing chronic pain care. These findings urge providers and payers to rethink referral pathways and benefit designs to prioritize timely PM specialist access, while offering policymakers a blueprint for reforming pain management practices to achieve cost savings and improve patient outcomes.

Table 10 Early PM Effect on Physical-Therapy and Chiropractic Sessions After the First Six Encounters (DiD–IV)

	Physical Therapy	Chiropractic
Early PM	1.25^{***}	-0.48
	(0.38)	(0.30)
Division FE	Yes	
Year FE	Yes	
Div × Demographic FE	Yes	
Controls	Yes	
Observations	53,153	
Model	OLS	

Notes: $^{*}\,p < 0.10$, $^{**}\,p < 0.05$, $^{***}\,p < 0.01$. Coefficients represent the marginal change in the cumulative number of sessions a patient receives *after* the first six clinical encounters, conditional on early PM specialist access. Standard errors, bootstrapped with clustering (M=1,000), are reported in parentheses.

7.1. Managerial and Operational Implications

From an operations management perspective, early engagement with PM specialists not only improves patient outcomes and streamlines care pathways, but also generates meaningful operational efficiencies across the healthcare system. By streamlining the patient care pathway and reducing unnecessary encounters, early PM engagement can reduce the strain on system resources. This includes reducing the use of expensive resources from high-intensity treatments like specialists and surgeons, thereby effectively serving as a gatekeeper (Shumsky and Pinker 2003, Freeman et al. 2021, Adida and Bravo 2023). A reduction in post-surgical visits compounds this effect, further improving efficiency. These benefits also extend to lower-intensity resources like primary care providers, as well as diagnostics, testing, and post-operative rehabilitation. While not as economically impactful, these shared resources affect treatment pathways for patients with other conditions and may result in positive spillover effects across the health system (Soltani et al. 2024, Kim et al. 2024).

For Accountable Care Organizations (ACOs) and others operating under global or capitated budgets, early PM engagement offers a valuable mechanism to reduce avoidable downstream use. Integrating PM specialists into care teams can also reduce fragmentation and improve care continuity, and thereby physician productivity (Kajaria-Montag et al. 2024). For payers, early PM referral can be a cost-effective lever to reduce unnecessary utilization and prevent escalation to high-cost interventions. Benefit design can be a primary tool to achieve this. Policies that waive co-payments or eliminate prior authorization requirements for initial PM consultations can directly incentivize patients and referring providers to select this care pathway (Möllenkamp et al. 2019). Beyond patient-facing incentives, payers can incorporate early PM benchmarks into provider contracts, such as shared-savings agreements with ACOs or performance incentives for primary care groups that meet targets for timely

PM referrals. Such contracts would provide direct financial support for providers to adopt what our study suggests is a more efficient care pathway.

However, implementing a shift toward earlier PM engagement would require increasing capacity to support access. In the short term, this may lead to additional congestion and delays in accessing pain management specialists, as well as increase the risk of provider burnout. Therefore, it is important to consider any shifts not only from a clinical perspective, but also from a capacity and resource planning one. This may necessitate investments in expanding PM fellowship training, integrating PM expertise into primary care teams, or leveraging technologies like e-consultations to manage demand efficiently. In the long run, the reduction in demand for downstream surgical and acute care services should enable a strategic reallocation of resources to meet the increased need for early assessment by PM specialists.

From a policy perspective, our findings have implications for both public health and quality measurement. Early PM lowers reliance on high-dose opioid regimens, which mitigates both clinical risk and the financial fallout of opioid-related complications. Given the role of prescription opioids as the root cause of the opioid crisis (Alpert et al. 2022) and the role of physicians (Bobroske et al. 2022) in creating excessive supply, the findings have significant implications for combating opioid abuse. Policymakers can also support this effort by including early PM referrals in quality measures by including "time to first PM consult" or "early PM referral rate" in programs like the Healthcare Effectiveness Data and Information Set (HEDIS). Finally, changes to clinical best practices would necessitate policymakers addressing specialist shortages. This may include expanding PM fellowship training or streamlining certification pathways in underserved areas.

7.2. Limitation and Future Research

While our quasi-experimental design provides a strong causal interpretation of our findings, there are a few limitations. First, the validity of our IV strategy depends on the exclusion restriction. We argue that the "zero-copay" policy for three PT/C visits affects downstream outcomes through its impact on the PM referral pathway for our sample of patients, and our robustness checks support this. However, we cannot definitively rule out that the direct therapeutic benefits of PT/C independently contribute some of the observed effect for the complier population. Thus, our estimates should be interpreted as the effect of early PM for the specific subpopulation whose treatment access was influenced by the policy.

Second, our study uses administrative claims data, which thoroughly captures utilization and costs. However, it lacks granular clinical details such as patient-reported pain scores, functional status, and other clinical data. Integrating the claims data with clinical data in future research would provide a more holistic and robust understanding of the impact of early PM on patient outcomes.

Third, our findings are based on commercially insured and Medicare Advantage patients. We expect that the findings are broadly applicable to other patient populations; however, there may be unobserved differences in other populations like Medicaid and Medicare that may lead to different responses. This offers an opportunity for future research into the heterogeneity of treatment effect across different patient subgroups that would improve the precision of patient guidelines.

References

Adelman RD, Greene JW, Moin T, Cunningham FW, Donovan RM (2013) Costs and quality implications of fragmented care for patients with multiple chronic conditions. *JAMA* 310(3):241–249.

Adida E, Bravo F (2023) Primary care first initiative: Impact on care delivery and outcomes. *Manufacturing & Service Operations Management* 25(4):1471–1488.

Agha L, Ericson KM, Geissler KH, Rebitzer JB (2022) Team relationships and performance: Evidence from healthcare referral networks. *Management Science* 68(5):3735–3754.

Ahuja V, Alvarez CA, Staats BR (2020) Maintaining continuity in service: An empirical examination of primary care physicians. *Manufacturing & Service Operations Management* 22(5):1088–1106.

Alam A, Gomes T, Zheng H, Mamdani M, Juurlink DN (2012) Long-term analgesic use after low-risk surgery: A retrospective cohort study. *Archives of Surgery* 147(8):887–894.

Alpert A, Evans WN, Lieber EM, Powell D (2022) Origins of the opioid crisis and its enduring impacts. *The Quarterly Journal of Economics* 137(2):1139–1179.

American Board of Pain Medicine (2023) Fourth annual survey of pain medicine specialists highlights. ABPM diplomates now number over 2,300.

Bobroske K, Freeman M, Huan L, Cattrell A, Scholtes S (2022) Curbing the opioid epidemic at its root: The effect of provider discordance after opioid initiation. *Management Science* 68(3):2003–2015.

Brummett CM, Waljee JF, Goesling J, Moser S, Lin P, Englesbe MJ, Bohnert A, Kheterpal S, Nallamothu BK (2017) New persistent opioid use after minor and major surgical procedures in us adults. *JAMA Surgery* 152(6):e170504.

Buchmueller TC, Carey C (2018) The effect of prescription drug monitoring programs on opioid utilization. *Journal of Health Economics* 56:229–243.

Centers for Medicare & Medicaid Services (2018) Medicare Program; Revisions to Payment Policies Under the Physician Fee Schedule and Other Revisions to Part B for CY 2019. Federal Register.

Chen C, Savva N (2018) Unintended consequences of hospital regulation: The case of the hospital readmissions reduction program. *Available at SSRN 3236983* .

Chou R, Fu R, Carrino JA, Deyo RA (2009) Imaging strategies for low-back pain: systematic review and meta-analysis. *The Lancet* 373(9662):463–472.

Chou R, Gordon DB, de Leon-Casasola OA, Rosenberg JM, Bickler S, Brennan T, Carter T, Cassidy CL, Chittenden E, Degenhardt E, Griffith S, Kristeller JL, Lally SE, McCarberg BH, Montgomery R, Murphy J, Perkal MF, Suresh S, Sluka KA, Strassels S (2016) Management of postoperative pain: A clinical practice guideline from the american pain society, the american society of regional anesthesia and pain medicine, and the american society of anesthesiologists. *Annals of Surgery* 264(2):282–293.

Chou R, Hashimoto R, Friedly J, Fu R, Bougatsos C, Dana T, Sullivan SD, Jarvik J (2015) Epidural corticosteroid injections for radiculopathy and spinal stenosis: a systematic review and meta-analysis. *Annals of internal medicine* 163(5):373–381.

Cohen SP, Vase L, Hooten WM (2021) Chronic pain: an update on burden, best practices, and new advances. *The Lancet* 397(10289):2082–2097.

Deyo RA, Martin BI, Kreuter W, Jarvik JG, Angier H, Mirza SK (2011) Revision surgery following operations for lumbar stenosis. *JBJS* 93(21):1979–1986.

Dowell D (2022) Cdc clinical practice guideline for prescribing opioids for pain—united states, 2022. *MMWR. Recommendations and reports* 71.

Dowell D, Haegerich TM, Chou R (2016) Cdc guideline for prescribing opioids for chronic pain — united states, 2016. *Morbidity and Mortality Weekly Report* 65(1):1–49, centers for Disease Control and Prevention.

Emery P, Breedveld F, Dougados M, Kalden J, Schiff M, Smolen J (2002) Early referral recommendation for newly diagnosed rheumatoid arthritis: evidence based development of a clinical guide. *Annals of the rheumatic diseases* 61(4):290–297.

Flynn TW, Smith B, Chou R (2011) Appropriate use of diagnostic imaging in low back pain: a reminder that unnecessary imaging may do as much harm as good. *journal of orthopaedic & sports physical therapy* 41(11):838–846.

Foster NE, Anema JR, Cherkin D, Chou R, Cohen SP, Gross DP, Ferreira PH, Fritz JM, Koes BW, Peul W, et al. (2018) Prevention and treatment of low back pain: evidence, challenges, and promising directions. *The Lancet* 391(10137):2368–2383.

Foster NE, Hill JC, OBrien S, Gste D, Dunn KM, Main CJ (2015) The start back approach to stratified care for low back pain: Complexity and impact. *Best Practice & Research Clinical Rheumatology* 29(1):9–13.

Freeman M, Robinson S, Scholtes S (2021) Gatekeeping, fast and slow: An empirical study of referral errors in the emergency department. *Management Science* 67(7):4209–4232.

Fritz JM, Brennan GP, Hunter SJ (2015) Physical therapy or advanced imaging as first management strategy following a new consultation for low back pain in primary care: associations with future health care utilization and charges. *Health services research* 50(6):1927–1940.

Fritz JM, Delitto A, Erhard RE (2012) Comparison of classification-based physical therapy and therapy-based classification of patients with low back pain: A randomized clinical trial. *Spine* 37(17):1501–1511.

Frogner BK, Harwood K, Andrilla CHA, Schwartz M, Pines JM (2018) Physical therapy as the first point of care to treat low back pain: an instrumental variables approach to estimate impact on opioid prescription, health care utilization, and costs. *Health services research* 53(6):4629–4646.

Gaskin DJ, Richard P (2012) The economic costs of pain in the united states. *The Journal of Pain* 13(8):715–724.

Harwood KJ, Pines JM, Andrilla CHA, Frogner BK (2022) Where to start? a two stage residual inclusion approach to estimating influence of the initial provider on health care utilization and costs for low back pain in the us. *BMC Health Services Research* 22(1):694.

Hayden JA, Van Tulder MW, Malmivaara AV, Koes BW (2005) Meta-analysis: exercise therapy for nonspecific low back pain. *Annals of Internal Medicine* 142(9):765–775.

Heatley SA (2009) Optimal referral is early referral. *Peritoneal dialysis international* 29(2_suppl):128–131.

Hill JC, Whitehurst DG, Lewis M, Bryan S, Dunn KM, Foster NE, Konstantinou K, Main CJ, Mason EE, Somerville S, Sowden G, Vohora K, Hay EM (2011) Comparison of stratified primary care management for low back pain: The start back randomized trial. *The Lancet* 378(9802):1560–1571.

Hollenbeck BK, Creanga AA, Khashan S, Wilson RA, Hollenbeck TK (2011) Episode-based cost of surgery for orthopaedic and urologic procedures. *Health Affairs* 30(6):1089–1098.

Hutchins TA, Peckham M, Shah LM, Parsons MS, Boulter DJ, Burns J, Cassidy RC, Davis MA, Holly LT, Hunt CH, Khan MA, Moritani T, Ortiz AO, O'Toole JE, Promes SB, Reitman C, Shah VN, Corey AS, on Neurological Imaging EP (2021) ACR Appropriateness Criteria® Low Back Pain: 2021 Update. *Journal of the American College of Radiology* 18(11S):S361–S379.

Institute of Medicine (2011) *Relieving Pain in America: A Blueprint for Transforming Prevention, Care, Education, and Research* (Washington, DC: National Academies Press), URL http://dx.doi.org/10.17226/13172.

Kajaria-Montag H, Freeman M, Scholtes S (2024) Continuity of care increases physician productivity in primary care. *Management Science* 70(11):7943–7960.

Kim SH, Zheng F, Brown J (2024) Identifying the bottleneck unit: Impact of congestion spillover in hospital inpatient unit network. *Management Science* 70(7):4200–4218.

Kjaer P, Kongsted A, Ris I, Abbott A, Rasmussen CDN, Roos EM, Skou ST, Andersen TE, Hartvigsen J (2018) Gla: D® back group-based patient education integrated with exercises to support self-management of back pain-development, theories and scientific evidence. *BMC musculoskeletal disorders* 19:1–21.

Lebecque P, Leonard A, De Boeck K, De Baets F, Malfroot A, Casimir G, Desager K, Godding V, Leal T (2009) Early referral to cystic fibrosis specialist centre impacts on respiratory outcome. *Journal of Cystic Fibrosis* 8(1):26–30.

Lumbreras B, Donat L, Hernández-Aguado I (2010) Incidental findings in imaging diagnostic tests: a systematic review. *The British Journal of Radiology* 83(988):276–289.

Maher C, Underwood M, Buchbinder R (2017) Non-specific low back pain. *The Lancet* 389(10070):736–747.

Manchikanti L, Falco F, Singh V, Pampati V, Parr AT, Benyamin RM, Fellows B, Hirsch JA (2012) Utilization of interventional techniques in managing chronic pain in the medicare population: analysis of growth patterns from 2000 to 2011. *Pain physician* 15(6):E969.

Manchikanti L, Singh V, Pampati V, Smith HS, Hirsch JA (2009) Analysis of growth of interventional techniques in managing chronic pain in the medicare population: A 10-year evaluation from 1997 to 2006. *Pain Physician* 12(1):9–34, ISSN 1533-3159.

Mannion AF, Brox JI, Fairbank JC (2013) Comparison of spinal fusion and nonoperative treatment in patients with chronic low back pain: long-term follow-up of three randomized controlled trials. *The spine journal* 13(11):1438–1448.

Martin BI, Mirza SK, Spina N, Spiker WR, Lawrence B, Brodke DS (2019) Trends in lumbar fusion procedure rates and associated hospital costs for degenerative spinal diseases in the united states, 2004 to 2015. *Spine* 44(5):369–376.

Martin BI, Mirza SK, Spina NL, Spence P, Lawrence BJ, Brodke D, Turner J (2011) Repeat surgery after lumbar fusion: The cost and complication burden. *Spine* 36(24):2021–2029.

Meara E, Horwitz JR, Powell D, McClelland R, Zhou W, O'Malley AJ (2016) State legal restrictions and prescription-opioid use among disabled adults. *New England Journal of Medicine* 375(1):44–53.

Mirza SK, Deyo RA (2007) Systematic review of randomized trials comparing lumbar fusion surgery to nonoperative care for treatment of chronic back pain. *Spine* 32(7):816–823.

Möllenkamp M, Zeppernick M, Schreyögg J (2019) The effectiveness of nudges in improving the self-management of patients with chronic diseases: a systematic literature review. *Health Policy* 123(12):1199–1209.

Murphy SL, Kochanek KD, Xu J, Arias E (2020) Mortality in the united states, 2019. NCHS Data Brief 395, National Center for Health Statistics, Centers for Disease Control and Prevention, Hyattsville, MD, accessed 16 June 2025.

National Center for Health Statistics (2024) Chronic pain and high-impact chronic pain in u.s. adults, 2023. Data Brief 518, 24.3% of adults had chronic pain; 8.5% high-impact chronic pain.

O'Connell NE, Cook CE, Wand BM, Ward SP (2016) Clinical guidelines for low back pain: a critical review of consensus and inconsistencies across three major guidelines. *Best practice & research Clinical rheumatology* 30(6):968–980.

Optum (2022) Optum's de-identified Clinformatics® Data Mart Database (Optum® CDM). Statistically de-identified administrative health claims data provided by Optum, URL https://www.optum.com/solutions/life-sciences/real-world-data/clinformatics-data-mart.html, data extracted for the period of January 1, 2007, through July 31, 2022. Includes longitudinal medical and pharmacy claims for commercially insured and Medicare Advantage beneficiaries across the United States. Eden Prairie, MN.

Shumsky RA, Pinker EJ (2003) Gatekeepers and referrals in services. *Management Science* 49(7):839–856.

Smart NA, Dieberg G, Ladhani M, Titus T (2014) Early referral to specialist nephrology services for preventing the progression to end-stage kidney disease. *Cochrane Database of Systematic Reviews* (6).

Soar H, Comer C, Wilby M, Baranidharan G (2022) Lumbar radicular pain. *BJA education* 22(9):343–349.

Soltani M, Batt RJ, Bavafa H (2024) Quality improvement spillovers: Evidence from the hospital readmissions reduction program. *Management Science* .

Srinivas SV, Deyo RA, Berger ZD (2012) Application of "less is more" to low back pain. *Archives of internal medicine* 172(13):1016–1020.

Staats PS, Patel K, Gillen EM, Bello T, Epple T, Bilder SM, Wisor D (2022) Healthcare utilization and costs associated with new-onset pain in a medicare population. *Pain Physician* 25(9):E1415–E1422, ISSN 1533-3159.

Staats PS, Vallejo R, Gasquet NC, Ricker CN (2023) Effect of provider specialty on medical resource utilization and costs in chronic spinal-pain management. *Pain Physician* 26(2):207.

Terza JV, Basu A, Rathouz PJ (2008) Two-stage residual inclusion estimation: addressing endogeneity in health econometric modeling. *Journal of health economics* 27(3):531–543.

Treede RD, Rief W, Barke A, Aziz Q, Bennett MI, Benoliel R, Cohen M, Evers S, Finnerup NB, First MB, et al. (2015) A classification of chronic pain for icd-11. *Pain* 156(6):1003–1007.

Turk DC (2002) Clinical effectiveness and cost-effectiveness of treatments for patients with chronic pain. *The Clinical journal of pain* 18(6):355–365.

Turk DC, Wilson HD, Cahana A (2011) Treatment of chronic non-cancer pain. *The Lancet* 377(9784):2226–2235.

US HHS (2019) Pain management best practices: Updates, gaps, inconsistencies, and recommendations. Technical report, U.S. Department of Health and Human Services, Washington, DC, URL https://www.hhs.gov/sites/default/files/pmtf-final-report-2019-05-23.pdf.

Von Korff M, Moore JC (2001) Stepped care for back pain: activating approaches for primary care. *Annals of internal medicine* 134(9_Part_2):911–917.

Wanderer JP, Ehrenfeld JM, Terekhov M, Philbrick M, Stanley BT, Lin L, Huang Y, Dy C (2018) Implementation of an enhanced recovery after surgery (eras) protocol for major spine surgery: a prospective observational study. *Anesthesia & Analgesia* 126(1):197–205.

Webster BS, Bauer AZ, Choi Y, Cifuentes M, Pransky GS (2013) Iatrogenic consequences of early magnetic resonance imaging in acute, work-related, disabling low back pain. *Spine* 38(22):1939–1946.

Weinstein JN, Lurie JD, Olson PR, Bronner KK, Fisher ES (2006) United states' trends and regional variations in lumbar spine surgery: 1992–2003. *Spine* 31(23):2707–2714.

Wooldridge JM (2015) Control function methods in applied econometrics. *Journal of Human Resources* 50(2):420–445.

Woolf CJ, Bennett GJ, Doherty M, Dubner R, Kidd B, Koltzenburg M, Lipton R, Loeser JD, Payne R, Torebjork E (1998) Towards a mechanism-based classification of pain?

Zajacova A, Grol-Prokopczyk H (2024) More americans have chronic pain than ever before. *Pain News Network (preprint)* .

Electronic Companion

EC.1. Supplemental Tables and Figures

Table EC.1 Descriptions of Pain-Related ICD-10 Codes

ICD-10 Code	Description
G Codes – Diseases of the Nervous System	
G43	Migraine
G44	Other headache syndromes
G50	Disorders of trigeminal nerve
G51	Disorders of facial nerve
G52	Disorders of other cranial nerves
G54	Nerve root and plexus disorders
G56	Mononeuropathies of upper limb
G57	Mononeuropathies of lower limb
G58	Other mononeuropathies
G59	Mononeuropathy in diseases classified elsewhere
G89	Pain, not elsewhere classified
M Codes – Diseases of the Musculoskeletal System and Connective Tissue	
M06	Other rheumatoid arthritis
M07	Psoriatic and enteropathic arthropathies
M13	Other arthritis
M15	Polyosteoarthritis
M16	Osteoarthritis of hip
M19	Other and unspecified osteoarthritis
M25	Other joint disorders, not elsewhere classified
M30	Polyarteritis nodosa and related conditions
M31	Other necrotizing vasculopathies
M32	Systemic lupus erythematosus
M33	Dermatopolymyositis
M35	Other systemic involvement of connective tissue
M40	Kyphosis and lordosis
M41	Scoliosis
M43	Other deforming dorsopathies
M46	Other inflammatory spondylopathies
M47	Spondylosis
M48	Other spondylopathies
M50	Cervical disc disorders
M51	Other intervertebral disc disorders
M53	Other dorsopathies, not elsewhere classified
M54	Dorsalgia
M60	Myositis
M62	Other disorders of muscle
M65	Synovitis and tenosynovitis
M70	Soft tissue disorders related to use, overuse and pressure
M72	Fibroblastic disorders
M79	Other soft tissue disorders, not elsewhere classified
R Codes – Symptoms, Signs and Abnormal Findings	
R07	Pain in throat and chest
R10	Abdominal and pelvic pain
R29	Other symptoms and signs involving the nervous and musculoskeletal systems
R52	Pain, unspecified
S Codes – Injury and Certain Other Consequences of External Causes	
S42	Fracture of shoulder and upper arm
S52	Fracture of forearm
S62	Fracture at wrist and hand level
S82	Fracture of lower leg, including ankle
S83	Dislocation and sprain of joints and ligaments of knee
S92	Fracture of foot and toe, except ankle

Table EC.2 Pain-Related Major Operations CPT-10 Codes and Descriptions

CPT-10	Description
Hip Arthroplasty	
27130	Total hip arthroplasty
27134	Femoral neck fracture fixation/replacement
Knee Arthroplasty / Revision	
27446	Unicompartmental knee arthroplasty
27447	Total knee arthroplasty
27486	Revision—tibial component only
27487	Revision—tibial & femoral components
Femur Fracture Fixation	
27236	Proximal femur fracture fixation
27244	Femoral shaft fracture fixation
27245	Femoral shaft fixation (nail/plate)
27506	Distal femur fracture fixation
27535	Supracondylar femur fracture fixation
Ankle / Foot Fixation & Replacement	
27759	Medial malleolus fixation
27814	Trimalleolar ankle fixation
27822	Distal tibia (pilon) fixation
27828	Pilon fixation & ankle fusion
27829	Pilon fixation with external fixator
27792	Total ankle replacement
Distal Radius / Ulna Fixation	
24685	Distal radius fracture fixation
25607	Distal radius OR ulna fixation
25608	Distal radius & ulna fixation (plate)
25609	Distal radius/ulna fracture fixation
Humerus & Upper Extremity Fixation	
23615	Proximal humerus fixation
24515	Proximal humerus hemiarthroplasty
25575	Midshaft humerus fracture fixation
Spine Decompression / Laminectomy	
63015	Lumbar laminectomy (single segment)
63030	L1–L2 laminectomy
63045	Two-level laminectomy
63047	Single-segment decompression
63048	Each additional decompression segment
63081	Lumbar foraminotomy (each segment)
63082	Additional lumbar foraminotomy
61312	Lumbar microdiscectomy
61313	Each additional lumbar microdiscectomy
62223	Lumbar nerve root injection
Spine Fusion / Arthrodesis	
22551	Anterior cervical fusion (single level)
22554	Anterior cervical fusion with interbody
22558	Posterior cervical fusion (single level)
22600	Posterior lumbar fusion (single level)
22610	Posterior lumbar fusion with instrumentation
22612	Combined anterior-posterior lumbar fusion
22633	Posterior lumbar interbody fusion
22856	Cervical disc replacement
22224	Lumbar osteotomy & fusion
Cranial Nerve / Head & Neck Pain	
61107	Trigeminal nerve decompression
61120	Gasserian ganglion neurolysis
61210	Glossopharyngeal nerve lesion excision
61510	Glossopharyngeal nerve ablation
61521	Sphenopalatine ganglion rhizotomy
Peripheral Nerve Ablation / Pain Procedures	
61458	Stereotactic radiosurgery (trigeminal)
61618	Genicular nerve ablation

Table EC.3 Control Variables Included in the Study

Patient-Level Controls	
ICD-10 diagnosis codes	Clinical status at baseline
Early-treatment specialties	Variance of specialty consulted within first 6 visits
Age	Age in years at index diagnosis
Gender	Patient's self-reported gender
Race	Race/ethnicity categories
Household Children	Number of Children in patient's household
Household Adults	Number of Adults in patient's household
Income	Household income level
Education	Highest educational attainment
Occupation Type	Occupation type indicator
Fixed effect Controls	
Year of index diagnosis	Dummy variables for calendar year of first pain diagnosis
Division fixed effects	Capture time-invariant geographic differences
Division × Demographics interactions	Allow demographic effects to vary by division

Table EC.4 Comparison of Pre-Policy Healthcare Utilization and Costs: South Atlantic vs. Control Divisions (Standard deviations shown in parentheses)

Metric	South Atlantic	Control Divisions	p-value (t-test)
Early PM Adoption			
Number of Patients (N)	261	404	
Adoption Rate (%)	1.2%	1.1%	
Overall Sample			
Number of Patients (N)	20,551	36,110	
Total Cost ($)	50,715.15	48,769.68	0.91
	(61,933.26)	(69,078.64)	
Total Number of Visits	13.02	12.77	0.10
	(17.80)	(16.54)	
Number of Surgeries	0.026	0.025	0.47
	(0.18)	(0.17)	
Patients Prescribed with Opioid Drugs			
Number of Patients (N)	4,135	7,586	
Percentage (%)	20.1%	21.0%	0.22
High-Dose Patients (≥50 MME)	1,973	3,671	
Percentage (%)	9.6%	10.1%	0.18
Total MME	1,120.4	1,175.6	0.22
	(1,332.7)	(1,408.3)	
MME | Rx	5,574.3	5,627.9	0.20
	(3,012.5)	(2,978.4)	
Among Surgery Patients			
Number of Patients (N)	474	798	
Post-Surgery Visits	6.20	5.89	0.45
	(6.14)	(6.15)	
Post-Surgery Cost ($)	50,187	48,051	0.60
	(56,755)	(81,187)	

Table EC.5 PM Procedure Cost Tiers and Representative CPT Codes (U.S. 2024)

Cost Tier	Examples of Procedures	Principal CPT Codes
1. Low-cost: needle-based injections and blocks (<$2,000)		
	Trigger-point injections; epidural steroid injections; facet or medial-branch blocks; SI-joint injection; sympathetic or peripheral nerve blocks; radio-frequency ablation	20552, 20553, 62321, 62323, 64479--84, 64490--95, 27096, 64510, 64520, 64530, 64517, 64405, 64415--17, 64430, 64450, 64625, 64633--36
2. Mid-cost: percutaneous & biologic ($2,000–$10,000)		
	Platelet-rich plasma or stem-cell injections; percutaneous disc decompression; intradiscal thermal annuloplasty; endoscopic discectomy; outpatient vertebroplasty; MILD (minimally invasive lumbar decompression)	0232T, 0263T, 20999, 62287, 22526, 22527, 62380, 22511, 22512, 0274T
3. High-cost: implantable / surgical (≥$10,000)		
	Spinal-cord or DRG stimulators; peripheral nerve stimulators; intrathecal drug-delivery pumps; balloon kyphoplasty; vertebral body stenting; spinal fusion; laminectomy or multi-level decompression	63650, 63655, 63685, 63688, 63663--66, 64555, 64590, 64595, 62362, 62365, 62368, 62370, 22513--15, 0200T, 22612, 22630, 22633, 22842--43, 63030, 63047, 63056

Notes: CPT codes listed are the ones most commonly billed for each intervention in U.S. practice; some payers require add-on codes (e.g., fluoroscopic guidance 77003). Dollar bands reflect typical 2024 cash or Medicare-allowable prices; local contracts vary.

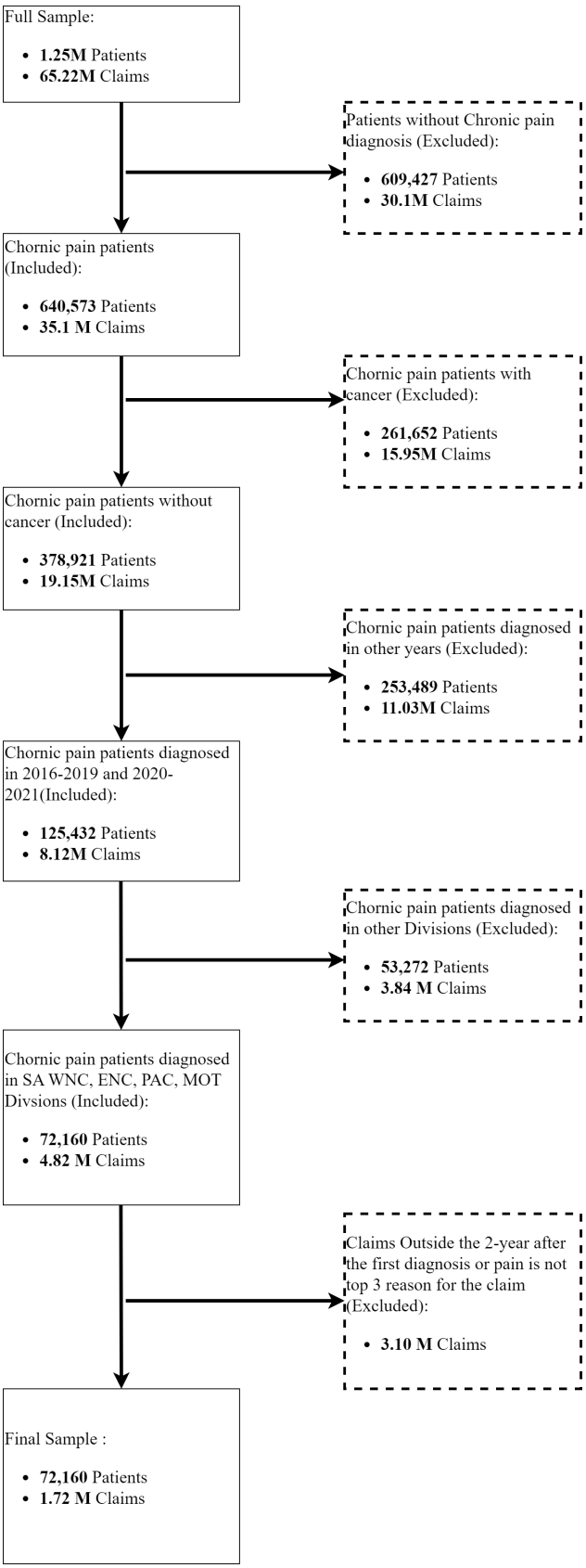

Figure EC.1 Sample Construction

EC.2. Robustness Checks

In this section, we perform comprehensive sets of robustness checks to validate our empirical approach. Sections EC.2.1 and EC.2.2 assess the *first-stage* instrument via timing and geographic placebo tests and a pre-policy trends analysis, while Section EC.2.3 examines the exclusion restriction by focusing on low PT/C users. Finally, Section EC.2.4 evaluates the stability of our *second-stage* 2SRI estimates under alternative definitions of early PM access and varying follow-up periods.

EC.2.1. Placebo Tests

To validate our identification strategy and rule out false correlations, we implement three placebo exercises: (i) a timing placebo that artificially reassigns the policy date within the pre-policy window, (ii) a geographic placebo that applies our first-stage instrument to divisions that have not implemented the policy, and (iii) an outcome placebo using a clinically unrelated cohort. In each case, we find no significant treatment effect, validating our main results.

Timing Placebo. We perform a timing placebo analysis by restricting the sample to the pre-policy period (2012–2019) and artificially redefining the "post" period to be in 2016–2019 and "pre" period to be in 2012–2015, while keeping the treatment division unchanged (SA division). Since no policy change occurred during this period, any observed treatment effect would likely reflect a false correlation rather than a causal relationship. Consistent with expectations, the first-stage association between the instrument and the treatment indicator ($\texttt{EarlyPM} = 1$) is statistically insignificant, supporting the validity of our identification strategy.

Table EC.6 Estimated `EarlyPM` Engagement with Redefined Pre- and Post-Period

Variable	Estimate (SE)
SA × Post	−0.064 (0.165)
p-value	0.695
Division FE	Yes
Year FE	Yes
Div × Demographic FE	Yes
Controls	Yes
Pseudo R^2	0.24
Observations	94,561
Log-Likelihood Ratio p-value	0

Notes: Coefficient estimated via logistic regression. Standard errors, bootstrapped with clustering (M=1,000), are reported in parentheses.

Geographic Placebo. As a placebo test on geographic assignment, we replicate our analysis using only data from the four control divisions (Pacific, Mountain, West North Central (WNC), and East North Central (ENC)), where the policy was not implemented. Since these divisions were unaffected by the policy, any significant treatment effect would suggest spurious correlation or confounding. We

estimate four models where each division, in turn, is designated as the treatment division, and the remaining three divisions are designated as controls. Consistent with the identifying assumptions, we find no significant effect of the instrument on the treatment indicator ($\texttt{EarlyPM} = 1$) in these control divisions.

Table EC.7 Estimated `EarlyPM` Engagement in Non-Treated Divisions

Variable	Pacific	Mountain	WNC	ENC
SA × Post	−0.44	0.17	−0.39	0.15
	(0.35)	(0.25)	(0.43)	(0.28)
p-value	0.210	0.490	0.370	0.600
Division FE		Yes		
Year FE		Yes		
Div × Demographic FE		Yes		
Controls		Yes		
Pseudo R^2	0.17	0.15	0.18	0.16
Observations		72,160		

Notes: Coefficients estimated via logistic regression for four non-treated Census divisions. Standard errors, bootstrapped with clustering (M=1,000), are reported in parentheses.

Outcome Placebo. Using a cohort of patients diagnosed with dermatitis, a condition unrelated to chronic pain management, we replicate the full DiD–IV procedure to predict $\hat{\Pr}(\texttt{EarlyPM}_i)$. We then regress total costs and visit counts on this placebo treatment probability. Table EC.8 reports coefficients of 653.11 (SE = 11,148.26) for costs and 1.02 (SE = 3.48) for visits, neither of which is statistically significant, indicating that our DiD-IV framework does not produce false positives in an irrelevant patient population in the targeted divisions during the policy period.

Table EC.8 Estimated `EarlyPM` on Dermatitis Treatment Outcomes

Outcome	Total Costs	Total Visits
Early PM	653.11	1.02
	(1,148.26)	(3.48)
Division FE	Yes	
Year FE	Yes	
Div × Demographic FE	Yes	
Controls	Yes	
Model	OLS	
Adjusted R^2	0.27	0.22
Observations	26,984	

Notes: Placebo test regresses early PM engagement (`EarlyPM`) along with residuals on dermatitis-related outcomes. Coefficients estimated using OLS with standard errors in parentheses, clustered at the division level and bootstrapped ($M = 1{,}000$).

EC.2.2. Pre-Policy Parallel Trends

We formally test for differential pre-trends in early PM adoption between treatment and control groups by interacting the South Atlantic indicator with year indicators for 2016 and 2017 (the pre-policy period), with 2018 serving as the omitted baseline year. As shown in Table EC.9, the interaction terms for 2016 and 2017 are small and statistically insignificant. This indicates that the pre-policy trend in early PM adoption in South Atlantic was not significantly different from that in the control group (relative to the 2018 baseline). These results support the assumption of parallel trends that underpins our DiD design.

Table EC.9 Test for Pre-Policy Parallel Trends in Early PM Adoption

Interaction Term	Estimate (SE)
2016 × `SA`	0.04
	(0.13)
2017 × `SA`	−0.011
	(0.10)
Division FE	Yes
Year FE	Yes
Div × Demographic FE	Yes
Controls	Yes
Model	Logit
Pseudo R^2	0.26
Observations	56,661

Notes: Estimates from a logistic regression of early PM adoption on year×`SA` division interaction terms, using pre-policy years only. Standard errors, clustered at the division level and bootstrapped ($M = 1{,}000$), are reported in parentheses.

EC.2.3. Addressing Exclusion Restriction Concerns

To investigate the validity of our instrument's exclusion restriction, we restrict attention to the 6,371 patients with fewer than three PT/C sessions prior to a potential PM referral. By focusing on a cohort that minimally engages with PT/C, we reduce the risk that these PT/C sessions influence downstream utilization rather than solely through early PM engagement.

Table EC.10 presents the first- and second-stage results for this low-PT/C subsample. The first-stage logistic regression remains strong (instrument coefficient = 0.15, $p < 0.05$). In the second-stage 2SRI results, early PM engagement indicates significant effects, closely aligned with our main estimates. Specifically, early PM reduces total healthcare costs by \$5,265 ($p < 0.05$), total visits decrease by 3.16 visits ($p < 0.01$), and surgery rates decline by around 11% (Poisson Coefficient = 0.114, $p < 0.10$). Among the 230 patients who undergo surgery, early PM is associated with \$5,647 in post-operative savings ($p < 0.05$) and 0.23 fewer follow-up visits ($p < 0.05$). The persistence of these magnitudes in a cohort with minimal PT/C engagement reinforces that the policy instrument operates primarily through early PM access, not through differential use of PT/C sessions.

Table EC.10 First- and Second-Stage Regression Results on Low PT/C Users

First-Stage Logistic Regression Predicting `EarlyPM` = 1		
Variable	Estimate (SE)	
SA × Post	0.15^{**} (0.07)	
Pseudo R^2	0.28	
Log-Likelihood Ratio p-value	0.000	
Observations	6,371	
Second-Stage 2SRI Cost & Utilization Results		
Outcome	Estimate (SE)	Model
Total Costs	$-5{,}265^{**}$ (2,234)	OLS, Adj. $R^2 = 0.28$
Total Visits	-3.16^{***} (1.21)	OLS, Adj. $R^2 = 0.31$
Surgeries	-0.114^{*} (0.06)	Poisson, Adj. $R^2 = 0.20$
Observations	6,371	
Second-Stage 2SRI Post-Surgery Results		
Post-Surgery Costs	$-5{,}647^{**}$ (2,418)	OLS, Adj. $R^2 = 0.26$
Post-Surgery Visits	-0.23^{**} (0.09)	Poisson, Adj. $R^2 = 0.25$
Observations	230	
Second-Stage 2SRI Opioid Outcomes		
Total MME	138.6^{**} (61.53)	OLS, Adj. $R^2 = 0.30$
Any Rx	0.282^{**} (0.126)	Logit, Pseudo $R^2 = 0.24$
MME | Rx	-113.1^{**} (48.61)	OLS, Adj. $R^2 = 0.28$
High-Dose	-0.093^{**} (0.042)	Logit, Pseudo $R^2 = 0.21$
Observations	2,108	

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. All models include Division FE, Year FE, Div×Demographic FE, and covariates included. Two-Stage Residual Inclusion (2SRI) IV estimation is used. Standard errors, clustered at the division level and bootstrapped ($M = 1{,}000$), are reported in parentheses.

EC.2.4. Alternative Specifications

In this section, we demonstrate that our core findings are not sensitive to modeling choices in two dimensions: (i) how "early" PM access is defined, and (ii) the length of the follow-up window. We show that the estimated cost and utilization effects persist when we vary both the timing threshold for PM referral and the observation horizon.

EC.2.4.1. Definition of Early Access. To assess the sensitivity of our findings to how we define "early" PM access, we re-estimate the DiD–IV models under six alternative definitions: using fixed time periods defined as a PM visit within 2, 4, or 6 months of diagnosis, and fixed number of visits defined as before the 4, 5, or 7 outpatient visits (see Table EC.11). Across all specifications, the estimated effects on total costs, surgery rates, and post-surgical outcomes remain consistently negative and, in most cases, statistically significant at the 5% level or higher.

While the point estimates for total visits are attenuated under the 4-visit threshold, they achieve significance once the cutoff is raised to 5 or 7 visits. These results confirm that our core conclusions are not driven by any single arbitrarily chosen definition of early PM access.

Table EC.11 Robustness Checks on Alternative Definitions of Early PM Access

	Specification					
	2 Mo.	4 Mo.	6 Mo.	4 Visits	5 Visits	7 Visits
Total Costs	$-5{,}052.16^{**}$ (2,404.15)	$-5{,}897.54^{**}$ (2,479.38)	$-6{,}314.63^{**}$ (2,637.07)	$-3{,}644.17^{*}$ (1,896.24)	$-5{,}766.62^{***}$ (1,580.21)	$-5{,}858.54^{***}$ (1,656.78)
Total Visits	-2.10^{*} (1.16)	-2.73^{**} (1.23)	-2.69^{*} (1.95)	-1.81 (1.23)	-2.67^{**} (1.06)	-2.74^{**} (1.08)
Surgeries	-0.073^{**} (0.03)	-0.096^{**} (0.04)	-0.085^{**} (0.04)	-0.065^{**} (0.03)	-0.095^{**} (0.04)	-0.093^{**} (0.04)
Post–Surgery Costs	$-3{,}066.87^{**}$ (1,359.13)	$-3{,}611.91^{**}$ (1,529.83)	$-3{,}846.64^{**}$ (1,614.06)	$-2{,}369.51$ (1,858.32)	$-4{,}010.21^{**}$ (1,664.53)	$-3{,}840.37^{**}$ (1,728.41)
Post–Surgery Visits	-0.070^{*} (0.038)	-0.084^{**} (0.042)	-0.099^{**} (0.044)	-0.056 (0.053)	-0.101^{**} (0.044)	-0.100^{**} (0.049)
Total MME	116.1^{**} (51.3)	124.8^{**} (57.0)	134.7^{**} (61.5)	82.5^{*} (44.4)	132.3^{***} (45.6)	128.4^{***} (47.1)
Any Rx	0.072^{**} (0.031)	0.083^{**} (0.036)	0.089^{**} (0.039)	0.051^{*} (0.028)	0.088^{***} (0.030)	0.086^{***} (0.030)
MME | Rx	-103.2^{**} (44.4)	-113.4^{**} (48.9)	-121.5^{**} (52.5)	-69.3^{*} (36.9)	-117.6^{***} (37.8)	-116.4^{***} (39.0)
High-Dose	-0.025^{**} (0.011)	-0.029^{**} (0.013)	-0.031^{**} (0.014)	-0.017^{*} (0.010)	-0.030^{***} (0.011)	-0.029^{***} (0.011)
Division FE	Yes					
Year FE	Yes					
Div×Demographic FE	Yes					
Controls	Yes					

Notes: * $p<0.10$, ** $p<0.05$, *** $p<0.01$. All models are estimated via DiD–IV 2SRI procedures with clustered standard errors bootstrapped ($M=1{,}000$).

EC.2.4.2. Follow-up Period. We further examine the robustness of the estimated effects by varying the follow-up window to 12 and 18 months post-diagnosis (see Table EC.12). Even with the shorter 12-month horizon, early PM access yields savings of \$2,532 ($p<0.05$) and reductions of 1.6 visits ($p<0.05$) and 8% surgeries per patient ($p<0.05$). Extending the window to 18 months amplifies these effects, with cost savings increasing to \$3,398 ($p<0.05$) and surgeries increasing to around 9% ($p<0.01$). Post-surgical cost and visit reductions follow the same pattern, demonstrating that the benefits of early PM access accrue steadily over time.

EC.2.4.3. Residual Specification. We further test the robustness of the residual specification used in the second stage. In addition to the raw residual proposed in Terza et al. (2008), recent work has shown that score-based residuals, including generalized residuals and Pearson residuals (Wooldridge 2015), may provide more robust estimations. The generalized and Pearson residuals are closely related. The former is obtained by rescaling the residual by the model-implied variance of the predicted treatment probability, while the latter is rescaled by the standard deviation. As shown in

Table EC.12 Robustness Checks by Follow-Up Period (DiD–IV)

	Follow-Up Period	
	12 Months	18 Months
Total Costs	$-2{,}532.21^{**}$ (1,017.45)	$-3{,}397.60^{**}$ (1,493.73)
Total Visits	-1.58^{**} (1.29)	-2.59^{**} (1.47)
Surgeries	-0.083^{**} (0.031)	-0.095^{***} (0.041)
Post-Surgery Costs	$-4{,}251.89^{**}$ (1,524.67)	$-5{,}857.64^{**}$ (2,010.88)
Post-Surgery Visits	-0.102^{**} (0.051)	-0.152^{**} (0.061)
Total MME	-112.2^{**} (52.8)	-125.7^{**} (57.9)
Any Rx	-0.081^{**} (0.035)	-0.093^{***} (0.031)
MME | Rx	-99.3^{**} (43.5)	-115.5^{**} (48.6)
High-Dose	-0.027^{**} (0.012)	-0.033^{***} (0.011)
Division FE	Yes	
Year FE	Yes	
Div × Demographic FE	Yes	
Controls	Yes	

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. All models are estimated via DiD–IV 2SRI procedures with clustered standard errors bootstrapped ($M = 1{,}000$).

Table EC.13, the results across these alternative residual specifications remain consistent with our main findings, further supporting the robustness of our estimates.

EC.2.5. Control Group Composition

To assess the robustness of our findings in settings where the control group includes different mixtures of "never PM" and "late PM" patients, we conduct a subgroup analysis focusing on clinical populations in which late PM intervention is rare. In our main specification, the control group may include patients who received no PM intervention at all, as well as those who received PM at a later stage. The relative proportion of these two subgroups could affect the estimated treatment effect of early PM access. To evaluate whether our findings are robust to this mixture, we examine diagnoses where late PM is infrequent, and its prevalence remains stable before and after the policy.

Specifically, we focus on patients with ICD codes R10 (abdominal and pelvic pain) and M25 (joint disorders). Among these patients, the share receiving late PM intervention remains low and stable, with late PM patients having less than 0.5% share both before and after the policy change. For R10,

Table EC.13 Robustness Check with Pearson & Generalized Residual in 2SRI

	Residual Type	
	Pearson	Generalized
Total Costs	$-5{,}310.45^{***}$ (1,605.43)	$-5{,}375.28^{***}$ (1,560.12)
Total Visits	-2.29^{**} (1.08)	-2.40^{**} (1.05)
Surgeries	-0.084^{***} (0.008)	-0.086^{***} (0.008)
Post–Surgery Costs	$-3{,}950.76^{**}$ (1,690.55)	$-4{,}020.88^{**}$ (1,655.00)
Post–Surgery Visits	-0.095^{**} (0.045)	-0.102^{**} (0.044)
Total MME	129.0^{***} (46.0)	133.5^{***} (45.0)
Any Rx	0.086^{***} (0.031)	0.089^{***} (0.030)
MME | Rx	-115.2^{***} (38.5)	-118.7^{***} (37.4)
High-Dose	-0.029^{***} (0.011)	-0.030^{***} (0.011)
Division FE	Yes	
Year FE	Yes	
Div×Demographic FE	Yes	
Controls	Yes	

Notes: * $p<0.10$, ** $p<0.05$, *** $p<0.01$. All models are estimated via DiD–IV 2SRI procedures with clustered standard errors bootstrapped ($M=1{,}000$).

the rate changed modestly from 0.352% to 0.410%, and for M25 from 0.423% to 0.467%. This design allows us to examine the effect of early PM access treatment in a setting where the control group approximates a "never PM" counterfactual, with minimal impact from late PM exposure.

Re-estimating our DiD–IV specification on this restricted sample, we find robust effects of early PM access on key outcomes (see Table EC.14 for results). Early PM is associated with a reduction of approximately \$5,910 in total costs ($p<0.01$) and 2.92 fewer total visits ($p<0.05$). The effect on surgical utilization remains negative (-0.102) and marginally statistically significant ($p=0.089$), despite reduced statistical power in the smaller subsample (21,487 observations). These results reinforce the robustness of our main findings and suggest that early PM access has consistent benefits in populations with low and stable late PM adoption.

Table EC.14 Robustness Check on Early PM Effect in Low and Stable Late-PM Populations (Patients with ICD Codes R10 and M25)

Outcome	**Total Costs**	**Total Visits**	**Surgeries**
Early PM	$-5{,}910.47^{***}$	-2.92^{**}	-0.102^{*}
	(2,031.56)	(1.18)	(0.060)
Division FE		Yes	
Year FE		Yes	
Div × Demographic FE		Yes	
Controls		Yes	
Model	OLS	OLS	Poisson
Adjusted R^2 / Pseudo R^2	0.29	0.24	0.19
Observations		21,487	

Notes: * $p<0.10$, ** $p<0.05$, *** $p<0.01$. Standard errors, clustered at the division level and bootstrapped ($M=1{,}000$), are presented in parentheses.