



Published in final edited form as:

J Health Econ. 2021 January ; 75: 102406. doi:10.1016/j.jhealeco.2020.102406.

Is Great Information Good Enough? Evidence from Physicians as Patients

Michael Frakes,

Duke University School of Law, 210 Science Drive, PO Box 90362, Durham, NC 27708

Jonathan Gruber,

Massachusetts Institute of Technology

Anupam Jena

Harvard Medical School; Massachusetts General Hospital

Abstract

We place an upper bound on the degree to which policies aimed at improving the information deficiencies of patients may lead to greater adherence to clinical guidelines and recommended practices. To do so, we compare the degree of adherence attained by a group of patients that should have the best possible information on health care practices—i.e., physicians as patients—with that attained by a comparable group of non-physician patients, taking various steps to account for unobservable differences between the two groups. Our results suggest that physicians, at best, do only slightly better in adhering to both low- and high-value care guidelines than non-physicians.

I. Introduction

Clinical practice guidelines have become pervasive in medicine. As defined by the Institute of Medicine in 1990, clinical practice guidelines are “systematically developed statements to assist practitioner and patient decisions about appropriate health care for specific clinical circumstances.” Among other uses, practice guidelines—along with care-recommendation lists and performance measures that derive from such guidelines—offer policymakers and institutions a means of diagnosing the current status of our health care system and a target towards which they may attempt to push provider and patient behavior through various reform initiatives.

corresponding author: 919-613-7185 (michael.frakes@law.duke.edu).

CRedit author statement

Michael Frakes: Conceptualization, Methodology, Writing—Original Draft, Writing—Review & Editing, Formal Analysis, Funding acquisition, Data Curation, Investigation

Jonathan Gruber: Conceptualization, Methodology, Writing—Original Draft, Writing—Review & Editing, Formal Analysis, Funding acquisition, Data Curation, Investigation

Anupam Jena: Conceptualization, Methodology, Writing—Original Draft, Writing—Review & Editing, Formal Analysis, Funding acquisition, Data Curation, Investigation

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

A range of studies, however, have demonstrated that medical encounters in the U.S. exhibit a low degree of adherence to established guidelines (McGlynn et al. 2003). There is a widespread belief in health care that one of the possible explanations for this deviation from recommended care is the deficiency in information and medical knowledge among consumers. An enormous number of health policies follow from this assumption, including efforts to educate patients about health, both at a population level through large-scale public health campaigns and at an individual level through provider-led patient education, decision-making support, and so on. Complementing these information-based policies are important demand-side approaches (e.g., high-deductible health plans) intended to improve the efficiency of health care by encouraging direct consumer engagement in the purchase of health care, relying heavily on the notion that informing patients about the costs and benefits of health care services can steer patients towards compliance with guidelines and recommended care.

Recent research has demonstrated the limited utility of demand-side approaches in this regard (Brot-Goldberg et al. 2017). If the information scarcity of patients is to blame for this limited success, it remains possible that further improvements could be made in encouraging compliance with recommended care should policymakers encourage even greater levels of disclosure to patients. In this paper, we attempt to shed light on whether eliminating (or, at least, nearly eliminating) the scarcity of information in medical knowledge – an extreme form of information disclosure – could plausibly lead to the delivery and receipt of care that is more likely to comply with certain recommended guidelines.

A natural way to explore the importance of scarcity of medical knowledge and information is to look at the care received by a group of patients that should have the best possible information on health care practices and clinical conditions: physicians. The decisions of physicians about which type of health care to receive—as patients themselves—would likely place an upper bound on how guideline-compliant non-physicians could become when selecting their health care treatments if fully informed about the costs and benefits of different types of health care interventions.

We do so by using a source of data that has been little used in economics research to date: data on physicians in the Military Health System (MHS). The MHS provides health insurance for all active duty military, their dependents, and retirees. Care is provided both directly on military bases and purchased from an off-base network of contracted providers. The MHS is one of the largest sources of health care spending in the U.S., with spending of over \$50 billion per year. We have gathered data on the complete claims records for all MHS enrollees over a ten-year period. Importantly, this includes the claims data for MHS physicians when they are treated as patients themselves (drawing from records of over 35,000 military physicians). The ability to observe physicians as patients provides us with a unique and powerful opportunity to answer the question: do especially well informed patients adhere more closely to the care recommendations made by the established guideline-promulgating community?

For these purposes, we focus on a set of services for which objective, evidence-based guidelines and standards exist and with respect to which the relevant expert community has

reached a consensus regarding such standards. We then assess whether physicians as patients receive more services deemed “high value” according to these standards and fewer services deemed “low-value,” in each case relative to the less-informed comparison group of similar non-physician patients. If deficiency of medical knowledge and information is an important reason why demand-side interventions to improve patient adherence with recommend care have met with limited success, we should expect that physicians-as-patients would exhibit markedly higher levels of services deemed high-value by the guidelines and lower levels of services deemed low-value, in each case compared to otherwise similar non-physician patients.

We evaluate two key areas of care that are deemed low-value—and thus over-utilized in practice—by the relevant expert community: (1) diagnostic testing prior to low-risk surgeries (e.g., chest x-ray before eye surgery) and (2) cesarean delivery (deemed low-value on the margin, as opposed to on average). We then evaluate the following forms of care that are deemed high-value—and thus under-utilized—by the relevant expert community: (1) Comprehensive Diabetes Care (e.g., eye-exam, hbA1c testing and attention to nephropathy) for patients with diabetes, (2) statin therapy for patients with diabetes, (3) statin therapy for patients with cardiovascular disease, (4) medication adherence for patients with new diagnoses of hypertension and hypercholesterolemia, and (5) childhood vaccinations (e.g., for Chicken Pox, MMR, etc.).

One concern with this analysis, of course, is that physicians may be of different health statuses and have different tastes for medical interventions as a general matter than non-physicians, which may confound an information-based-interpretation of any observed differences in guideline adherence between physician and non-physician patients. We take a range of steps to address this concern. First, by examining both low- and high-value care, we can rule out one-sided bias; e.g., if physicians are unobservably healthy, they will get less low-value care but also less high-value care. Second, we control for a rich set of health indicators, including prior year medical spending, that correlate with any underlying health differences across groups. Third, we compare physicians to other military officers, to control for underlying tastes. Fourth, we compare dependents of physicians with dependents of military officers, two groups that are likely even more similar along unobservables, but one of which—dependents of physicians—potentially benefits from the informational advantages of their physician family member.

Our results suggest that physicians do only slightly better than non-physicians in adhering to the recommendations that we identify – but not by much and not always. Across most of our low-value settings, physician-patients do adhere more closely to low-value-care recommendations than do non-physician patients, but the differences are modest, and generally amount to less than one-fifth of the gap between what is received by non-physicians and recommended guidelines. The results are slightly more mixed in the case of the high-value care analysis, with some evidence suggesting that physician-patients adhere to the high-value-care recommendations at roughly the same rate as do non-physicians and some evidence suggesting that physician-patients adhere at slightly higher rates than non-physicians.

In addition to illuminating the impacts of information by comparing physician and non-physician patients, we also compare the levels of adherence reached by physician patients with recommended levels of adherence, finding substantial gaps in these latter comparisons. Altogether, we note that both groups—physicians and non-physicians—fail to adhere to the studied guidelines a substantial portion of the time—e.g., upwards of 35% of the sponsors in the low-risk surgery sample receive some form of pre-operative diagnostic care.¹

If one assumes that physician-patients agree with the merits of the guidelines that we investigate, then our results suggest that the information deficiencies of patients may not explain much of the deviations that we observe from recommended practices. Accordingly, one might expect only limited impacts from reforms providing additional disclosures (beyond prevailing levels) to patients.²

We acknowledge that one possible explanation behind our observation that physician-patients deviate from the studied guidelines may be that they have yet to embrace the merits of the guidelines. This may be due to a situation in which the physician-patients either affirmatively disagree with the guidelines or simply do not practice in the relevant clinical area and have not received the day-to-day dosage of information that may push them to adopt the guidelines. We do set forth limited evidence to support this latter failure-to-adopt possibility in the case of our investigation into diagnostic testing prior to low-risk surgeries, whereby we find slightly larger treatment effects for those physician-patients who also perform low-risk procedures as providers. Even this group of informed providers with relevant practice-area experience, however, fails to adhere to the indicated guidelines much of the time, suggesting that this explanation cannot fully account for the observed deviations from recommended care by physician patients.

Taking this analysis one step further, we do not find evidence of a correspondence in the rate by which physician-patients adhere to the pre-operative testing guidelines when they are patients and the rate by which their own patients adhere when these physicians perform these procedures as providers. This latter observation cuts against an interpretation of our findings as arising from affirmative disagreements with the guidelines. Altogether, our findings suggest that factors other than information and clinical beliefs contributes to the bulk of the deviations from recommended practices that we observe.

In any event, even if some amount of the observed guideline-deviations derives from the failure of the physician-patients in our sample to have updated their priors in accordance with the guidelines, the policy implications arguably remain the same. That is, if one assumes that physician-patients have been exposed to the recommendations but simply lack the day-to-day experiences that may lead them to adopt the guidelines, our results still speak

¹Moreover, as discussed in the Online Appendix, the levels of adherence by the patients in our sample roughly correspond with those found in other samples (and from other patient populations) in the literature, appeasing concerns that the limited information effects implied by our analysis can be explained by a generally high level of adherence at baseline among the military population.

²Even if one assumes that both the practicing provider community and the guideline-promulgating community agree on the merits of these guidelines, we emphasize that our analysis aims to speak to the question of whether information alone is sufficient to encourage individuals to make *medically optimal* decisions, not necessarily welfare improving decisions. In any event, we take this caveat one step further and also acknowledge that the guidelines may, in fact, not be medically optimal (despite any assumed agreement within the community). In this case, our analysis still speaks to the role of information in inducing adherence to care recommendations made by influential guideline-promulgating organizations.

to an upper bound of the impacts on guideline-adherence that might arise from *feasible* information-disclosure initiatives aimed at lay patients.

Our paper proceeds as follows. Part II reviews the existing literature on the impacts of health care information initiatives and on informed providers. Part III provides a background on the MHS and discusses our data and empirical methodology. Part IV presents the results of our analysis and Part V fleshes out the implications of those findings. Part VI concludes.

II. Literature Review and Conceptual Overview

II.A. Previous Literature on Information-Sharing Policies

An essential component of most modern proposals for a consumer-driven, patient-centered approach to health care is the provision of sufficient information to patients to facilitate informed medical decisions that are adherent to recommended care. Our study will build on others that have attempted to shed light on whether information disclosure may reach this goal.

Arguably illuminating the potential for prevailing levels of disclosure to improve decision-making, a number of studies show that patients typically enter and proceed through their treatment regimens with relatively little information on their treatment options. Several studies survey patients with particular conditions or undergoing particular procedures and “quiz” them on their knowledge of the relevant options, generally documenting very low levels of understanding (Pope 2017). For instance, Weeks et al. (2012) found that only 19% of patients with colorectal cancer understood that chemotherapy was not likely to cure their cancer.

A range of similar studies have endeavored to explore the capacity of patients to understand their scenarios in the first place. A representative study of this latter nature can be found in Herz et al. (1992), in which the researchers provided pre-treatment teaching sessions by a neurosurgeon to a group of 106 patients receiving either anterior cervical fusion or lumbar laminectomy. Immediately after the session, patients were given a basic written test and, on average, scored only 43.5 percent. Given the simplicity of the questions, the authors concluded that greater information disclosure to patients cannot necessarily ensure accurate comprehension.³

Another literature bearing on the potential for information disclosure to improve patient choice examines policy initiatives aimed at facilitating choice of providers—for instance, policies requiring hospitals to publicly report on various patient outcomes, whether at the state level (e.g., CABG report card programs in New York and Pennsylvania) or at the federal level (e.g., Medicare Hospital Compare program). Studies suggest, however, that the vast majority of patients do not access this information in the first place (Kaiser Family Foundation 2008; Associated Press—NORC Center for Public Affairs Research 2014). Moreover, even when patients are aware of the disclosures, it is unclear whether it affects

³Meisel and Roth (1983) survey additional early studies on the degree to which patients understand disclosures by their physicians.

their decisions. A number of studies have evaluated the effect of report cards on provider market shares, generally finding mixed results.⁴

II.B. Overview of Physicians-as-Patients Approach

Even with robust physician disclosure of information, patients may face inherent limitations in their ability to process this information—e.g., due, in some cases, to the lack of intellectual capacity or educational background, or due to the fact that patients are often in diseased states during the time of disclosure, impeding their ability to absorb this information. In our analysis, we look to the decisions of a set of patients—e.g., physicians as patients themselves—that are less likely to suffer from at least some of these limitations. As such, our approach is designed to provide an upper bound on the gains to guideline adherence that may arise from providing patients with greater sources of information at the clinical decision-making stage.

This bounding approach is similar to several recent studies on the impacts of medical expertise, beginning with Bronnenberg et al.'s (2015) analysis of consumer behavior using scanner data. Bronnenberg et al. compare the use of store brand versus generic products—e.g., over-the-counter medications, pantry supplies, etc.—in the case of expert versus non-expert shoppers. Of closest relevance to our piece is their finding that pharmacists are substantially more likely to choose generic headache medicines (roughly 91% of the time) compared with the average consumer (choosing generic headache medicines roughly 74% of the time).

Johnson and Rehavi (2016) perform a similar analysis, though one that focuses more on an actual patient-physician encounter. They explore cesarean-section delivery. Using vital statistics data from California matched with licensure information from physicians, they find that informed patients—i.e., physician mothers—are roughly 10 percent (relative to the mean) less likely to deliver children via cesarean section in comparison with non-physician mothers. Johnson and Rehavi's focus is on the implications of their findings for physician induced demand, arguing that their findings provide evidence that excess care is provided. Another interpretation of their results, however, is as a bound on the effects of a perfectly informed consumer. In our analysis below, we will similarly compare physician and non-physician patients but will do so across a much broader array of clinical settings than just the cesarean decision.

Our analysis is further similar to that of Chen et al. (2019), which draws on Swedish data and uses randomized medical school lotteries and other identification strategies to assess the effects of having a health professional in the family on various outcomes. Much of their analysis focuses on the consequences for certain physical health outcomes—e.g., the incidence of suffering from lung cancer or diabetes—effects that may arise from a combination of better lifestyle choices and better medical decision-making. In this light,

⁴Among those studies finding some impact of report cards on provider choice include Bundorf et al. (2009), Wang et al. (2011), and Cutler et al. (2004). Among those studies finding either no or minor impacts include Kolstad (2013), Romano and Zhou (2004), Mukamel et al. (2010), Schneider and Epstein (1998). Even in those studies documenting report card effects, however, such responses need not necessarily emanate from individual patient choice itself but may instead reflect either employer or managed-care responses. See Madison (2017) for a recent discussion of information forcing initiatives at the state and federal level.

much of their analysis is distinct from our own in that our analysis aims to focus specifically on this latter medically-centric mechanism. Nonetheless, Chen et al. do extend their analysis to consider preventive medical measures similar to those that we explore below, including utilization of diabetes and cardiovascular medicines among those with affected conditions (though, it is not clear how their at-risk samples are formed or what guidelines they are using to frame their adherence analysis). We note that Chen et al. estimate larger physician effects on these preventive measures than we will estimate below. For instance, they find that having a health professional as a family member is associated with a 32 percent relative increase in the likelihood of using statins.⁵ Ultimately, the differences in the magnitudes between our study and that of Chen et al. may stem from the institutional differences between these countries and/or from differences in the inherent tastes for medical care between the U.S. and Sweden, including perhaps differences in the degree of information disclosure to patients at baseline.

There are limitations in using the results from Johnson and Rehavi (2016) and Chen et al. (2019) as a measure of the bounded effects of information. While the methodologies employed in these studies—including the medical school lottery in Chen et al.—may help address concerns over inherent differences in those individuals that do and do not become physicians, their approaches are unable to address the general concern that the act of becoming a physician may generally alter their tastes for medical care. That is, consider the possibility that once becoming a physician, an individual will desire more care of all types—both recommended and not recommended care. In this case, by observing stronger adherence to recommended preventive care (where recommendations call for more care), one may be overstating the influence of information. By exploring adherence to guidelines of both high-value and low-value varieties, our analysis will build on the approaches of these other studies and attempt to separate the impacts of differential knowledge between physician and non-physician patients from the impacts arising from any underlying health care preferences (or health differences) between these groups.

Finally, as we emphasize below, a key component to our empirical analysis and to our attempt to explore the impacts of patient information on clinical decision-making will extend beyond a comparison of physician and non-physician patients and will also emphasize a comparison of physician patients with the levels of care recommended by the guidelines themselves. Put simply, even if physician effects happen to be large in relative terms, if those effects are still small relative to the gap between prevailing and recommended levels of adherence, such results may nonetheless imply a limited role for information disclosure initiatives in encouraging guideline compliance. In this light, our analysis expands upon both

⁵Chen et al.'s (2019) use of admissions lotteries in Sweden is similar to a study using admission lotteries in Netherlands to explore physician effects on various health outcomes (Artmann et al. 2019), though the Artmann et al. analysis does not focus on specific clinical decisions like the present analysis or the preventive analysis of Chen et al. (2019). Specifically, Artmann et al. find little differences in various aggregate spending / utilization metrics and health outcomes (e.g., mortality) between parents of physicians and parents of non-physicians. Our analysis is also related to a vignette-based survey literature in medicine that compares the decisions that physicians might make for themselves with what they would recommend for their patients in the face of a hypothetical choice over treatment options that pose various tradeoffs. See, for instance, Ubel et al. (2011). Such studies are distinct from the present study in that they are not comparing the actual clinical decisions made by informed and less-informed patients. Rather, they are comparing what an informed patient might decide for themselves—a decision that may be affected by a range of influences—and what they might recommend for another person (where this recommendation may not be subject to the kinds of influences that one faces with respect to themselves).

Johnson and Rehavi (2016) and Chen et al. (2019), which focus only on the physician / non-physician comparison.

III. Background on Military Health System, Data and Methodology

III.A. Background on Military Health System

The Military Health System (MHS) is the primary payer of health care services for all active duty military, their dependents and retirees through the TRICARE program. TRICARE is not involved in health care delivery in combat zones and operates separately from the Department of Veterans Affairs' Veterans Health Administration health service delivery system (Schoenfeld et al. 2016). The MHS actually consists of two systems. For some beneficiaries, the MHS directly delivers health care at Military Treatment Facilities (MTF) on military bases (i.e., the "Direct Care" system). For other beneficiaries, the MHS purchases care from private providers who are within a contracting administrator's network (i.e., the "Purchased Care" system), similar to most privately insured in the U.S. Whether MHS beneficiaries receive care on the base or off the base is largely a function of where they live. Those living close to the MTFs are expected to receive care from the MTF and those living farther away are expected to go off the base.

The MHS offers alternative insurance plans with different cost-sharing and other terms. All active-duty personnel are required to enroll in TRICARE Prime plans. While facing other alternatives, nearly 90% of the non-active duty likewise choose to enroll in TRICARE Prime. TRICARE Prime beneficiaries who are active duty or dependents of active duty face no out-of-pocket costs whether they go on the base for care or off the base for care.⁶

III.B. Data

The data for our analysis come from the Military Health System Data Repository (MDR), which is the main database of health records maintained by the MHS. Broadly, it provides incident-level claims data across a range of clinical settings and contexts, with data on inpatient stays, outpatient stays, pharmaceutical records, and radiological and laboratory testing, in all cases for both the Direct Care and Purchased Care settings. Each record provides details regarding the encounter—primarily the diagnosis and procedure codes associated with the event and various other utilization metrics. Furthermore, the MDR database also contains separate files with coverage, demographic, geographic and other information on each MHS beneficiary, which we link to the claims records.

Critically, these data provide Department-of-Defense-specific identifiers for each MHS beneficiary. The Direct Care records use those same identifiers to acknowledge the identity of the provider associated with each record. The fact that these same identifiers are used to identify both patients and providers presents us with a unique opportunity—that is, the ability to observe active-duty MHS physicians as patients themselves.⁷ Large-scale health

⁶Matters are slightly different for prescription cost-sharing, however. Active duty patients have no co-payments regardless of their care location. Non-active-duty patients likewise pay no co-payments when filling their prescriptions at MTF pharmacy facilities, though they do pay small co-payments when filling prescriptions via mail order (\$7 for generics) or at civilian retail pharmacies (\$11 for generics).

care claims databases of this sort rarely provide information on the profession of the patient, let alone with enough specificity to identify those who are also physicians. The key methodological thrust of our paper is to take advantage of this opportunity and to compare the care received by physicians as patients with the care received by otherwise similar non-physicians in an effort to elucidate the role that information plays in encouraging compliance with high-value and low-value care recommendations.

One of the empirical challenges in this exercise is to ensure that the physician and non-physician comparison groups are otherwise equal, such that the differences that we observe between these groups can be attributed to the informational advantages of the physician-as-patients group. Non-comparability concerns arise in at least two key dimensions. First, physician patients and non-physician patients may need health care services to different degrees given differences in underlying health statuses. Second, even aside from need, physician patients and non-physician patients may prefer health care services to different degrees—e.g., perhaps physicians have especially little time to seek care for their own maladies. We take a number of approaches to address this concern throughout the analysis below.

At the outset, we emphasize several sample selection choices that are helpful in addressing comparability concerns between physician and non-physician patients. First, we remove any differences in financial incentives among the patients studied by limiting the analysis to those beneficiaries with TRICARE Prime coverage. Second, in each clinical context that we explore, we attempt to construct comparison groups among MHS beneficiaries that are similar to active-duty physicians in terms of socio-economic status and of the demands of their time, both of which may bear on a patient's inclination to receive particular forms of medical care. For these purposes, we limit our attention to non-physician military officers.

In alternative specifications, we focus on dependents—that is, we compare dependents of physicians with dependents of officers. This approach is premised on the idea that dependents of active duty physicians may benefit from the knowledge of their physician family members, to the extent the physician family members are involved in the relevant medical decisions.⁸ A key potential benefit of this approach is that the dependents on both sides may be otherwise more similar to each other—especially in terms of the demands on their time—relative to the similarity between active duty physicians and officers. As discussed further below, it is true in the case of observable patient characteristics that the treatment-control differential is generally smaller in the case of the dependents sample relative to the sponsors sample.

⁷Secondarily, the MDR records contain occupation codes for each beneficiary, which likewise allows us to identify military physicians as patients.

⁸This approach of drawing on family members of physicians to illuminate the impacts of health literacy is consistent with the approach taken in Chen et al. (2019) which focuses almost entirely on this perspective. We caution that our dependent results may not be interpreted as the effect of information on guideline adherence should it not be true that family members turn to their physician family members for a source of information when making medical decisions.

III.C Low vs. High Value Care

In order to assess the role of informed consumers in making medical decisions, we compare physicians-as-patients with non-physician patients in their compliance with both high-value care guidelines and low-value care guidelines. “High-value” care guidelines represent recommendations to receive certain services with respect to which the relevant guideline promulgating organization perceives the clinical benefits to justify the costs and harms. “Low-value” care guidelines represent recommendations to avoid certain services with respect to which the relevant organization perceives the costs and harms to outweigh the clinical benefits. In this sub-section, we briefly describe the various clinical samples that we explore and the guidelines that we investigate. We provide even greater details in the Online Appendix.

Low-Value Care: Labor and Delivery—The first low-value care guideline that we consider calls for the reduction of cesarean delivery on the margin. The low-value label applied by the health care community—primarily, the World Health Organization or WHO—in this cesarean context is not premised on the idea that cesareans should never be performed, but rather that they should be performed at less than half of the rates that currently prevail (WHO, 2015). To explore the potential impact of greater information disclosure to patients on cesarean rates, we compare cesarean use between physician and non-physician mothers using the sub-sample of deliveries in the MDR. We consider the incidence of any cesarean delivery, along with the incidence of a “primary” cesarean delivery as defined by the Agency for Health Care Research and Quality. Primary cesarean rates remove from the relevant delivery sample breech deliveries, multiple deliveries (e.g., twins) and previous cesarean deliveries—i.e., deliveries with respect to which physicians have less discretion in the cesarean decision.

Low Value Care: Pre-operative care for low-risk surgeries—While cesareans are low-value in the sense of being over-used, other forms of health care are deemed “low value” by the relevant guideline-promulgating organization in the sense that the organization recommends that such services not be performed at all. We next consider a low-value measure of this nature inspired by the Choosing Wisely Campaign: pre-operative diagnostic testing prior to low-risk surgeries. For these purposes, we consider the sample of low-risk surgeries specified by Schwartz et al. (2012)—e.g., cataract removal and hernia repair—and thereafter flag the incidence of the following unnecessary tests prior to—i.e., within 30 days of—the relevant low-risk surgery: chest radiography, complete blood count, coagulation panel or comprehensive metabolic panel. We estimate specifications that consider these tests individually, while also estimating specifications that pool them.⁹ As a robustness exercise, we also consider an alternative to the Schwartz et al. list of low-risk surgeries where we instead use the inpatient and outpatient MDR records to identify a set of surgeries with low

⁹We drop from this sample patients that also receive a nearby high-risk surgery, along with patients over the age of 75. We acknowledge that some of the tests that we identify through this method may represent tests unrelated to the surgery—e.g., annual visit tests—that spuriously fall into this 30-day window prior to the low-risk surgery. We nonetheless verify in the Online Appendix that the vast majority of those tests received within this 30-day window are likely connected with the low-risk surgery itself and thus deemed low-value according to the motivating guideline.

mortality rates (mortality of less than 1 per 1,000 surgeries within 30 days following the surgery/procedure).

High-Value Care: Diabetes—In our first high-value-care guideline analysis, we focus on a sample of patients with diabetes and follow them throughout the course of the relevant sample year to determine if they received “Comprehensive Diabetes Care” (CDC), as that term is specified by the Healthcare Effectiveness Data and Information Set (HEDIS), which is created by the National Committee for Quality Assurance. To better ensure comparability in health status across patients, we limit the sample to those who have had a diabetes diagnosis flagged in their medical records for at least two years prior to the focal year. We identify compliance with CDC by receipt of all of the following: (1) hemoglobin A1c (HbA1c) testing, (2) retinal eye exam, and (3) medical attention for nephropathy. Following HEDIS, we also assess whether the focal diabetes patient receives statin therapy over the observation year (subject to certain additional sample restrictions).

High-Value Care: Cardiovascular Care—In our next high-value care analysis, we assess whether patients comply with the HEDIS protocols for cardiovascular care, focusing on the subsample of patients with a previous atherosclerotic cardiovascular disease diagnosis (CD sample). Within this sub-sample, we follow patients over a year-long observation period and identify high value care by observing (a) whether the affected patient received statin therapy (of high or moderate intensity) at least once over this period or, alternatively, (b) the number of days over the year in which patients filled a prescription for statin therapy, in addition to an indicator, following HEDIS guidelines, for whether they received statin therapy for at least 80% of the observation year.

High-Value Care: Medication Adherence—The above statin-adherence analysis focuses on adherence on an annual basis for those that have the indicated condition. To complement that steady-state medication-adherence analysis, we also explore adherence during the critical first year following the diagnosis of a new condition. For these purposes, we follow influential papers in the medication adherence literature (for instance, Briesacher et al. 2008) in selecting the following two conditions / samples: (1) patients with a new diagnosis for hypertension and (2) patients with a new diagnosis of hypercholesterolemia. In both cases, we follow the affected patient over the first year following the diagnosis of this new condition and determine the patient’s Medication Possession Ratio, which equals the share of the year (in days) in which the patient has filled her prescription.

High-Value Care: Vaccination / Immunization—In a final high value care analysis, we explore the extent to which children of physicians receive by two years of age the following vaccinations (and dose frequencies) recommended by HEDIS: (1) four vaccines for diphtheria, tetanus and acellular pertussis (DTaP), (2) three polio vaccines (IPV), (3) one measles, mumps and rubella (MMR), (4) three vaccines for haemophilus influenza type B (HiB), (5) one vaccine for hepatitis A, (6) three vaccines for hepatitis B, (7) one vaccine for chicken pox, (8) four vaccines for pneumococcal conjugate (PCV), and (9) two or three vaccines for rotavirus.¹⁰

III.C. Methodology

To assess whether greater patient information is likely to lead to greater compliance with both low- and high-value-care recommendations, we estimate the following specification for each of the relevant samples:

$$Y_{ite} = \alpha + \theta PHYSICIAN_i + \delta \mathbf{X}_{ite} + \mu_{jt} + \varepsilon_{ite} \quad (1)$$

Where i denotes the individual patient, j denotes the patient's assigned military base and t denotes the relevant observation year. The care location—either on-base at an MTF or off-base at a civilian facility—is captured by I (discussed further below). The encounter or episode over which we are evaluating the care provided to the patient is captured by e . In the case of the cesarean analysis, e represents the individual delivery. For the pre-operative care analysis, e represents the 30-day period prior to an individual low-risk surgery. For the HEDIS diabetes and cardiovascular disease samples, e represents the full sample year over which we observe the care provided to the relevant patient. For the first-year drug adherence analyses, e similarly represents the full sample year over which we observe the care provided to the affected patients. For the immunization analysis, e represents the first two years of the relevant child's life.

Y_{ite} represents the relevant outcome variable for the particular sample at issue—e.g., the incidence of cesarean delivery, the receipt of a chest radiography in the 30 days prior to a low-risk surgery, etc. $PHYSICIAN_i$ is an indicator for whether the patient in question is an active-duty physician. This indicator equals 0 for our key control group—i.e., non-physician military officers. In separate specifications, as discussed above, we focus on an analysis of dependents only—e.g., dependents of active duty physicians ($PHYSICIAN_{it} = 1$) compared with dependents of non-physician military officers ($PHYSICIAN_{it} = 0$).

We also attempt to control for other differences between physicians and non-physicians by including a rich set of control variables, \mathbf{X}_{ite} , which includes patient age-by-sex dummies, patient race dummies (white, black, and other), and pay-grade-level dummies for the relevant patient's sponsor (junior officer and senior officer). \mathbf{X}_{ite} also includes four additional metrics reflective of the health status of the relevant patient. Three are measures of patient ex-ante resource utilization: (1) the number of inpatient bed-days over the preceding year, (2) the patient's aggregate "Relative Value Units" (RVU) for the previous calendar year,¹¹ and (3) the patient's aggregate "Relative Weighted Product" for the previous calendar year; the latter is an MHS-created measure which captures the overall intensity of patient treatment in inpatient settings for the preceding year (Frakes and Gruber 2020, Frakes and Gruber 2019). The fourth is a direct measure of patient health, the Charlson comorbidity score.¹² In the pre-operative testing analysis, where we pool across

¹⁰One concern with this analysis respects possible under-reporting / recording of vaccinations in the MDR database. We address this concern in the Online Appendix along with our associated response.

¹¹Developed for use with Medicare's reimbursement system, RVUs are a measure specified on a common scale across health care services that are meant to be reflective of the resources used in providing the relevant service, where resources accounted for reflect (1) physician work efforts involved (time, technical skill, mental effort, etc. involved in providing the service), (2) practice expenses involved (accounting for the nonphysician clinical and nonclinical labor of the practice, in addition to other office expenses) and (3) medical liability expenses. For an overview of RVUs, see <https://www.cms.gov/apps/physician-fee-schedule/overview.aspx>.

different types of low-risk surgeries, we also include fixed effects for the given surgeries comprising this sample.¹³

Further, in specification (1), we include a rich set of fixed effects for military base-by-year-by-care-location groups, μ_{jil} . Care location controls—i.e., flagging on-base versus off-base care—are important to the extent that physician patients are more or less likely to receive care on the base versus off the base and to the extent that the nature of health care services differs at MTFs versus civilian facilities. With this rich set of effects, we effectively compare physician patients and non-physician patients in their receipt of high- and low-value medical services (as deemed by the relevant guidelines) within a given base, within a given year, and within a given on-base-versus-off-base site—e.g., within Evans Army Community Hospital in Fort Carson, Colorado in 2009. Since patient locational choices are potentially endogenous, we also estimate models excluding these care-location controls, and the results are virtually unchanged.¹⁴

For some of the clinical settings—e.g., in the case of labor and delivery—it is straightforward to assign the on-base indicator variable. For other settings—e.g., drug adherence over the observation year—it is less straightforward in that patients may receive medical services both on the base and off the base over the observation year. In these latter instances, we determine for each patient, the share of the RVUs associated with their medical care that is rendered on the base and set an indicator variable for the receipt of on-base care equal to “1” for above-median on-base RVU shares (and 0 for below-median on-base RVU shares).¹⁵

Table 1 shows the sample balance. Since we have many different samples corresponding to our different high/low value analyses and many different covariates along which to assess balance, we evaluate, for each sample, the balance of predicted outcomes. That is, we form predictions for the key outcome variables—e.g., predicted cesarean delivery—based on the relevant set of covariates¹⁶ and thereafter test for balance in these predicted treatment measures between the physician and non-physician groups. This omnibus approach provides

¹²The Charlson score is a weighted index of 17 comorbidity conditions (Deyo et al. 1992). In the Online Appendix, we also present a full set of results for our estimated physician effect across all clinical contexts where we do not control for previous medical spending or for the Charlson score to reflect a potential concern that the physician status of the respective patient (and the information advantages associated with that status) also contributes to a generally healthier state (Leuven et al. 2013), which may be associated with possibly better guidelines adherence decisions in the contexts under investigation. As demonstrated in the Online Appendix, the results are nearly identical when dropping these health status controls. The only modest impacts that arise from this modification occur in the case of the vaccinations analysis, where if anything, the physician effect gets even smaller, further reinforcing the main conclusions of the paper.

¹³For those specifications, where it is possible for there to be repeat observations over time for the same patient—i.e. for all specifications other than the first year- drug adherence and childhood immunization analysis—we cluster standard errors at the patient level.

¹⁴Given that the choice of on-base versus off-base care is largely driven by geography (Frakes and Gruber 2018), we treat those specifications that drop care-location variation as specification checks only.

¹⁵The results presented below are nearly identical when including base-by-year fixed effects and when simply controlling for this on-base RVU share. For the pharmaceutical-use specifications, care-location measures capture their choice of on-base versus off-base care in the sense of what medical providers they visit, not what type of pharmacy they visit—MTF pharmacy versus retail pharmacy. As discussed above, this pharmacy-location distinction is potentially relevant for the dependents sample given that dependents face higher co-payments for prescriptions filled outside of MTF pharmacies. This creates a potential concern for the dependents sample to the extent the physician group differentially uses MTF pharmacies. While the vast majority of prescriptions are nonetheless filled at MTFs, we include an additional control for the dependents sample (in the pharmaceutical specifications) for the location of the pharmacy (on base versus off).

¹⁶We form these predictions over the whole sample—physician and non-physician patients—such that we use the same coefficients of the covariates for both types of patients.

us with a means of exploring the degree of covariate imbalance in a collective sense. In the Online Appendix, we show a comparison of individual covariates one by one.

Generally, with the sponsors comparison, we find modest imbalance across individual covariates, but with no particular pattern – and therefore there are only small differences in the predicted outcomes summary measure. For example, we find that physician sponsors are predicted, based on their covariates, to deliver via cesarean section at a roughly 1 percentage-point lower rate relative to non-physician officers. In each sample, the degree to which these predictions differ between physicians and non-physicians is small in magnitude, generally between 0 and 2 percentage points but upwards of 4 percentage points in the case of predicted comprehensive diabetes care. Moreover, across most of these samples, the covariate imbalance appears to shrink when comparing dependents of physicians with dependents of non-physician officers, consistent with our expectations that the treatment and control groups in the dependents sample are likely to be more similar. In the childhood immunization analyses, naturally, we show imbalance only for the dependents sample. The evidence suggests at most only a minor degree of imbalance between the physician and non-physician group in the immunization sample.

As one final methodological note on patient comparability, we highlight the importance of the fact that, in some contexts, we are testing for the incidence of care that is deemed high-value by the relevant guidelines, whereas in other contexts, we are testing for the incidence of care that is deemed low-value by the relevant guidelines. If there is any bias in our approach created by unobservable differences in health status or tastes for receiving medical interventions between our treatment and control groups, one would expect that bias to work in one consistent direction—e.g., physician patients should consistently receive less low- and high-value care than their non-physician counterparts (for instance, perhaps physicians always desire more care, whether recommended or not). If any such bias would indeed be one-sided in nature (if it exists at all) and if the findings happen to demonstrate small physician effects in *both* the high- and low-value care settings, then one might confidently infer that the information level of patients plays a small role in determining the use of high- and low-value services.¹⁷

IV. Results

IV.A. Low-Value Care Analysis

In Table A2 of the Online Appendix, we show uncontrolled differences between physicians and non-physicians in the means of our various outcome measures. Across many (though not all) of the low-value care specifications, we find that the uncontrolled mean low-value care rate across the respective measures are lower for the physician group relative to the non-physician group—that is, the raw means suggest somewhat greater compliance with the low-value-care recommendations for the physician-as-patients group. In this section we turn to

¹⁷That is, if information does substantially alter compliance with guidelines and if physicians are of weaker health—which may hypothetically explain small physician effects in low value-care—then one would tend to expect large positive physician effects in the case of the high-value care analysis (since both information and unobservable health status would elevate physician’s receipt of treatment). In this case, if we observe modest physician effects in the high-value care setting, this would tend to suggest a small, if any, degree of bias and a small information effect.

regression estimation to assess how these physician / non-physician comparisons hold up to controlling for various measures that may also differ between these groups.

Cesarean Delivery—We begin in Table 2 by presenting the results from our first low-value care analysis, where we explore whether physicians as patients receive fewer cesarean deliveries relative to military officers. In Panel A, we focus on the sponsors comparison (physicians versus non-physician officers), whereas in Panel B, we focus on the dependents comparison. Given the gender differences between the sponsor and dependent samples, we have a notably larger sample size in the dependents analysis. We present results separately for specifications that evaluate the incidence of any cesarean delivery among all deliveries and that evaluate the incidence of “primary” cesareans among a more restricted set of deliveries.

Each row in Table 2 represents a different regression using the indicated dependent variable (in the specified sample). Point estimates and standard errors for the estimated physician effect coefficient are provided in Columns 4 and 5. To help place the magnitudes of the physician effects in perspective, we indicate in Column 3 the baseline mean rate for the relevant cesarean rate. For conciseness purposes, we follow this format throughout when presenting results across our various clinical contexts given the large number of resulting specifications estimated.

For sponsors, we find that the physician status of the mother has an insignificant effect on the odds of having any cesarean, but for the more discretionary set of “primary” cesareans, there is a significant 2.9 percentage-point reduction associated with being a physician mother, which is about 15 percent of the mean. For dependents of physicians, a much larger sample, we find significant reductions in the incidence of any cesarean—at an amount equal to roughly 6 percent of the mean cesarean rate—but not in the case of the incidence of a primary cesarean delivery.

Though the findings vary across samples and across cesarean measures, the magnitude of the estimated physician effect—in those specifications with significant effects—corresponds somewhat closely with the roughly 10-percent effect (relative to the mean) found in Johnson and Rehavi (2016). Johnson and Rehavi’s analysis uses vital statistics data from California, limiting their ability to consider the full set of clinical contexts explored in our analysis. The correspondence in our findings, however, perhaps provides some support for the generalizability of our findings—including our non-cesarean-related findings—to individuals beyond the beneficiaries of the military health system.

Even in those specifications with significant effects, this cesarean finding is arguably still a modest effect. The World Health Organization (WHO 2015) suggests that the overall cesarean rate should be 10–15 percent, a relative reduction of more than 50 percent in light of the nearly 30 percent prevailing rate of cesarean deliveries. By these standards, our results demonstrate a fairly modest difference between physicians and non-physicians in the overall cesarean rate—i.e., a reduction of 6 percent of the mean among physicians compared to the 50-percent-plus reduction called for by the WHO.¹⁸ These results therefore suggest that, at most, we may only observe a minor improvement in adherence to the WHO

recommendation when providing patients with substantially greater sources of information. In Part V below, we further flesh out the implications of these findings, along with those from the other clinical samples studied below.

Low-Risk Surgeries—Building on this cesarean analysis, we turn now to an evaluation of diagnostic testing prior to low-risk surgeries, another form of care that has been routinely deemed over-used across the various medical specialty societies that have banded together as part of the Choosing Wisely campaign.¹⁹ Unlike cesareans, however, this service is recommended to be avoided altogether. That is, tests such as chest radiography or a complete blood panel are never recommended before procedures such as cataract surgery, yet they occur with either modest or alarming frequency. We present results using both the sample of low-risk surgeries set forth in Schwartz et al. (2014) and the sample of surgeries that are low-risk according to our MDR sample (derived low-risk sample).

Table 3 shows the results of this preoperative testing analysis, with Panel A focusing on the sponsors comparison and Panel B focusing on the dependents comparison. As above, each row in this table reflects a different specification. In the sponsors sample, we find a smaller rate of pre-operative chest radiography for physicians—at a roughly 2 percentage-point lower level in the case of the Schwartz et al. sample (relative to a mean rate of 10 percent) and at a roughly 0.3 percentage-point lower level in the case of the derived low-risk sample (relative to a mean rate of 4.8 percent). However, in the case of each of the other pre-operative measures in the sponsors sample—and in the case of the aggregate any-preoperative testing measure—we estimate near-zero point estimates for the physician effects (each statistically indistinguishable from zero). For instance, in the sponsors derived-low-risk sample, we estimate a 95-percent confidence interval for the estimated physician effect of the any-preoperative-testing specification that spans -0.4 to 0.5 percentage-points, or roughly -2.4 to 2.6 percent relative to the mean rate of any pre-operative testing (we find a similar range in the case of the Schwartz et al. sample). Accordingly, based on the results of the sponsors sample, we can rule out a large physician impact and thus a large influence of superior patient information in the choice to receive any low-value testing prior to low-risk surgeries.

In the case of the dependents sample, we generally find evidence of modestly lower use of pre-operative testing for the physician group. For instance, with respect to the Schwartz et al. (2014) sample, we find a 1.7 percentage-point lower use of chest radiography prior to low risk surgeries and a roughly 1.0 percentage-point lower rate of coagulation panels. With respect to any preoperative testing, however, our findings are small and indistinguishable from zero, largely due to the lack of a difference in complete blood count testing between physicians and non-physicians. When drawing on the low-risk surgery sample derived from the MDR records, we find more consistent evidence of a negative physician effect. In the

¹⁸We note that the WHO does not specify a recommended “primary” cesarean rate, limiting our ability to compare the observed values of this measure with the WHO recommendations, though we suspect that the WHO recommended rate there would be substantially less than 10–15% considering that the “primary” rate removes from the sample those instances in which cesareans are almost uniformly performed.

¹⁹For a recent discussion of the Choosing Wisely campaign’s recommendations (and the supporting medical literature) regarding preoperative testing, see Siddaiah et al. (2020).

case of any preoperative test within this sample, we find that physician patients receive such tests at a roughly 1.2 percentage-point lower rate, amounting to a 5.4 percent effect relative to the mean testing rate of 21.9 percent. However, even in this case, physician patients receive unnecessary preoperative testing close to 20 percent of the time prior to low-risk surgeries.²⁰ Accordingly, even when taking results from this pre-operative testing analysis that suggest the strongest physician effect, our findings imply that the upper bounds of information disclosure are not likely to close the considerable gap between prevailing rates of care and the rates recommend by the Choosing Wisely campaign.

As discussed in Part III, we acknowledge the possibility that these modest findings may be attributable to omitted-variables bias. That is, there could indeed be a large information effect that drives down the use of low-value services. However, masking this information effect could be unobservable factors that drive physicians to receive more care generally and that thus partially wash out this hypothetically strong information effect. For instance, perhaps physicians, at baseline, have a fundamentally higher taste for more medical care of any kind—recommended or non-recommended. While we have included a rich set of controls (including resource-use measures from previous years) and also employ our dependents analysis to appease these concerns, one may nonetheless still have residual concerns. To further address such matters, we now invert our focus and explore the role of information in the utilization of services deemed high-value by the relevant guideline-promulgating entities.²¹

IV.B. High-Value Care Analysis

In Table A2 of the Online Appendix, we show uncontrolled differences between physician and non-physician patients in the use of certain high-value medical interventions that guideline-promulgating organizations unambiguously recommend for patients. We do not find any consistent pattern in uncontrolled mean differences between these groups, except in the case of the immunization analysis, where we tend to find a minor degree of stronger adherence by the physician group. In Table 4, we use our regression approach to control for observable differences between physicians and non-physicians. We focus our discussion on these controlled results.

²⁰As discussed above, while we take steps to rule out situations in which these tests are conducted prior to nearby high-risk surgeries, we arguably cannot rule out that some amount of these tests reflect routine—e.g., annual—tests that spuriously fall in this 30-day window. Accordingly, given our calculation of the preoperative testing rate, we acknowledge that perfect compliance with the Choosing Wisely recommendations would not reflect a 0% rate for our measure. However, Figure A3 of the Online Appendix—which shows a frequency distribution of the relevant tests in the period leading up to and following a low-risk surgery and which demonstrates a large spike prior to the surgery—suggests that the majority (if not vast majority) of the tests that we are capturing in the 30-day window are not simply a reflection of standard likelihoods of testing over time but are associated with the surgery itself. That is, to assess the gap between observed rates and recommended rates of preoperative testing, one can effectively consider the difference between the observed frequency of testing in this window and what one would expect based on general likelihoods of testing in other periods (which can be visualized in this figure). In this case, we still find a meaningful gap between observed rates of preoperative testing and recommended rates. Our observation of small physician effects suggests that this is true for even our informed patients.

²¹Another possibility is that physicians may be in closer proximity to the Military Treatment Facilities (MTF), leading to generally more care selection. Our investigation into both high- and low-value appeases concerns of any such one-way bias, as discussed below. Nonetheless, we can also alleviate this distance-based concern by controlling for a series of distance bins based on the distance between patient residence and the closest MTF (based on zip-code centroids). The results for each of our subsamples explored in this paper are nearly identical when including these additional controls.

Diabetes care—We next explore adherence by diabetes patients to HEDIS guidelines for Comprehensive Diabetes Care (CDC) within the observation year, where each observation is a patient-by-year cell. We also show physician effects for compliance with each of the components of CDC. Furthermore, we test for physician effects in the receipt of statin therapy for diabetes patients.

In the case of both the dependents analysis and the sponsors analysis, we do not find evidence that physicians are more likely to follow the CDC guidelines than non-physicians or more likely to receive statin therapy. For instance, in the case of compliance with CDC care for sponsors, the 95-percent confidence interval for the physician effect spans from -3.8 to 5.0 percentage points or from roughly -6.5 to 8.6 percent relative to the mean CDC rate. As such, even if we focus on the end of this interval, these results would still suggest that physician patients fail to receive CDC over 30 percent of the time. The statin therapy results imply a similar conclusion. Accordingly, these results suggest only modest potential impacts, at best, resulting from greater information disclosure to diabetes patients.

Cardiovascular Care—Table 4 also explores the role of information in the receipt of HEDIS-recommended care for cardiovascular disease. In one specification, the dependent variable captures the receipt of the indicated therapy at all over the annual observation period. In another specification, we specify a dependent variable indicative of the degree of adherence to the relevant medication over the observation period, either based on the days supplied over the observation year or the incidence of meeting at least 80% adherence over the year.

Both in the case of the sponsors sample and the dependents sample, we estimate a physician effect that is indistinguishable from zero. For instance, in the sponsors sample, the 95-percent confidence interval for the physician effect on the incidence of at least 80-percent adherence to statin therapy throughout the year spans from a -6.2 to 2.8 percentage-point change, or from roughly -13.3 to 6.0 percent relative to the mean. We find similar effects for the dependents analysis. Thus, even if the true effect were at the upper range of the confidence interval—where physicians attain 80-percent adherence at only a 2.8 percentage-point higher rate than non-physicians—only 49.5 percent of physicians would themselves receive recommended care.

Other Medication Adherence Analysis—In Table 5, we find mixed evidence regarding a physician effect on adherence to medication within the first year following a new diagnosis for hypertension and hypercholesterolemia. For the sample of sponsors with a new diagnosis of hypertension, we estimate a 1.9 percentage-point physician effect on the Medication Possession Ratio (MPR) for hypertensive therapy. This absolute effect corresponds to a modest 2.6 percent increase relative to a baseline rate of adherence of 72.3 percent. The point estimate for this effect in the dependents sample is similar, though the standard errors rise slightly such that the dependents effect is not distinguishable from zero.

For the case of new diagnoses of hypercholesterolemia, we now find a stronger physician effect in the dependents sample than the sponsors sample. In the dependents sample, we estimate a modest 2.2 percentage-point (or 3 percent relative to a baseline rate of 63.8%)

physician effect on the MPR for hypercholesterolemia therapy (though only marginally significant). This estimate is closer to and indistinguishable from zero in the sponsors sample.

The conclusion from this new-diagnosis adherence analysis is much the same as in the settings above. While physician patients may adhere slightly more closely to the recommended care than non-physicians, the results suggest that physicians nonetheless continue to fall far short in meeting those recommendations. For instance, physicians fail to adhere to hypertension medication roughly 26 percent of the time and to hypercholesterolemia medication over 30 percent of the time.

Vaccination / Immunization Analysis—In Table 6 we present estimated physician effects in the case of the childhood immunization analysis. Across several of the individual measures—e.g., chicken pox, pneumococcal conjugate (PCV), and rotavirus—we estimate statistically significant physician effects suggestive of small improvements in compliance with recommended HEDIS guidelines for dependents of physicians relative to dependents of physician officers. We find marginally significant effects in the case of measles/mumps/rubella (MMR) immunizations. With the remainder of measures, we estimate insignificant physician effects that are relatively tightly bound around zero.²² Overall, consistent with the other high-value care analyses considered above, we only find that physicians receive more of recommended immunizations than non-physicians in select cases. Moreover, even in those instances, the magnitudes of the physician effect are modest and the physician patients continue to fall far short of the care recommended by the relevant guidelines.

IV.C. Proportion of Deviation from Recommended Care Explained by Information Effect

Critical to our assessment thus far, however, has not simply been a comparison between physicians and non-physicians. Of relevance is also an independent description of the rate by which physicians adhere to the guideline—in other words, of relevance is also a comparison between the behavior of physicians-as-patients and the care recommended by the relevant guidelines. This comparison is arguably less subject to the concerns over omitted variable biases in estimating the difference in care between physician and non-physician patients.

As discussed throughout in the above analysis, even in those situations where we do estimate a difference between the physician and non-physician group, the gap between the physician's treatment rate and the guideline-recommended rate is still considerable. In Figures 1 and 2, we attempt to summarize the above results graphically and in a way that brings together both of the key forms of comparison driving our analysis—that is, comparisons of physicians and non-physicians and comparisons of physicians against the guideline-recommended care. Figure 1 focuses on the sponsors samples and Figure 2 on the dependents samples. In each figure, we depict a series of bar graphs for a select group of our

²²In the Online Appendix, we demonstrate the robustness of this analysis to a determination of whether the child has received the relevant vaccine by age 3 instead of age 2. As discussed in the Online Appendix, our base specifications exclude certain bases that record vaccinations at very low rates (raising potential under-reporting concerns). In the Online Appendix, we nonetheless present results where we include all bases, even those that report very low rates of immunizations. The conclusion is similar—if anything, the physician group receives slightly more immunizations, but in many case, there is no discernable difference between the physician and non-physician group. While we do not have strong reason to believe why this might happen, we do acknowledge the possibility that reporting issues may differ by the physician status of the child's sponsor.

key outcome measures, the heights of each bar reflecting the degree to which the utilization of the respective measure deviates from the level of care recommended by the relevant guideline (which is 100% utilization in the case of our high-value measures, 0% in the case of pre-operative testing and 15% in the case of the overall cesarean rate—a conservative take on the WHO recommendation). In other words, from the perspective of those advocating for guideline compliance, this bar height represents the degree of over-utilization in the case of low-value care and under-utilization in the case of high-value care. For the purposes of brevity and so as to not crowd this figure, we show this for just one representative measure across each of our classes of samples.

Immediately evident from such graphs—and from the height of the relevant bars—is that the degree of deviations from recommendations—whether for informed or uninformed patients—is considerable across each of the clinical settings. We note that this average degree of guideline deviation among the MHS population is arguably consistent with that found elsewhere in the literature on low- and high-value care, as we discuss in the Online Appendix.

Within each bar graph, we also show the fraction of this deviation-from-recommended-care that our analysis suggests can be explained by the inferior information of the non-physician patient group, using our estimated physician effects for these purposes (focusing only on the estimated point estimates for these purposes). The consistent implication arising from each of the bar graphs across these two figures is that the share of inappropriate treatment that can be explained by information relative to other factors is small in some cases and negligible in others.

Even though the estimated physician effects are small and even if information disclosure may not meaningfully close the gap between prevailing rates and recommended rates of care, one may nonetheless wonder whether information nonetheless imposes a meaningful effect relative to other determinants and sources of variation in such adherence rates. To shed light on this question, we present a series of graphs in the Online Appendix in which we compare the magnitude of the estimated physician effects across our various clinical settings with the magnitude of the variation in the respective adherence measure that we observe across the different catchment areas associated with our sample (which may presumably be driven by variation in certain institutional factors across regions). In each case, the estimated physician effect is negligible in comparison with the extent of regional variation observed (with the possible exception of cesarean delivery).

Overall, the findings summarized in Figures 1 and 2 suggest that much of the explanation behind the deviation from recommended services likely arises from factors other than informational deficiencies of patients. In Part V below, we offer further context regarding what these other factors might be, including the possibility that physician-patients are deferential to their treating providers who may be subject to certain institutional forces and the possibility that our physicians-as-patients may simply disagree with the guidelines.

IV.D. Mechanism Analysis

Before turning to this discussion of the possible sources behind the deviation from recommended practices among informed physicians, we note briefly that we have set forth an analysis in the Online Appendix that attempts to shed light on the mechanisms that may underlie the small differences in adherence between physician and non-physician patients that we do observe. In particular, we assess whether the modest physician effects are due to (1) informed patients making more adherent choices at the time of care selection itself, (2) informed patients choosing treating physicians that are more adherent to the recommendations or (3) informed patients negating the ability of financially-motivated physicians from inducing demand (where induced demand might otherwise lead to guideline non-adherence). The various results set forth in the Online Appendix do not support a robust physician-selection or induced-demand explanation for the observed differences between physician and non-physician patients. While it is difficult to conclude with certainty, this analysis tends to support a story in which informed patients simply make more guideline-adherent choices at the time of care delivery.

V. IMPLICATIONS

Our analysis above has demonstrated that (1) physicians-as-patients fall quite short in adhering to certain recommendations set forth by influential guideline organizations, and (2) that physicians do not appear to adhere to such guidelines at a meaningfully higher rate than a control group of similar non-physician patients. In this section, we discuss what these findings imply for the potential of information disclosure initiatives to induce greater adherence to guidelines and what they imply regarding those factors other than information that may explain adherence failures.

Ultimately, the implications that one draws from our analysis may depend on certain assumptions one makes about the beliefs held by our physicians-as-patients treatment group. Accordingly, we organize this section around two fundamental, alternative assumptions: (1) that the physicians-as-patients group agrees with the merits of the relevant guidelines, (2) that the physicians-as-patients group does not hold firm beliefs in the merits of the guidelines.

V.A. Agreement with Guidelines

Suppose that the physicians-as-patients in our sample believe in the validity of the guidelines that we study. In this instance, our findings suggest that those desiring to encourage greater adherence with the relevant guidelines may find little gains from policy initiatives aimed at disclosing greater amounts of information to patients about their medical conditions, their treatment options and/or about the recommendations set forth in the relevant guidelines. If the vast amounts of information that physicians have is not enough to induce their adherence, information disclosure to lay patients is perhaps not likely to have substantial effects.

If physicians do believe in these guidelines, why aren't they following them in their care? First of all, it may be that patients remain generally deferential to the care recommendations

of their treating physicians, even in the case of near fully-informed patients that believe in the merits of the guidelines. This interpretation would be consistent with other recent findings in the health economics literature regarding the strength of patient deference to treating physicians, including recent research by Chernew et al. (2018) showing that referring physicians are dramatically stronger determinants of where patients receive MRIs relative to patient cost-sharing factors.

Of course, that simply raises the question of why the treating physicians are not following low-value-care guidelines. One natural candidate is the set of incentives posed by fee-for-service reimbursement. As demonstrated in the Online Appendix, we find that the differences in adherence rates between physician- and non-physician patients are similar off-the-base (fee-for-service environment) and on-the-base (salaried environment), suggesting that the small differences observed between these groups likely does not arise from a reduced ability to induce demand among informed patients. Of course, that specific finding does not necessarily speak to whether induced demand contributes to the use of low-value care by physician patients more generally. After all, providers may take advantage of deferential patients and induce demand among informed and uninformed patients alike. To shed light on this matter—even if only suggestively—the Online Appendix also explores whether physician-patients alone are more adherent to the low-value-care guidelines on the base, where induced-demand motivations are weaker. Our findings on this question are mixed, with notably lower rates of c-sections on the base but either similar or higher rates of pre-operative testing on the base. Ultimately, whether induced demand—or related mechanisms such as defensive medicine—are contributing to the utilization of low-value medical services, even among informed patients, is an important area of future and ongoing research.²³

As to why patients may not be following high-value-care guidelines, the question is more difficult. It may be that in this case, the decisions are more under the control of the patient, who can ignore the recommendation of their provider. Patients, even when physicians themselves, may suffer from the same type of behavioral biases that lead the general population to underuse preventive care, perhaps due to the presence of hyperbolic discounting or related behaviors.²⁴

V.B. Lack of Firm Belief in Guidelines

Another possible explanation behind the deviations from recommended care that we observe among physician-patients is that such patients may not hold firm beliefs regarding the validity of the relevant guidelines. This may be due to a situation in which at least some of the physician-patients have been exposed to the guidelines through their presence in the medical community but lack the repeated day-to-day dosage of the information in those

²³Of course, we acknowledge that determinants of low-value-care utilization may go beyond mere financial- and liability-related influences. For instance, providers may be inclined to deliver via cesarean-section on the margin due to certain convenience and scheduling-related considerations (Evans et al. 1984).

²⁴With respect to high-value care, it could also be the case that informed patients believe in the clinical merits of the recommendation, but still find that those clinical merits are outweighed by the particular compliance costs that they may face. In other words, we acknowledge that—even if the guidelines happen to be meritorious from a *clinical* perspective, compliance is not necessarily synonymous with welfare improvement in all cases.

guidelines and the day-to-day experience in the relevant medical specialty that may be necessary to induce their full embracement of the guidelines. Alternatively, this may be due to a situation in which physicians have all of the experience they need to make an informed decision but simply have an affirmative disagreement with the merits of the recommendations.

Of course, the policy implications are unlikely to turn on whether the physician-patients hold firm beliefs in the guideline merits. That is, as long as we assume the physicians have had exposure to the recommendations, their failure to adhere to said guidelines is still informative regarding the possible gains in adherence that may ensue from an information disclosure campaign directed at lay patients. After all, if exposure to the information among our physician-patients is still not enough to induce adherence, it is unlikely that it will for patients generally.

Nonetheless, policy implications aside, the remainder of this sub-section will explore whether some of our observed deviations from recommended care may arise from lack of beliefs in the guidelines. We begin by turning to a new empirical exercise in which we estimate variants of the above specifications that focus our treatment group on physician-patients who themselves practice in the associated clinical setting—i.e., physicians with more day-to-day experience in this arena. Unfortunately, we cannot perform this exercise across all of the clinical samples that we explore above given that, in most such instances, we observe few observations from this more limited physician set. Accordingly, we focus this final exercise on the incidence of chest radiography prior to low-risk surgery – where samples are large and where there is the strongest effect among the relevant sample in Table 3.

In Table 7, we build on the corresponding analysis from Table 3 and present estimated physician effects across 4 different constructions of the physician treatment group, broken up into different bins of the number of procedures in this low-risk sample that the treatment group of physician-patients has performed as providers. As reported in Table 7, we find that the estimated physician effect gets stronger when we focus on a treatment group with greater experience in the relevant clinical setting. When using the most experienced providers as the treatment group, we estimate a physician effect that is over twice as large as the baseline effect from Table 3 and over three times as large as the physician effect when focusing on a treatment group of physician-patients with no experience as a provider in the relevant procedure set. Interestingly, however, even the most informed physician-patients continue to deviate from the guideline at a notable rate—e.g., nearly 4% are receiving unnecessary chest radiographies prior to low-risk surgeries. In other words, when using a treatment group of informed patients that practices (as providers) in the same clinical area as the care choice under investigation, we still find that the gap between observed and recommended care is only closed by less than 20% in the face of superior information.²⁵

²⁵As discussed above, one challenge with our pre-operative testing analysis is that some amount of routine testing may fall within our 30-day pre-surgery window by chance (though likely representing only a minority of the tests we are finding), in which case compliance with the guidelines may not manifest itself as a 0 percent testing rate in the 30-day period prior to a low-risk surgery. This concern is arguably less paramount in the case of chest radiography, given that it is less likely to be performed on a routine basis (as confirmed by the trend depicted in Figure A3 of the Online Appendix). As such, we can say with greater confidence that the prevailing

In part, this analysis offers a welcome rationality check on the link between adherence to the selected guidelines and the underlying information channel. Moreover, the results suggests that some amount of the observed guideline deviations might be due to the fact that some of the physician-patients in our sample lack the day-to-day exposure to and experience with the relevant guidelines and with the practice of medicine in the relevant clinical area.²⁶

Importantly, however, our analysis suggests that this lack-of-routine-exposure explanation can only account for a small share of the observed deviations.

Finally, we consider the possibility that our physician patients simply disagree with the guidelines. While we have tried to minimize this possibility by selecting guidelines with respect to which there appears to be agreement within the guideline-promulgating community, we acknowledge that guideline-expert beliefs may differ from those of practicing physicians. We consider one final empirical exercise to shed light—even if only suggestively—on this possibility. This exercise will build on the chest radiography scenario just discussed in Table 7.

For these purposes, we attempt to avoid conflating disagreement with the guidelines and insufficient experience in the relevant clinical space and thus focus on a set of physician-patients receiving low-risk procedures of the sort that they also implement as providers. We categorize these physician-patients into different bins based on their propensities to perform pre-operative testing *as providers*. Specifically, we group providers based on whether their own patients receive chest radiography prior to low-risk surgery less than 5%, 5–10%, 10–15%, 15–20%, or more than 20% of the time. Across these categories, we then determine the rates by which these physician-patients receive chest radiography prior to low-risk surgery *as patients*. As demonstrated by Table 8, there appears to be little relationship between the rate by which these providers receive chest radiography as patients and the rate by which their own patients receive chest radiography. Even those providers whose own patients receive chest radiography over 20% of the time receive chest radiography prior to their own low-risk surgeries (as patients themselves) at nearly the same rate as those providers whose own patients receive chest radiography less than 5% of the time.

If one thought that affirmative disagreements with guidelines played a large role in explaining why our informed patients were deviating from guidelines, then one might expect to see a stronger relationship between a provider's propensity to call for pre-operative testing on their own patients and their own likelihood of receiving pre-operative testing. Failing to find such a relationship, our results suggests that something other than beliefs and disagreement—including those factors discussed in subpart A above—is likely to explain the

rates of pre-operative chest radiography prior to low-risk surgery is indicative of deviations from the Choosing Wisely recommendation over 4% of the time.

²⁶We acknowledge that an extreme variant of this possibility is one in which our providers simply have not been exposed to the guideline recommendations at all. Unfortunately, we do not have a simple way to distinguish between this possibility and the lack-of-frequent-exposure possibility. We note, however, that complete lack of exposure to the recommendations is perhaps unlikely for most of the clinical settings that we estimate—including such things as childhood vaccinations—given the long-standing messaging surrounding these recommendations. The one scenario where this is perhaps more likely to be true is the pre-operative testing context, considering that the Choosing Wisely campaign did not commence until 2012. Even in this case, however, the pre-operative testing recommendation was not so much a creation of the campaign as a reflection of a nearly universally held, existing view among the large set of medical societies comprising the campaign (Siddaiah et al. 2020).

variation in the degree to which these providers call for their own patients to receive pre-operative chest radiography.²⁷

V.C. Concerns about Generalizability

Despite our efforts to ensure comparability across our treatment and control groups, concerns may remain over the generalizability of our findings, along two dimensions. First, our focus is on a particular patient population, beneficiaries of the MHS. In particular, one may be concerned that this patient population may be fundamentally more adherent to guidelines at baseline, leaving little room for marginal improvement with more information. Mediating against this final concern, however, is evidence that we present in the Online Appendix demonstrating that the rates of adherence to the various guidelines we study are no stronger in this population relative to a corresponding population of individuals with private coverage.²⁸ To be clear, our analysis is not necessarily subject to a related concern that we are limited only to an analysis of care delivered by military providers. On the contrary, while our patient population is limited to the MHS, we observe care delivered both on *and* off the base. With respect to this latter point, the fact that we estimate similar physician effects on and off the base (as discussed in Part IV.D) suggests that our findings arguably generalizes across health care systems.

A second generalizability concern may stem from the fact that we test the impacts of information in an environment characterized by low cost-sharing. Consider, for instance, pre-operative testing. Even though such tests may render insufficient benefits to justify their full social costs and even if patients are aware of this, informed patients may nonetheless undertake the tests anyway given the negligible financial implications. In contrast, informed patients may decide to forego these tests when cost-sharing is more substantial.

This consideration does not necessarily threaten the external validity of our low-value care findings—at least not overly so—in light of the fact that the costs to undergoing low-value care extend well beyond financial outlays. This is clearest in the case of cesarean delivery given the physical implications and necessary recovery period. It is also likely true in the case of pre-operative testing as well. We demonstrate in the Online Appendix that much or most of these instances of pre-operative testing necessitate an entirely separate visit to a medical provider on a separate day, an outcome that likely imposes costs and inconveniences of another nature on the affected patient.²⁹ Accordingly, if we still estimate modest physician effects in the face of low-value care decisions that carry potentially significant

²⁷On a final note, we acknowledge that disagreement with guidelines can take one of two forms: a wholesale disagreement with the merits of the recommendations or a disagreement that the recommendation is relevant for the circumstances of a given patient. Not only have we tried to minimize the scope for wholesale disagreements by selecting guidelines over which the guideline community has reached a consensus, we have also tried to minimize the scope for idiosyncratic disagreements by selecting guidelines whose clinical recommendations are meant to be followed by all of those meeting the inclusion criteria—e.g., HEDIS suggests that all of those with diabetes should indeed receive CDC. We acknowledge that the one setting where this may be a viable concern is the case of pre-operative testing. However, appeasing such concerns, in Table A12 of the Online Appendix (and surrounding text), we demonstrate that the deviations from the pre-operative testing guidelines that we observe do not appear to be strongly related to patient comorbidities and clinical characteristics (with deviations instead perhaps arising from institutional factors).

²⁸Relatedly, we acknowledge that additional generalizability concerns may arise considering that marginal information effects may be stronger among lower income patients (Chen et al. 2019) and considering that our analysis has focused both its treatment and control groups on military officers as opposed to lower pay-ranks within the military. Nonetheless, any such concern is less paramount that it would otherwise be if we compared the care received by a higher-paid civilian physicians with a similarly situated control group of non-physicians.

non-cost-sharing consequences, such modest findings may indeed generalize to populations facing higher levels of cost-sharing. It is important to note that generalizability concerns stemming from the low-cost-sharing environment do not extend to the high-value care analysis.

VI. CONCLUSION

We have explored the care choices of physicians when acting as patients themselves in an effort to estimate the upper bounds of the gains to guideline adherence that may ensue from consumer-driven health care approaches that draw upon greater information-disclosure to patients. For these purposes we compared the clinical choices of physicians to the care recommended by certain consensus guidelines, while also comparing their choices to those of a set of non-physician patients, taking various approaches to ensure comparability between the physician and non-physician groups. Our results consistently suggest that, at most, superior patient knowledge is associated with only modest improvements in adherence to recommended care.

Of course, our results do not necessarily suggest that patient information is universally of minor relevance in adhering to recommended care, such that if we unraveled all of the existing sources of information dissemination to patients, we would observe no meaningful consequences. Rather, our results more conservatively imply that if we were to provide additional information to patients on top of the current information-revelation mechanisms in place, we may only observe minor additional improvements in guideline adherence.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We are grateful to the Military Health Agency for providing the data used in this project, to the NIA for funding, and to Ian Calaway and Tim Simmons for excellent research assistance.

REFERENCES

- Artmann Elisabeth, Oosterbeek Hessel, and van der Klaauw Bas. 2019 “Do doctors improve the health care of their parents? Evidence from admission lotteries,” CEPR Discussion Paper No. DP14061. Associated Press and NORC. 2017 Finding Quality Doctors: How Americans Evaluate Provider Quality in the United States, 7 2017.
- Briesacher Becky, Adrade Susan, Fouayzi Hassan, and Chan Kinwei Arnold. 2008 “Comparison of Drug Adherence Rates among Patients with Seven Different Medical Conditions,” *Pharmacotherapy* 28(4): 437–43. [PubMed: 18363527]
- Bronnenberg Bart J., Dube Jean-Pierre, Gentzkow Matthew, and Shapiro Jesse. 2015 “Do Pharmacists Buy Bayer? Informed Shoppers and the Brand Premium,” *Quarterly Journal of Economics*: 1669–1726.

²⁹Pre-operative tests may also impose certain physical costs. For instance, in the case of chest radiography, there are known-but-minor carcinogenic risks associated with exposure to chest x-rays. This consideration may explain why our estimated physician effects are of greater magnitudes (relative to the respective baseline mean) in the case of chest radiography relative to blood testing, an implication which may further support an information interpretation of our pattern of findings.

- Brot-Goldberg Zarek C., Chandra Amitabh, Handel Benjamin R., and Kolstad Jonathan T.. 2017 “What does a Deductible Do? The Impact of Cost-Sharing on Health Care Prices, Quantities, and Spending Dynamics,” *Quarterly Journal of Economics* 132(3): 1261–1318.
- Bundorf M. Kate, Chun Natalie, Gopi Goda, and Kessler Daniel. 2009 “Do Markets Respond to Quality Information? The Case of Fertility Clinics,” *Journal of Health Economics* 28:718–27 [PubMed: 19328568]
- Chernew Michael, Cooper Zach, Larsen-Hallock Eugene, and Morton Fiona Scott. 2018 “Are Health Care Services Shoppable? Evidence from the Consumption of Lower-Limb MRI Scans,” NBER Working Paper No. 24869.
- Cutler David, Huckman Robert, and Landrum Mary Beth. 2004 “The role of information in medical markets: An analysis of publicly reported outcomes in cardiac surgery,” *American Economic Review* 94:342–346.
- Deyo Richard, Cherkin Daniel, and Ciol Marcia. 1992 “Adapting a clinical comorbidity index for use with ICD-9-CM administrative databases.” *Journal of Clinical Epidemiology* 45:613–19. [PubMed: 1607900]
- Evans MI, Richardson DA, Sholl JS, Johnson BA. 1984 “Cesarean Section. Assessment of the Convenience Factor,” *Journal of Reproductive Medicine* 29(9): 670–6.
- Frakes Michael. 2012 “Defensive Medicine and Obstetric Practices,” *Journal of Empirical Legal Studies* 9(3): 457–81.
- Frakes Michael, and Gruber Jonathan. 2019 “Defensive Medicine: Evidence from Military Immunity,” *American Economic Journal: Economic Policy* 11(3): 197–231. [PubMed: 31632594]
- Frakes Michael, and Gruber Jonathan. 2020 “Defensive Medicine and Obstetric Practices: Evidence from the Military Health System,” *Journal of Empirical Legal Studies*, forthcoming.
- Gruber Jonathan, John Kim, and Mayzlin Dina. 1999 “Physician fees and procedure intensity: the case of cesarean delivery,” *Journal of Health Economics* 18(4): 473–90. [PubMed: 10539618]
- Herz David A., Looman Jan E., and Lewis Sharon K.. 1992 “Informed consent: Is it a myth?,” *Neurosurgery* 30(3):453–8. [PubMed: 1620316]
- Johnson Erin M., and Rehavi M. Marit. 2016 “Physicians Treating Physicians: Information and Incentives in Childbirth,” *American Economic Journal: Economic Policy* 8(1): 115–41
- Kaiser Family Foundation. 2008 2008 Update on Consumers’ Views of Patient Safety and Quality Information.
- Kolstad Jonathan. 2013 “Information and Quality When Motivation Is Intrinsic: Evidence from Surgeon Report Cards,” *American Economic Review* 103(7): 2875–2910.
- Leuven Edwin, Oosterbeek Hessel, and de Wolf Inge. 2013 “The effects of medical school on health outcomes: evidence from admission lotteries,” *Journal of Health Economics* 32(4): 698–707. [PubMed: 23665538]
- Madison Kristin. 2017 “Legal & Policy Issues in Measuring and Improving Quality,” in *THE OXFORD HANDBOOK OF U.S. HEALTHCARE LAW* (Cohen I. Glenn, Hoffman Allison K. & Sage William M. eds.).
- McGlynn Elizabeth A., Asch Steven M., Adams John, Keesey Joan, Hicks Jennifer, DeCristofaro Alison, and Kerr Eve A.. 2003 “The Quality of Health Care Delivered to Adults in the United States,” *New England Journal of Medicine* 348(26): 2635–45.
- McGuire Thomas. 2000 “Physician agency,” in *THE HANDBOOK OF HEALTH ECONOMICS, VOL. 1* (Culyer Anthony and Newhouse Joseph, eds.), pp. 462–536.
- Meisel Alan, and Roth Loren H.. 1983 “Toward an informed discussion of informed consent: a review and critique of the empirical studies,” *Arizona Law Review* 25: 265–346. [PubMed: 11658677]
- Mukamel Dana, Spector William, Jacqueline Zinn, Weimer David, and Ahn Richard. “Changes in Clinical and Hotel Expenditures Following Publication of the Nursing Home Compare Report Card” *Medical Care*. 2010;48(10):869–74 [PubMed: 20733531]
- Pope Thaddeus Mason Pope. 2017 “Certified Patient Decision Aids: Solving Persistent Problems with Informed Consent Law,” *Journal of Law, Medicine & Ethics* 45: 12–40.
- Romano Patrick S., and Zhou Hong. 2004 “Do well-publicized risk-adjusted outcomes reports affect hospital volume?,” *Medical Care*, 42(4):367–77. [PubMed: 15076814]

- Schneider Eric, and Epstein Arnold. 1998 “Use of public performance reports: a survey of patients undergoing cardiac surgery,” *JAMA* 27: 1638–42.
- Schoenfeld Andrew J., Jiang Wei, Harris Mitchell B., Cooper Zara, Koehlmoos Tracey, Learn Peter A., Weissman Joel S., and Haider Adil H.. 2017 “Association Between Race and Postoperative Outcomes in a Universally Insured Population Versus Patients in the State of California,” *Annals of Surgery* 266(2): 267–73. [PubMed: 27501169]
- Schwartz Aaron L., Landon Bruce E., and Elshaug Adam G.. 2014 “Measuring Low-Value Care in Medicare,” *JAMA Internal Medicine*, 174(7):1067–1076. [PubMed: 24819824]
- Siddaiah Harish, Patel Shilpadevi, Shelvan Anitha, Kenneth Philip Ehrhardt Cain W. Stark, Ulicny Kenneth, Ridgell Sasha, Howe Austin, Cornett Elyse M., Uruman Richard D., and Kaye Alan D.. 2020 “Preoperative laboratory testing: Implications of “Choosing Wisely” guidelines,” *Best Practice & Research Clinical Anaesthesiology* (in press).
- Uber Peter A., Angott Andrea M., Zikmund-Fisher Brian J.. 2011 “Physicians Recommend Different Treatments for Patients Than They Would Choose for Themselves,” *Archives of Internal Medicine* 171(7):630–634. [PubMed: 21482835]
- Wang Justin, Hockenberry Jason, Chou Shin Yi, and Yang Muzhe. 2011 “Do bad report cards have consequences? Impacts of publicly reported provider quality information on the CABG market in Pennsylvania,” *Journal of Health Economics* 30(2): 392–407. [PubMed: 21195494]
- Weeks Jane C., Catalano Paul J., Cronin Angel, Finkelman Matthew D., Mack Jennifer W., Keating Nancy L., and Schrag Deborah. 2012 “Patients’ Expectations about Effects of Chemotherapy for Advanced Cancer,” *New England Journal of Medicine* 367: 1616–25
- World Health Organization. 2015 WHO Statement on Caesarean Section Rates.

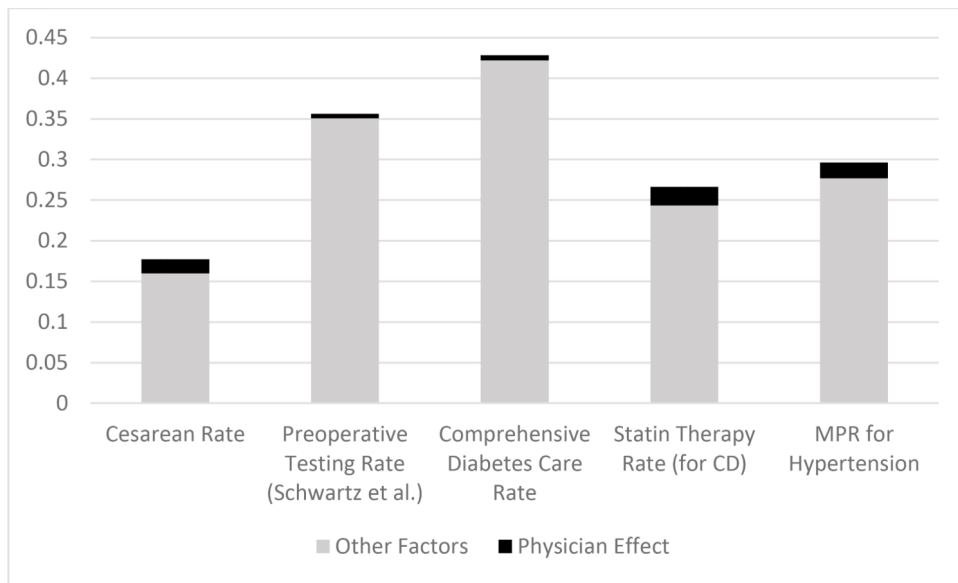


Figure 1. Degree of Over- or Under-utilization of Various Measures explained by Physician Effect—i.e., Information Effect—and by Other Factors (Sponsors Samples)
 Notes: data from the Military Health System Data Repository. The height of each bar represents the degree of over-utilization (for the case of cesarean delivery and preoperative testing prior to low-risk surgeries) and under-utilization (for the case of diabetes and cardiovascular care adherence and medication adherence for hypertension) represented in the respective samples, assessed with reference to non-physician patients. The black portion of this bar represents the degree of this over- or under-utilization that can be explained by the weaker information state of the non-physician patients. The height of the black portion is derived from the estimated physician effects from Tables 2–5.

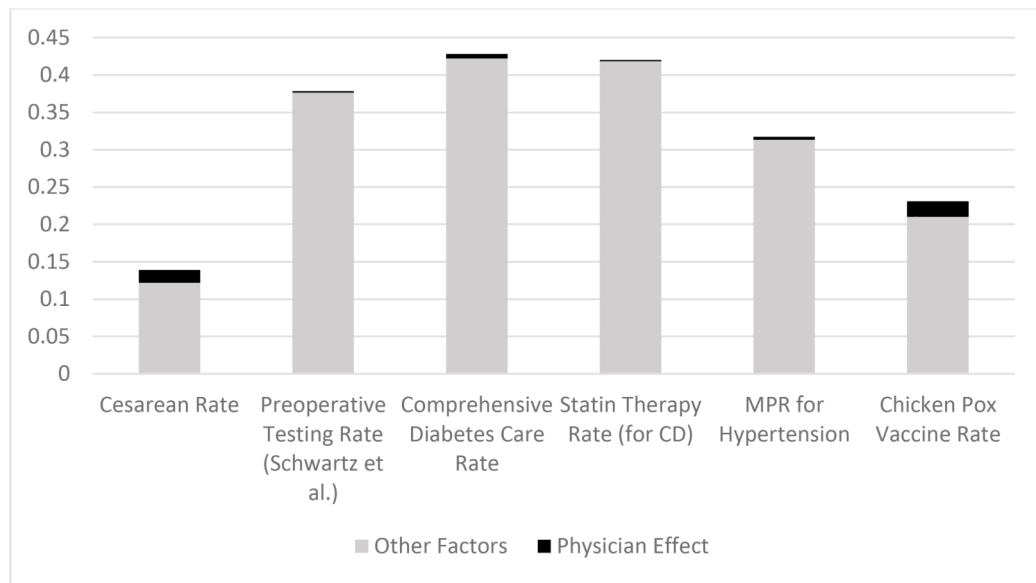


Figure 2.

Degree of Over- or Under-utilization of Various Measures explained by Physician Effect—i.e., Information Effect—and by Other Factors (Dependents Samples)

Notes: data from the Military Health System Data Repository. The height of each bar represents the degree of over-utilization (for the case of cesarean delivery and preoperative testing prior to low-risk surgeries) and under-utilization (for the case of diabetes and cardiovascular care adherence, medication adherence for hypertension, and chicken pox vaccination) represented in the respective samples, assessed with reference to non-physician patients. The black portion of this bar represents the degree of this over- or under-utilization that can be explained by the weaker information state of the non-physician patients. The height of the black portion is derived from the estimated physician effects from Tables 2–6.

Table 1

Covariate Balance between Treatment and Control Groups

	(1)	(2)	(3)	(4)	(5)	(6)
	Physicians (as patients)	Officers	Difference	Dependents of Physicians	Dependents of officers	Difference
Predicted Incidence of Cesarean Delivery	0.300 (0.148)	0.312 (0.165)	-0.015 ^{***} (0.003)	0.266 (0.093)	0.272 (0.099)	-0.006 ^{***} (0.001)
Predicted Incidence of Preoperative Testing	0.184 (0.105)	0.176 (0.111)	0.005 ^{***} (0.001)	0.221 (0.131)	0.219 (0.129)	0.003 (0.002)
Predicted Incidence of Comprehensive Diabetes Care	0.616 (0.214)	0.577 (0.190)	0.039 ^{***} (0.010)	0.517 (0.204)	0.538 (0.189)	-0.020 ^{***} (0.008)
Predicted Incidence of Statin Use	0.758 (0.142)	0.757 (0.126)	0.002 (0.008)	0.608 (0.214)	0.581 (0.194)	0.028 (0.018)
Predicted Medication Possession Ratio for Hypertension	0.713 (0.091)	0.723 (0.095)	-0.011 ^{***} (0.003)	0.685 (0.120)	0.687 (0.129)	-0.002 (0.0050)
Predicted Medication Possession Ratio for hypercholesterolemia,	0.621 (0.089)	0.642 (0.104)	-0.022 ^{***} (0.003)	0.632 (0.112)	0.638 (0.115)	-0.006 (0.004)
Predicted Incidence of DTaP Immunization Compliance	-	-	-	0.658 (0.119)	0.648 (0.126)	0.011 ^{***} (0.003)
Predicted Incidence of IPV Immunization Compliance	-	-	-	0.627 (0.113)	0.631 (0.121)	-0.004 (0.003)
Predicted Incidence of HiB Immunization Compliance	-	-	-	0.687 (0.153)	0.695 (0.156)	-0.008 ^{**} (0.004)
Predicted Incidence of Hepatitis A Immunization Compliance	-	-	-	0.862 (0.100)	0.865 (0.103)	-0.003 (0.002)
Predicted Incidence of Hepatitis B Immunization Compliance	-	-	-	0.739 (0.112)	0.740 (0.122)	0.001 (0.003)
Predicted Incidence of Chicken Pox Immunization Compliance	-	-	-	0.783 (0.096)	0.795 (0.100)	-0.012 ^{***} (0.002)
Predicted Incidence of PCV Immunization Compliance	-	-	-	0.787 (0.112)	0.801 (0.115)	-0.014 ^{***} (0.003)
Predicted Incidence of Rotavirus Immunization Compliance	-	-	-	0.317 (0.209)	0.317 (0.209)	0.000 (0.004)

Notes: standard deviations are reported in Columns 1, 2, 4 and 5. Standard errors are reported in the differencing columns (Columns 3 and 6) and are clustered at the individual beneficiary level. Predicted values of the indicated outcome variables are from regressions of the indicated measure (within the relevant sample) on the key covariates employed in the analysis: patient age-by-sex dummies, patient race dummies, patient pay-grade dummies, previous year RVU, previous year RWP, previous year inpatient days, Charlson comorbidity index, and base-by-year-by-care-location fixed effects. Data are from the Military Health System Data Repository, 2003–2013.

^{***} Significant at the 1 percent level

^{**} Significant at the 5 percent level

^{*} Significant at the 10 percent level.

Table 2

Estimated Physician Effects in Cesarean Utilization

	(1)	(2)	(3)	(4)	(5)
	N: Total	N: Physicians	Baseline Rate	Physicians Coefficient	Standard Error
Panel A. Sponsors Analysis: Comparison of Physician Patients with Non-Physician Military Officers					
<i>Cesarean Incidence (among all deliveries)</i>	18,092	2,395	0.3100	-0.0172	0.0127
<i>Primary Cesarean Incidence (among restricted delivery sample)</i>	13,859	1,784	0.1844	-0.0288 ***	0.0108
Panel B. Dependents Analysis: Comparison of Dependents of Physician Patients with Dependents of Non-Physician Military Officers					
<i>Cesarean Incidence (within all deliveries)</i>	96,436	5,421	0.2716	-0.0173 **	0.0075
<i>Primary Cesarean Incidence (among restricted delivery sample)</i>	73,106	4,078	0.1251	-0.0080	0.0055

Notes: standard errors are clustered at the individual beneficiary level. Each row represents results from a different specification using the specified dependent variable and the specified sample. Primary cesarean specifications exclude multiple births, breech presentations and previous cesarean deliveries from the sample. All specifications control for patient age dummies, patient race dummies, patient pay-grade dummies, previous year RVU, previous year RWP, previous year inpatient days, Charlson comorbidity index, and base-by-year-by-care-location fixed effects. Data are from the Military Health System Data Repository, 2003-2013.

*** Significant at the 1 percent level

** Significant at the 5 percent level

* Significant at the 10 percent level.

Table 3

Estimated Physician Effects on Preoperative Testing Prior to Low-Risk Surgeries

		(1)	(2)	(3)	(4)	(5)
		N: Total	N: Physicians	Baseline Rate	Physicians Coefficient	Standard Error
Panel A. Sponsors Analysis: Comparison of Physician Patients with Non-Physician Military Officers						
<i>Schwartz et al Sample</i>	Chest Radiography	88,560	4,813	0.0959	-0.0215***	0.0039
	Complete Blood Count	88,560	4,813	0.3068	-0.0007	0.0073
	Comprehensive Metabolic Panel	88,560	4,813	0.0940	0.0010	0.0038
	Coagulation Panel	88,560	4,813	0.1041	-0.0076	0.0065
	Any Preoperative Care	88,560	4,813	0.3506	-0.0054	0.0075
Derived Low-Risk Sample	Chest Radiography	1,313,701	48,963	0.0452	-0.0030**	0.0012
	Complete Blood Count	1,313,701	48,963	0.1386	0.0017	0.0020
	Comprehensive Metabolic Panel	1,313,701	48,963	0.0608	0.0023	0.0012
	Coagulation Panel	1,313,701	48,963	0.0241	-0.0008	0.0009
	Any Preoperative Care	1,313,701	48,963	0.1786	0.0002	0.0022
Panel B. Dependents Analysis: Comparison of Dependents of Physician Patients with Dependents of Non-Physician Military Officers						
<i>Schwartz et al Sample</i>	Chest Radiography	99,994	4,614	0.0591	-0.0165***	0.0029
	Complete Blood Count	99,994	4,614	0.3470	0.0004	0.0077
	Comprehensive Metabolic Panel	99,994	4,614	0.0857	-0.0052	0.0038
	Coagulation Panel	99,994	4,614	0.0668	-0.0104***	0.0055
	Any Preoperative Care	99,994	4,614	0.3762	-0.0024	0.0078
Derived Low-Risk Sample	Chest Radiography	609,011	18,980	0.0416	-0.0062***	0.0018
	Complete Blood Count	609,011	18,980	0.1764	-0.0070**	0.0036
	Comprehensive Metabolic Panel	609,011	18,980	0.0878	-0.0060**	0.0025
	Coagulation Panel	609,011	18,980	0.0351	-0.0032*	0.0020
	Any Preoperative Care	609,011	18,980	0.2188	-0.0124***	0.0040

Notes: standard errors are clustered at the individual beneficiary level. Each row represents results from a different specification using the specified dependent variable and the specified sample. All specifications control for patient age-by-sex dummies, patient race dummies, patient pay-grade dummies, previous year RVU, previous year RWP, previous year inpatient days, Charlson comorbidity index, and base-by-year-by-care-location fixed effects, along with fixed effects indicating the relevant surgery within the respective sample of low-risk surgeries. Data are from the Military Health System Data Repository, 2003–2013.

*** Significant at the 1 percent level

** Significant at the 5 percent level

* Significant at the 10 percent level.

Table 4

Estimated Physician Effect on High-Value Diabetic and Cardiovascular Care

	(1)	(2)	(3)	(4)	(5)
	N: Total	N: Physicians	Baseline Rate	Physicians Coefficient	Standard Error
Panel A. Sponsors Analysis: Comparison of Physician Patients with Non-Physician Military Officers					
<i>HEDIS Diabetes Measures</i>					
hbA1c Testing	69,343	1,095	0.7631	0.0002	0.0180
Eye Exam	69,343	1,095	0.7855	-0.0155	0.0190
Attention to Nephropathy	69,343	1,095	0.8550	0.0125	0.0152
Comprehensive Diabetes Care	69,343	1,095	0.5778	0.0061	0.0218
Statin Therapy	38,623	623	0.7084	0.0469	0.0346
<i>HEDIS Cardio Measures</i>					
Any Statin Therapy	62,838	1,217	0.7565	0.0228	0.0220
Days Supplied of Statin Therapy	62,838	1,217	225.4458	2.1289	7.5821
Incidence of at least 80% Adherence	62,838	1,217	0.4666	-0.0170	0.0225
Panel B. Dependents Analysis Comparison of Dependents of Physician Patients with Dependents of Non-Physician Military Officers					
<i>HEDIS Diabetes Measures</i>					
hbA1c Testing	73,122	1,592	0.7310	-0.0022	0.0185
Eye Exam	73,122	1,592	0.7862	0.0148	0.0173
Attention to Nephropathy	73,122	1,592	0.8048	-0.0106	0.0163
Comprehensive Diabetes Care	73,122	1,592	0.5380	0.0064	0.0197
Statin Therapy	32,677	658	0.6121	0.0224	0.0363
<i>HEDIS Cardio Measures</i>					
Any Statin Therapy	26,099	503	0.5814	0.0017	0.0408
Days Supplied of Statin Therapy	26,099	503	163.2406	-1.9929	12.9882
Incidence of at least 80% Adherence	26,099	503	0.3193	-0.0141	0.0324

Notes: standard errors are clustered at the individual beneficiary level. Each row represents results from a different specification using the specified dependent variable and the specified sample. All specifications control for patient age-by-sex dummies, patient race dummies, patient pay-grade dummies, previous year RVU, previous year RWP, previous year inpatient days, Charlson comorbidity index, and base-by-year-by-care-location fixed effects. Data are from the Military Health System Data Repository, 2003–2013.

*** Significant at the 1 percent level

** Significant at the 5 percent level

* Significant at the 10 percent level.

Table 5

Estimated Physician Effect on Medication Adherence during First Year of Diagnosis for Hypertension and Hypercholesterolemia

	(1)	(2)	(3)	(4)	(5)
	N: Total	N: Physicians	Baseline Rate	Physicians Coefficient	Standard Error
Panel A. Sponsors Analysis: Comparison of Physician Patients with Non-Physician Military Officers					
Medication Possession Ratio: Hypertension	39,435	1,018	0.7229	0.0191 **	0.0093
Medication Possession Ratio: Hypercholesterolemia	52,017	1,322	0.6423	0.0030	0.0084
Panel B. Dependents Analysis: Comparison of Dependents of Physician Patients with Dependents of Non-Physician Military Officers					
Medication Possession Ratio: Hypertension	23,856	595	0.6868	0.0040	0.0137
Medication Possession Ratio: Hypercholesterolemia	30,014	719	0.6381	0.0218 *	0.0120

Notes: standard errors are clustered at the individual beneficiary level. Each row represents results from a different specification using the specified dependent variable and the specified sample. All specifications control for patient age-by-sex dummies, patient race dummies, patient pay-grade dummies, previous year RVU, previous year RWP, previous year inpatient days, Charlson comorbidity index, and base-by-year-by-care-location fixed effects. Data are from the Military Health System Data Repository, 2003–2013.

*** Significant at the 1 percent level

** Significant at the 5 percent level

* Significant at the 10 percent level.

Table 6

Estimated Physician Effect on Childhood Immunization Rates

	(1)	(2)	(3)	(4)	(5)
	N: Total	N: Physicians	Baseline Rate	Physicians Coefficient	Standard Error
DTaP	34556	2030	0.67	-0.0132	0.0109
IPV	34556	2030	0.64	0.0149	0.0111
MMR	34556	2030	0.80	0.0166*	0.0088
HiB	34556	2030	0.70	0.0163	0.0100
Hepatitis A	34556	2030	0.85	0.0109	0.0076
Hepatitis B	34556	2030	0.72	0.0140	0.0097
Chicken Pox	34556	2030	0.79	0.0209**	0.0091
PCV	34556	2030	0.80	0.0187**	0.0088
Rotavirus	34556	2030	0.32	0.0199**	0.0098

Notes: standard errors are clustered at the individual beneficiary level. Each row represents results from a different specification using the specified dependent variable and the specified sample. All specifications control for patient age-by-sex dummies, patient race dummies, patient pay-grade dummies, previous year RVU, previous year RWP, previous year inpatient days, Charlson comorbidity index, and base-by-year-by-care-location fixed effects. Data are from the Military Health System Data Repository, 2003–2013.

*** Significant at the 1 percent level

** Significant at the 5 percent level

* Significant at the 10 percent level.

Table 7

Estimated Physician Effect on Chest Radiography Prior to Low-Risk Surgery (from Derived Low-Risk Surgery Sponsors Sample) across Different Constructions of Physician-Patient Group

	(1)	(2)	(3)	(4)	(5)
	N: Total	N: Physicians	Baseline Rate	Physicians Coefficient	Standard Error
Physician-Patient Group: No Restrictions (Baseline, Replicating Corresponding Row from Table 3)	1,313,701	48,963	0.0452	-0.0030**	0.0012
Physician-Patient Group: Those Performing (as Providers) None of Procedures in Sample	1,292,613	27,875	0.0453	-0.0019	0.0016
Physician-Patient Group: Those Performing (as Providers) between 0 and 100 of Procedures in Sample	1,281,251	16,513	0.0452	-0.0036**	0.0018
Physician-Patient Group: Those Performing (as Providers) at least 100 of Procedures in Sample	1,269,313	4,575	0.0451	-0.0077**	0.0031

Notes: standard errors are clustered at the individual beneficiary level. Each row represents estimated physician effects from a different specification using the specified construction of the physician treatment group, where groups vary depending on the volume of the procedures in this low-risk surgery sample that were presided over—as providers—by the relevant physician-patient. All specifications control for patient age-by-sex dummies, patient race dummies, patient pay-grade dummies, previous year RVU, previous year RWP, previous year inpatient days, Charlson comorbidity index, and base-by-year-by-care-location fixed effects. Data are from the Military Health System Data Repository, 2003–2013.

*** Significant at the 1 percent level

** Significant at the 5 percent level

* Significant at the 10 percent level.

Table 8

Relationship between Physicians-as-Patients use of Chest Radiography Prior to Low-Risk Surgery for themselves and the Rate by Which their Own Patients Receive Pre-operative Chest Radiography

	Percentage of Physician-Patients within indicated Provider Bin Receiving Chest Radiography prior to their own Low-Risk Surgery (as a patient)	N	P-value of test of no difference in Chest Radiography Rate for Provider-as-Patient at indicated Bin Relative to Bin 1
Provider Bin 1: Providers whose Own Patients' Pre-operative Chest Radiography Rate is < 5%	0.039 (0.193)	7,109	-
Provider Bin 2: Providers whose Own Patients' Pre-operative Chest Radiography Rate is 5–10%	0.040 (0.196)	3,247	0.138
Provider Bin 3: Providers whose Own Patients' Pre-operative Chest Radiography Rate is 10–15%	0.052 (0.222)	1,711	0.245
Provider Bin 4: Providers whose Own Patients' Pre-operative Chest Radiography Rate is 15–20%	0.053 (0.223)	816	0.363
Provider Bin 5: Providers whose Own Patients' Testing Rate is 20%+	0.052 (0.222)	2,257	0.199

Notes: standard errors are reported in parentheses. The reported statistics are from the sponsors, derived low-risk surgery sample. Specifically, the statistics are from the set of those physician-patients within this sample that have performed (on their own patients) at least one of the low-risk procedures in the sample. Within this group, each row presents statistics from a sub-sample of these physicians according to the rate by which their own patients receive pre-operative chest radiography. Data are from the Military Health System Data Repository, 2003–2013.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript