Impact of Policy

The Impact of Medicare Coverage Policies on Health Care Utilization

Susan Bartlett Foote, Beth A. Virnig, Robert J. Town, and Lacey Hartman

Objective. To determine whether Medicare coverage policies affect utilization of services in Medicare.

Data Sources. We constructed an analysis data set for eight different procedures using secondary data obtained from Medicare claims (1999–2002) and Medicare coverage policies posted on Center for Medicare and Medicaid Services website.

Study Design. We analyzed the impact of coverage policies using difference-in-difference approach in a regression framework.

Principal Findings. We found that in only one case (transesophageal echocardiography) out of eight did utilization change (reduced by 13.6 percent) after the effective date of the local policies. There is no systematic pattern that policies affect utilization, and the type of coverage policy does not seem to play an important role in its impact.

Conclusions. Coverage policies have the potential but do not consistently impact utilization as policy makers intend and expect them to do. These findings raise significant policy questions about the effectiveness of Medicare coverage policies, which deserve further study.

Key Words. Medicare coverage, Medicare services, local coverage decisions (LCDs), Medicare utilization

Medicare finances health care for nearly 40 million elderly and disabled Americans. The Medicare statute authorizes reimbursement only for items or services that are “reasonable and necessary.” Medicare administrators have struggled to interpret and apply the “reasonable and necessary” standard (Foote 2002). Over the years, Medicare has developed a body of coverage policies, formal evidence-based assessments of selected items or services, to clarify if these specific services qualify for payment, and if so, under what clinical conditions (Foote, Halpern, and Wholey 2005). While only a small percentage of Medicare services are subject to coverage policies, these policies represent a potentially important tool to manage utilization. However, little is

known about if and how coverage policies actually impact provider behavior. In this paper, we address two related questions: (1) Do Medicare's coverage policies change provider behavior? (2) Do coverage policies increase or decrease utilization depending on their type?

BACKGROUND ON MEDICARE COVERAGE POLICIES

The Medicare statute defines the program's benefits, specifically excluding certain categories, such as personal care items or hearing aids, and explicitly including other broad categories, such as physician and hospital services. Within the categories, however, the statute provides that Medicare will only pay for items and services considered "reasonable and necessary." The original statute delegated to the private local contractors—Part A Fiscal Intermediaries (FIs) and Part B Carriers—the responsibility to process claims. By 2003, there were approximately 40 local contractors.¹ Medicare payment is based on a complex set of payment methodologies that depend upon a standard set of procedure and diagnostic codes. The contractor reviews each coded claim to determine whether it will be paid.

In Medicare's first decade, interpretation of the reasonable and necessary provision presented few problems. Evaluation of items and services occurred informally; contractors and physicians mediated disputes on a case-by-case basis (Foote 2002). In general, experimental products and services were not considered "reasonable and necessary." Food and Drug Administration (FDA) approval signaled that a device or drug was no longer experimental. Procedures, however, do not receive FDA review, making the determination of experimental status uncertain. The advent of new, complex, and expensive technologies, such as heart transplantation in the late 1970s, prodded the Health Care Financing Administration (HCFA), predecessor agency to Center for Medicare and Medicaid Services (CMS), to develop specific limitations and conditions on a few high-profile technologies (Foote 2002). Coverage policy was born.

Address correspondence to Susan Bartlett Foote, J.D., M.A., Professor, Division of Health Policy and Management, School of Public Health, University of Minnesota, 420 Delaware Street, S.E., MMC 729, Minneapolis, MN 55455. E-mail: foote003@umn.edu. Beth A. Virnig, Ph.D., Associate Professor, and Robert J. Town, Ph.D., Associate Professor, are with the Division of Health Policy and Management, School of Public Health, University of Minnesota, Minneapolis, MN. Lacey Hartman, M.P.P., Senior Research Scientist, is with the Minnesota Department of Health, St. Paul, MN.

Over time, CMS developed explicit administrative procedures for national coverage determinations (NCDs) and currently issues approximately 20/year (McClellan and Tunis 2005; Neumann et al. 2005). CMS can trigger an NCD based on its own internal judgment or upon request of an external party. The final NCD is transmitted to local contractors for implementation. Although CMS can issue noncoverage NCDs, such as acupuncture in 2004 and artificial lumbar spinal disk replacement in 2006, it rarely does so. The majority of NCDs establish evidence-based conditions of use.

Contractors acquired authority in 1990 to issue local coverage determinations (LCDs) applicable in their own jurisdictions if no NCD applies (Foote 2003). Contractors independently decide if and when to develop policies, often in consultation with their coverage advisory committees (CACs). LCDs must follow specific procedures for development, and conform to a prescribed format, including a description of the covered service, diagnosis codes to support medical necessity, indications of coverage, reasons for denial, supporting data, coding and documentation requirements and the effective date. Some LCDs cover new technologies, signaling providers that reimbursement is available under certain conditions. A majority of LCDs focus on appropriate utilization of widely used services that are subject to overuse such as routine chest X-rays. There are over 2,000 LCDs across the contractor regions, with considerable variation in the number of LCDs and their effective dates, among other differences in each region (Foote et al. 2004).

In spite of the availability of NCDs and LCDs, most claims do not have an applicable coverage policy. If there is no policy, contractors approve or deny claims based on information in the claims form. Once a claim is approved, payment flows based on the application of CMS' complex payment formulas.

RESEARCH OBJECTIVES

Our goal was to estimate the effect of coverage policies on the provision of services to Medicare beneficiaries. Using LCDs promulgated at different times in different contractor jurisdictions, we can exploit the variability across contractors to assess the impact of coverage decisions on utilization before and after the policies were in place. We track eight case studies that reflect policy types and use a straightforward difference-in-difference estimator as an empirical approach.

Classification by Policy Type

A formal classification process for these LCDs is described elsewhere (Foote, Wholey, and Halpern 2006). Briefly, using two physician consultants, policies were classified into three types. New technology (NT) policies provide guidance for, and limitations on, the use of new clinical interventions. These policies set the conditions under which the NT will be considered reasonable and necessary (i.e., if provided in certain settings, by specifically trained providers, or after alternatives had failed) and send signals to providers that the technology is now "covered" if provided as outlined. Technology extension (TE) policies provide for coverage for new uses of procedures already covered for other uses. This type of policy sends signals that new uses of approved technologies are also "covered." Another set of policies focuses on widely used services that providers know are "covered." Utilization management (UM) policies circumscribe the clinical indications for widely diffused procedures and signal that contractors are alert to misuse and overuse (Foote, Wholey, and Halpern 2006).

Hypotheses

We hypothesize that NT policies and the related TE policies can increase utilization by signaling coverage for new technologies or new uses. However, the interrelationship between coverage decisions and utilization can be significantly more complex. As we have noted, many procedures can be reimbursed without a coverage policy using existing codes, temporary codes, or case-by-case contractor review. Thus, subsequent NT or TE policy may limit the use based on clinical evidence.

UM policies that specifically limit utilization of widely used procedures may affect different dimensions of utilization. For example, toenail debridement or chest X-ray policies put a cap on the number of procedures a patient receives in a given period. For procedures likely to be repeated many times per month or quarter, implementation of this type of policy is most likely to shift the range of number of procedures by lowering the maximum number received in a given time period. Because there are many theoretically grounded reasons to expect that coverage policies could increase or decrease utilization we test two-tailed hypotheses.

DATA/METHODS

We use data from two sources: Medicare claims from 1999 to 2002 and the LCDs posted to the CMS website on May 31, 2001, www.cms.hhs.gov/mcd.

CMS Website

Coverage policies were obtained from a download of all posted policies on our snapshot date (May 31, 2001). To confirm that our database contains all the LCDs issued on our snapshot date, we downloaded “missing” policies on April 2, 2002 (Foote, Wholey, and Halpern 2006).

Selection of Case Studies

We considered a case to be a group of policies focused on a particular technology or procedure developed and implemented by different local Carriers. We used only Carrier policies because Carrier jurisdictional boundaries are fixed, in contrast to FIs who are selected by hospitals (Foote et al. 2004). Cases had to meet several criteria to be considered. First, we required least 15 different local policies per case at the outset of our baseline year. This allows for enough variation across Carriers between those with and without a policy over time. Second, we included all types of policies (NT, TE, and UM) in order to measure potential differences in effect due to policy type. There were 80 technologies or procedures that met our criteria. We selected 11, including four NT, five TE, and two UM and obtained from CMS all Carrier claims for the period 1999–2002 that contained one or more associated Common Procedural Terminology (CPT)/Healthcare Common Procedure Code System (HCPCS) codes associated with the 11 case studies. After receipt of the data, we excluded three case studies—carcinoembryonic antigen, percutaneous transluminal balloon angioplasty, and insertable loop recorders—because we determined that we had not requested from CMS all relevant CPT codes. Being unable to assess the bias associated with partial ascertainment of utilization, we elected not to include the cases in our analysis.

Table 1 provides a summary of dates on which the underlying drugs or devices embedded in the procedures received FDA approval, and the number of policies at the beginning and end of the sample period. Notice that there is significant variation in the issuance of coverage policies across Carriers over our sample period. The number of changes in policy status ranges from a minimum of five for toenail debridement to 13 for EPO treatments. The policies are briefly described below.

NTs

Helicobacter Pylori Breath Test (HBT). *H. pylori* is a bacterium discovered in 1982. It is found in the gastric mucus layer of the stomach lining and is responsible for 90 percent of duodenal and 80 percent of gastric ulcers, and

Table 1: Approval Dates, Coverage Policy Summaries, and Number of Coverage Policies

<i>Procedure</i>	<i>Year FDA Approved/ Introduced</i>	<i>Number of Local Coverage Policies at Beginning of Sample</i>	<i>Number of Local Coverage Policies at End of Sample</i>
New technologies			
Urea breath test for <i>Helicobacter pylori</i> (HBT)	1996	21	27
Deep brain stimulation (DBS)	1997	23	31
Ocular photodynamic therapy (OPT)	2000	16	28
Technology extension			
Epoetin (EPO)	1989 for ESRD; 1993 extension to non-ESRD	33	46
Transesophageal echocardiography (TEE)	1976	33	40
Abdominal and pelvic CT scans	1970s	20	28
Utilization management			
Toenail debridement (TD)	Pre-1970s	29	34
Chest X-rays	Pre-1970s	35	42

ESRD, end stage renal disease; CT, computed axial tomography.

is also linked to chronic gastritis, gastric cancer, and gastric lymphoma. In contrast to invasive tests to detect the presence of the bacteria, such as biopsy sampling and esophagogastroduodenoscopy, the breath tests are noninvasive and use analysis of breath samples. FDA approved the first breath test in 1996. The LCDs in the case study describe limited circumstances in which the breath test is the clinically appropriate choice, such as after recurrence of symptoms following appropriate treatment, and if repeated within 30 days, or as a screening test without associated symptoms.

Deep Brain Stimulation (DBS). DBS is a neurosurgical procedure where deep brain structures are stimulated via implanted electrodes. The therapeutic maneuver targets selected regions of the thalamus or the basal ganglia to control essential or Parkinsonian tremors. The FDA approved the relevant device in 1997. The LCDs describe the preconditions necessary for patient selection for DBS, primarily that the patient is refractory to drug therapy, and

contraindications for this invasive procedure. The policies make clear that careful patient selection is necessary to justify this effective, but highly invasive intervention.

Ocular Photodynamic Therapy (OPT). Age-related macular degeneration is the leading cause of new severe vision loss in patients over age 50. Little is known about the condition and there are no effective cures. OPT treats certain types of subretinal neovascularization that is a secondary effect of visual loss. Verteporfin, a photosensitivity drug agent, is administered intravenously and then low-energy laser light is directed at the lesion to activate the drug to seal the leaking vessels. The FDA approved Verteporfin in 2000 and thus our analysis sample for OPT begins in 2001. The LCDs describe the patient selection criteria required to qualify for treatment to be sure that only patients with conditions where there is evidence of effectiveness receive the treatment.

TE

Transesophageal Echocardiography (TEE). Cardiac ultrasound provides structural, functional, and hemodynamic information, including anatomic information of the proximal great vessels. Transthoracic echocardiography (TTE) is noninvasively applied to the anterior thorax to examine the heart. FDA approved the device in 1984. TEE has been available since 1976, but was made more effective based on significant advances in imaging technology. TEE positions an ultrasound generator in the esophagus to obtain additional information, but is more invasive than TTE with higher potential morbidity. The LCDs clarify that TEE is not medically necessary if TTE is technically adequate, and that TEE is covered only when it significantly augments TTE or contributes to clinically relevant decisions. The LCDs enumerate the many conditions for which TEE is considered clinically contraindicated.

Abdominal and Pelvic Computerized Axial Tomography (CT) and Magnetic Resonance Imaging (MRI) Scans. CT is a form of X-ray, which creates cross-sectional images via computer. MRI uses a powerful and highly uniform static magnetic field to produce images on multiple planes. FDA approved CT scans in the 1970s and MRI machines were approved in the mid-1980s for limited uses. Over time, FDA expanded indications for use. The LCDs in this case require physicians to document their choice of imaging modality,

reasons for combining modalities, and frequency of imaging of the abdomen and/or pelvis. The LCDs demonstrate concern that there is overuse of these tools.

Epoetin α (EPO) for Non-End-Stage Renal Disease (ESRD) Uses. Erythropoietin is naturally produced in the kidneys to stimulate red blood cell production in the bone marrow. EPO is a glycoprotein produced by recombinant DNA technology. EPO received FDA approval for ESRD patients in 1989. In 1993 EPO was approved for non-ESRD patients who had other conditions, such as chronic anemia due to HIV or for certain chemotherapy patients. The LCDs define all the clinical preconditions necessary to qualify for non-ESRD EPO.

UM

Toenail Debridement. Toenail debridement involves the reduction of a thickened dystrophic nail resulting from severe systemic conditions using specialized equipment such as forceps or a rotary drill. There is one HCPCS code for debridement of one to five nails and a second code for six or more, suggesting evidence of repetitive use. The LCDs include extensive documentation requirements, such as the tools used, the number of nails treated, the dates of treatment, names of treating physicians, and often require that the patients' charts include dated photographs of affected nails. These policies clearly seek to reduce utilization.

Chest X-rays. Radiologic examination of the chest (chest X-ray) facilitates the detection, diagnosis, staging, and management of pathophysiological processes. The LCDs in this case include numerous HCPCS codes that support medical necessity sufficient to justify payment. The policies define coverage limitations, such as screening X-rays, routine preoperative X-rays, as well as limitations on interpretation (one per study), frequency, and type of facility performing the service.

Claims Files

We obtained all physician supplier/carrier records containing the relevant HCPCS codes from 1999 to 2002. The Medicare physician supplier/carrier file contains claims, which are further separated into line items, for procedures billed to Medicare. For each procedure, we identified paid line items. For procedures that can be billed separately for the professional and technical

components (e.g., one bill for an X-ray and another for the reading of the X-ray by a radiologist) or for the two components combined, we counted only one procedure per day to avoid over-counting by mistakenly considering two components of a single procedure to be two procedures. For EPO, we used the unit field in the line-item to count the number of units received per dose. The individual claims were then summarized into person-level counts per month or quarter. Table 2 includes information for each procedure on how the claims were summarized.

Dependent Variables and Data Frequency

Table 2 provides the means (untransformed) and standard deviations of the dependent variable for each of our procedures. The NT and TE procedures

Table 2: Description and Summary Statistics of Sample Procedures

<i>Procedure</i>	<i>CPT/HCPCS Codes</i>	<i>Data Freq. Type</i>	<i>Mean Dep. Variable (SD)</i>
New technologies			
Urea breath test for <i>Helicobacter pylori</i> (HBT)	83013, 83014, 78267, 78268	M	74.9
Deep brain stimulation (DBS)	E0752, E0756, 61862, 61880, 61886, 61888, 95961, 95962, 61885	Ln(N)	(135.1)
Ocular photodynamic therapy (OPT)	Q3013, J3490, J3395, 67221, 67225, G0184	M	11.7
		Ln(N)	(12.2)
		M	163.0
		Ln(N)	(166.8)
Technology extension			
Epoetin (EPO)	Q0136	Q	1,452.2
		Ln(N)	(1,767.2)
Transesophageal echocardiography (TEE)	93312, 93313, 93314, 93318	M	246.5
Abdominal and pelvic CT scans	72192, 72193, 72194, 74150, 74160, 74170	Ln(N)	(228.2)
		Q	13,797
		Ln(N)	(11,982)
Utilization management			
Toenail debridement (TD)	11720, 11721	Q	2.04
		99th	(0.77)
Chest X-rays	71010, 71015, 71020, 71021, 71022, 71023, 71030, 71034	M	7.17
		99th	(1.95)

Notes: “Data Freq” refers to the observational frequency of the data with “M” denoting monthly and “Q” denoting quarterly. “Type” refers to the formulation of the dependent variable with “Ln(N)” denoting the logarithm of the number of procedures and “99th” denoting the number of procedures at the 99th percentile. “Mean Dep. Variable” refers to the untransformed mean of the dependent variable. The sample timeframe for all procedures, except OPT, begins in 1999 and end in 2002. The sample time frame for OPT is 2001–2002.

CT, computed axial tomography.

range from the rare to the common. For example, for the average Carrier, DBS is performed approximately only 12 times a month while CT scans are performed over 13,000 times a month. For the UM procedures, the 99th percentile of the per-patient uses range from approximately 2.0 per quarter for toenail debridement to 7.2 per month for chest X-rays. We use the 99th percentile because the UM policies are designed to change extreme not average behavior. Thus, we do not expect major changes in mean usage in response to the policy.

The dependent variable in our analysis is medical care utilization—but we quantify the utilization differently for each of our cases. For NT and TE procedures, the dependent variable is based on the number of procedures processed by the Carrier within the time window (either quarter or month depending on the procedure). We use a logarithmic transformation to the dependent variable to make it scale-free in the fixed effects framework and to allow for meaningful comparisons of changes in utilization across Carriers. The structure of the claims data does not allow us to identify the precise at-risk population needed to calculate a procedure rate. We do, however, include both Carrier fixed effects and Carrier-specific time trends to control for differences in the mean and trend over time in the local providers' propensity to perform procedures.

For utilization management procedures, the dependent variable is the 99th percentile of the distribution of the procedure frequency (conditional on a beneficiary receiving the procedure) in a month (X-ray) or quarter (toenail debridement).² For example, for the Carrier in Arizona, in January 2000, the 99th percentile of the number of X-rays performed for a beneficiary is nine.

ANALYTIC APPROACH

To understand the relationship between the implementation of an LCD and Medicare utilization, we use a difference-in-difference approach. We compare the change in utilization for Carriers who issued a coverage policy to the change in utilization for those Carriers who did not issue a policy. For example, our dependent variable for NT and TE polices is the logarithm of the number of procedures performed in a Carrier's geographic region in a given quarter. The dependent variable for UM procedures is the measure of utilization is the number of procedures at the 99th percentile of the procedure distribution. In our regression specification, local jurisdictions are indexed by i and time is indexed by t . For each of the six NT and TE procedures that we

study here, we will estimate the parameters from the following regression equation:

$$y_{it} = \alpha_i + \delta_t + \beta A_{it} + \gamma_i t + \varepsilon_{it} \quad (1)$$

In this specification, the parameter of interest is β that measures the impact of an approval by contractor i in period t on the measure of utilization, y_{it} . The variable A_{it} takes on the value of 1 if the procedure has been approved by local Carrier i in period t and 0 otherwise. Thus, β is the percentage increase in utilization that results from the approval of the procedure by the local Medicare Carrier. Carrier-fixed effects, α_i , capture differences in procedure usage across jurisdictions and δ_t are period fixed effects and controls for the national utilization trends over time. The specification includes Carrier-specific time trends, γ_i , allowing for the possibility that there are different trends in utilization across the Carriers that might be correlated with the implementation of a policy. The error term, ε_{it} , is mean zero and may have an autocorrelation structure. The parameters are estimated using a standard, mean differenced fixed effects estimator clustering the standards errors at the Carrier (Bertrand, Dufflo, and Mullainathan 2004).

The model specification assumes that the impact of the policy is the same across Carriers. Previous analysis has shown NT policies are virtually identical and that there is significant overlap in content of all policies across Carriers (Foote, Halpern, and Wholey 2005). Thus, in our context this assumption is appropriate.

Identification

We identify the impact by comparing changes in utilization between those Carriers that implemented a policy and those Carriers that did not for a given procedure. This approach controls for differences in average utilization across Carriers that may be correlated with the issuance of a policy. We include Carrier-specific time trends so our approach generates consistent estimates even if there are differences in the trends of utilization across Carriers that might be correlated with the implementation of the policy. Such a correlation might happen if an increase in utilization causes the Carrier to issue a policy clarifying criteria for use of the procedure.

Table 3 lists the specific Carriers that implemented new policies for each of the procedures over the span of our data. Carriers are listed by state in which the Carrier operates (Carrier geographic boundaries correspond to states except as noted). The specific Carriers that issued new policies over the sample period are different across the procedures. Over 40 different Carriers issued

Table 3: Distribution of New Medicare Coverage Decisions by Procedure and Carrier

New Technologies		Technology Extension			Utilization Management		
HBT	DBS	OPT	EPO	TEE	CT Scans	TD	X-Rays
IA, DC, UT, IL, MI, Northern CA (31140)	LA, IN, MO, NE, KY, PA, MA, NH	FL, MT, SD, AK, AZ, HI, OR, WA, CO, WY, IA, VA	AL, AR, CT, IN, KY, NY (801), VA, TN, Northern CA (31140), ME, MA, VT, Southern CA (31146)	IN, KY, MT, PA, OH, WV, Northern CA (31140)	AL, GA, IN, KY, TX, MD, DE, DC	NY (801) SD, ND, CO, WY	IN, KY, NY (801), NY (803), DE, DC, Northern CA (31140)

Note: Most states are covered by just one Carrier. For the states that have multiple Carriers we list the state and Carrier code. Table 1 defines the abbreviations.

policies covering one of the eight procedures during our study time frame and southern California, the most active Carrier, issued policies for four procedures.

Carriers have discretion as to whether or when to develop LCDs, based on their assessment of utilization problems in their region or by physician or manufacturer requests. Carriers also vary significantly in the number of claims processed, resources for policy development, total policies issued, and other factors (Foote, Halpern, and Wholey 2005). Our empirical strategy could lead to misleading inferences if there are Carrier-specific unobservable factors that are systematically correlated with the issuance of a policy. While it is difficult to directly test that proposition, the fact that many different Carriers issued policies and that the most active Carrier issued policies for only half of our procedures is strong, albeit indirect, evidence that Carrier-specific unobservables do not bias our results. If utilization shocks are correlated with a policy, it seems likely that they would be correlated across procedures. That, in turn, would imply that the Carriers that issue policies would be very similar across the procedures. That is not the case here. However, in an attempt to control, at least in part, for the correlation between use and the issuance of a policy that we include a Carrier main effect term in our model.

RESULTS

Table 4 presents our results from the estimation—the coefficient estimate and standard errors for the estimate of β from equation (1). For only one procedure, TEE, a TE policy, is the coefficient on β significantly different from zero (t -statistic = 3.88; p -value = .00001). The coefficient implies that the policies reduce utilization by 13.6 percent.³ Coefficients for five of the other seven procedures are small in magnitude and not significantly different from zero. In only two cases, HBT and Chest X-rays, are the coefficients large (0.371 and 0.471, respectively) but imprecisely estimated. The fit of our models (as measured by the within carrier R^2) for the NT and TE are consistently high, indicating that the Carrier-specific time trends are capturing much of the within Carrier variance in utilization. The standard errors for the UM coefficients in particular are quite large, suggesting that we may not have sufficient power to reject the null hypotheses when, in fact, the alternative hypothesis is true. This is particularly true for the NT procedures.

In an attempt to improve the power of our study, we combine the data from procedures within a category and estimate the coverage parameters. This approach assumes a constant coverage parameter across procedures within a

Table 4: Impact of Medicare Coverage Decisions on Utilization

<i>Procedure</i>	<i>Coverage Policy Coefficient (Standard Errors in Parentheses)</i>	<i>Within Carrier R² Number of Observations</i>
New technologies (NT)		
Urea breath test for <i>Helicobacter pylori</i> (HBT)	0.371 (0.443)	0.63 598
Deep brain stimulation (DBS)	− 0.158 (0.231)	0.53 677
Ocular photodynamic therapy (OPT)	0.0035 (0.352)	0.58 1,175
Combined all new technologies	− 0.046 (0.149)	0.35 2,450
Technology extension (TE)		
Transesophageal echocardiography (TEE)	− 0.136* (0.035)	0.61 2,352
Abdominal and pelvic CT scans	− 0.025 (0.016)	0.93 802
Epoetin	− 0.038 (1.19)	0.93 790
Combined epoetin and CT scans	− 0.052 (0.044)	0.84 1,592
Utilization management (UM)		
Toenail debridement (TD)	− 0.040 (0.070)	0.31 801
Chest X-rays	0.471 (0.349)	0.27 2,402

Note: Regressions include year/month (or quarter) dummies and Carrier specific time trends. Combined all new technologies and combined epoetin and CT scans regressions include Carrier/procedure fixed effects, procedure time dummies and Carrier time trends. For this analysis, we form two groups by coverage policy category: (1) HBT, DBS, and OPT from NT and (2) EPO and CT scans from TE. We do not include UM procedures in this analysis as the dependent variables from the two procedures cannot be meaningfully combined. We do not include TEE in the TE analysis, as that procedure yielded precise coefficient estimates.

*Significant at < 1% level.

grouping and the regressions control for Carrier/procedure fixed effects, and procedure and Carrier-specific time trends. Table 4 includes the results of this analysis. The estimated impact of coverage for both NT and TE groupings is small and consistent with our procedure-specific analysis. The precision of the coefficient for the NT grouping is modest and greatly improved over the procedure-specific average precision for the category. If one is willing to accept the assumption of a constant policy coverage impact across procedures within a grouping, these results suggest that our lack of significance is not solely due to lack of power. This assumption is supported by the generally small magnitude of effect as is seen in the coefficients themselves.

The pattern of the coefficients in Table 4 is robust to a number of different specifications. We have estimated the coefficients without carrier specific time trends and in only one of the procedures did the policy affect utilization.⁴ We also examined different measures of procedure frequency for UM policies (95th percentile, 90th percentile, and logarithm of the number of procedures at the different percentiles and kurtosis) and in no case did the magnitude or statistical significance of our model change.

DISCUSSION

Policy makers assume that Medicare coverage policies impose significant limitations on physician behavior. How else to explain the controversies surrounding the development and implementation of Medicare coverage decisions? Our findings suggest that coverage policies alone can, but generally do not, impact provider behavior. We found no systematic evidence that policies affect utilization, regardless of the type of coverage decision.

Only in TEE do our findings indicate a 13.6 percent reduction in its use after the policy's effective date. TEE policies specifically describe a preference for noninvasive alternatives, limiting its use to cases where there is documented need for the information it provides. The policies clearly intend to limit use to appropriate clinical situations. Carriers report anecdotally that the reduction in use occurs because the CPT codes and International Classification of Diseases, Ninth Revision (ICD-9) codes are sufficiently specific that an electronic edit can identify proper and improper claims. However, other cases in our set, such as the HBT, also appear to have similarly specific codes without similar effect.

We recognize important limitations in our study. Many factors may influence a physician's decision to utilize technologies, including reimbursement, physician preferences, and professional norms. Differences in the timing of policy development, resources, and priorities among contractors may also have some effect. Because many services and procedures can be reimbursed before a coverage policy affecting the service is issued, the effect of the policy may be obscured. For example, a policy may codify existing utilization patterns based on knowledge about the procedure that is also disseminating, leading to no measurable impact of the policy. Another limitation is the challenge of generalizing across interventions. In many cases, the specifics of the technology, the availability of alternatives, the acceptance by physicians, or the adequacy of payment may impact utilization for reasons specific to the

technology rather than due to the coverage policy alone. Differences in coding specificity, as indicated by the TEE case, suggest that further research, including additional policies to replicate our results across a broader range of interventions, is warranted. Testing the hypothesis that the specificity of the coding affects the impact of a policy would contribute greatly to the policy debate.

There are limitations to the conclusions that can be drawn from this quantitative research alone. However, there is qualitative research that is consistent with these conclusions (Foote and Town 2007). There are significant limitations to enforcement of coverage policies that may impede their effect on provider behavior. Contractors must rely on limited information in the claims form; the forms do not include all details necessary to apply the coverage policy. Many policies require review of patient records to determine compliance. While contractors can request them, manual reviews are expensive and time-consuming given the millions of claims to be processed. These qualitative findings, along with the quantitative results in this paper, raise significant issues about the effectiveness of coverage policies in Medicare.

ACKNOWLEDGMENT

We would like to thank The Robert Wood Johnson Foundation's Changes in Health Care Financing and Organization (HCFO) Initiative for supporting this work.

Disclosures: This paper reflects the views of the authors and not those of the Minnesota Department of Health.

NOTES

1. The Medicare Prescription Drug Improvement and Modernization Act of 2003 contains contractor reform provisions that will merge Parts A and B contractors, now called Medicare Administrative Contractors or MACs, and reduce the number of jurisdictions to 15. Transition to new contractors will occur from 2006 to 2009, but does not affect the analysis here.
2. We have examined other measures of the size of tail of the distribution including the logarithm of the 99th percentile, 95th percentile, the 90th percentile, and kurtosis. Our results are robust to these different measures.
3. The probability that at least one of the eight coefficients generates a *t*-statistic of three under the null hypothesis that the coefficients are zero is .007.

4. In this case, the coefficient in the DBS regression was significant. However, the specification with carrier specific time trends is a more general specification and thus it is our preferred model.

REFERENCES

- Bertrand, M., E. Duflo, and S. Mullainathan. 2004. "How Much Should We Trust Differences-in-Differences Estimates?" *Quarterly Journal of Economics* 119 (1): 249–75.
- Foote, S. B. 2002. "Why Medicare Cannot Promulgate a National Coverage Rule: A Case of *Regula Mortis*." *Journal of Health Politics, Policy and Law* 27 (5): 707–30.
- . 2003. "Focus on Locus: Evolution of Medicare's Local Coverage Policy." *Health Affairs* 22 (4): 137–46.
- Foote, S. B., R. Halpern, and D. R. Wholey. 2005. "Variation in Medicare's Local Coverage Policies: Content Analysis of Local Medical Review Policies." *American Journal of Managed Care* 11 (3): 181–7.
- Foote, S. B., and R. J. Town. 2007. "Implementing Evidence-Based Medicine through Medicare Coverage Decisions." *Health Affairs* 26 (6): 1634–42.
- Foote, S. B., D. R. Wholey, and R. Halpern. 2006. "Rules for Medical Markets: The Impact of Medicare Contractors on Coverage Policies." *Health Services Research* 41 (3): 721–42.
- Foote, S. B., D. R. Wholey, T. Rockwood, and R. Halpern. 2004. "Resolving the Tug of War between Medicare's National and Local Coverage." *Health Affairs* 23 (4): 108–23.
- McClellan, M. B., and S. R. Tunis. 2005. "Medicare Coverage of ICDs." *New England Journal of Medicine* 352 (3): 222–4.
- Neumann, P. J., N. Divi, M. T. Beinfeld, B. Levine, P. S. Keenan, E. F. Halpern, and G. S. Gazelle. 2005. "Medicare's National Coverage Decisions, 1999–2003: Quality of Evidence and Review Times." *Health Affairs* 24 (1): 243–54.

SUPPLEMENTARY MATERIAL

The following supplementary material for this article is available online:
Appendix SA1. Author matrix.

This material is available as part of the online article from: <http://www.blackwell-synergy.com/doi/abs/10.1111/j.1475-6773.2008.00836.x>
(this link will take you to the article abstract).

Please note: Blackwell Publishing is not responsible for the content or functionality of any supplementary materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.