

Variation in Medicare's Local Coverage Policies: Content Analysis of Local Medical Review Policies

Susan Bartlett Foote, JD; Rachel Halpern, MPH; and Douglas R. Wholey, PhD

Objective: To assess variation in the content of Medicare's local medical review policies.

Study Design: Six case studies to compare differences in coverage policies by diagnosis codes, procedure codes, and indications for use.

Methods: All carrier policies from 48 carrier contracts (n = 5213) posted to the Centers for Medicare & Medicaid Services Web site were downloaded on May 31, 2001. All policies in the data set were coded based on a typology: new technology (NT), extensions of new technology (TE), and utilization management (UM) of widely used procedures. We identified policies addressing the same procedure or technology. We required at least 20 separate policies in each case study to allow meaningful comparisons. We randomly selected 1 case study of a diagnostic and 1 for a treatment modality from each policy type (NT, TE, and UM).

Results: Given previous research on local carriers, we expected to find variations among policies in each case study. We found substantial similarity, however, among policies covering the NT and TE types. We found significantly more variation among our UM-type case studies.

Conclusions: Medicare legislation has called for greater coverage policy consistency in Medicare. This analysis on variation in policy content, part of a larger study on variation in Medicare's local coverage process, provides data on policy content differences. Policy reform should reflect the nature of and reasons for policy variation as suggested by the findings of this research.

(*Am J Manag Care.* 2005;11:181-187)

Local and regional variations in healthcare have been well documented and raise important issues about quality of care in the system.¹ A payer's coverage policies are intended to define whether a particular service or technology will be offered and, if so, set forth the conditions of use. There is growing interest among payers, providers, and policy makers about the use of scientific and economic evidence in healthcare coverage decision making to encourage more appropriate use and reduce variation.^{2,3}

Managed care organizations have a strong interest in understanding practice variations and encouraging appropriate use to improve outcomes and reduce costs. They have an additional interest in regional variation within the Medicare program because the calculation of the Medicare Advantage premium is tied to fee-for-serv-

ice adjusted average per capita costs. Regional variation strongly influences the amount of a Medicare Advantage premium. Because Medicare Advantage plans must follow the coverage policy in the county where a beneficiary resides, plans offering services over a large geographic region or nationally face operational challenges when Medicare coverage policies differ. Moreover, recently enacted Medicare provisions are moving the program toward new regional models, raising additional concerns about the implications of policy variations and practice patterns on participating health plans.⁴

Although Tunis⁵ has recently described efforts to enhance Medicare's national coverage process, most Medicare coverage decisions are decentralized. Medicare relies on a patchwork of nearly 50 local contracting organizations (called *carriers* for Part B and *fiscal intermediaries* for Part A) that develop thousands of local medical review policies (LMRPs) applicable only within their own jurisdictions. The 1965 Medicare statute authorized local contractors to process claims as a buffer between providers and government.⁶ Contractors subsequently acquired the coverage policy making function through administrative action in the 1990s.

The Medicare Payment Advisory Commission and the US General Accounting Office have recommended elimination of Medicare's local coverage process,^{7,8} and the 2003 Medicare Prescription Drug Improvement and Modernization Act directs the Centers for Medicare & Medicaid Services to work to achieve greater coverage policy consistency. However, little empirical work has been done to examine the extent of, reasons for, or consequences of variation in local coverage.

From the Division of Health Services Research and Policy, School of Public Health, University of Minnesota, Minneapolis.

This study was supported by Changes in Health Care Finance and Organization project 042086 from The Robert Wood Johnson Foundation, Princeton, NJ.

Address correspondence to: Susan Bartlett Foote, JD, Division of Health Services Research and Policy, School of Public Health, University of Minnesota, 420 Delaware Street SE, MMC 729, Minneapolis, MN 55455. E-mail: foote003@umn.edu.

Our research team has been involved in a multiyear study of Medicare's local coverage processes. The variations in the size, resources, and stability of the diverse contracting organizations have been reported elsewhere.⁹ Significant variations in contractor productivity (number of policies each contractor posted), use of evidence (based on evidence cites in LMRPs), and effective dates of policies among local contractors have also been found.¹⁰ These findings document variation across a wide range of variables.

There has been no systematic analysis of variation in policy content, to our knowledge. For example, while we know when local policies covering deep brain stimulation (DBS) were issued, how those policies differ in specific detail one from another has not been analyzed. What can the analysis tell us about variation that would inform payers, providers, and policy makers? To answer these questions, we developed 6 case studies that reveal patterns of policy variation with important implications for Medicare.

CASE STUDIES

Selection Methods

Medicare carriers are required to post LMRPs on the Centers for Medicare & Medicaid Services Web site with prescribed data fields.¹¹ We downloaded all LMRPs on May 31, 2001. We observed that all LMRPs are not the same. We identified 3 types of LMRPs: "new technology" (NT), "technology extension" (TE), and "utilization management" (UM) of widely diffused interventions, defined as follows: NT policies provide guidance for, and limitations on, the use of new clinical interventions; TE policies focus on new uses of procedures or technologies already covered for other uses; and UM policies focus on widely diffused procedures to avoid misuse or overuse. Two physician consultants coded all policies in our database. More than 85% of LMRPs focus on widely diffused technologies (UM); the rest (< 15%) evaluate new technologies (NT) or extensions of technologies (TE). A complete discussion of the coding methods has been published previously.¹⁰

For our case studies, we established the following criteria: (1) We chose only policies from among 48 carrier files (n = 5213) because carriers are more active policy makers than fiscal intermediaries and their regions are more specifically drawn. A complete discussion of the complex structure of Medicare contracting organizations has been previously published.⁹ (2) We selected only cases in which our physician consultants agreed on the coding by type. (3) To enable meaningful comparisons, we selected only case studies with at least 20 posted policies.

Using these criteria, we selected 6 case studies from a pool of 80 cases that met our criteria. We selected 2 case studies from each of our 3 policy types, 1 diagnostic and 1 treatment or procedure in each type. The 6 case studies represent a detailed analysis of 195 separate coverage policies. Although additional case studies might confirm or challenge our findings, the clear patterns we found suggest that our results will be replicated with further research. The case descriptions are drawn from posted policies on the Centers for Medicare & Medicaid Services Web site.¹¹

Deep Brain Stimulation (NT, Treatment). Deep brain stimulation is a neurosurgical procedure that uses electrical stimulation of subcortical structures (the thalamus or the basal ganglia) to control tremors.

Helicobacter pylori Breath Test (NT, Diagnostic). *Helicobacter pylori* is a gram-negative rod that has been causally linked to chronic gastritis, peptic ulcers, gastric cancer, and gastric lymphoma. *Helicobacter pylori* breath tests are noninvasive diagnostic procedures to determine the presence of active infection.

Urethral Stents (TE, Treatment). Male urethral obstructions may result from infections, an enlarged prostate, prostate cancer, prostatitis with fibrosis, and other constrictions. Elimination of the obstruction includes medical and surgical options. Urethral stenting is an extension of the growing use of stents to hold open occluded vessels.

Transesophageal Echocardiography (TE, Diagnostic). Transesophageal echocardiography (TEE) applies an ultrasound generator to the exterior chest wall or in the esophagus to obtain additional cardiovascular information. The instrumentation is invasive with potential for serious complications.

Toenail Debridement (UM, Treatment). Toenail debridement involves the reduction of a thickened dystrophic nail resulting from mycosis or a severe systemic condition using specialized equipment, such as forceps or a rotary drill.

Cardiovascular Stress Test (UM, Diagnostic). Cardiovascular stress testing uses cardiac physiological monitoring during and after stress, with or without subsequent cardiac imaging. A diseased heart responds abnormally, allowing a diagnostic determination.

Analysis

The case studies include 36 DBS policies, 39 *H pylori* breath test policies, 33 urethral stent policies, 33 TEE policies, 23 toenail debridement policies, and 31 stress test policies. There were 41 foot care policies, but only 23 of them addressed toenail debridement and other toenail problems exclusively, while 18 embedded debridement with several other

foot care procedures. Similarly, we identified 43 policies covering transesophageal echocardiography; of those, 33 focused specifically on TEE, and 10 included TEE within a wider range of echocardiography. We excluded the broader policies from the analysis because we could not clearly identify which components in the policy applied to the specific procedure and which referred to the broader conditions. Coding all data points would artificially inflate the diversity of the policies.

RESULTS

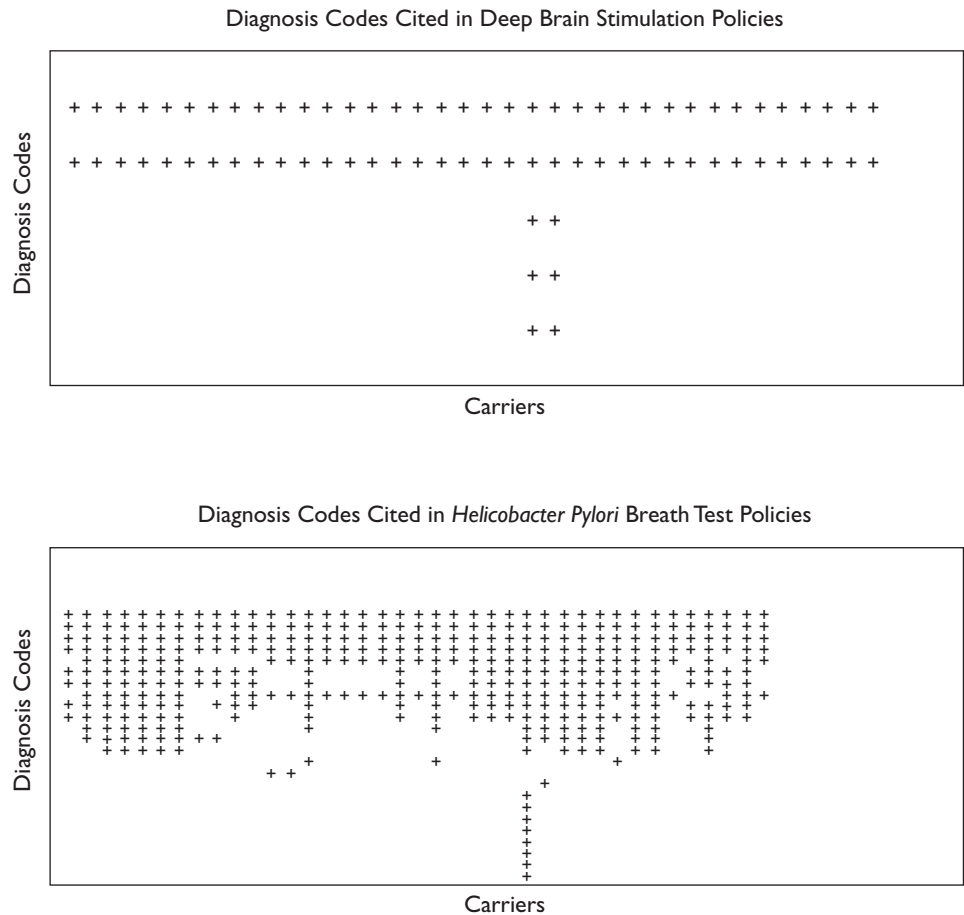
Our policy content comparisons focus on 3 areas: diagnosis codes, procedure codes, and indications.

Diagnosis Code Concentration

One key measure of variation among policies is the array of diagnosis codes in each policy. The *International Classification of Diseases, Ninth Revision, Clinical Modification* is the principal system for coding patient diagnoses and conditions. Local medical review policies specify which diagnosis codes are required for coverage. We found substantial variation in the ways that carriers cited diagnosis codes. Some policies listed codes as ranges (ie, 427.0-427.9); others listed all codes separately. We coded all ranges as including the most comprehensive range; we coded single diagnosis codes separately.

The scattergrams in Figures 1, 2, and 3 illustrate degrees of uniformity or concentration by policy type. Each point on the x-axis reflects 1 carrier; the y-axis shows the frequency of specific codes cited. Figure 1 demonstrates the concentration in our 2 NT case studies, DBS and the *H pylori* breath test. Of the 36 DBS policies, all contain 2 codes, idiopathic Parkinson dis-

Figure 1. Concentration in Diagnosis Codes Among NT Policies

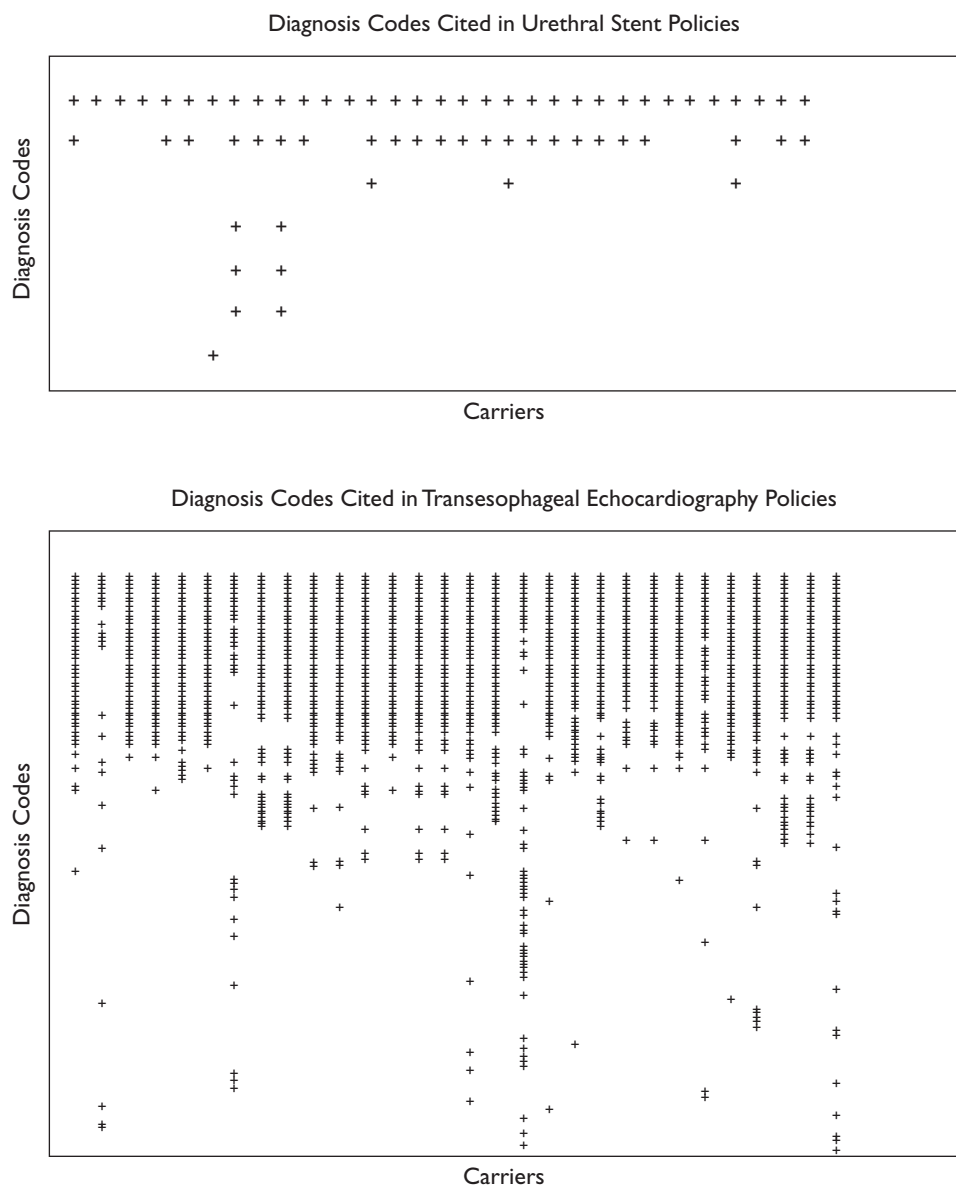


ease and essential tremor. Only 2 carriers include 3 additional codes, all of which refer to more generalized spasmodic conditions. In other words, 34 of the 36 carriers have identical diagnosis codes; only 2 are marginally more inclusive.

There is also considerable consistency among NT policies applicable to the *H pylori* breath test. There are 24 different codes cited across all 39 policies. Four appear in all policies (gastric ulcer, duodenal ulcer, gastrojejunal ulcer, and gastric ulcer), 32 policies include the diagnosis code for peptic ulcer, 29 cover *H pylori* infection and general abdominal pain, followed by a few other abdominal disorders.

Figure 2 shows significant concentration in our 2 TE cases, urethral stent and TEE. Of the 33 LMRPs for urethral stent, there are only 7 cited codes or code ranges. All the policies include a code range that includes different urethral strictures (eg, unspecified infection or trauma). Twenty-three policies cover stents for benign prostatic hypertrophy, and only 2 cover strictures due to shistosomiasis, syphilis, or gonococcal infection.

Figure 2. Concentration in Diagnosis Codes Among TE Policies



Therefore, for the wide variety of strictures, the policies are similar. However, in some jurisdictions, other specific strictures can be treated with stenting.

Transesophageal echocardiography includes 130 different diagnosis codes or code ranges. The mean number of codes or code ranges per policy is 39.9. Figure 2 shows that there is high concentration for many codes, but there is variation among policies for a wide range of other cardiac disorders.

Finally, Figure 3 demonstrates significant variation among UM policies covering toenail debridement and stress testing. These are widely used procedures in the conventional medical arsenal. The UM policies tend to

emerge when carriers identify overuse or misuse in their claims data and want to set clearer limits. Because there is substantial variation in practice patterns across regions, it is not surprising to find these corresponding variations in UM policies.¹²

Procedure Code Concentration

The Healthcare Common Procedural Coding System specifies which provider activities will be reimbursed as part of the covered service. There are 3 levels of codes: Level I *Current Procedural Terminology*, Healthcare Common Procedural Coding System national codes for services and supplies not included in the *Current Procedural Terminology*, and local codes if no national code is yet available. Typically, policies listed discrete procedure codes; occasionally, related codes were listed in ranges. The **Table** demonstrates that patterns of concentration among

procedures (Healthcare Common Procedural Coding System) vary from those related to diagnosis codes. The concentration is measured by the percentage of policies that included the 5 most commonly cited codes. While there are several ways to measure concentration (eg, the Herfindahl index), we used the percentage of policies citing the most commonly cited codes because it is the most straightforward method and is not confounded with the number of procedures cited. A high percentage of policies citing the 5 most common procedure codes indicates a high degree of homogeneity among policies.

We found highest levels of concentration in TE, with significant but somewhat lower levels of concentration

in NT. On closer examination, we found that NT variation was related to new procedures. For example, the variation in the DBS policies may be explained by the presence of temporary (E) codes in some policies and by variation in the listing of ancillary procedures, such as subcortical mapping or preliminary monitoring. These variations do not appear to affect beneficiary access to the core procedure but might affect total provider reimbursement.

Some of this incidental variation presumably will ameliorate as NT procedures proliferate, temporary codes become permanent codes, and policies are revised to reflect evolving patterns of use. The concentration in procedure codes for TE policies probably reflects a greater understanding of a technology being extended to a new use.

By contrast, the classic UM policy in our case study, stress testing, presents a significant lack of concentration. Stress testing is an important diagnostic, but carriers vary considerably on what procedures they accept for payment. Because these policies arise from review of claims data, it is not surprising that some carriers have tried to manage use with careful specification of acceptable procedure codes.

Another interesting anomaly is the high degree of concentration of procedure codes for toenail

Figure 3. Concentration in Diagnosis Codes Among UM Policies

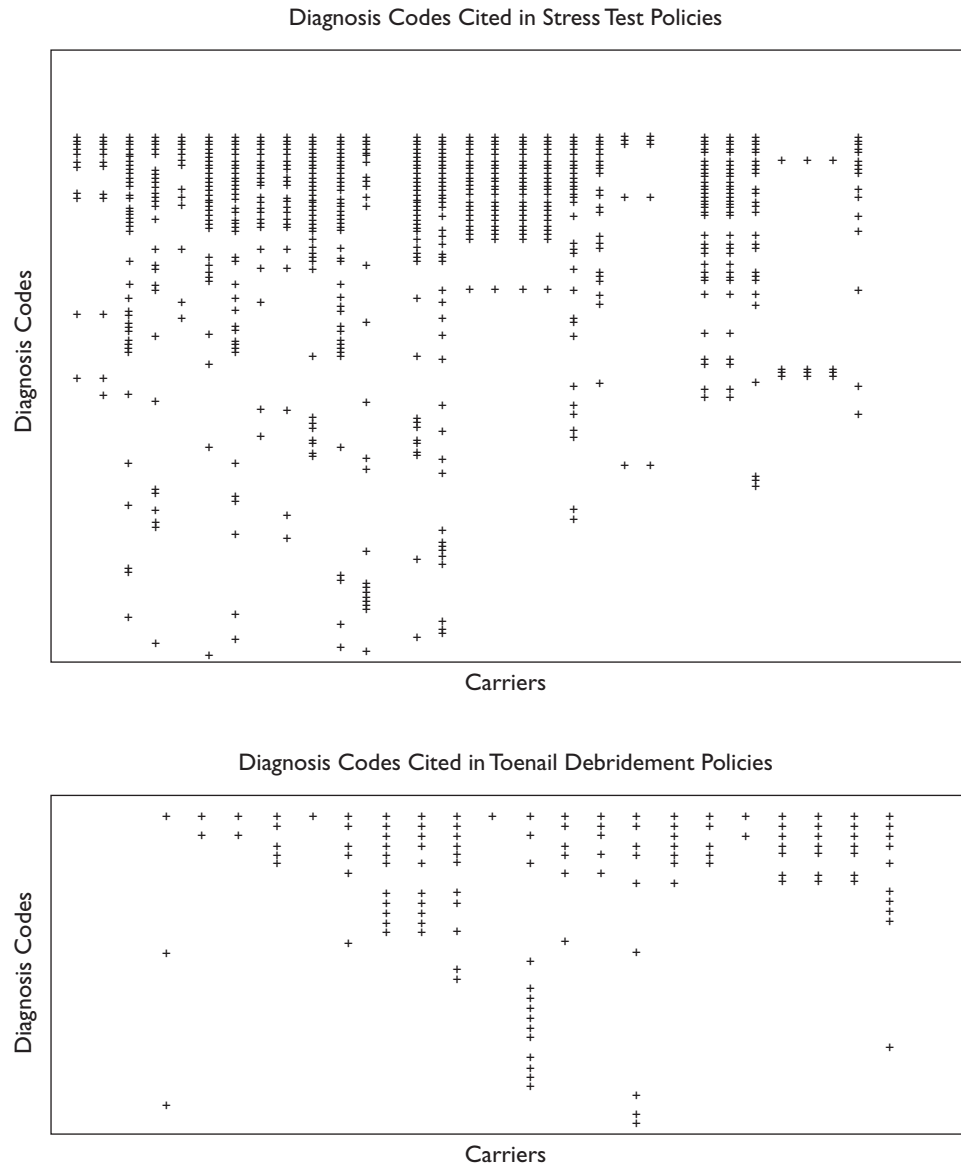


Table. Concentration Measures for Procedure Codes

Category	Case	No. of HCPCS Codes Cited	% in Top 5
NT	Deep brain stimulation	20	66.04
	<i>Helicobacter pylori</i> breath test	18	64.26
TE	Transesophageal echocardiography	9	84.32
	Urethral stents	6	97.96
UM	Cardiovascular stress test	69	38.81
	Toenail debridement	15	64.65

HCPCS indicates Healthcare Common Procedural Coding System.

debridement. In part, this occurs because there are only a few procedures to treat the condition. The policy variations are more likely to occur among permissible diagnoses or allowable indications. Use concerns have been widespread for this common condition. Concerned carriers have worked together to develop a common template for toenail debridement (Kathleen Brooks, MD, Carrier Medical Director for Wisconsin Physician Services; oral communication; Seventh Annual Minnesota Health Services Research Conference; March 4, 2003; Minneapolis). As a result, carriers have adopted a more uniform set of allowable codes than might have been expected otherwise.

Concentration by Indications

Each policy includes sections labeled "Indications and Limitations of Coverage" and "Reasons for Denial" that provide the clinical, procedural, and administrative guidelines that determine reimbursement. We allocated all indications as "for inclusion" or "for exclusion" to prevent redundancy in coding. Criteria for inclusion document who or what is covered; criteria for exclusion record who or what is not covered. In this analysis, we use criteria for inclusion as the measure of comparison.

We found that DBS indications are concentrated, with the first and second most cited indications appearing in 100% and 91%, respectively, of all DBS policies. The third DBS indication, appearing in only 60% of policies, warns providers about the seriousness of the procedure. This type of "admonitory" policy should have little or no effect on beneficiary access.

There appears to be less consistency among the *Hpylori* breath test LMRPs. Of the 21 cited indications, the first and second most cited appeared in only 76% of the policies. On further examination, however, it was clear that many indications were variations on a theme, such as recounting symptoms, levels of symptom severity, or plans for or timing of endoscopy. All focused on careful evaluation of the patient before performing the procedure.

The TE policies for urethral stent and TEE also are concentrated, with the top 3 indications appearing in 90% and 97%, respectively, of all the policies. By contrast, the UM policies show substantial variation across the indications. For example, the most frequently cited indication for stress testing appeared in 58% of the policies. Some policies stated indications generally; others included specific lists of complications. The toenail debridement policies exhibited similar variation, with the most cited indication appearing 77% of the time and the second most cited indication appearing 43% of the time. The indications seek to control unnecessary use, with requirements that the medical

record contain proof of medical necessity in the form of laboratory tests and dated photographs. Other policies included lists of required findings consistent with the diagnosis and indicative of severe peripheral involvement. These UM policy variations demonstrate efforts by specific carriers to control problem use arising in their regions.

DISCUSSION

Key Findings

We would expect some variation across all policy types based on the differences among carriers in terms of resources, size, and output. The results confirmed our hypothesis that NT and TE policies would be most similar, given the limited scientific and clinical data available for evaluation of new technologies. By contrast, we assumed that UM policies were more likely to vary. The UM policies come to the attention of carriers as they review their claims data. Variations in practice patterns are likely to trigger different responses from local carriers. Therefore, if a carrier wants to control overuse of stress tests, it will draft a policy to do so; if there are no perceived abuses, there is less incentive for a carrier to develop a policy or to address any unobserved misuse.

Limitations

Generalizations from 6 case studies should be made with caution. There are many unique characteristics of medical technologies. Although we randomly selected different interventions, additional case studies can confirm or qualify our conclusions. However, the consistency of the variation patterns we identified in our cases suggests a strong likelihood that further studies will be consistent with our findings. More research is also necessary to determine whether and how variations in coverage policies might affect clinical practice. We are looking at the relationship between local policies and Medicare claims data, but these issues are beyond the scope of this case study analysis.

Policy Implications

Because there is little variation in policy content in our NT and TE case studies, we suggest that these policy types are not likely to encourage practice variations. However, the decentralized policy environment means that there is significant variation in effective dates of these policies. The DBS policies, for example, took more than 2½ years from the 1st policy to the 36th policy. A detailed aggregate analysis of the timing of policies has been published previously.¹⁰ Some carriers never issue

policies; their absence is also a sign of variation. These differences mean that access will vary depending on whether a policy is in place. Are these variations all clinically important? We have commented on the variations we found in our cases. However, it is worth noting that clinical importance is in the eye of the carrier; we must assume that, if carriers include specific codes in their policies, they must view them as relevant or important. Clinicians may have opinions about the importance of some variations in specific instances, but conclusions on this point are beyond the scope of this study.

Why allow multiple carriers to engage in numerous assessments of the same technologies when they reach nearly identical results? This duplicative process is inefficient. There are issues regarding quality of the assessments as well. Carriers often lack the resources and expertise to perform formal technology assessments. Therefore, the policies may be technically flawed and inefficiently developed.

Most LMRPs are UM policies. The UM policies respond to local conditions. Given disparate practice patterns across the country, it is not surprising to have local policies reflect different use issues. In fact, policy variations may encourage greater conformance to accepted practices. One carrier may require specific controls to prevent misuse, another may not observe misuse and require no policy, and a third might identify wholly different practice issues calling for an alternative policy response. All may lead to similar use results. A more flexible and responsive local carrier is perhaps better suited to adapt policies to local needs rather than rely on a national policy that might not address specific local problems. Carriers also consult local practitioners, and policies may reflect local concerns. Although one could argue for greater standardization in policies embodying best practice in theory, we must balance the value of standard utilization management approaches against creative, locally crafted policy responses.

CONCLUSIONS

Medicare's decentralized local coverage process results in a complex array of organizations engaged in the development of multiple policies. The patchwork policy environment presents challenges for providers and beneficiaries in the delivery of care. It also presents challenges to managed care organizations that offer plans subject to policies developed by multiple carriers. The objective of this study was to look carefully at how LMRPs vary in content and to assess the implications of that variation on the Medicare program. From a policy perspective, our research suggests that there is room for more policy consistency through greater use of the national process and closer collaboration among carriers. However, reform efforts should reflect the nature of and reasons for policy variation to most effectively address practice variations across the country.

REFERENCES

1. Fisher ES, Wennberg DE, Stukel TA, Gottlieb DJ, Lucas FL, Pinder EL. The implications of regional variations in Medicare spending: part 1: the content, quality, and accessibility of care. *Ann Intern Med.* 2003;138:273-287.
2. Neumann PJ. Why don't Americans use cost-effectiveness analysis? *Am J Manag Care.* 2004;10:308-312.
3. Bloom BS. Use of formal benefit/cost evaluations in health system decision making. *Am J Manag Care.* 2004;10:329-335.
4. *Medicare Prescription Drug, Improvement, and Modernization Act of 2003: Law and Explanation.* 2nd ed. Chicago, Ill: CCH Inc; 2004.
5. Tunis SR. Economic analysis in healthcare decisions. *Am J Manag Care.* 2004;10:301-304.
6. Feder JM. *Medicare: The Politics of Federal Hospital Insurance.* Lexington, Mass: Lexington Books; 1977.
7. *Report to the Congress: Reducing Medicare Complexity and Regulatory Burden.* Washington, DC: Medicare Payment Advisory Commission; 2001:27.
8. *Medicare: Divided Authority for Policies on Coverage of Procedures and Devices Results in Inequities.* Washington, DC: US General Accounting Office; 2003:16. Publication GAO-03-175.
9. Foote SB. Focus on locus: evolution of Medicare's local coverage policy. *Health Aff (Millwood).* 2003;22(4):137-146.
10. Foote SB, Wholey D, Rockwood T, Halpern R. Resolving the tug-of-war between Medicare's national and local coverage. *Health Aff (Millwood).* 2004;23(4):108-123.
11. Centers for Medicare & Medicaid Services Web site. Available at: <http://www.cms.hhs.gov>. Accessed September 1, 2004.
12. Wennberg JE, Fisher ES, Skinner JS. Geography and the debate over Medicare reform. *Health Aff Web Exclusive* [serial online]. February 13, 2002;W96-W114. Available at: <http://content.healthaffairs.org/cgi/reprint/hlthaff.w2.96v1.pdf>. Accessed September 1, 2004.