E/M Coding Problems Plague Physicians, Coders

by Mitchell S. King, MD, Martin S. Lipsky, MD, and Lisa Sharp, PhD

As the government turns its high beams on fraudulent billing, physician E/M coding is raising questions. With several studies spotlighting the difficulty physicians have in applying CPT E/M codes, the authors wanted to know if credentialed coders had the same problem. Here's what they found.

Over the last several years, controlling healthcare costs and fighting fraudulent billing practices have been major government priorities. As a result, physician Current Procedural Terminology (CPT) coding has been the subject of great debate—in particular, the accuracy of physician coding for evaluation and management (E/M) services.

Recent data from several studies suggest that physicians code inaccurately. In a study published in the *Journal of Family Practice*, trained observers using current CPT guidelines found that physicians agreed with the observers' E/M codes for established patients 55 percent of the time, with equal occurrence of undercoding and overcoding.[1] In the *Archives of Family Medicine*, a study using retrospective chart review to compare coding with documentation revealed similar results for established patient visits but found that for new patient visits, physicians coded accurately only 48 percent of the time, with a tendency to overcode.[2] Chart audits reviewing established patient visits in a later study in the *Journal of Family Practice* also produced similar results.[3]

These studies suggest that despite revisions in the E/M coding guidelines, physicians still have trouble using them correctly. Because coding is tied to reimbursement, there is concern that financial incentives might lead to these coding inaccuracies. However, they may instead stem from the complexity of the revised coding systems.

By proposing that physicians are performing E/M coding inaccurately, we are assuming that the CPT E/M coding guidelines can be applied uniformly by auditors or coding specialists. However, in the latter *Journal of Family Practice* study, the three auditors (physician faculty, resident, and professional coder) agreed with each other only 31 percent (according to 1995 guidelines) and 44 percent (according to 1998 guidelines) of the time, which is similar to the findings in the study in the *Archives of Family Medicine*.[4,5] One explanation for inaccurate coding may be that the system is too complex and subjective to be applied uniformly.[6,7] However, despite the financial and legal implications of the assumption that the coding system can be applied consistently by coding specialists, there is little research examining the agreement among coding specialists in their interpretation of Centers for Medicaid and Medicare Services (CMS) guidelines.

In this study we examined the level of agreement in using current CPT E/M coding guidelines when applied by coding specialists to outpatient visits. In addition, we sought to determine if characteristics such as specific coding certification status, educational training, years of experience in coding, time per week spent coding, number of records coded per week, and type or location of practice are associated with coding accuracy. The results might help to determine a natural background error rate for coding, as well as coder characteristics associated with more accurate coding. We found that credentialed coders frequently disagree on the proper E/M code, which suggests that the E/M coding guidelines may not provide enough guidance.

Survey Targets CCS-Ps

The study group consisted of 300 certified coding specialists, physician-based (CCS-Ps), selected randomly by AHIMA from active members. CCS-P-certified coders were chosen because the credential represents training and certified competency through testing in physician office-based coding. AHIMA members were chosen because they represent a heterogeneous group, including coding specialists from urban, suburban, and rural settings, as well as from different practice models. In addition, AHIMA endorsed the study and provided a mailing list of the 300 randomly selected active members.

Six cases presented as hypothetical progress notes were developed representing different levels of service, as well as new and established patient visits. The problems, selected because they are commonly encountered by family physicians, were:

- pneumonia
- leg cramps/hypertension
- deep vein thrombosis (follow-up)
- exercise-induced asthma
- gastroenteritis
- sinusitis/hypertension

The patient cases were labeled as "new" or "established," and only the appropriate CPT E/M codes were provided as choices for selection. For example, codes 99201 through 99205 were provided for cases of new patients and codes 99211 through 99215 were provided for cases of established patients. In addition, a table of the individual coding elements was provided for the coding specialists to indicate the specific level of history, physical examination, and medical decision assigned for each case. These cases were then peer-reviewed by family physician faculty at Northwestern University Medical School for completeness and to assess the authenticity in representing actual patient cases.

Further, we developed a brief survey with demographic and practice characteristics that might be associated with coding ability. Items were generated using information derived from related literature and expert opinions. For example, practice location was included in the survey because a prior study indicated that practice location influences physician coding.[8] The survey instrument was tested among coding specialists from AHIMA for content validity and reliability. Feedback from the coders was then incorporated into a final survey instrument.

The survey instrument and cases were mailed with a cover letter that briefly described the project and contained AHIMA's endorsement. Due to the potentially sensitive nature of coding errors, complete anonymity was assured. Instructions were provided to complete the survey and to code the office visit cases with a CPT E/M code based on the documentation found in the sample progress notes, using the 1997 CPT E/M coding guidelines. In addition, participants were asked to indicate the level assigned for the individual components (history, physical examination, medical decision making) used to determine the final E/M code. Participants were allowed to use whatever resources they might typically use in their own practice (such as coding books or articles) to code the sample notes. A $25 stipend was offered to individuals completing the survey.

The "correct" or consensus CPT E/M code was defined as the coding level most commonly agreed on for each case. Coding accuracy was determined by subtracting the consensus coding level from the coding specialist's coding level on each of the six cases. A negative score reflected undercoding, a positive score reflected overcoding, and a zero reflected agreement with the defined consensus code. To compare the coding specialist's responses on the new cases versus the established, a frequency count of the cases coded correctly, overcoded, and undercoded was completed across the three new cases and across the three established cases. Individual performances were evaluated by adding the number of the six cases coded in agreement with the consensus code.

Results Reveal Undercoding, Overcoding

A total of 136 of the 292 coders eligible for study returned the survey for a response rate of 46 percent. The characteristics of the study group are summarized in "Survey Coder Characteristics." The results of the coding of the six cases are shown in"Coder Agreement on Hypothetical Cases," below. The agreement among the coders in assigning CPT codes for the cases ranged from 50-71 percent. The overall level of agreement for all of the cases was 58.7 percent. New patient progress notes were overcoded in 33 percent of cases, which is four times the rate of undercoding for new patients and twice the rate of overcoding of established patients. Established patient progress notes were undercoded in 25 percent of cases, which represents three times the rate of undercoding for new patient cases. Thus, undercoding occurred significantly more often for established patients and overcoding occurred significantly more often for new patients.

Survey Coder Characteristics

Experience	Average

Years coding experience	10.9
Years experience coding in physician office	8.3
Hours per week coding	24.9
Number of records coded per week	278

Coding certification	**Total**
CCS-P only	70 (52%)
CCS-P and CPC	32 (23%)
CCS-P and RHIT	19 (14%)
CCS-P and RHIA	13 (9.7%)

Educational degrees	**Total**
High school graduate	51 (38%)
Two-year Associate's degree	43 (32%)
Four-year Bachelor's degree	40 (30%)

Type of Practice	**Total**
Primary care	19 (15%)
Specialty	54 (42%)
Both	56 (43%)

Practice Location	**Total**
Urban	46 (35%)
Suburban	38 (29%)
Rural	21 (16%)
Mixed/other	25 (18%)

Coder Agreement on Hypothetical Cases

Established patient cases

	Case 1	**Case 2**	**Case 3**	**All established patients**

Largest percentage in agreement (consensus)	59%	50%	67%	58.6%
Percentage overcoded	0.7%	30%	20%	17%*
Percentage undercoded	40%	20%	12%	25%*
Percentage within one coding level of consensus	95%	98%	100%	97.5%

New patient cases

	Case 4	Case 5	Case 6	All new patients
Largest percentage in agreement (consensus)	50%	71%	56%	58.9%
Percentage overcoded	38%	29%	32%	33%*
Percentage undercoded	13%	0%	13%	8.5%*
Percentage within one coding level of consensus	87%	95%	95%	92.5%

*p<0.001

Seven percent of the coders were in agreement for all of the cases and 26 percent were in agreement for five or more of the six cases. Twenty-eight percent of the certified coders were in agreement for less than 50 percent of the cases; however, 97.7 percent and 92.5 percent of responses were within one coding level of the consensus code for established patients and new patients respectively.

A breakdown of the individual elements assigned by the coding specialists is presented in "Coder Agreement for Individual Elements,"below.

Coder Agreement for Individual Elements	
History	49.4%
Physical examination	65.8%
Medical decision making	55.7%
Total for all elements	57.0%

No correlations were found between educational degree, number of years of coding experience, number of years coding in physicians' offices, or practice type or location. Negative correlations were found between coding accuracy and hours spent coding and number of records coded per week. Coders holding the CPC credential in addition to the CCS-P were significantly more likely to be in agreement with the consensus coding response than respondents with CCS-P alone or with any other additional certifications.

E/M Coding Tough for Coders, Too

The study results suggest that physicians are not alone in their struggle to select the proper CPT E/M code: credentialed coders also have difficulty agreeing on the proper CPT E/M code using current CPT E/M guidelines. The codes for both new and established patients were in agreement for 58 percent of the cases coded. These findings are similar to the earlier *Journal of Family Practice* study in which physicians' codes for established visits agreed with that of a trained observer 55 percent of the time, and slightly better than physician coding accuracy for new patients (48 percent) in the *Archives of Family Medicine* study. These data would suggest that physicians and coding specialists both have difficulty with the current CPT E/M guidelines.

The level of agreement of the coding specialists on the individual coding elements (history, physical examination, medical decision making) was 57 percent, very close to the level of agreement for the assigned CPT E/M codes. The area of highest

agreement was for the physical examination, most likely due to the "counting" involved in assigning this element. The history was the area of most disagreement, which may have been due to the subjectivity involved in determining what constitutes the various parts of the history. Some interpretation is required about what specifically constitutes the different parts of the history of present illness, as well as what is included in the history of present illness or the review of systems.

For example, a phrase such as "A patient presents with a three-day history of a fever to 103 degrees...the patient also has noted a productive cough" may be interpreted in different ways. Some may view the height of the fever as indicating severity in the history of present illness, but others may not come to the same conclusion. The "cough" may be considered an associated symptom or as a review of system. These differences could potentially lead to disagreement on the final codes assigned.

The patterns of errors for the final assigned CPT E/M code also were similar to prior physician studies.[9] For established patients, the most common error was undercoding. For new patients, overcoding occurred significantly more often than undercoding. One reason for this discrepancy could be a tendency to apply the same guidelines to all patients and not recognizing or applying the different criteria for new patients. Coding criteria are stricter for new patients; more documentation is required to establish the same level of service. In addition, all three elements must support the assigned code for new patients. Given the difficulty the coding specialists had in agreeing on the individual elements (history, physical examination, medical decision making), a high level of disagreement on new patient visits would be expected. Because the consensus codes for the new patient progress notes represented lower CPT E/M codes (99201, 99202), the errors or disagreements would tend toward overcoding. Finally, coding specialists may recognize that caring for new patients requires more effort and that there is more uncertainty in providing this care than for established patients. Thus, physicians and coders may feel that new patients are more difficult and their coding levels may reflect this feeling.

Although it was expected that coding experience or higher educational degrees would improve coding accuracy, this study found no such correlation. Further, no correlation between coding accuracy and type of practice or practice location was found. Interestingly, a negative correlation was found between coding accuracy and number of hours spent coding and numbers of records coded per week. This suggests that excessive time spent coding and higher volumes of coding may actually compromise accuracy.

In this study, individuals with the CPC credential in addition to the CCS-P credential coded more accurately. Individuals with other types of degrees or credentials may be devoting time to other areas of coding or performing administrative functions, whereas individuals with CCS-P and CPC credentials may be more likely to be exclusively devoting their time to physician office-based coding. Though time and number of records coded per week were negatively correlated with coding accuracy, the study suggests that staying focused on one skill may improve accuracy.

The format of the study was designed to test the coding specialists' accuracy in coding using hypothetical cases. The design removed any financial or legal incentives for incorrect coding. Further, all coders used the same typewritten progress notes, thus removing discrepancies from attempting to interpret handwritten progress notes or apply the guidelines to different cases of the same coding level.

Despite removing these potential sources of coding disagreement, only 55 percent of coding specialists agreed with the consensus code on four or more of the six cases, and 8 percent agreed with the consensus code on one or none of the cases. However, only 3 percent of established patient codes and 8 percent of new patient codes were more than one coding level different from the consensus code. Therefore, although there appears to be disagreement among coding specialists in assigning CPT E/M codes, many of the differences are within one level of the consensus code. Interestingly, this finding is consistent with findings from another study in which physicians' codes for established patients differed from reviewers' codes by more than one level in less than 4 percent of cases.[10]

Our results suggest that the E/M guidelines themselves are overly complex and open to subjective interpretations that create a high inherent error rate for both coding specialists and physicians. Having separate sets of guidelines for new and established patients may be a contributing factor. Standardizing the coding criteria into one set of guidelines for all patients could reduce the error rate. Further, decreasing the number of potential codes for each office visit as well as the number of steps required to arrive at a code would limit the potential for error and subjective interpretations. Using time and new versus established patient status as the deciding factors in arriving at the level of service provided might also reduce E/M coding errors.[11] Given the

complexity of the current CPT E/M guidelines, it seems that errors may be unavoidable. Clearly, further study of the CPT E/M coding guidelines is needed.

Note: Copies of the progress notes are available from the authors upon request.

Acknowledgment

A similar study by the authors was published in the *Archives of Internal Medicine* 162 (2002): 316-20. Copyrighted 2002, American Medical Association.

The authors would like to thank the American Academy of Family Physicians Foundation for its generous grant support and AHIMA for its assistance.

Notes

1. Chao, J. et al. "Billing for Physician Services: A Comparison of Actual Billing with CPT Codes Assigned by Direct Observation." *Journal of Family Practice* 47, no. 1 (1998): 28-32.

2. Kikano, G.E., M.A. Goodwin, and K.C. Stange. "Evaluation and Management Services: A Comparison of Medical Record Documentation with Actual Billing in a Community Family Practice." *Archives of Family Medicine* 9, no. 1 (2000): 68-71.

3. Zuber, T.J. et al. "Variability in Code Selection Using the 1995 and 1998 HCFA Documentation Guidelines for Office Services." *Journal of Family Practice* 49, no. 7 (2000): 642-45.

4. "Variability in Code Selection Using the 1995 and 1998 HCFA Documentation Guidelines for Office Services."

5. "Evaluation and Management Services: A Comparison of Medical Record Documentation with Actual Billing in a Community Family Practice."

6. Lasker, R.D. and M.S. Marquis. "The Intensity of Physicians' Work in Patient Visits: Implications for Coding of Patient Evaluation and Management Services." *New England Journal of Medicine* 341, no. 5 (1999): 337-41.

7. Iezzoni, L.I. "The Demand for Documentation for Medicare Payment." *New England Journal of Medicine* 341, no. 5 (1999): 365-7.

8. Purvis, J.R. and R.D. Horner. "Billing Practices of North Carolina Family Physicians." *Journal of Family Practice* 32, no. 5 (1991): 487-91.

9. Horner, R.D. et al. "Accuracy of Patient Encounter and Billing Information in Ambulatory Care. *Journal of Family Practice* 33, no. 6 (1991): 593-598.

10. "Evaluation and Management Services: A Comparison of Medical Record Documentation with Actual Billing in a Community Family Practice."

11. "The Intensity of Physicians' Work in Patient Visits: Implications for Coding of Patient Evaluation and Management Services."

Mitchell S. King *(m-king1@northwestern.edu) is assistant professor, Martin Lipsky is professor and chair, and Lisa Sharp is assistant professor, department of family medicine at Northwestern University Medical School.*

Article citation:
King, Mitchell S., et al. "E/M Coding Problems Plague Physicians, Coders." *Journal of AHIMA* 73, no.7 (2002): 62-66.
