Spring 2001

Accuracy of CPT Evaluation and Management Coding by Family Physicians

Mitchell S. King

Lisa Sharp

Martin S. Lipsky
Portland State University

Let us know how access to this document benefits you.

Citation Details

King, M. S., Sharp, L., & Lipsky, M. S. (2001). Accuracy of CPT evaluation and management coding by family physicians. The Journal
of the American Board of Family Practice, 14(3), 184-192.

Accuracy of CPT Evaluation and Management Coding by Family Physicians

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

Background: Limited data are available on physicians' accuracy in coding for their services. The purpose of this study was to determine the current procedural terminology (CPT) evaluation and management coding accuracy of family physicians and define demographic variables associated with coding accuracy.

Methods: Six hundred randomly selected active members of the Illinois Academy of Family Physicians were sent six hypothetical progress notes of office visits along with a demographic survey. The study group assigned CPT evaluation and management codes to each of the progress notes and completed the demographic survey. Five expert coders also assigned codes to each of the cases. The accuracy of family physicians in determining CPT E/M codes was determined relative to that of expert coders.

Results: Family physicians agreed with the experts' CPT evaluation and management codes for 52% of established patient progress notes, the most common error being undercoding. In contrast, for new patient progress notes, family physicians agreed with the experts only 17% of the time, the predominant error being overcoding. No surveyed demographic variable was associated with coding accuracy.

Conclusions: The error rate for physician CPT coding is substantial and occurs more commonly with new patients. The complexity of the CPT coding guidelines, along with limited physician training in CPT coding, likely account for these results. (J Am Board Fam Pract 2001;14:184–92.)

Currently there is a great deal of interest by the government and in the lay press on current procedural terminology (CPT) coding. CPT guidelines mirror the change from a charge-based system to a fee schedule that reflects the resources used in providing care. Clinicians use CPT codes to bill Medicare and other payors for their services. Despite the importance of the CPT codes, however, there is limited research regarding physician coding. For example, only three studies evaluated coding using current CPT evaluation and management guidelines, and only one study included new patient visits.[1–3] As a result, there is little information regarding physician coding accuracy, physician training in CPT coding, or characteristics associated with accurate or inaccurate coding.

There are several reasons why coding is an important topic for physicians. The most obvious is financial, because coding determines reimbursement for physician services. Available data suggest that physicians code improperly, with conflicting data on the net economic impact of this inaccuracy. Information from Health Care Finance Administration (HCFA) and the American Academy of Family Physicians (AAFP) indicate that family physicians often undercode for their services, resulting in a loss of potential revenues.[4] Conversely, the Office of the Inspector General recently issued a release citing $20 billion of Medicare overpayments with 29% due to improper coding for physician's services.[5] A recent study using trained observers and current CPT guidelines found that physicians' CPT codes for established patients agreed with the observers' codes 55% of the time, with approximately an equal amount of undercoding and overcoding.[1] A retrospective chart review by Zuber and associates[3] had similar results. In the only study to include new patient visits, a retrospective chart review by Kikano et al,[2] 46% of physician CPT evaluation and management codes agreed with the reviewers, but with a greater tendency to overcode new patients.

Despite revisions, many believe that the CPT evaluation and management coding guidelines are clinically irrelevant and overly complex.[6,7] One explanation for inaccurate coding could be the complexity of the coding system and a poor understand-

ing on the part of physicians regarding their use. Informal discussions with physicians suggest that inaccurate coding is more likely due to the difficulty of transferring coding guidelines into clinical practice rather than to fraudulent activity.

Coding properly has important legal as well as financial implications. For example, accurate evaluation and management coding can help protect physicians from the financial and legal ramifications of a Medicare audit. In training programs, the coding level documents a faculty member's involvement with a patient's care, touching on both legal and more complicated financial concerns. Against a background of several well-publicized audits of academic medical centers, these issues are critically important.

In this study we examined how accurately a sample of family physicians code outpatient visits. In addition, we sought to find out whether certain characteristics, such as practice setting, charges for different office visit levels, and training in CPT coding, are associated with coding accuracy. The results might help to target interventions to selected groups or settings where the coding guidelines might be applied less accurately. The data might also be helpful in determining a natural background error rate for coding. This natural error rate could help distinguish between fraudulent billing practices and the impossibility of applying a complex system with perfect accuracy.

Methods

The study group consisted of 600 family physicians selected randomly by the Illinois Academy of Family Physicians (IAFP) from their active members. Family physicians in Illinois were chosen because they represent a heterogeneous group, including physicians from urban, suburban, and rural settings, as well as from different practice models. In addition, the IAFP endorsed the study and provided a mailing list of the 600 randomly selected active members. The study took place during the years 1999 to 2000.

Six cases presented as hypothetical progress notes were developed to represent different levels of service, as well as new and established patient visits. The following six cases were chosen for these progress notes: pneumonia, leg cramps and hypertension, deep vein thrombosis (follow-up), exercise-induced asthma, gastroenteritis, and sinusitis

and hypertension. These cases were selected because they represent common problems encountered by family physicians. The progress notes are presented in Appendix I.

The patient cases were labeled as new or established, and only the appropriate CPT 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. These cases then underwent peer review by family physician faculty at Northwestern University Medical School (NUMS) to determine completeness and to assess the authenticity in representing actual patient cases.

The cases were then sent to 5 expert billing coders, selected and recruited through the Coding and Medical Information Systems Department at the American Medical Association. All the experts were certified coding specialists, had at least 12 years of experience with coding, and had served as faculty in programs designed to teach others how to code properly. These experts assigned CPT evaluation and management codes to these cases. The experts were not in full agreement for all the assigned CPT evaluation and management codes.[8] In cases where the experts were not in full agreement on the CPT code, the code provided by the majority of coders was assigned as the experts' CPT code for that case.

In addition, a brief survey was developed to elicit coding practices along with demographic and practice characteristics that might be associated with coding ability. Items were generated using information derived from the literature and expert opinion. For example, practice location was included in the survey because a previous study indicated that practice location influences physician coding.[9] The survey instrument was pilot-tested among the family physician faculty at NUMS for content validity and reliability. Feedback from the physicians was then incorporated into a final survey instrument.

The survey instrument and cases were mailed, with a self-addressed return envelope and cover letter, to the study participants. The cover letter briefly described the project and contained the endorsement of the IAFP. Because of the potential sensitive nature of coding errors, complete anonymity was assured. Instructions were provided for the physicians to complete the survey and to code the office visit cases with a CPT evaluation and

management code based on the documentation found in the sample progress notes. Participants were allowed to use whatever resources they might typically use in their own practice (eg, books, articles) to code the sample notes. After 1 month, nonresponders received a second mailing. Two additional mailings were sent to nonresponders.

The correct CPT code was defined as the coding level that the majority of the 5 expert coders agreed on for each case. Coding accuracy for the family physicians was determined by subtracting the correct coding level from the physician's coding level on each of the six cases. A negative score reflected undercoding, a positive score reflected overcoding, and a zero score reflected agreement with the experts. To compare physicians' responses on the new cases with the established patient cases, a frequency count of cases coded correctly, overcoded, and undercoded was completed across the three new cases and across the three established cases. Individual physicians' performances were evaluated by summing the number of the six cases coded correctly.

Descriptive statistics were used to summarize the sample characteristics. An analysis of variance was used to compare groups when appropriate. Scoring on the new cases compared with the established cases was analyzed using the nonparametric Wilcoxon matched-pairs test.

Results

Of the 600 surveys sent, 107 were returned as undeliverable, leaving a study group of 493. A total of 205 of the 493 eligible for study returned the survey and completed the coding for the cases, yielding a response rate of 42%. Thirty-three of these responders gave incomplete demographic information; however, coding was completed for the cases. Five responders completed all but one of the cases. The results of these responders are included in the analysis of CPT coding accuracy. Surveys returned with completion or partial completion of demographic data but without coding of the cases were not used and are not counted among the 205 responders.

Table 1 displays demographic data collected for the study group. The group averaged 31.9 hours in the office per week seeing an average of 93.6 patients. Nearly all the respondents were residency trained and board certified in family medicine. Sev-

Table 1. Physician Characteristics (n = 198).

Characteristic	Mean (SD)
Years in practice	4.3 (3.6)
Age, years	36.2 (5.6)
Hours per week seeing patients	31.9 (9.5)
	Number (%)
Board-certified in family medicine	182 (91.9)
Residency trained	190 (96.0)
	Total No. (%)
Type of practice	
Private, solo	16 (9.0)
Private group	66 (38.0)
Managed care	3 (2.0)
Hospital group	53 (30.0)
Faculty	23 (13.0)
Other	14 (8.0)
Practice location	
Urban	57 (33.0)
Suburban	74 (43.0)
Rural	39 (22.0)
Mixed	4 (2.3)

enty-four percent of physicians either had a productivity-based salary or a base salary with a productivity incentive.

Most physicians (63%) reported receiving CPT coding instruction in the past, 6 hours on average. Of those with previous CPT training, however, 37% had received no additional training since completing their residency. The person responsible for coding patient visits varied within the study groups' offices. Sixty-three percent of the physicians personally determined CPT coding for their patients' office visits. An additional 18% performed this coding in conjunction with a nurse. The remaining physicians had office or billing personnel determine coding. Eighty-two percent of physicians reported using resources, such as books, office staff, or coding references, to assist with determining the code.

The results of the physicians' coding of the six cases are shown in Table 2 along with the experts' coding. As shown in Table 3, the physicians' coding of the new cases was more inaccurate than coding of the established cases (P < .001). For established patient visits, physicians agreed with the experts' codes in 52% of the cases, overcoded in 16% of the cases, and undercoded in 33% of the cases. In contrast, for new patients physicians agreed with the experts' codes in only 17% of the cases, over-

Table 2. Number and Percentage of Family Physician CPT Coding of Six Hypothetical Cases Compared with Expert Consensus, by CPT Coding Level.

Case	No. (%)	No. (%)	No. (%)	No. (%)	No. (%)
Coding level, established patient	*99211*	*99212*	*99213*	*99214*	*99215*
1. Pneumonia	2 (1)	13 (6)	132 (64)	**55 (27)**	3 (1.4)
2. Cramps, hypertension	2 (1)	41 (20)	**137 (66)**	25 (12)	0 (0)
3. Deep vein thrombosis follow-up	11 (5)	**125 (60)**	65 (31)	3 (1.4)	0 (0)
Coding level, new patient	*99201*	*99202*	*99203*	*99204*	*99205*
4. Asthma	3 (1.4)	**39 (19)**	114 (55)	143 (21)	6 (3)
5. Gastroenteritis	**18 (9)**	128 (62)	53 (26)	5 (2)	0 (0)
6. Sinusitis, hypertension	4 (2)	**49 (24)**	119 (58)	30 (15)	0 (0)

CPT—Current procedural terminology.
Note: Resposes in boldface indicate the number and percentage of physicians agreeing with the experts' consensus coding.

coding in 82% of the cases, and undercoding in 1% of the cases. Table 4 displays data on the physicians' overall scores relative to the experts'. The number of cases in agreement with the experts was tabulated for each physician. Only 28% of the physicians agreed with the experts' codes for three or more of the six cases.

No statistically significant relations were found between physician accuracy in coding and the following variables: years in practice, physician age, type of practice, formal training, hours of training, patient care time, charges for office visits, practice location, or physician determination of the codes during their patients' visits.

Table 3. Percentage of Total Responses by Agreement With Experts.

Case	Undercode (%)	In Agreement (%)	Overcode (%)
Established patients	32.7*	51.6*	15.6*
New patients	1.1*	17.3*	81.5*

*P < .001.

Table 4. Number of Six Cases Coded Correctly by Physicians.

Cases Correct No.	Physicians No. (%)
6	1 (0.5)
5	5 (2.4)
4	16 (7.7)
3	36 (17.4)
2	83 (40.1)
1	53 (25.6)
0	11 (5.3)

Discussion

This study suggests that family physicians have difficulty determining the proper CPT code using current CPT guidelines. Physicians agreed with the experts' codes for established patients only 52% of the time, findings similar to a recent study in which physicians' codes for established visits agreed with that of a trained observer 55% of the time.[1] Perhaps the most important and surprising study result was that for new patients, the physicians' CPT codes for the office visits corresponded with that of the experts only 17% of the time. This lower coding accuracy for new patients is consistent with the findings of Kikano et al,[2] but the rate of inaccurate coding was much higher in the present study.

The patterns of errors also differed markedly for established and new patients. For established patients, the most common error was undercoding. This finding is consistent with statements by the AAFP and HCFA that family physicians tend to undercode for their services. In addition, previous studies[3,10] reviewing physicians' progress notes also found that compared with services documented in the progress notes for established patients, family physicians undercoded for their professional services. For new patients, the predominant error was that of overcoding, lending some credibility to concerns that physicians might be overcoding for some visits.

To our knowledge, the different patterns of errors for new patients compared with established patients has been reported in only one previous study.[2] One reason for this marked discrepancy could be that physicians apply the same guidelines to all patients, not recognizing the different criteria for new patients. Coding criteria are stricter for

new patients, requiring more documentation to establish a higher service level. Another factor that could account for overcoding new patients is the sense that new patients require more effort, and there might be more uncertainty in providing care for new patients than for established patients. Thus, physicians might think new patient visits are more difficult, and their coding levels could reflect this assumption.

Although one might predict that experience and training would improve coding accuracy, this study found no such association. In addition, no associations between coding accuracy and age, practice experience, time spent seeing patients, type of practice, charges for office visits, source of physicians' salaries, physician determination of CPT codes, or practice location were found. In a previous study,[9] an association was found between undercoding and a rural practice location. Since the time of that study, the coding guidelines have undergone considerable changes, which could account for lack of similar findings in the present study.

Although coding errors might conceivably relate to financial incentives, the format of the study was designed to test the physicians' accuracy in coding using hypothetical cases. This design removes any financial incentive to overcode or undercode. All physicians used the same progress notes, thus removing the discrepancies from the perceived level of the work performance compared with what was actually documented. Despite removing these potential sources of coding inaccuracy, the error rate was still high. Only 28% of physicians agreed with the experts on three or more of the cases, and 31% agreed with the experts on one or none of the cases. Only 3% of established patient codes were more than one coding level different from that of the experts, however. Thus although there appears to be a high background error rate for CPT coding among physicians, many of the errors are within one level of that of the experts.

One limitation to this study is the response rate of 42%. Even so, the study group characteristics were similar to the IAFP membership at large in terms of hours in patient care and type of practice. The responders were slightly younger than the total IAFP average age (36 vs 44 years of age), were more likely to be residency trained (96% vs 80%), and were more likely to be board certified (92% vs 84%). The percent in rural practice is similar to AAFP data (22% vs 23%) (IAFP data are not available for this comparison). Accordingly, although the response rate was only 42%, these survey respondents are representative of Illinois family physicians in type and location of practice and time spent in patient care. Busier physicians might not have taken time out to complete the survey and arguably could be more experienced and thus better coders. In this study, however, we found no association between number of patients seen, years in practice, or time spent in patient care and coding accuracy. Physicians who responded to the survey might actually be more comfortable with their coding abilities and have more interest in this subject.

If there were a selection bias, then the participants most likely were better coders than the nonresponders, which would result in an underestimation of the error rate.

Another potential bias that could result in a higher overcoding rate for new patients is our selection of new patient visits that had lower coding levels. Any errors would more likely occur on the side of overcoding. Even with this potential bias, however, the data still suggest that overcoding occurs for these lower levels of documented service.

From our results, it appears that the error rate with CPT coding is substantial and that it is highest with new patients. Having separate sets of guidelines for new and established patients might be a contributory factor. One possible solution to minimizing the error rate with CPT coding would be to standardize the coding criteria into one set of guidelines for all patients. For example, a level 3 new patient and a level 3 established patient would require the same criteria of documented services to arrive at the level 3 designation. A modifier could be added to each patient (eg, N or E) to indicate whether the patient is new or established. Another proposed solution, described in detail by Lasker and Marquis,[6] involves using time and new vs established patient status as the deciding factors in arriving at the level of service provided. Finally, given the complexity of the current CPT guidelines, another potential solution is to accept an inherent error rate. In an era of audits and physicians being charged with coding fraud, however, coding guidelines must be easy to apply objectively to achieve reliable, reproducible results. The current results, along with previous physician and expert studies,[1–3,8] indicate that the current guidelines do not meet these criteria. Clearly, further

revision and study of the CPT coding guidelines would appear to be warranted.

References

1. Chao J, Gillanders WG, Flocke SA, Goodwin MA, Kikano GE, Stange KC. Billing for physician services: a comparison of actual billing with CPT codes assigned by direct observation. J Fam Pract 1998;47:28–32.

2. Kikano GE, Goodwin MA, Stange KC. Evaluation and management services. A comparison of medical record documentation with actual billing in a community family practice. Arch Fam Med 2000;9:68–71.

3. Zuber TJ, Rhody CE, Muday TA, et al. Variability in code selection using the 1995 and 1998 HCFA documentation guidelines for office services. Health Care Financing Administration. J Fam Pract 2000;49:642–5.

4. Mastering Medicare's new documentation guidelines. Fam Pract Manage 1995;Jan:60–7.

5. Martin S. OIG. $20 billion in "improper" Medicare payments. Am Med News 1998;41(18):11–13.

6. Lasker RD, Marquis MS. The intensity of physicians' work in patient visits—implications for coding of patient evaluation and management services. N Engl J Med 1999;341:337–41.

7. Iezzoni LI. The demand for documentation for Medicare payment. N Engl J Med 1999;341:365–7.

8. King MS, Lipsky MS, Sharp L. Current procedural terminology coding: do the experts agree? J Am Board Fam Pract 2000;13:144–6.

9. Purvis JR, Horner RD. Billing practices of North Carolina family physicians. J Fam Pract 1991;32:487–91.

10. Horner RD, Paris JA, Purvis JR, Lawler FH. Accuracy of patient encounter and billing information in ambulatory care. J Fam Pract 1991;33:593–8.

Appendix I

Coding Survey Cases

The following six progress notes represent hypothetical patients seen in a practice. We realize that physicians vary in the extent of the history, examination, and testing procedures and medications prescribed. The following cases are meant to be analyzed, not for their clinical decisions, but for the content as it relates to office billing for physician services. Based on the progress notes, please circle the level of code for each visit based on the CPT evaluation and management level of service you believe this visit represents.

Case 1. Established patient, chief complaint: fever, fatigue

99211
99212
99213
99214
99215

A 45-year-old previously healthy man has 3-day history of feeling feverish and fatigued. He notes fever to 103°F and fatigue to the point of not being able to perform his normal exercise and home chores. He has also noted a slight cough worsened with activity. He denies pain, shortness of breath, upper respiratory tract congestion, sore throat, ear pain, emesis, diarrhea, dysuria, or rash. + smoker - 35 pack-years. Temperature 101.2°F, heart rate 100 beats per minute, respiratory rate 24/min, blood pressure 130/80 mm Hg

Alert, in no acute distress
Head, ears, eyes, nose, throat: Tympanic membranes and throat clear, nose without significant congestion or rhinorrhea
Neck: supple without LA
Lungs: left lower lobe with bronchial breath sounds, E to A changes and dullness to percussion, no wheezing
Heart: regular tachycardia without murmurs
Abdomen: Bowel sounds present, no masses, organomegaly, CVA, or other tenderness
Chest radiograph with left lower lobe pneumonitis
Pulse oximetry 96% on room air

Assessment: left lower lobe pneumonia
Plan:
1. Clarithromycin 500 mg bid for 10 days
2. Follow-up in 7 days or as needed for worsening symptoms
3. Will need follow-up chest radiograph to show resolution

Case 2. Established patient, chief complaint: hypertension follow-up
99211
99212
99213
99214
99215

A 55-year-old man comes in for follow-up of his hypertension with complaints of occasional leg cramps. He has had the cramps for the past 2 months and has noted them bilaterally. They awaken him from sleep. He denies any recent change in activity level or diet. He denies any leg pains or cramps associated with activity. He denies any other leg symptoms, chest pains, shortness of breath, PND, or peripheral edema. He takes atenolol 50 mg/d and hydrochlorothiazide 25 mg/d for his hypertension. He is a nonsmoker. Blood pressure 126/86 mm Hg, pulse 82 beats per minute, weight 190 pounds

Lungs: chest clear to auscultation and percussion
Heart: RRR without murmurs
Extremities: without cyanosis, clubbing, edema; distal pulses 2+ bilateral lower extremities
Neurologic examination: lower extremities with intact sensation to touch, pinprick, and vibration. Motor 5/5 throughout the lower extremities

Assessment: Hypertension- well controlled, leg cramps- possibly secondary to hydrochlorothiazide
Plan:
1. Continue atenolol 50 mg daily
2. Discontinue hydrochlorothiazide and monitor blood pressure and symptoms
3. Sequential Multiple Analyzer-7 today
4. Follow-up in 2 weeks for recheck

Case 3. Established patient, chief complaint: hospital follow-up
99211
99212
99213
99214
99215

A 68-year-old white woman returning for follow-up after hospitalization for deep vein thrombosis. Notes less leg swelling, and no further pain. Denies shortness of breath. On coumadin 2 mg daily, prothrombin time INR 2.2 on discharge 1 week ago. Blood pressure 120/76 mm Hg, heart rate 80 beats per minute and regular, respirations 16/min, temperature 98.8°F

Lungs: clear
Heart: RRR without murmur
Extremities: without edema, tenderness

Assessment: LLE deep vein thrombosis- stable
Plan:
1. Continue coumadin 2 mg daily
2. Check prothrombin time INR today
3. Follow-up 2 weeks or prn

Case 4. New patient, chief complaint: wheezing
99201
99202
99203
99204
99205

A 12-year-old girl complains of coughing and wheezing whenever she plays sports. She denies shortness of breath except with exercise. No animals in the house. Symptoms began 2 months ago when she began playing on the school soccer team. She denies other problems and has not been sick in any other way.

Surgeries: none
Medical illnesses: admitted for wheezing at age 2 years, no problems since
Medications: none
Allergies: none
Family history: asthma in father
Social history: 7th grade, lives with parents and younger brother, no pets or smokers in house. Patient denies smoking.
Review of systems: all other systems negative.
Temperature 98.6°F, blood pressure 100/65 mm Hg, heart rate 70 beats per minute and regular, respiratory rate 16/min, height 60 inches

Ears, nose and throat: tympanic membranes and throat clear, no nasal discharge
Lungs: clear, no wheezing, rales, or rhonchi; peak flow 330 L (100% predicted)
Heart: RRR without murmurs

Assessment: exercise-induced asthma
Plan:
1. Monitor peak flow with activity
2. Proventil MDI 2 puffs 30 minutes before exercise
3. Explained re: warm-up
4. Follow-up in 2 weeks for further discussion and recheck, and remainder of physical examination

Case 5. New patient, chief complaint: diarrhea
99201
99202
99203
99204
99205

A 24-year-old man complains of a 1-day history of nausea, abdominal cramps, and loose stools. He has had four loose, brown bowel movements without blood or melena. He has had no known exposures. He denies fever or lightheadedness. No previous gastrointestinal tract problems, no medications. Blood pressure 130/80 mm Hg, heart rate 82 beats per minute and regular, temperature 98.8°F

Lungs: clear
Heart: RRR without murmur
Abdomen: active bowel sounds, soft nontender

Assessment: gastroenteritis, likely viral
Plan:
1. Instructed re: diet, clear liquids
2. Advised re: expected course and to follow-up if signs and symptoms of dehydration occur, fever, blood

Case 6. New patient, chief complaint: sinus infection
99201
99202
99203
99204
99205

A 50-year-old man with 2-week history of upper respiratory tract congestion, cough, and headache. The drainage is green in color. The headache is increased with bending over and coughing, and is located over the frontal and maxillary regions. The cough is occasionally productive. He denies fever, shortness of breath, or chest pain. He has noted ear pressure, but no pain.

Review of systems: otherwise negative
Surgeries: none
Medical illnesses: hypertension
Medications: Actifed (over-the-counter), captopril 25 mg po bid
Allergies: no known drug allergies
Nonsmoker, works as accountant

Blood pressure 128/70 mm Hg, heart rate 72 beats per minute and regular, respiratory rate 16/min, temperature 99.1°F

Tympanic membranes clear; throat clear; + maxillary sinus tenderness
Neck: supple without LA
Lungs: clear
Heart: RRR without murmur

Assessment: sinusitis, hypertension
Plan:
1. Amoxicillin 500 mg tid × 14 days
2. Sudafed over-the-counter
3. Continue current captopril
4. Follow-up in 2 weeks if needed or as needed if new or worsening symptoms