Medical Decision Making: Guide to Improved CPT Coding

Jim Holt, MD, Ambreen Warsy, MD, and Paula Wright, CPC-E/M, CPMA

**Background:** The Current Procedural Terminology (CPT) coding system for office visits, which has been in use since 1995, has not been well studied, but it is generally agreed that the system contains much room for error. In fact, the available literature suggests that only slightly more than half of physicians will agree on the same CPT code for a given visit, and only 60% of professional coders will agree on the same code for a particular visit. In addition, the criteria used to assign a code are often related to the amount of written documentation. The goal of this study was to evaluate two novel methods to assess if the most appropriate CPT code is used: the level of medical decision making, or the sum of all problems mentioned by the patient during the visit.

**Methods:** The authors—a professional coder, a residency faculty member, and a PGY-3 family medicine resident—reviewed 351 randomly selected visit notes from two residency programs in the Northeast Tennessee region for the level of documentation, the level of medical decision making, and the total number of problems addressed. The authors assigned appropriate CPT codes at each of those three levels.

**Results:** Substantial undercoding occurred at each of the three levels. Approximately 33% of visits were undercoded based on the written documentation. Approximately 50% of the visits were undercoded based on the level of documented medical decision making. Approximately 80% of the visits were undercoded based on the total number of problems which the patient presented during the visit. Interrater agreement was fair, and similar to that noted in other coding studies.

**Conclusions:** Undercoding is not only common in a family medicine residency program but it also occurs at levels that would not be evident from a simple audit of the documentation on the visit note. Undercoding also occurs from not exploring problems mentioned by the patient and not documenting additional work that was performed. Family physicians may benefit from minor alterations in their documentation of office visit notes.

**Key Words:** CPT audits, CPT coding, medical decision making

---

**Key Points**
- All previous studies of CPT coding have audited the written encounter note.
- Medical decision making (MDM) is the most appropriate basis for selecting the CPT code for an office visit, as long as the history or the physical exam documentation also support that level.
- Using MDM to retrospectively audit office visit notes showed that 50% of visits were undercoded. A small amount of additional documentation would allow the higher code.
- Addressing all patient-mentioned problems during the visit, although clearly more time-consuming, would allow a higher CPT code to be used for 80% of the audited visits.
larger practices. Some professional coders attain auditor status through further study as a Certified Evaluation and Management Coder. Author PW is an auditor for the East Tennessee State University (ETSU) medical practices.

Despite the critical importance of appropriate CPT coding to the financial viability of a practice, the medical literature contains little on the topic. Only three studies directly examined family physicians’ selection of CPT codes. Each of these studies showed only about half of the CPT codes selected were the same as those recommended by professional coders. In addition, family physicians undercoded visits of established patients (established visits), and overcoded new patient visits. In the most recent study, family doctors committed more undercoding errors than overcoding errors. In the other two studies, the errors were more evenly split between overcoding and undercoding.

A few articles indicated that no group using CPT visit codes consistently agreed on the level of complexity: on six sample office notes, physicians agreed on the appropriate CPT code 55% of the time—similar to the results of prior studies—but 300 certified coders agreed only 57% of the time. One article on family physician visits, and one on physicians caring for patients with Alzheimer disease, suggested that the written documentation underestimated the complexity of the visit—ie, physicians were not writing down all the historical information, the physical exam elements, and the thought processes involved in their medical decision making. Since the written visit note supports the level of the CPT code, inadequate documentation supports only a lower-complexity CPT code. Such a discrepancy could result in substantial underpayment.

All but two of the studies directly examining coding involved retrospective chart audits, to match the written visit note to the appropriate CPT code. However, these last two articles, which suggested undercoding due to incomplete documentation, were exclusively qualitative studies.

Professional coders may more consistently agree on codes for procedures. A single study assessed accuracy of coding for interventional radiology encounters: only 44% were correctly coded, though professional coders assigning corrected codes agreed in 91% of the encounters.

The coding literature is distinct from the medical literature, and in fact is not listed in MEDLINE searches. The literature contains reviews of overutilized codes, and articles on the accuracy of coding for office procedures, visits, and surgeries. The literature also acknowledges the difficulty of documenting consistent, accurate codes for office visits; office and hospital visits are among the categories with the highest error rates, in the survey by the Centers for Medicare & Medicaid Services. However, the authors found no studies evaluating office visit coding accuracy.

Guidelines for CPT coding break down a visit note into three components: the history—including the 8 defined elements of the history of present illness (HPI), the physical exam (PE), and the medical decision making (MDM). According to the Center for Medicare and Medicaid Services, the agency which administers Medicare, MDM should be the most important determinant in selecting the CPT code for any visit. However, of the three components of the visit, medical decision making is clearly the most complicated and poorly understood. The coding subprograms in electronic health records (EHRs) tend to rely on the history and physical exam; the subprogram simply counts the entries made in the history and physical exam templates to estimate the level of complexity. In contrast, to accurately describe the level of MDM, the subprogram must be told if each condition is new or old, stable or unstable, and in the process of workup or not (Fig. 1). Accordingly, EHR coding programs are at risk of substantial overcoding. This is not a good strategy for increasing practice income, given hefty penalties for systematic overcoding! In contrast, provider knowledge of the criteria for medical decision making could safely increase both coding accuracy and practice income.

As noted above, Medicare emphasizes using MDM as the usual basis for selecting the CPT code. However, other factors may affect which CPT code is appropriate:

1. Inadequate documentation of both history and physical exam. In an established visit, two of the three components—history, physical exam, and MDM—must reflect the same or a higher level of complexity than the CPT code charged. Higher MDM does not guarantee a higher CPT code; a poorly documented history and physical exam would still necessitate a less complex CPT code.

2. Inadequate documentation of tests ordered, medications prescribed (or recommended over-the-counter medications), problems addressed, or workup planned. These factors reduce the level of MDM (Fig. 1).

It may be common that MDM is more complex than the CPT code charged—whether due to poor documentation of the history and physical exam, or due to poor understanding of the coding guidelines. Most likely, if the physician addressed each problem which the patient mentions, it would frequently increase both MDM and the CPT code charged. The goal of this study is to quantitatively
**Fig. 1 Audit form.**
examine the hypothesis that the written note does not fully reflect the complexity of an office visit, leading to frequent undercoding.

The question for this study was: how did use of either medical decision making, or of all problems the patient mentioned, compare with the use of a retrospective audit of visit note documentation to determine the appropriate CPT code?

Methods

We obtained approval for the project from the ETSU Institutional Research Board. A power analysis demonstrated that an audit of at least 160 charts from each of the two residency programs studied would uncover 15% differences in CPT codes. Since residents graduate at the end of June, we chose the first two weeks in May as a period likely to reflect consistent coding. Since both residency programs use paper charts, and "lost" charts are common, 200 office visits (of 551 and 595 total office visits in Johnson City (Res A) and Bristol (Res B), respectively) from May 1, 2005 to May 15, 2005 were randomly selected via computer program in each of the two East Tennessee State University family medicine residency programs.

We used the medical decision making auditing template currently employed by the compliance officer at ETSU to evaluate appropriate coding on 3 levels: on the first level, the CPT code appropriate for the written visit note; on the second level, the CPT code most appropriate for the documented level of medical decision making; and on the third level, the CPT code most appropriate for the number and complexity of the problems mentioned by the patient, regardless of the documentation supplied about these problems (Appendix A). The third level estimated the most complex office visit which could be produced from any given encounter (CPT-max). In addition, the author who serves as residency faculty (JH) also documented when undercoding may have been associated with Medicare teaching guidelines: moderately and highly complex codes can only be charged by a resident when the preceptor also sees the patient. (Unfortunately, in a retrospective chart review, an auditor cannot ascertain the reasons for undercoding.)

All three authors—a professional coder (PW), a PGY-3 family medicine resident (AW), and a residency faculty member who teaches a coding curriculum (JH)—reviewed 181 notes from Res A and 170 notes from Res B. The randomization program selected 41 notes (14 from Res A and 27 from Res B), which could not be found; these were excluded from the study. Also, 5 notes from Res A and 3 from Res B were available for at least one reviewer, but were “lost” before all 3 authors could audit them; these charts were also excluded.

The authors completed a medical decision-making template (Fig. 1) for each visit note; the Family Medicine Research Division at ETSU scored and compiled the results. Each author completed each medical decision-making tem-

Fig. 2 Documentation relationship to code charged: comparison of actual CPT code to code recommended by audit, based on written note. *D>C: documentation supports a higher CPT code than the code charged; †D<C: documentation supports a lesser CPT code than the code charged; ‡D=C: same.

Fig. 3 MDM relationship to code charged: comparison of CPT codes with audited levels of MDM, Medical decision making.
Holt et al • Medical Decision Making: Guide to Improved CPT Coding

signed the same level of medical decision making for 23% of the visit notes; two of three authors assigned the same level of medical decision making for 62% of the visit notes (Table 2). Agreement with the actual CPT codes charged was low: two or more authors agreed with the assigned code for only 39.3% of the audited visit notes.

Of the visits classified as undercoded, nearly all were undercoded by only one CPT coding level: 97.7% of undercoded visits based on the written note, 96.8% based on MDM, and 99% based on CPT-max.

Discussion

This study breaks new ground: it confirms that undercoding is worse than was previously believed. Two studies evaluated family physicians’ coding accuracy with the current documentation guidelines. Kikano et al² found 21% undercoding and 19% overcoding; Chao et al³ found 21% and 24%, respectively. Both studies reviewed written office visit notes (1st level in this study). King et al used 6 fabricated visit notes; 205 family physicians randomly selected from the membership roll of the Illinois Academy of Family Physicians coded those 6 visit notes. For established patient visit notes, 33% were undercoded; for new patient visit notes, 82% were overcoded.¹ (In all studies, physicians tended to overcode new visits, as noted in the above studies. Our study audited only established visits. Second, no other study evaluated our “second and third levels.” Our second level, a CPT code recommendation based on MDM, labelled half of the office visits undercoded. (Note that the severity of the problem, as well as the extent of the workup planned, were the major considerations in determining the level of MDM. These two factors should prompt a provider to consider a higher code.) Our third level, CPT-max, suggested that 80% of 99213 codes should be 99214 codes, if all problems which the patient mentioned were addressed. If only our first level, the retrospective audit based on the written note, is selected, our level of undercoding for established patients is nearly identical to the rate found in the King et al¹ study.

The rule requiring a preceptor to see a patient to charge a 99214 or 99215 in a teaching program could have been a contributor to undercoding in up to 60% of office notes coded 99213, but audited as 99214.

As noted above, the literature suggested some undercoding of established office visits, counterbalanced by overcoding of new office visits. A standard audit cannot elicit the extent of undercoding based on MDM and CPT-max. Our study suggests that undercoding is much more prevalent than overcoding, but that poor documentation obscures its extent.

As noted already, appropriate CPT coding is critical for a family physician’s (FP’s) financial viability, and for avoiding huge financial penalties. These penalties could equal three times the government’s damages, plus civil penalties of $5,000 to $11,000 per claim¹¹ However, both this study and the existing literature confirm that both professional coders and medical providers disagree when a given CPT code is appropriate. Medicare has consistently refused to publish any percentage as “an acceptable error rate,” instead stating that their auditors will investigate those who significantly deviate from current bell curves of code usage. Informally, local Medicare representatives and the practice attorney both verbally suggested to author PW that 20% would be an “acceptable error rate,” but Medicare reserves the right to audit and to penalize providers with lower error rates.

Finally, since all published literature suggests error rates among both providers and coders is more than 20%, it is clear that the current coding system is suboptimal: it exposes nearly all providers to risk of penalties from an audit, in the absence of any attempts at fraud.

Nevertheless, FPs depend on the current flawed system for their incomes. It is likely that most family physicians would perform and document a small number of additional HPI or PE elements to permit them to charge a higher CPT

Table 2. Inter-rater agreement

<table>
<thead>
<tr>
<th>Reviewer agreement</th>
<th>MDM (%)</th>
<th>Documentation (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>None agree</td>
<td>46 (13.1)</td>
<td>17 (4.8)</td>
</tr>
<tr>
<td>Two of three agree</td>
<td>226 (64.4)</td>
<td>219 (62.4)</td>
</tr>
<tr>
<td>Faculty and resident</td>
<td>57 (25.2)</td>
<td>52 (23.7)</td>
</tr>
<tr>
<td>Faculty and professional coder</td>
<td>122 (54.0)</td>
<td>105 (47.9)</td>
</tr>
<tr>
<td>Resident and professional coder</td>
<td>47 (20.8)</td>
<td>62 (28.3)</td>
</tr>
<tr>
<td>All agree</td>
<td>79 (22.5)</td>
<td>115 (32.8)</td>
</tr>
</tbody>
</table>

¹MDM, medical decision making.
code. Therefore, much of the 50% of undercoding based on MDM is likely due to a poor understanding of the MDM criteria. For the first time, family physicians have a way to estimate the cost of suboptimal coding.

Finally, we estimate that 80% of 99213 visits could be coded higher, simply by addressing each problem which the patient brings up, ie, using the CPT-max. The FP can evaluate all mentioned problems to maximally reimburse any individual office visit. This knowledge can be very useful when building a practice, or when experiencing a “slow day” in an established practice. Of course, addressing each problem which a patient mentions, even if a relatively cursory evaluation of many minor problems is used, may lengthen an FP’s workday unacceptably. Yet offering this more thorough visit, when possible, is likely to increase patient satisfaction and quality of care. The 80% potential for upcoding may indeed be of some practical use to FPs.

**Limitations of the Study**

The study’s performance in two residency programs may affect the results. One rule specific to teaching programs may cause excess undercoding by MDM: a preceptor must see the patient to permit coding a 99214 or 99215. However, the preceptor can use the primary care exclusion for 99213 or 99212 visits, and need not see the patient face-to-face. Also, one would expect coding accuracy to be lower at baseline among residents, since they are still learning how to code. The study’s external validity may be limited, as it was restricted to academic practices.

**Suggestions for Further Research**

The family medicine resident author (AW) reported a much enhanced understanding of MDM after assisting with the audits for this study. Accordingly, enhanced education on coding, especially coding the level of MDM, may reduce undercoding. Following the results of this study, author JH has instituted a much more thorough coding curriculum at Residency A. Assessing the effect of this enhanced education may be interesting.

Also, there is no definite way to eliminate uncertainty about potential bias due to conducting this research in training programs, unless a similar study were done in a private practice.

**Conclusions**

We have established that undercoding in family medicine practices is very much underestimated. This is especially true if only the entire written visit note is considered. Documenting the history and physical exam more thoroughly, to match the MDM, can increase revenue substantially without a great deal of extra work. In addition, family physicians can moderately increase their work to address a patient’s other stated problems in order to obtain additional reimbursement in the preponderance of their office encounters (Appendix B). Therefore, it behooves family physicians to thoroughly understand the criteria for setting the level of MDM.

**Appendix A: Evaluation of an Office Visit Note—Mary**

Mary (not her real name), a 67-year-old woman, presents to your office two days after a syncopal episode. She had walked to the bathroom at 4 a.m., and as she walked back, she rapidly became lightheaded. She grabbed at the near corner of the bed, but ended up falling against it, landing on her left side on the floor. She thinks she lost consciousness briefly, then struggled up by grabbing the corner of the bed. She has a history of hypertension, hypercholesterolemia, and tobacco abuse.

Physician X sees Mary, reviews her old records, and obtains the social history that Mary’s husband died 3 months ago. Mary has lost ten pounds since his death. Dr. X then completes an extensive review of systems (11 systems), and a physical exam with orthostatic pulse and blood pressure, and thorough exams of the respiratory, cardiac, and neurologic systems (24 physical exam elements). Mary has borderline-positive orthostasis, and the remainder of her exam is normal. Dr. X offers Mary a choice between immediate hospital admission with a partial syncopal workup done inpatient, or prompt scheduling of a full outpatient syncopal workup. Mary chooses to go home with close followup, and promises to increase her intake of fluids and calories. Dr. X orders an echocardiogram, a magnetic resonance imaging (MRI) of the head, carotid ultrasound, and several blood tests.

**Coding commentary:** Although many providers would admit this patient for a syncopal workup, clearly qualifying for high medical decision making (MDM) (Figure 1), after two days without symptoms, it is acceptable to offer an outpatient workup. Dr X’s visit satisfies the requirements for a 99215 code in all 3 areas: history, physical exam, and MDM:

- History: Four history of present illness (HPI) elements (duration, context, severity, associated signs and symptoms) for the chief complaint of syncope; 11 systems reviewed; and review of both past and social history.
- Needed for 99215: Four HPI elements, >9 systems other than HPI system, 2 of 3 past/family/social.
- Physical exam: Twenty-four physical exam elements documented.
- Needed for 99215: At least 18 physical exam elements documented.
- Medical decision making: Syncope is a new problem for which workup is scheduled: 4 points in Box A (supports high MDM). In Box B, blood tests are clinical labs, the echocardiogram is a medical test, and the MRI is a radiology test: 1 point each, for 4 points (supports high MDM).
- Needed for 99215: 2 of 3 in Boxes A to C supporting high MDM.

*Southern Medical Journal* • Volume 103, Number 4, April 2010

321
However, many family physicians, not confident in their coding skills, would charge a 99214. If 99214 were charged, our audit process would look as follows:

Level 1: (Actual documentation): undercoded—supports 99215
Level 2: (Level of MDM): undercoded—supports high MDM/99215
Level 3: (All problems mentioned): syncope and grief, with weight loss—99215

It is conceivable that a physician may perform a less thorough evaluation of this problem:

Dr. Y, Mary’s long-standing physician, knows Mary often gets dehydrated with viral gastroenteritis. He questions her about any concurrent symptoms when she blacked out. She reports having vomited the morning of the previous day, followed by 4 stools the night before. She got up at 4 a.m. to have a diarrheal stool, and blacked out on the way back to bed. She successfully rehydrated with copious fluids since her syncopal episode, and feels “much steadier on my feet” when she presents. The HPI has 3 elements, the review of systems covers 3 additional systems. Dr. Y checks orthostatic pulse and blood pressure, with an exam of the chest and abdomen, with cursory neurologic exam (10 exam elements).

Dr. Y reduces her antihypertensive meds. He charges a 99213. It is conceivable that a physician may perform a less thorough evaluation of this problem:

Level 1: Correctly coded. The MDM supports a 99214 (Figure 1), but the history and physical exam both support 99213. (A 99214 would require 4 HPI elements or 12 physical exam elements.)
Level 2: Undercoded—MDM moderate, supports 99214.
Level 3: Undercoded—syncope, viral gastroenteritis, orthostatic hypotension, and hypertension are all addressed, and Dr. Y. changes the medications—supports 99214.

Appendix B: Financial Implications of Undercoding

The “average” family physician may see 120 patients per week, and code 12 visits (10%) 99212, 60 visits (50%) 99213, 45 visits (37%) 99214, and 3 visits (3%) 99215, for a total reimbursement of $7916.70. Our results would suggest the following financial consequences from undercoding:

Level 1: 25% net undercoding of 99212s and 99213s (33% undercoded minus 8% overcoded). By the 2007 Medicare rates quoted above, 9 × 99212, 48 × 99213, 60 × 99214, and 3 × 99215 yields an additional $498.99 per week ($8415.69 total per week).
Level 2: 40% net undercoding of 99212s and 99213s (50% undercoded minus 10% overcoded). By 2007 Medicare rates, 7 × 99212, 41 × 99213, 69 × 99214, and 3 × 99215 yields $802.78 more than the basic weekly total, an increase of over 10% (total $8719.48).
Level 3: 84% net undercoding of 99212s and 99213s (85% undercoded minus 1% overcoded).

By 2007 Medicare rates, 2 × 99212, 20 × 99213, 95 × 99214, 3 × 99215 yields $1663.64 more than the baseline weekly total, an increase of over 21% (total $9580.34).

Acknowledgments

The authors gratefully acknowledge the assistance of Ivy Click and the ETSU Research Department in the scoring and statistical analysis used in this study, and for reviewing the manuscript. Also, the authors appreciate Drs. Zia Rahman and Max Bayard for reviewing the manuscript.

References