CLINICAL CODING

Outpatient evaluation and management coding after the 2021 revision: medical decision making, time, and the evidence on upcoding

ABSTRACT

On January 1, 2021, the American Medical Association and CMS retired the 1995 and 1997 history-and-examination bullet tables for office and outpatient evaluation and management (E/M) visits (CPT 99202–99215), allowing clinicians to choose visit levels solely by medical decision making (MDM) or total date-of-service time. The reform carried two explicit hopes: shorter notes with less clerical bloat, and narrower payment gaps between cognitive and procedural specialties. Five years of claims and EHR audit-log studies show a sharper, less flattering result. In a national pre–post study of 303,547 clinicians across 389 organizations, level 3 visits fell by 2.41 percentage points immediately while note length and EHR time stayed effectively unchanged. In a 100% Medicare Carrier File analysis of 180,624 physicians, high-intensity (level 4 and 5) coding rose in nearly every specialty — by 17.8 percentage points in psychiatry and 10.1 points in orthopedic surgery — while the primary-care-to-specialist payment gap narrowed by just 2.0% ($825.1). Meanwhile, CMS added visit-complexity code G2211 in 2024 (slated to transition to modifiers B1 and B2 in CY 2027), FY 2025 CERT improper payments reached 6.55% ($28.83 billion), and commercial payers introduced automated pre-review E/M downcoding edits. This review examines the rules, the empirical record and what AI coding suggestions in an EMR must prove.

Type Evidence and regulatory review **References** 25 **Reading time** 15 min **Last reviewed** November 2026

1 The end of bullet counting, and what actually changed

For a quarter century, outpatient office visits were graded like tax forms. Under CMS's 1995 and 1997 Documentation Guidelines, billing an established-patient level 4 visit (CPT 99214) drove clinicians to pack notes with review-of-systems checklists and multi-organ physical exam templates — not because those systems mattered clinically, but because missing one bullet point could drop a claim on audit.[1,2]

Effective January 1, 2021, the AMA CPT Editorial Panel and CMS overhauled office and other outpatient E/M codes (99202–99215; code 99201 was deleted).[1,3] History and physical examination still belong in the chart only "as medically appropriate" and no longer contribute points to code selection. Instead, the clinician selects the visit level using either *medical decision making* (MDM) or *total time* spent by the billing practitioner on the date of the encounter, including non-face-to-face chart review, order entry, care coordination and documentation.[1,3] At the

same time, CMS raised work Relative Value Units (wRVUs) for established-patient visits by up to 23.5% for 99215 while lowering the Physician Fee Schedule conversion factor from $36.09 to $34.89 to maintain statutory budget neutrality.[4]

−2.41 pp
immediate drop in level 3 E/M visits across 303,547 clinicians in 2021

−2.0%
shrinkage in median Medicare payment gap between PCPs and specialists

6.55%
FY 2025 Medicare FFS CERT improper payment rate ($28.83 billion)

0.33 wRVU
Medicare work RVU assigned to complexity add-on HCPCS G2211

Five years of claims data, EHR audit logs and regulatory updates now show how the revision played out in practice: billing levels shifted immediately, note bloat barely budged, payers answered higher E/M intensity with automated downcoding edits, and Medicare added (then reshaped) a longitudinal complexity add-on.[2,4,5,6]

2 How MDM and total time govern code selection

Under the post-2021 framework, an office visit qualifies for a level when it meets or exceeds two of the three MDM elements — number and complexity of problems addressed, amount and/or complexity of data to be reviewed and analyzed, and risk of complications and/or morbidity or mortality of patient management — or meets the total-time threshold on the date of service.[1] Table 1 sets out the established-patient and new-patient thresholds as revised by AMA CPT.

TABLE 1

Post-2021 office and outpatient E/M levels (CPT 99202–99215): MDM level (two of three elements required) and total date-of-service time thresholds (reflecting the 2024 CPT alignment from ranges to minimum thresholds).[1,3]

Level	New / Established CPT	MDM complexity	Minimum total time (New / Est.)	Representative pain-clinic MDM anchor
Level 2	99202 / 99212	Straightforward	15 min / 10 min	1 self-limited minor problem; minimal risk
Level 3	99203 / 99213	Low	30 min / 20 min	1 stable chronic illness (e.g., controlled osteoarthritis) with low management risk
Level 4	99204 / 99214	Moderate	45 min / 30 min	1 chronic illness with exacerbation, or ≥2 stable chronic illnesses, plus prescription drug management
Level 5	99205 / 99215	High	60 min / 40 min	Chronic illness with severe exacerbation or threat to bodily function, plus drug therapy requiring intensive toxicity monitoring or decision on high-risk intervention

Three recurring audit traps in the MDM table

Because the MDM grid uses clinical language rather than bullet counts, three specific misunderstandings account for a disproportionate share of outpatient coding disputes.[1,7]

- **"Addressed" is not "listed."** Copying a patient's entire 14-item problem list into an Assessment section does not earn credit for multiple chronic conditions. AMA CPT guidelines state explicitly that a problem is addressed only when it is evaluated or treated at the encounter; a notation that another specialist manages a condition, without clinical assessment or impact on today's plan, does not count.[1]
- **Prescription drug management versus high-risk monitoring.** Managing a prescription medication places the risk column in *Moderate* (supporting 99214 when paired with a moderate problem or data element). Reaching *High* risk for 99215 requires drug therapy necessitating *intensive monitoring for toxicity* — monitoring directed at a serious adverse effect of the drug itself (not routine therapeutic-effectiveness checks or annual lab panels), documented with clinical rationale.[1]
- **Double-counting separately billed tests and procedures.** Reviewing or ordering a laboratory test, imaging study or urine drug screen that the practice bills independently (with its own CPT/HCPCS code) cannot simultaneously be counted in the MDM data column, and time spent performing a separately reportable procedure (such as a trigger-point injection or joint aspiration) must be carved out of E/M total time.[1,3]

3 Note length did not fall — why clerical bloat survived

CMS projected that dropping mandatory history and exam bullets would trim clinical documentation burden substantially. Large-scale EHR telemetry told a different story.

Apathy, Hare, Fendrich and Cross analyzed weekly provider-level E/M billing and Epic Signal EHR usage metadata for 303,547 physicians and advanced practice providers across 389 US health care organizations from September 2020 through April 2021.[2] Coding moved immediately: level 3 visits fell by 2.41 percentage points (95% CI, −2.48 to −2.34 pp) to 38.5% of all E/M visits, a 5.9% relative decline from fall 2020. Yet note length and total EHR time remained "effectively unchanged" across the post-policy window; statistically detectable differences were too tiny to be clinically meaningful.[2]

A companion cross-sectional analysis of 203,728 US ambulatory physicians over the same 2020–2021 period explained why note bloat matters and why templates did not fix it.[8] Physicians in the top decile of overall note length spent 4.2 more minutes per visit in notes than median physicians (14.8 vs. 8.1 minutes; B = 4.22, 95% CI 3.71–4.74; P < .001), spent 39% more time in the EHR outside scheduled hours (7.3 vs. 4.0 minutes per visit), and closed 5.6 percentage points fewer encounters on the same day (64.5% vs. 73.5%).[8] Copy-paste text and manual text both drove higher note time — top-decile manual writers spent 10.9 more minutes per visit in notes — while institutional templates, medico-legal habit and lingering audit anxiety kept auto-populated review-of-systems blocks glued inside notes long after Medicare stopped asking for them.[8,9]

4 Coding intensity, specialty redistribution and the 2021 payment shift

The second policy goal of the 2021 E/M overhaul was redistributing Medicare Physician Fee Schedule dollars toward cognitive and primary care specialties. Neprash, Golberstein, Ganguli and Chernew evaluated that hypothesis using 100% Medicare Carrier File claims for 180,624 US office-based physicians comparing July–December 2020 with July–December 2021.[4]

First, holding visit volume constant by repricing 2020 services at 2021 rates simulated gains up to 11.0% ($3,683.1) for the median family physician and losses of −3.3% (−$4,557.0) for the median radiologist.[4] Second, observed coding intensity — the share of E/M visits billed at level 4 or level 5 (99204, 99205, 99214, 99215) — rose across nearly every specialty. The five largest jumps in median high-intensity share occurred in psychiatry (46.0% to 63.8%, +17.8 percentage points), orthopedic surgery (24.5% to 34.6%, +10.1 pp), urology (47.7% to 54.8%, +7.1 pp), family practice (62.5% to 67.5%, +5.0 pp) and otolaryngology (28.7% to 33.3%, +4.6 pp).[4] A separate specialty study of Medicare urology office visits from 2010 through 2021 documented the same inflection: 99213 had been the modal established visit for a decade until 2021, when 99214 overtook it at 46% of all established visits.[10]

Because procedural and cognitive specialists also billed large volumes of office visits and shifted up in coding intensity, the observed median Medicare payment gap between primary care physicians ($41,193.3 to $45,962.4) and specialists ($81,453.1 to $85,397.1) shrank by only $825.1 — a 2.0% relative reduction.[4]

TABLE 2

Measured effects of the 2021 outpatient E/M revision in national EHR and Medicare claims studies.

Study	Data source and sample	Measured change after Jan. 2021
Apathy et al., 2022[2]	Epic Signal metadata, 303,547 clinicians across 389 organizations	Level 3 share fell 2.41 pp to 38.5%; note length and EHR time effectively unchanged
Neprash et al., 2023[4]	100% Medicare Carrier File, 180,624 office-based physicians	High-intensity (L4/L5) share rose in nearly all specialties (psychiatry +17.8 pp; orthopedics +10.1 pp; family practice +5.0 pp); PCP–specialist payment gap shrank 2.0% ($825.1)
Apathy et al., 2023[8]	Epic Signal cross-section, 203,728 ambulatory physicians	Top-decile note length tied to +4.2 min/visit in notes, +39% after-hours EHR time and −5.6 pp same-day visit closure
Urology Medicare series, 2023[10]	CMS Physician/Procedure Summary, 2010–2021	Mean visit payment rose from $99.42 (2020) to $110.95 (2021); 99214 overtook 99213 as modal visit at 46%

5 G2211, modifier 25, and the CY 2027 shift to modifiers B1 and B2

Because the 2021 wRVU increase closed little of the primary-care pay gap, CMS implemented HCPCS add-on code **G2211** (0.33 work RVUs) in the CY 2024 Physician Fee Schedule final rule (88 FR 78974–78975) after a three-year statutory delay.[11,12] G2211 pays for visit complexity inherent to an ongoing, longitudinal clinical relationship —

either serving as the continuing focal point for all of a patient's needed health care services, or providing ongoing care for a single, serious or complex condition (a category that explicitly reaches longitudinal chronic-pain and specialty care, not exclusively primary care).[11,13]

Two payment edits govern G2211 operationally. First, Medicare denies G2211 automatically whenever the base E/M code carries **modifier 25** (a significant, separately identifiable E/M on the same day as a minor procedure), subject to narrow CY 2025 preventive-service exceptions finalized in 89 FR (CMS-1807-F).[12,14] Second, in the CY 2027 Medicare Physician Fee Schedule final rule released in November 2026, CMS finalized replacing standalone code G2211 for dates of service starting January 1, 2027 with claim-line **modifier B1** (and **B2** for Shared Savings Program and LEAD ACO clinicians) on the base E/M line itself, preserving the modifier 25 exclusion.[13,15] Commercial insurers and Medicare Advantage plans set independent policies on G2211 and its successors; assuming commercial parity leads directly to routine contractual denials.

6 Upcoding, downcoding, CERT audits and automated payer edits

The post-2021 climb in 99214 and 99215 claims does not by itself prove fraud: relaxing rigid exam-bullet penalties legitimately reclassified visits that were clinically moderate or complex under MDM all along.[4,16] Yet payers and federal auditors have responded aggressively to the upward shift.

Under CMS's Comprehensive Error Rate Testing (CERT) program, the FY 2025 Medicare fee-for-service improper payment rate stood at 6.55%, or $28.83 billion, driven predominantly by insufficient documentation rather than proven intentional fraud.[6] Historically, an HHS Office of Inspector General review of Medicare E/M claims (OEI-04-10-00181) found roughly 55% incorrectly coded or insufficiently documented — about 26% upcoded and about 15% downcoded, the latter underpaying physicians by an estimated $1.8 billion in a single year.[6,17] While those OIG figures predate the 2021 guidelines, they illustrate an enduring reality: coding accuracy fails in both directions. Upcoding carries False Claims Act exposure (31 U.S.C. §§ 3729–3733) and OIG outlier audits when a clinician's 99215 curve sits far above specialty peers; chronic defensive downcoding creates silent revenue loss that never shows up in a denial queue because the claim pays cleanly at the lower level.[6,18]

A newer operational pressure emerged in late 2025: automated *pre-review payer downcoding*. Cigna's R49 Evaluation and Management Coding Accuracy policy, effective October 1, 2025, automatically reduces submitted 99204–99205, 99214–99215 and 99244–99245 claims by one level before reviewing the chart, shifting the burden onto the clinic to submit records and appeal — prompting formal objection from the AMA and specialty societies in August 2025.[6,19] How automated payer edits interact with clinic revenue-cycle automation is analyzed in this library's companion paper on revenue cycle automation.

LIMITATIONS OF CLAIMS-BASED CODING STUDIES

Claims datasets such as the Medicare Carrier File observe the CPT code billed, not the underlying clinical note.[4] They therefore cannot separate appropriate reclassification under the 2021 MDM grid from genuine overcoding without paired chart audits. Conversely, pre-2021 OIG E/M error percentages cannot be quoted as post-2021 baselines because the rules changed fundamentally on January 1, 2021.[17]

7 What AI-assisted E/M coding in an EMR must — and must not — do

Because an ambient note generator can turn a ten-minute visit into a crisp, richly structured page of prose, generative AI introduces a distinct compliance hazard: *synthetic documentation inflation*. If a language model pulls inactive chronic problems from the chart header into the Assessment, embellishes risk language beyond what was discussed, or suggests a level 4 or 5 code to maximize RVUs, the resulting note can appear superficially audit-ready while misrepresenting the encounter.[20,21]

An EMR that drafts E/M level suggestions inside a human-in-the-loop architecture should enforce four structural safeguards, consistent with the library's approval-gates specification and automation-bias review:

1. **Count only problems actually evaluated in the visit.** Bind every MDM problem credit to an explicit statement or transcript anchor from the encounter, never to passive chart carry-forward.[1]
2. **Display the exact 2-of-3 MDM evidence trail.** Show the clinician which specific problems, data items and management risks (or documented date-of-service minutes) generated the draft code, rather than presenting a bare 99214 badge that encourages rubber-stamping.[1,22]
3. **Enforce NCCI, modifier 25 and time-carve-out rules at sign-off.** Flag same-day minor procedures automatically so procedure time is excluded from E/M time and modifier 25 / G2211 (and CY 2027 modifier B1/B2) conflicts are caught before submission.[12,13]
4. **Require clinician confirmation and monitor internal distributions.** Keep the suggested code strictly draft-only until the clinician attests it, and audit periodic internal E/M curves against peer specialty benchmarks — checking both outlier high-level coding and systematic payer downcoding.[6,23,24,25]

REFERENCES

Entries 1, 3, 11, 12, 14, 15, 17, 18 and 24 are primary AMA CPT guidance, federal statutes, Federal Register rules, CMS fact sheets and HHS OIG reports. Entries 2, 4, 8, 9, 10, 16, 20, 21 and 25 are peer-reviewed

empirical studies in Annals of Internal Medicine, JAMA, JAMA Network Open, Health Services Research, JAMIA and Urology Practice; they carry the primary empirical weight here. Entries 5, 6, 7, 13, 19, 22 and 23 are compliance, payer-policy and trade analyses, cited for operational and 2025–2026 policy updates.

1. American Medical Association. *CPT Evaluation and Management (E/M) Office or Other Outpatient (99202–99215) and Prolonged Services Code and Guideline Changes*. Effective 1 January 2021; time-threshold alignment effective 1 January 2024. ama-assn.org GUIDELINE

2. Apathy NC, Hare AJ, Fendrich S, Cross DA. *Early Changes in Billing and Notes After Evaluation and Management Guideline Change*. Annals of Internal Medicine. 2022;175(4):499–507. doi:10.7326/M21-4402 COHORT

3. Centers for Medicare & Medicaid Services. *Medicare Program; CY 2020 Revisions to Payment Policies Under the Physician Fee Schedule (adopting 2021 E/M framework)*. Final Rule, 84 Fed. Reg. 62568 (15 Nov. 2019). federalregister.gov FINAL RULE

4. Neprash HT, Golberstein E, Ganguli I, Chernew ME. *Association of Evaluation and Management Payment Policy Changes With Medicare Payment to Physicians by Specialty*. JAMA. 2023;329(8):662–669. doi:10.1001/jama.2023.0879 OBSERVATIONAL

5. Centers for Medicare & Medicaid Services. *Calendar Year (CY) 2026 Medicare Physician Fee Schedule Final Rule*. 90 Fed. Reg. (5 Nov. 2025). federalregister.gov FINAL RULE

6. Medicodio. *Upcoding vs Downcoding: Risks, Examples and Audit Triggers*. 8 November 2026. Summarizes FY 2025 CMS CERT improper payment rate (6.55%, $28.83B), OIG OEI-04-10-00181 and Cigna R49 policy. medicodio.ai SECONDARY

7. Centers for Medicare & Medicaid Services. *Evaluation and Management Services Guide*. Medicare Learning Network (MLN006764). cms.gov GUIDANCE

8. Apathy NC, Holmgren AJ, Cross DA. *Documentation Dynamics: Note Composition, Burden, and Physician Efficiency*. Health Services Research. 2023;58(3):674–685; PMCID: PMC10154172. pmc.ncbi.nlm.nih.gov CROSS-SECTIONAL

9. Apathy NC, Hare AJ, Fendrich S, Cross DA. *I Had Not Time to Make It Shorter: An Exploratory Analysis of How Physicians Reduce Note Length and Time in Notes*. Journal of the American Medical Informatics Association. 2022;29(7):1203–1211. doi:10.1093/jamia/ocac060 OBSERVATIONAL

10. Changes in Billing and Reimbursement for Urology Office Visits Following the 2021 E/M Revisions. *Urology Practice*. 2023;10(3):225–232. doi:10.1097/UPJ.0000000000000394 OBSERVATIONAL

11. Centers for Medicare & Medicaid Services. *HCPCS Code G2211 Frequently Asked Questions*. Summarizing CY 2024 PFS Final Rule (88 Fed. Reg. 78974–78975) and MLN Matters MM13272. cms.gov GUIDANCE

12. Centers for Medicare & Medicaid Services. *Medicare and Medicaid Programs; CY 2025 Payment Policies Under the Physician Fee Schedule*. Final Rule (9 Dec. 2024; 0.33 wRVU for G2211 and modifier 25 exceptions). federalregister.gov FINAL RULE

13. AMS Solutions. *G2211 and Modifier B1: When It Applies and Denies*. 11 November 2026. Summarizes Medicare transition from G2211 to modifiers B1/B2 for CY 2027. ams-solutions.com TRADE PRESS

14. Centers for Medicare & Medicaid Services. *Calendar Year (CY) 2025 Medicare Physician Fee Schedule Final Rule Fact Sheet*. 1 November 2024. cms.gov AGENCY RELEASE

15. Centers for Medicare & Medicaid Services. *Calendar Year (CY) 2027 Medicare Physician Fee Schedule Final Rule Fact Sheet*. 4 November 2026. cms.gov AGENCY RELEASE

16. Association of Time-Based Billing With Evaluation and Management Revenue for Outpatient Visits. *JAMA Network Open*. 2022;5(8):e2229504. doi:10.1001/jamanetworkopen.2022.29504 MODELING STUDY

17. Office of Inspector General, HHS. *Improper Payments for Evaluation and Management Services Cost Medicare Billions in 2010*. Report OEI-04-10-00181. May 2014. oig.hhs.gov OIG REPORT

18. False Claims Act, 31 U.S.C. §§ 3729–3733. United States Code. uscode.house.gov STATUTE

19. American Medical Association. *Letter to Cigna Expressing Strong Concerns Over R49 Evaluation and Management Coding Accuracy Downcoding Policy*. 28 August 2025 (reported in Medicodio, 8 Nov. 2026). medicodio.ai SECONDARY

20. Apathy NC, Holmgren AJ, Cross DA. *Physician EHR Time and Visit Volume Following Adoption of Team-Based Documentation Support*. JAMA Internal Medicine. 2024; PMCID: PMC11348094. Notes progress note length increased +431 characters/note after adopting documentation support. pmc.ncbi.nlm.nih.gov COHORT

21. Landon BE. *A Step Toward Protecting Payments for Primary Care*. New England Journal of Medicine. 2019;380(6):507–510. doi:10.1056/NEJMp1810848 PERSPECTIVE

22. Kumetz EA, Goodson JD. *The Undervaluation of Evaluation and Management Professional Services: The Lasting Impact of Current Procedural Terminology Code Deficiencies on Physician Payment*. Chest. 2013;144(3):740–745. doi:10.1378/chest.13-0381 ANALYSIS

23. Frakt AB, Chernew ME. *The Importance of Relative Prices in Health Care Spending*. JAMA. 2018;319(5):441–442. doi:10.1001/jama.2017.20047 VIEWPOINT

24. Office of Inspector General, HHS. *General Compliance Program Guidance*. November 2023. Treats periodic internal coding auditing as a core compliance element. oig.hhs.gov GUIDANCE

25. Coumoundouros C, et al. / Coustasse A, et al. *Upcoding Medicare: Is Healthcare Fraud and Abuse Increasing?* Perspectives in Health Information Management. 2021;18(4); PMCID: PMC8649706. pmc.ncbi.nlm.nih.gov REVIEW

RELATED PAPERS

Automation in the revenue cycle

What is measurable about autonomous coding, denial rates and prior authorization.

Interventional spine procedures and LCDs

Documentation specifications under Palmetto GBA's facet and ESI determinations.

Approval gates specification

Why suggested billing codes stay draft-only until clinician sign-off.