AI in Health Care Documentation & Coding

GAO Report Signals Coming Oversight for AI Scribes and Medical Coding Tools

July 16, 2026

Overview

U.S. clinicians average a 57-hour workweek, including about 7 hours of administrative work such as drafting patient visit notes and reviewing billing paperwork, which contributes to clinician burnout. A growing number of health care providers are turning to AI tools to help, including AI "scribes" that draft clinical documentation during a visit and AI-enabled medical coding tools that help generate insurance claims afterward. These tools may save time and reduce burnout, but the accuracy of the tools can be difficult to verify, and their overall effect on health care spending remains uncertain.

On July 16, 2026, GAO released *AI for Medical Notes and Coding (GAO-26-109116)*, examining how these tools work, the opportunities and challenges they present, and what policymakers need to know to provide effective oversight. GAO found that policymakers need more information about the performance of these tools to determine the appropriate level of oversight needed to help minimize mistakes and ensure proper billing, creating a gap that signals formal oversight requirements are to come.

How AI Is Affecting Clinical Documentation & Medical Coding

AI scribes record the conversation between a patient and clinician in the background of a visit (known as ambient listening) and use conventional and generative AI to draft a written summary for the clinician to review for accuracy. AI medical coding tools analyze patient records and suggest codes for a human coder to review; newer tools are beginning to use generative and agentic AI to assign codes autonomously, a capability that could make coders faster or, eventually, replace parts of that work entirely.

Adoption is accelerating. Clinician use of AI for documentation and coding **grew from 21% to 28% between 2024 and 2026**, according to the American Medical Association. In cited studies, clinicians using AI scribes reduced documentation time by **~20%**, or about two minutes per appointment, and one AI medical coding vendor reported generating codes with more than **95% accuracy** while helping a five-hospital health system **reduce annual coding costs by more than $1 million**.

GAO's Key Findings

Opportunities

GAO identifies three primary benefits:

1. Reduced administrative burden for clinicians.
2. More detailed clinical summaries than manual notetaking.
3. Increased operational efficiency that can reduce the staff needed for administrative tasks.

Challenges

Few studies have evaluated the accuracy of these tools, and some do not retain the underlying patient recordings or transcripts needed to verify their output — inaccuracies could result in patient harm or improper reimbursement. Cost remains a barrier for under-resourced hospitals, health centers, and small practices. GAO also flags a reimbursement risk: tools that capture more diagnoses or services than manual documentation could increase costs borne by Medicare, insurers, employers, and patients. Data privacy and patient consent are additional concerns, since patients are not always informed that a recording is occurring.

Oversight Gap

GAO's central conclusion is that policymakers currently lack sufficient information about tool performance to determine the right level of oversight. This is a precursor to future requirements likely covering accuracy verification, data retention, and governance of increasingly agentic AI coding tools, consistent with oversight frameworks emerging throughout Federal AI policy.

Regulatory & Policy Context

AI adoption in clinical settings does not replace existing obligations. Providers must continue to meet HIPAA privacy and security requirements, CMS billing and documentation rules, and applicable state health-data privacy laws. Because GAO explicitly calls out reimbursement risk and accuracy verification as open questions, health IT vendors should expect future Federal guidance, potentially from CMS, ONC, or HHS, addressing how AI-assisted documentation and coding tools are validated, audited, and reimbursed.

What Does This Mean for Vendors?

GAO's findings create a window of opportunity for vendors ahead of formal procurement standards. Federal health systems including the VA/Veterans Health Administration, DoD's Military Health System, and HHS's Indian Health Service face the same burnout and administrative-burden pressures GAO documents, and adoption is already climbing quickly. However, GAO's oversight gap signals that vendors who can demonstrate accuracy validation, data privacy and retention compliance, and responsible agentic AI governance will be best positioned as formal requirements take shape.

Federal health systems are actively adopting AI scribes and medical coding tools, but GAO's oversight gap means procurement standards are still forming. Vendors that can prove accuracy, protect patient data, and support human oversight of increasingly autonomous coding will be best positioned as requirements take shape.

Strong areas for vendors include:

- AI scribes and ambient listening tools
- AI-assisted and agentic medical coding
- Accuracy validation & audit tools
- Data privacy & retention compliance
- Agentic AI governance
- Administrative workflow automation

Vendors should show clear accuracy evidence, strong data privacy and retention practices, human oversight of autonomous coding, and measurable cost and time savings. They should also monitor GAO, CMS, and HHS/ONC guidance closely, since oversight requirements for these tools appear likely to emerge as this fast-growing category matures.

Contact Us:

Email: Research@carahsoft.com

See more from the Carahsoft Team:
To explore our catalog of Federal, state, and local technology policies, executive orders, and directives shaping public sector modernization scan this QR code.