Computer-Assisted Coding: What's Here, What's Ahead

by James R. Flanagan, MD, PhD, FACP, and Karen Doyle, RN, MSM

Automated coding is already expediting billing. That may be just the first step in applying the technology to information-based medicine.

More and more you hear and read about computer-assisted coding (CAC). What is it, and why do we need it? What should we expect in the years to come?

CAC is already bringing change to coding professionals, introducing new roles and diminishing some old ones. Further, the technology behind CAC has the potential to enable larger changes beyond the billing department. CAC represents a small step in achieving information-based medicine in the years to come.

What It Is, How It Works

CAC is the "use of computer software that automatically generates a set of medical codes for review, validation, and use based upon clinical documentation provided by healthcare practitioners."[1] This definition encompasses both coding for billing purposes as well as coding for a broad range of uses that go well beyond.

Reasons behind automation of the coding process include the dramatically increasing financial pressure to file claims as efficiently as possible. In addition, the rules of coding change over time, introducing new coding requirements. The correct coding initiative and payer-specific coverage policies, for instance, expand the set of rules that a coder must apply. Compliance liability for erroneous claims has increased.

Yet despite these pressures, we cannot afford to have highly trained (and continually retrained) coders analyze every document. For many types of care, the expected level of reimbursement is low and the volume is high. And, as always, there is an effort to further increase accuracy and reliability. In principal, automation can address all of these concerns.

To what extent does CAC succeed now, and what is its potential for the future? To answer this, we go a bit deeper into how CAC works. Whatever the purpose of the coding, both the automated and manual processes have two steps: getting the data from the documentation and applying the rules to assign a code. The first of these steps, getting the data, divides the world of CAC into two approaches. If you already have electronic structured documentation, you have the data. That is, if you have an electronic health record (EHR) in which providers complete online forms using pick lists of predefined choices, then a subset of these data is what you need to apply the rules for coding.

A good example of a setting where encounters can be completely structured in the EHR is the Department of Defense's EHR, which is known as AHLTA. An entire encounter can be described using pick lists, coding the information using MEDCIN vocabulary. A set of computer rules turn this MEDCIN information into an evaluation and management billing code.

AHLTA also provides a good example of the approach's weakness. Even in the military it has proved impossible to get providers to use completely structured documentation for more than a small fraction of clinical encounters. In both military and civilian centers, only certain kinds of encounters (e.g., some subspecialties and selected procedures) lend themselves to structured documentation. For the vast majority of encounters, providers prefer to use natural language text to describe most, or at least parts, of an encounter.

If you don't have structured documentation-that is, if you have only text-then you need a way to turn the text into structured and defined data before you can apply rules of coding. You need automated natural language processing (NLP).

Natural Language Processing and Beyond

One example of NLP involves automated coding radiology reports for ICD-9 and CPT billing. Many organizations attest to the utility of CAC using NLP in this setting. These companies are trying to branch out to apply CAC to other specialties, so far with limited success, for reasons discussed below. The return on investment for radiology customers can be as much a result of reengineering processes as it is from the CAC application itself.[2] Still, it does seem clear that NLP in certain settings can generate codes accurately enough for real-world purposes.

How accurate is accurate enough? Accuracy is how often you get the "right" answer. However, in a field such as coding, there is no gold standard other than those coding results on which two or more coders can agree. We know there can be significant disagreement even among professional coders.

"Reliability," in the technical sense, is a term for reproducibility of a measurement. Reliability is used to measure the extent of agreement among experts. The important thing to realize about CAC is that a computer will reliably follow whatever instructions it receives. An automated application can't have any higher percentage of accuracy than the percentage of agreement (reliability) among human experts. Once the automated process is vetted (i.e., it becomes as accurate as it can get), it will be reliable, producing the same results for the same or similar cases. For this reason, use of CAC could result in reduced need for auditing individual documents. In place of this, the audit will focus on adherence to a process that has been vetted as having acceptable accuracy.

The drawback of the currently marketed NLP methods is the use of statistical approaches, such as artificial neural networks, which need to be trained on precisely the type of document and type of coding that is to be done. Even seemingly small changes in the language of documentation from one healthcare setting to another (let alone from one specialty to another) result in a severe drop in accuracy. This brittleness is hard to explain, because a statistically trained process is a "black box," its inner workings a mystery. Statistical modeling results in a set of mathematical equations that render the answers as the output. Working backward to understand how the equations generated the answer can be difficult.

This drawback is the reason for the limited success that has been experienced using this approach in new areas outside of radiology. Using the statistical NLP approach requires a completely new engine that is as much work as was done to create the first engine. Using the statistical approach will eventually succeed in any give new area but at a large cost in development for each new area. Even companies that market successfully in the radiology domain are looking for new technologies that reduce the effort to move the technology to new areas.

A more robust approach that holds promise for the near future is a newer approach known as natural language understanding (NLU). This technology more closely emulates a human's approach to understanding text. NLU analyzes each sentence much the way we do and identifies the concepts found using a medical knowledge model. This is in sharp contrast to the black box approach that statistical NLP now takes. NLU is not commercially available in CAC as of yet, but it should be within the next year. NLU will be critical to achieving the interoperability of data needed for the future of healthcare.

Preparing for Change

In select settings, CAC already provides advantages. These include increased coding productivity, freeing professionals from mundane tasks, consistent application of rules that can be updated as needed, and reengineering of processes. All of these benefits can also provide return on investment. Wider acceptance of structured EHR documentation would remove one of the barriers to CAC-improvements of NLP in the direction of NLU would improve the ability to robustly handle text documents. To remove another current barrier, costs must come down. The initial investment in hardware, software, and process change can be large.

One of the biggest implications of CAC is the role change for coding professionals. Using CAC, coders will be called upon to review the performance of the automation using sampling techniques. They will review only the most difficult cases that are not handled by the automation. In one sense this is not such a large change from current practice, in that even now there are insufficient numbers of trained coders to review documentation for all encounters.

Automation such as CAC should help an already overtaxed professional group deliver better service. As HIM professionals look at the years ahead, a number of strategies are useful both in adopting CAC today and preparing for the growth of

information-based medicine tomorrow:

- Educate the organization regarding data standards
- Evaluate the organization's needs for management of a data dictionary
- Advocate for interoperability standards as a strategic aim of the organization
- Include interoperability standards in selection criteria for new applications
- Work with vendors toward standards compliance
- Participate in regional health information organizations
- Become the domain expert for health information standards

Demosthenes said, "Small opportunities are often the beginning of great enterprises." The opportunity to use CAC for billing is a small step toward preparing for a future of information-based-medicine.

Wider Applications for CAC Technology

To quote journalist George Will, the future has a way of arriving unannounced. In the year 2000 researchers in the human genome project announced completion of a "working draft" DNA sequence of the human genome. The announcement led to predictions that by 2010 information technology would enable personalized healthcare. Imagine a future where genetic profiling and integrated patient data help identify patients at risk for adverse drug reactions, tailor individualize treatments for a variety of diseases, and improve clinical trials and drug discovery.

Is that future now? The demand for information-based medicine already requires interoperability: the ability to organize, share, and apply information across the healthcare system. Providers collect and manage an increasing amount of clinical data to satisfy the requirements related to quality, cost, and patient care delivery. Healthcare information is the key factor in the personalized medicine equation. Problems integrating data, sharing resources, and transmitting data with the same meaning all limit the adoption of the personal health record in 2006. The need for interoperability is as real today as it will be in 2010.

Automated tools for clinical data abstraction represent another type of CAC that can help meet interoperability demands in 2006 and the emerging demand for information-based medicine. This includes computer-assisted data abstraction and interoperability standards.

Computer-Assisted Data Abstraction

Abstracting is necessary to extract coded data elements from patient records for regulatory, quality, biosurveillance, and research purposes. The same technology used to support CAC for billing can be applied to these tasks.

Several examples demonstrate the use of various schemes, codes, and controlled clinical terminology. The Joint Commission on Accreditation of Healthcare Organizations and the Centers for Medicare and Medicaid Services are working to align a set of hospital quality measures common to both organizations. SNOMED CT provides a standardized means of describing clinical findings and concepts. It also is mapped to other terminologies, facilitating the translation of clinical concepts to different coding systems, enabling them to meaningfully exchange patient information, reimburse claims, and report outcomes. Biosurveillance systems read the ICD-9 codes from physician visit records and emergency department discharge data. Other systems abstract data from emergency department logs, 911 calls, and nurse call-line data. Cancer registries collect information about the occurrence of cancer-the types diagnosed, their locations within the body, the extent at the time of diagnosis, and the treatments prescribed.

Standards to Enable Interoperability

The drive for interoperability will only pick up speed. HIPAA's implementation began with requirements for interoperability in billing codes. Health Level Seven is also working with the Accredited Standards Committee X12N on behalf of the Department of Health and Human Services to create the HIPAA-mandated standard claims attachments for emergency departments, ambulance services, medications, laboratory results, rehabilitative services, and clinical reports. These attachments are clinical information submitted to third-party payers in support of a claim.

IT adopted to support the interoperability needs for HIPAA claims attachments can be applied to research, quality, and biosurveillance data abstraction efforts, making the abstraction and coding processes far more accurate, productive, and secure. As demands increase for coded clinical data elements from patient records to support administration and billing, it is a natural next step to use CAC technology.

The process redesign and technology investment required for CAC is substantial, for either billing or clinical abstracting. Fortunately, there is a large degree of overlap in the technology needed for both types of information extraction. It only makes sense to coordinate the efforts of both types of CAC in order to recover the most information from the investment.

Notes

1. AHIMA. "Delving into Computer-assisted Coding." 2004. Available online in the FORE Library: HIM Body of Knowledge at www.ahima.org.

2. Ibid.

James R. Flanagan *(james@landc.be) is medical director of informatics at Language and Computing.* ***Karen Doyle*** *(karen@landcglobal.com) is healthcare informatics specialist at Language and Computing.*

Article citation:
Flanagan, James R.; Doyle, Karen. "Computer-Assisted Coding: What's Here, What's Ahead." *Journal of AHIMA* 77, no.7 (July/August 2006): 36-40.