Fighting Fraud Automatically: How Coding Automation Can Prevent Health Care Fraud

by Jennifer Hornung Garvin, PhD, RHIA, CPHQ, CCS, CTR, FAHIMA; Valerie Watzlaf, PhD, RHIA, FAHIMA; and Sohrab Moeini, BSIS

Automation helps streamline the coding process. When combined with sophisticated analytic software, it can reduce fraud, too.

Automating coding software is generating interest for many reasons--an important one is its ability to speed the turnaround time between patient encounter and reimbursement. But automated coding has interesting potential in another important, though less-discussed, aspect of healthcare: fraud management.

In combination with anti-fraud software, automated coding software has the potential to reduce fraudulent activity by preventing code reporting errors, increasing the accuracy of coded data, and detecting false claims. The combination is all the more relevant as the industry moves to electronic health records (EHRs).

An Overview of Coding Software

Current coding software products span a range of automation. Supplementary coding tools increase the consistency of the diagnostic and procedural codes assigned. These products consist of tools from the very basic to the more complex.

Basic coding tools include software that prompts correct code assignment based on official guidelines and reporting rules. This software may be coupled with other functionality such as bar codes, pick or lookup lists, and automated physician fee tickets or "super bills." It can also be used in conjunction with preprogrammed codes for tests, drugs, and supplies incorporated into physician practice management systems or institutional financial systems.

More robust tools include complex prompting based on reference rules, color-coded references and edits, and software that allows coding from a remote, secure location. Additional prompts can be added onto existing software to help the coder fully code each case and view the reimbursement results. This is usually referred to as coding optimization software. Codes may be inconsistently assigned because of variability of coding education and training, the degree of accuracy of the coding tools used, and error introduced into the workflow because of incomplete documentation and interrupted workflows. Through the use of coding software, the healthcare industry seeks to increase the accuracy of code assignments and minimize the potential for error in the associated processes that affect coding.

Automated Coding

Coding software that offers the greatest amount of automation evaluates electronic clinical documentation and automatically generates a set of medical codes. These initial code assignments are then reviewed and validated by the clinician or a coding professional. The editing process can be simple or more complicated, depending upon the accuracy of the coding software and the complexity of the encounter.

This process is usually referred to as automated coding or computer-assisted coding. It must be noted, however, that no product on the market today is able to automate code assignments completely, because of the multiple variables in coding that complicate machine processing. A coding professional or clinician makes the final determination of the codes reported or stored.

Automated coding is commonly used in settings with limited variability of documentation, such as when performing endoscopies, in the emergency, outpatient surgery, and radiology departments of a hospital, and in specialty physician offices. A minimal number of coding software programs code inpatient documents, and these programs are not yet widely used. In primary care settings, the creation of text can be mapped to associated codes for physician validation. Software companies are rapidly responding to the marketplace and are planning to expand into new areas as the technology becomes more familiar with the more complex clinical and surgical scenarios.

When combined with the EHR or electronic documents, automated coding can streamline the way that healthcare organizations gather data and submit claims for services. It can help organize work and make documents easier to find. It can also provide a way to analyze health data and coding patterns to perform continuous auditing prior to billing and claims submission.

Despite advances in automation, coded data are typically reviewed by qualified coding staff prior to use in the billing process. Vendors and users note the continuing need to train coding professionals to evaluate and validate coded data. For the most part, users report that automated coding can enhance workflows so that coding staff can be better utilized.

How Automated Coding Selects Codes

Automated coding software is based on two methods: natural language processing (NLP) and structured, protocol, or template-based text. NLP is a software technology applied to a text-based electronic document. It uses computational linguistics and artificial intelligence to extract pertinent data and terms from the text and convert them into a set of medical codes. NLP can read any type of electronic text or EHR.

Structured input or text (a.k.a. codified input) differs from NLP. Structured data are discrete data using controlled vocabulary rather than narrative text. Structured input is a form of data entry that captures data in a structured manner, such as point-and-click fields, pull-down menus, structured templates, and macros. (For more on structured and unstructured data input, see Fenton, S. "Structured or Unstructured?" *Journal of AHIMA* 77, no.3 [2006]: 52.)

Automated coding products then use one of two methods to assign the code: statistics-based and rules-based. Some products use a combination of both.

In a statistics-based approach, the software predicts what code might apply for a given word or phrase based on what statistics have shown in the past. The statistics are generated from a pool of previously coded aggregate data. This approach is also known as a data-driven approach.

A rules-based approach applies coding rules such as those in logic or rules-based encoders, groupers, and imaged coding applications to electronic clinical documents. The initial codes are then processed through associated rules that have been established to further refine the code assignment. The coded data are then forwarded to a professional coder who verifies the codes prior to final assignment for use in billing. The rules-based approach is also known as a knowledge-driven approach.

Most NLP-based automated coding software use a combination of statistics-based and rules-based approaches. In most cases, the statistics-based approach is applied first. If errors are detected, the rules-based approach is applied. An extensive quality check is usually incorporated. As always, it is necessary to have experienced human coders check or edit the final codes.

Anti-Fraud Software

Automated coding products can incorporate and analyze patient data generated from a variety of sources. They can also evaluate record-specific information. Both of these aspects can help prevent fraud in reimbursement claims. It should be noted that some basic text to code mapping products may not provide anti-fraud features and may contribute to improper coding if not properly designed.

The sophistication of anti-fraud tools and software varies across products. Products can include basic tools, such as post-payment audits, or more complex data-mining techniques and machine learning. An example of the latter is artificial neural networks (ANNs).

ANNs can predict the potential for fraud in a specific claim based on the data in the claim and in the EHR. ANNs do not need constant updating; they learn continuously by analyzing certain pieces of information. Much like the text analytics in NLP, the medical data in ANNs are analyzed for any given claim and provide a statistical estimate that the data will either match or not match the desired output. Training the system to detect fraud is improved by using examples of fraudulent cases. Once this is completed, the system uses its prior knowledge to determine if a medical claim or data is falsified. These systems can be used for both prepayment and post-payment fraud detection.

Three mechanisms help an ANN system deal with fraud detection: data profiling, advanced analytic models, and rank scoring. Data profiling works by taking all relevant historical information and condensing it into a file that the program understands. The incoming claim is compared with the historical information, and the data are compared to advanced analytic models to determine whether the new claim matches the past information or differs in some way. The model is then updated with new information from the current processed claim. The more data profiling is used, the better and more learned the system becomes. Pattern recognition is also used in ANN systems. The data are compared to multiple sources of information to eventually seek patterns that may suggest possible fraud or abuse.

Rank scoring is used to identify claims with a high risk of fraud. A high number on the claim signifies a statistical high fraud risk. The scoring is checked by managers or staff, who can use the data to examine patterns across providers, settings, diagnoses, and procedures.

Since anti-fraud software uses a combination of the three systems described above, it continues to learn about the characteristics and patterns of legitimate and illegitimate claim behavior, becoming more intelligent and increasingly accurate in its detections over time.

A Combination to Deter Fraud

Just as there is a range of automated coding products, there is also a range of EHR products, from basic to sophisticated. In the primary care setting, there are software programs that suggest potential codes as patient records are generated. The practitioner often must select or validate the appropriate codes. In this model, there may also be edits or prompts to help the practitioner select the correct code. The code is not automatically assigned without validation, and there are limited, if any, anti-fraud functions.

The capabilities to combat healthcare fraud are possible when several types of technology are used together. The greatest potential comes from combining automated coding with NLP (employing both rules-based and statistics-based approaches) with ANNs and predictive modeling to detect fraud within an EHR.

Audit trails are also vital in order to assess the patterns of use within the EHR as well as the patterns of coding and billing. This type of technology is available and very promising. However, it needs extensive testing, because the combination of these technologies is very new. Together, software developers, users, payers, consumers, and government agencies can work to deter fraud through the recommendations appearing in the sidebar below.

As the HIM profession moves toward the routine use of electronic records and automated coding, it will need to maintain many traditional skills and continue many current practices, such as the use of current coding tools, working within an established compliance program, and evaluating products for accuracy. HIM professionals will also need new skills, certified products, the ability to work with aggregate data analysis, and familiarity with the use of ANNs. The role of skilled coding professionals will not be diminished but enhanced.

Background on the Report

This article is excerpted from "Automated Coding Software: Development and Use to Enhance Anti-Fraud Activities," a report prepared by AHIMA's Foundation of Research and Education under contract with the Department of Health and Human Services' Office of the National Coordinator for Health Information Technology.

The report offers full detail on a descriptive research project that examines the types of automated coding software available across healthcare settings and patient types. The research identifies the characteristics that have the potential to detect improper coding and the components of the coding process that have the potential to minimize improper or potentially fraudulent coding practices when using automated coding. It concludes with recommendations for software developers and users of coding products on how to maximize anti-fraud practices. The full report may be read online.

Recommendations for Deterring Fraud

The design and optimal use of automated coding and anti-fraud software can help manage fraud. Payers, consumers, and government agencies also have roles in anti-fraud education and initiatives. Recommendations include:

- Automated coding products should use a combination of statistics-based and rules-based systems.
- A standardized database (rather than a facility-specific database) should be used to train the statistics-based rules engine that will determine codes.
- Audit trails should be used to preserve the changes made to a given document. This will also ensure that the codes assigned are based on specific documentation by clinicians.
- Machine learning such as artificial neural networks should be available for predictive modeling to reveal trends and scores to detect fraud and abuse and should be used before bills are sent to the payer whenever possible.
- An appropriate compliance program should monitor for potential patterns of abuse, employing appropriately trained coding professionals, using current coding references and appropriate coding practice standards.
- Qualified coding staff should be hired and retained.
- Up-to-date references must be used by both providers and payers.
- Payers should use standard coding practices and conventions based on CPT Assistant and Coding Clinic as well as Medicare guidelines.
- Products should be certified verifying that products accurately map text to codes.
- Payers and providers should work together to increase the effectiveness of combating fraud.
- Education programs should help consumers detect fraud.
- Education should be directed toward generating the most accurate claims possible and detecting criminal activity.
- Staff who deal with the revenue cycle should evaluate new software products pertaining to billing and coding prior to implementation

Jennifer Hornung Garvin *(ccsreview@netcarrier.com) is medical informatics postdoctoral fellow at the Veterans Administration in Philadelphia, PA.* ***Valerie Watzlaf*** *is an associate professor in the department of health information management at the University of Pittsburgh.* ***Sohrab Moeini*** *is a graduate student in the Department of Health Information Management at the University of Pittsburgh.*

Article citation:
Garvin, Jennifer Hornung; Watzlaf, Valerie J.M.; Moeini, Sohrab. "Fighting Fraud Automatically: How Coding Automation Can Prevent Health Care Fraud" *Journal of AHIMA* 77, no.3 (March 2006): 32-36.