Automated Coding Software:

Development and Use to Enhance Anti-Fraud Activities

Contract Number:
HHSP23320054100EC

July 11, 2005

Submitted by:

Foundation of Research and Education
American Health Information Management Association
233 North Michigan Avenue, Suite 2150; Chicago, Illinois 60601-5800
(312) 233-1100 ■ www.ahima.org

ACKNOWLEDGMENTS

This report was prepared by The Foundation for Research and Education (FORE) of the American Health Information Association (AHIMA) and the University of Pittsburgh under contract to the Department of Health and Human Services, Office of the National Coordinator for Health Information Technology. Staff contributing to this report included: Jennifer Hornung Garvin, PhD, CCS, CPHQ, FAHIMA, RHIA, CTR; Valerie J.M. Watzlaf, PhD., RHIA, FAHIMA, and Sohrab Moeini, BSIS. We also acknowledge the vendors, law enforcement professionals, staff and members of Executive Committee of the Healthcare Anti-Fraud project, staff from AHIMA, and informatics professionals who were interviewed and provided guidance to our report. Finally, we acknowledge the cooperation and assistance of David Brailer, MD, PhD, Coordinator for Health Information Technology, and Kathleen Fyffe of the Office of the Coordinator for Health Information Technology.

Contents

Executive Summary 1
- Objectives 1
- Methodology 1
- Overview of Current Coding Processes 2
  - Coding Staff and Tools 3
  - Evolution of Coding Products 3
  - Automated Coding and EHR 4
  - Anti-Fraud Software 4
  - Anti-Fraud Software and Automated Coding in EHR 5
- Conclusions and Major Recommendations 6
- Limitations of This Research, Future Research, and Next Steps 7

Introduction 8
- Objectives 8
- Methodology 8

Overview of Coding 10
- Manual Coding Process 11
- Coding Staff and Tools 11
- Evolution of Coding Products: Automated Coding 12
  - Example: Automated Coding with NLP Using a Statistics-based Approach 13
  - Rules-based Approach 13
  - Combination of Statistics-based and Rules-based Approach 13
  - Automated Coding with Structured Text 13
  - Example: Automated Coding with Structured Text 14
- Automated Coding and the EHR 16
- Anti-Fraud Software 16
- Use of Anti-Fraud Software with Automated Coding 17
- Effectiveness of Anti-Fraud Software and Automated Coding 17
- Anti-Fraud Software and Automated Coding in EHR 18
- Automated Coding Capabilities and Data Quality 19
- Trends in Healthcare and Automated Coding 20

Issues Related to Fraud and Abuse 22

Results 24
- Vendor Information Matrices 24

Coding Process Flowcharts....24
Automated Coding Impact Table....28
Weak Links....28
Anti-Fraud Model....29
Features....29
Processes....29
Staffing....29
Detailed Recommendations....31
Software Developers....31
Users....32
Consumers....32
Payers....32
Government....33
Conclusions....34
Limitations of This Research, Future Research, and Next Steps....35
Appendix A: Government Interview Form....37
Appendix B: Vendor Interview Form....38
Appendix C: User Interview Form....39
Appendix D: Product Information Form....40
Appendix E: Vendor Information....42
Appendix F: Product Costs....43
Appendix G: Coding Optimization Tools....46
Appendix H: Automated Coding Impact....48
Appendix I: Weak Links....50
Appendix J: Anti-Fraud Model....52
Glossary....53
Footnote References....55
Bibliography....56

Executive Summary

The deliberate submittal of false claims to private health insurance plans and/or tax-funded public health insurance programs, such as Medicare and Medicaid, is a serious and increasing nationwide crime occurrence.

In 2003 alone, The National Healthcare Anti-Fraud Association (NHCAA) estimates that at least 3 percent of the nation's healthcare expenditures, or $51 billion, was lost to outright fraud. Other estimates by government and law enforcement agencies place the loss as high as 10 percent of our annual expenditures, or $170 billion each year[1]. According to the Centers for Medicare and Medicaid (CMS)[2], fraud may take different forms, such as incorrect reporting of diagnoses or procedures to maximize payments, fraudulent diagnosis, and billing for services not rendered. In addition, patterns of inaccurate claims that may be interpreted as fraudulent can unknowingly be submitted.

The impact of fraud can be mitigated, however, with appropriate technology, fraud prevention and detection processes, and ongoing educational efforts. This study examines automated coding software as an evolving technology and describes it across healthcare settings and patient types as well as its ability to reduce fraudulent activities. This study also examines how automated coding software can help a healthcare organization to enhance anti-fraud activities, detect errors, increase the accuracy of coded data, and detect false claims. The results are presented in this report.

Objectives

The objectives of this descriptive research project are:

- To identify the characteristics of automated coding systems that have the potential to detect improper coding.
- To identify the components of the coding process that have the potential to minimize improper or fraudulent coding practices when using automated coding and to relate them to the role of the electronic health record (EHR).
- To develop recommendations for software developers and users of coding products to maximize anti-fraud practices.

Methodology

This study consisted of the following general tasks:

- A review of the literature on automated coding software, anti-fraud software within automated coding systems, and the extent of fraud and abuse related to automated coding.
- Interviews with federal agencies to gather information about instances of improper reimbursement or potential fraud involving automated coding software.
- Completion of a product information form by vendors describing their coding products, use across settings, and approximate cost.

- Interviews with vendors and users of automated coding and anti-fraud software.
- Development of product information matrices regarding use of anti-fraud software and automated coding software across healthcare settings, cost of these systems, and the use of coding optimization software and other coding tools across healthcare settings.
- Creation of flowcharts to demonstrate how automated coding and anti-fraud software tools are used.
- Creation of an automated coding impact table that summarizes the impact of automated coding tools on coding and billing accuracy.
- Development of a table describing weak links in fraud and abuse software, user education, and compliance practices.
- Creation of an anti-fraud model that summarizes features, processes, and staffing for the ideal anti-fraud system.

Overview of Current Coding Processes

The coding of medical conditions and procedures involves the translation of medical words into codes or numbers that accurately reflect the care patients receive. The types of code sets that can be used are: International Classification of Diseases, 9th Revision, Clinical Modification (ICD-9-CM), Current Procedural Terminology (CPT), and Healthcare Common Procedure Coding System (HCPCS). Code sets typically include the codes and code descriptions and the rules, conventions, and guidelines for proper use of the codes within them. It should be noted that payers do not always abide by such standards for proper application of the medical code sets. One can see the complexity of the coding process by referring to the article "Internet Resources for Accurate Coding and Reimbursement Practices" AHIMA Practice Brief, 2004[3]. In this article alone, over 100 Internet resources are provided to assist with the coding process.

As discussed in the AHIMA 2002 Position Statement on Consistency of Healthcare Diagnostic and Procedure Coding[4], coded clinical data is used by healthcare providers, payers, researchers, government agencies, and others for:

- Measuring the quality, safety, and efficacy of care;
- Managing care and disease processes;
- Tracking public health and risks;
- Providing data to consumers regarding costs and outcomes of treatment options;
- Designing payment systems and processing claims for reimbursement;
- Conducting research, epidemiological studies, and clinical trials;
- Designing healthcare delivery systems and monitoring resource utilization;
- Identifying fraudulent practices; and
- Setting health policy.

New uses of healthcare data are constantly evolving, further demanding that careful attention be paid to accurate and consistent application and reporting of coded data. Code sets must be sufficiently flexible to meet these changing needs, while maintaining stability and continuity over time to ensure data comparability.

Coding Staff and Tools

Coding can be performed manually or with assistance from software. In either case, a qualified individual must assign or verify the final code assignment based on coding conventions and reporting guidelines. The qualifications of coding personnel vary based on region of the country, healthcare setting, and job position. Those responsible for coding clinical data must be educated and trained to apply coding standards correctly and uniformly.

The use of supplementary coding tools increases the consistency of the diagnostic and procedural codes assigned. These products consist of a full range 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 can 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. Simultaneous goals of the healthcare industry are to increase the accuracy of code assignment and to minimize the potential for error in the associated processes that impact coding.

Evolution of Coding Products

Software exists that can generate codes from electronic text. As in manual coding, errors can occur and fraud can be perpetrated. The development of automated coding has occurred for many reasons, but most important is the need to decrease the time between when the patient was treated and the time the healthcare encounter is reimbursed. Automated coding can speed this turnaround time. Products vary and can have different names, but two common names are computer-assisted coding (CAC) and automated coding. In this report both terms are used interchangeably with the distinction that no product on the market today is able to automate code assignments completely due to multiple variables that complicate machine processing. Computer Assisted Coding is best defined as the use of software that automatically generates a set of medical codes for review, validation and use based upon clinical documentation provided by health care practitioners.

Automated coding uses either natural language processing (NLP) of electronic text or structured, protocol, or template-based text. The software then links codes to segments of text. Some coding products incorporate a mix of automated coding with NLP or automated coding with structured text.

Automated coding products use algorithms to generate the codes. A statistics-based approach uses previously coded aggregate data from a pool of data to determine the proper code. A rules-based approach uses the coding rule similar to the algorithms found in many encoders traditionally used to determine the proper code. Many products use a combination of both.

Automated coding assignment differs from the manual process of coding in a significant way — it evaluates electronic text and determines the initial codes, rather than having a human user (coder or practitioner) assign codes from the start. However, in both methods the final determination of codes reported or stored by a coding professional or the clinician. The human validates the results and makes appropriate edits. The editing process can be a simple one, or more complicated depending upon the accuracy of the coding software and the context required to reflect the details of the encounter

Automated Coding and EHR

When combined with the electronic health record (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.

Automated coding is commonly used in settings where there is 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. There is a minimal number of coding software programs that 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 NLP coding engines become more familiar with the more complex clinical and surgical scenarios. In the short time since the development of the AHIMA Practice Brief “Delving into Computer-assisted Coding” (AHIMA, 2004), there have been additional developments. These include the evolution of the use of statistics-based NLP or rules-based NLP used alone in an automated coding product, to the use of a combination of both methods in many products.

Further, the majority of automated coding software companies interviewed reported that coded data is reviewed by qualified coding staff prior to use in the billing process. Of note, all parties interviewed discussed the continuing need for training coding professionals to evaluate and validate coded data. For the most part, users reported that the processes utilized by automated coding can enhance workflows so that coding staff can be better utilized.

Anti-Fraud Software

Automated coding products can incorporate patient data generated from a variety of sources and analyze it. It 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 it if not properly designed. The sophistication of the anti-fraud tools and software varies across products and 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, but rather continuously learn by analyzing certain pieces of information. Much like the text analytics in NLP, the medical data in ANNs is analyzed for any given claim and provides a statistical estimate that the data will either match or not match 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 whether a medical claim or data is falsified. These systems can be used for both pre-payment and post-payment fraud detection[8-9].

Three mechanisms that help the ANN system deal with fraud detection include:

- Data profiling
- Advanced analytic models
- Rank scoring

Data profiling works by taking all relevant historical information and condensing it into a file that the program can understand. The incoming claim is compared with the historical information and data analysis is performed to determine whether the new claim matches the past information or whether it is different in some way. The file is then updated with new information from the current processed claim. The more the data profiling is used the better and more learned the system becomes.

Advanced analytical models that perform pattern recognition are also used in ANN systems. The data is compared to multiple sources of information to eventually try to find patterns that may suggest possible fraud or abuse.

Rank scoring is used to identify which claims have a high fraud risk. A high number on the claim would signify a statistical high fraud risk and a low number would signify a statistical low fraud risk. The rank scoring is checked by managers or staff and they can use the data to examine patterns across providers, settings, diagnoses, and procedures.

Since the 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[8-11].

Anti-Fraud Software and Automated Coding in EHR

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 code(s). 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 algorithms. The capabilities to combat healthcare fraud are possible when several types of technology are used together. Automated coding with NLP (rules-based and statistics-based combination) combined with ANNs and predictive modeling to detect fraud within an EHR is ideal. However, audit trails are also vital in order to continue 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 but it needs extensive testing since the combination of these technologies is very new and is not readily found throughout the settings examined in this research study.

Conclusions and Major Recommendations

This research supports recommendations for software developers, users, payers, consumers, and government agencies. The following is a summary of major recommendations; detailed recommendations by stakeholder are provided in "Detailed Recommendations" on page 31.

- Computer-assisted coding software should utilize a combination of statistics-based and rules-based automated coding and a standardized national database (as opposed to a facility-specific database) to train the statistics-based engine. Audit trails are essential in all coding and billing software and EHR application to ensure that codes are based on documentation by clinicians. Machine learning such as ANNs should be available for predictive modeling to reveal trends and scores to detect fraud and abuse before it happens.
- Users of automated coding should have an appropriate compliance program that includes: continuous data analysis to detect potential patterns of abuse prior to claims submission and payment, appropriately trained coding professionals, use of current coding references and appropriate coding practice standards.
- Product certification for computer-assisted coding products should be instituted. Certification should be based on criteria assessing the accuracy with which health record documentation is converted to codes based on standard coding principles and guidelines.
- Payers and providers must work more closely to prevent fraud. Adherence to standard coding conventions and rules is essential, as is aggregate data analysis and continuous monitoring enabled by computerization.
- When making any software purchase, providers and payers should evaluate the potential impact on the accuracy of coding, billing and claims processing so there are no unintended consequences. Coding experts should participate in the selection and implementation processes.
- Consumer education can help in detecting fraud. Information regarding claims accuracy could be included in a quality measures reporting and consumers might be alerted to potential billing problems. This effort could be assisted by widespread use of patient-friendly billing formats.

- There should be greater cross industry collaboration to prevent fraud. This would involve multi-stakeholder collaboration including payers, billing organizations, and providers with the aim of fewer inaccurate claims, and reduced cost associated with the currently complex and often antagonistic processes. Joint education is needed on methods of prevention of fraud.

Limitations of This Research, Future Research, and Next Steps

This research was based on data gathered from selected vendors of automated coding products and a limited number of users. It consisted of Web-based product demonstrations and telephone interviews with vendors, users, and government personnel. Many of the technologies described are newly applied to clinical code assignment, and are not yet in widespread use for this purpose.

Generally, more thorough evaluation is needed regarding how these tools perform in a variety of settings with different types of health records. Particular attention should be directed to the coding features of primary care EHRs that prompt for evaluation and management (E & M) code assignment. It is necessary to develop some agreed upon measures so that these technologies can be evaluated over time.

Short-term research and action plans suggested by this descriptive study are:

- Institute programs to improve national adherence to standard coding guidelines and rules by all stakeholders. This will require education about the consequences of local policy and practice and incentives to drive compliance. Standardization is a necessary prerequisite to improving data quality. It will also make it less costly to develop automated coding solutions, will permit more reliable trending for fraud detection and facilitate adoption of updated code sets.
- Evaluate the use of computer assisted coding technologies in production EHR settings. Compare and contrast the benefits in terms of data integrity, productivity and compliance monitoring for EHRs that feature structured versus unstructured text and those that are based on a reference terminology.
- Create use cases and test databases on which to evaluate the capability of computer software to generate codes according to standard coding guidelines, conventions and rules. Many of these tools are new and not widely used in production settings. Early laboratory-based research could provide useful insight while broad-based field research is not feasible. This will permit assessing how best to certify these technologies in the future.
- Evaluate the potential of automated code generation and anti-fraud software used in conjunction with the EHR to relieve coding workforce shortages. Research is needed to better understand this potential and what skills and competencies will be needed by coding experts in the future.

Introduction

For purposes of this study, healthcare fraud is defined as the deliberate submittal of false claims to private health insurance plans and tax-funded public health insurance programs such as Medicare and Medicaid. Since the early 1990's, healthcare fraud has become a serious and escalating nationwide crime. In fact, in 2003 alone, The National Healthcare Anti-Fraud Association (NHCAA) estimates that at least 3 percent of the nation's healthcare expenditures, or $51 billion, was lost to outright fraud. Other estimates by government and law enforcement agencies place the loss as high as 10 percent of our annual expenditures, or $170 billion each year[1].

According to the Centers for Medicare and Medicaid (CMS)[2], fraud may take on many forms, the most frequent of which is incorrect reporting of diagnoses or procedures to maximize payments. A breakdown of common fraudulent activities by CMS demonstrates that 77 percent include fraudulent diagnosis (43%) and billing for services not rendered (34%).

Inaccurate claims that are submitted unintentionally can also be problematic because they can be interpreted as potentially fraudulent and often result in incorrect payment for services. In addition to inadvertently submitted inaccurate claims, intentional misrepresentation of services rendered can also occur resulting in false claims. Both patterns of inaccuracy that occur unintentionally and patterns of falsified claims that consistently benefit the provider are considered fraudulent and can result in law enforcement action and fines.

Objectives

This study examines the types of automated coding software available across healthcare settings and patient types and describes its ability to reduce fraudulent activities by preventing code reporting errors, increasing the accuracy of coded data, and detecting false claims.

The objectives of this descriptive research project are:

- To identify the characteristics of automated coding systems that have the potential to detect improper coding.
- To identify the components of the coding process that have the potential to minimize improper or fraudulent coding practices when using automated coding and to relate them to the role of the electronic health record (EHR).
- To develop recommendations for software developers and users of coding products to maximize anti-fraud practices.

Methodology

This descriptive research study was organized into four parts.

In the first part of the study, an extensive literature review was performed with assistance from the AHIMA professional practice staff, graduate students, and the University of Pittsburgh's reference librarians. Appropriate federal agencies were interviewed regarding

experience with improper reimbursement or potential fraud involving automated coding software. An interview form was developed by members of the research team (Appendix A: Government Interview Form).

The second part of the research study included an evaluation of automated coding software, coding optimization software, anti-fraud software, and coding application tools such as bar codes, pick or lookup lists, and so forth to determine its use as well as the cost of these systems. The number of vendors that participated in the development of coding optimization, coding automation and anti-fraud software was determined by sending each vendor a product information form to complete (Appendix D: Product Information Form). Then, three product matrices that demonstrate the extent of use and cost of these systems were developed (Appendixes E, F, and G).

The third part of the research study included describing the available automated coding software tools, how these tools are being used in the coding and billing process, the impact of these tools on coding and billing accuracy, and the characteristics and limitations of anti-fraud features now available in automated coding software. Special attention was paid to "weak links" in automated coding and fraud and abuse software, user education, and compliance practices. An extensive search of coding vendors and users was conducted via both the Internet and telephone interviews. Approximately 40 vendors were contacted and given the opportunity to complete the product information form. Once the form was received, it was reviewed by the research team who then determined whether an interview was needed. Interviews (Appendix B: Vendor Interview Form) were conducted with all vendors who had the specific coding-related software necessary for this research.

The fourth part of the research study included an extensive search of users of automated coding systems. Users were interviewed by phone to augment the information that was found via the Internet and literature searches and to determine the effectiveness of the automated systems currently in use. (Appendix C: User Interview Form)

Qualitative analysis of the interview data was performed and results were categorized into common themes related to strategies for reducing the fraud and abuse risk. Then, guidelines for developers and "best practices" for users of automated coding products were developed. Also, detailed recommendations regarding the development of automated coding tools and their use were compiled.

Before any part of the study was conducted, it was submitted to the University of Pittsburgh's Institutional Review Board (IRB) for review and approval. It received approval at the exempt level.

Overview of Coding

The correct coding of diagnoses and procedures is important because it provides a source of data that is used for many purposes that influence the healthcare system, including research and policy decisions. In terms of the healthcare revenue cycle, coding professionals fuel healthcare revenue and reimbursement cycles by providing quality data.

The coding of medical conditions and procedures involves the translation of the medical words describing conditions and services into codes or numbers that accurately reflect the care patients receive. The types of administrative code sets that can be used are: International Classification of Diseases, 9th Revision, Clinical Modification (ICD-9-CM), Current Procedural Terminology (CPT), and Healthcare Common Procedure Coding System (HCPCS). Code sets typically include the codes and code descriptions and the rules, conventions, and guidelines for proper use of the codes within them. It should be noted that payers do not always abide by such standards for proper application of the medical code sets and may influence the choice of codes to trigger health plan coverage. One can see the complexity of the coding process by referring to the article "Internet Resources for Accurate Coding and Reimbursement Practices" AHIMA Practice Brief, 2004[3]. In this article alone, over 100 internet resources are provided to assist with the coding process. The variability and interpretive nature of code reporting requirements inhibits automation of the process at this time due in part to the conventions of an outdated code set for diagnosis reporting and procedure designation for inpatients (ICD-9-CM) and the rules and guidelines required for its role in prospective payment and health plan coverage.

As discussed in the AHIMA 2002 Position Statement on Consistency of Healthcare Diagnostic and Procedure Coding[4], coded clinical data is used by healthcare providers, payers, researchers, government agencies, and others for:

- Measuring the quality, safety, and efficacy of care;
- Managing care and disease processes;
- Tracking public health and risks;
- Providing data to consumers regarding costs and outcomes of treatment options;
- Designing payment systems and processing of claims for reimbursement;
- Conducting research, epidemiological studies, and clinical trials;
- Designing healthcare delivery systems and monitoring resource utilization;
- Identifying fraudulent practices; and
- Setting health policy.

New uses of healthcare data are constantly evolving, further demanding that careful attention be paid to accurate and consistent application and reporting of coded data. Code sets must be sufficiently flexible to meet these changing needs, while maintaining stability and continuity over time to ensure data comparability.

Manual Coding Process

The manual process of coding includes reading the medical document and evaluating the information available for diagnoses, medical procedures, and elements of recorded facts that result in the translation of the written words into numbers. The numbers represent diagnoses and procedures that are submitted, through the claims process, to third party payers, including government contractors who subsequently reimburse providers for services based on the codes on the claim forms submitted for payment

There are several points during the reimbursement process where verification of the documentation or codes assigned may be compared and validated so that potential errors in the reporting process are prevented. If these evaluation processes are not completed, errors can occur and inaccurate claims can be submitted. Although it is important to avoid intentional false claim submission, it is equally important to reduce erroneous or inaccurate claims. The mechanisms used to decrease inaccuracies interface with the coding process at several intervals by:

- Evaluating the documentation resulting from a patient-provider encounter and requesting clarification if the documentation is ambiguous or contains conflicting data elements.
- Assigning codes that represent the written word accurately and consulting the practitioner when questions arise.
- Utilizing coding tools such as up-to-date code books, coding references, and encoder software that assists the coder in determining the correct code assignment through text prompts.
- Having appropriately trained coding personnel assign codes or oversee code assignment.
- Undertaking periodic compliance audits to detect errors or potential patterns retrospectively.
- Working with billing personnel to evaluate rejected claims to determine why the rejection occurred and to remediate processes that result in inaccurate claims.

Coding Staff and Tools

Coding can be performed manually or with assistance from software. In either case, a qualified individual must assign or verify the final code assignment to assure required context is applied and reporting requirements are met. The qualifications of coding personnel vary based on region of the country, healthcare setting, and job position. Those responsible for coding clinical data must be educated and trained to apply coding standards correctly and uniformly.

The use of supplementary coding tools increases the consistency of the diagnostic and procedural codes assigned. These products consist of a full range of tools from the very basic to the more complex. Basic coding tools include software that prompts correct code assignment based on code set conventions and rules. This software may be coupled with other functionality such as bar codes, pick or lookup lists, and automated physician bills. It can also be used in conjunction with preprogrammed codes for tests, drugs, and supplies. More robust tools can 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. 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. Simultaneous goals of the healthcare industry are to increase the accuracy of code assignment and to minimize the potential for error in the associated processes that impact code set reporting for payment or quality of care indicators..

Evolution of Coding Products: Automated Coding

In the past decade we have seen an outpouring of advancements in technology to automate and streamline various healthcare related processes. Automated coding is the use of computer software to automatically generate a set of medical codes for review or validation and use based upon clinical documentation provided by healthcare practitioners[5].

Automated coding has emerged for many different reasons, including inconsistencies in the quality of coding, cost, and high turnover rates of coders. Automated coding products have various names, and functionality can vary based upon the type of system used. Automated coding can also be referred to as computer-assisted coding (CAC). Automated coding is normally used with either natural language processing (NLP), which can read any type of electronic text or electronic health record (EHR), or structured, protocol, or template based text. Some automated coding products incorporate a mix of automated coding with NLP and automated coding with structured text.

Natural language processing (NLP) is software technology that is applied to a text-based document that uses computational linguistics and artificial intelligence to extract pertinent data and terms and convert them into a set of medical codes. NLP-based applications may use either a statistics-based (aka data-driven) or rules-based (aka knowledge-driven) approach to assign the code. Often a hybrid, or combination of both is employed in the NLP system architecture. With 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 rules-based approach uses programmed rules, or algorithms

Automated coding products, whether they are automated coding with NLP or automated coding with structured text, use either a statistics-based or rules-based approach to assign the code and some use a combination of both. With a statistics-based approach, the software makes predictions of what a proper code should include for a given word or phrase based on what statistics have shown in the past. The rules-based approach includes incorporating coding rules such as those in logic or rules-based encoders, groupers, imaged coding applications, and so forth and applying them to electronic clinical documents so that the best clinical code is chosen[6].

Following the initial determination of the codes, well developed automated coding programs process the initial codes through associated rules that have been established to further refine the code assignment. Following these steps, the coded data is forwarded to an appropriate expert coder who verifies the codes prior to final assignment for use in the billing process.

Example: Automated Coding with NLP Using a Statistics-based Approach

Automated coding software is used to assign a clinical code to a specific claim. After analyzing any documentation provided by the physician or hospital staff, the analysis finds "fractured knee."

The software may continue to check the context of this phrase within the entire sentence and, once finished, conclude with a statistical analysis to see which code best matches this condition.

The system determines that Code 800.00 is the best statistical match for "fractured knee" based on earlier findings analysis with other clinical documents. In this statistics-based approach the software makes predictions of what a proper code should be for a given word or phrase based on what statistics show regarding the number of times this code has been applied to this term in the past.

This system works well when there is a large accumulated body of data, and it can be easily determined which code to use. When there is a small body of data, and only two previous code instances come up, for example, it becomes much more difficult for the software to choose a code[6].

Rules-based Approach

Another approach that is used with NLP is called the "rules-based" approach. Using this method, coding rules embedded in logic or rules-based encoders, groupers, imaged coding applications, and so forth are applied to electronic clinical documents and the best clinical code is chosen[6].

Combination of Statistics-based and Rules-based Approach

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

Automated Coding with Structured Text

Another way automated coding works is by using structured input or text, or codified input, which is different from NLP. Structured data is discrete data using controlled vocabulary rather than narrative text. Structured input is a form of data entry that captures data in a structured manner (for example, point-and-click fields, pull-down menus, structured templates, macros)[5].

Example: Automated Coding with Structured Text

As an individual menu item is chosen, a narrative text phrase is produced that becomes part of the health record documentation.

Each menu item that affects coding is directly mapped to its relevant code.

For example, the pre-op diagnosis menu item of "acute tear lateral anterior horn of the meniscus" is directly mapped to the applicable ICD-9-CM diagnosis code (836.1).

The physician chooses the applicable clinical menu item, and the ICD-9-CM code is automatically produced to be used or edited by the coding professional[6].

In some systems, certain custom menu items can be added. In general, structured input is not supported by a rules-based or statistics-based approach. It is a simpler approach in which the coding module creates an internal mapping between the clinical content and the codes. In some cases, quality edits are built into the system to help assure that an accurate code is provided. Again, it is still necessary to have an experienced human coder check or edit the final codes generated.

According to the E-HIM Work Group on Computer Assisted Coding[6], there are many advantages to automated coding. Based on our research, we agree with the following advantages:

- Increase in coding productivity
- Increase in coding consistency
- Availability of a coding audit trail
- Data query ability
- Potential for more comprehensive code assignment
- Potential increase in coding accuracy
- Potential decrease in coding costs
- Improved documentation (unique to structured input based automated coding tools)
- Decreased documentation costs (unique to structured input based automated coding tools)
- Creation of ancillary documentation (unique to structured input based automated coding tools)
- Use of free text for recording documentation (unique to NLP based automated coding tools)
- System improvements through feedback (unique to NLP based automated coding tools)

However, the authors of the workgroup also cite many disadvantages that this research study confirms:

- User-specific integration
- User acceptance and change management
- High cost (initial purchase and ongoing maintenance)
- Potential for coding errors or fraudulent claims
- Use of structured input (unique to structured input-based automated coding tools)
- Extensive software development efforts (unique to NLP-based automated coding tools)
- Potential NLP coding mistakes (unique to NLP-based automated coding tools)
- Reliance on electronic documents (unique to NLP-based automated coding tools)
- Complexity, quality, and format of health record documentation
- Technological limitations
- Lack of industry standards

Automated coding products are currently in use in a variety of healthcare settings. The NLP systems are most commonly found in radiology, cardiology, general medicine, and emergency medicine. The structured input-based automated coding tools are more commonly found in specialty settings such as gastroenterology, orthopedics, urology, and pulmonary medicine[7].

The mechanisms used to decrease inaccuracies interface with the coding process at several intervals. The automated coding process undertakes the following processes almost simultaneously:

- The coding engine evaluates the documentation resulting from a patient-provider encounter and identifies incomplete documentation. It does this be evaluating patterns of documentation that are statistically different from the average documentation for similar claims, so that physicians can be queried. It also assigns codes that represent the classification or description of the written word accurately based on recorded health record information. The physicians should be provided with education regarding documentation variances so that the clinical pertinence of the permanent health record and utility of the data recorded and reported for the encounter can be improved.
- Initial code assignment is sent to the appropriate personnel for verification.
- The reviewer validates the code assignment created through the prior steps using coding tools such as up-to-date code books, coding references, and encoder software applications that assist in determining the correct code assignment through text prompts. The expert coders or practitioners validating system code suggestions also apply knowledge gained through appropriate training or experience in their area of expertise.

The use of automated coding can result in both increased errors and decreased errors, depending upon the associated processes that are undertaken with the NLP engine. With appropriate statistical evaluation and expert coding validation, errors can be decreased and new aggregate analysis similar to an audit can take place concurrently. This can result in improved documentation and increased accuracy in claims.

Automated Coding and the EHR

When combined with the electronic health record (EHR) or electronic documents, automated coding can speed the coding process and reduce variability of coding assignment. This conceivably could reduce the time between provision of services and reimbursement. Further, the associated processes can streamline workflow, increase administrative and functional operations, and provide a mechanism to continuously analyze and audit documentation and coding patterns prior to billing rather than retrospectively.

Automated coding assigns ICD-9-CM, CPT, and HCPCS codes following the analysis of electronic documents. The assignment of codes commonly takes place in settings where there is limited variability of documentation such as in endoscopies, emergency department, radiology, specialty physician areas, and outpatient surgery. Based on our research, we found there is a limited number of coding software programs that address inpatient coding and these have not been widely deployed and their use is limited in clinical sites. In primary care settings, the creation of text can be mapped to associated codes. Codes or questions are presented to the physician and they must select the appropriate code. The code is not automatically assigned. Basic systems that map text to codes blur the lines of demarcation between automated coding and coding assistive tools. In these systems, software prompts the practitioner generating the text to determine various aspects of a diagnostic, procedural, or evaluation and management code assignment.

Software companies are responding to the need to increase reliability of code assignment and are planning to expand into new areas as the NLP coding engines become more familiar with the more complex clinical and surgical scenarios. This rapid evolution of automated coding products is based on the industry's response to the market demands.

In the short time since the development of AHIMA's Practice Brief "Delving into Computer-assisted Coding"[5-7], there have been additional developments in automated coding including the evolution of the use of statistics-based NLP or rules-based NLP to a combination of both methods. All of the NLP software companies reported that all coding should be reviewed prior to use in the billing process. The initial expectation that automated coding would eliminate manual coding has evolved into an increased use of highly trained coding staff for editing and verification of computer generated codes.

Vendors advise 100% review and this is AHIMA's best practice recommendation, however there are anecdotal reports that users may skip the review process in limited instances based on specific parameters where high level of confidence in computer generated codes has been demonstrated consistently over time and circumstances (eg, normal mammograms).

All parties interviewed for this study stressed the importance of coding professionals and discussed the continuing need for training them to evaluate and validate coded data.

Anti-Fraud Software

Some automated coding products incorporate anti-fraud software that provides aggregate data analysis and record-specific edits. Some basic text-to-code-mapping products may not provide any anti-fraud software. The sophistication of anti-fraud software varies across products and can include basic tools such as post-payment audits or more complex data mining techniques such as ANNs that predict the potential for fraud based on the data in the claim and in the EHR.

Because of the variability of the types of automated coding systems available, the use of these systems raises questions about the accuracy and efficiency of coding diseases and related procedures with minimal human intervention. All constituents in the healthcare reimbursement process can benefit from using coding edits and anti-fraud software. Using these software tools increases coding and billing accuracy, which results in timely and appropriate reimbursement.

Use of Anti-Fraud Software with Automated Coding

Without a proper fraud detection system, an automated coding tool will simply do what it has been designed to do. The document will be processed and possibly checked for inaccuracies at a later time. However, by adding fraud prevention software, processed documents will be checked digitally either before or after the automated coding process.

Other aspects of fraud detection include prepayment fraud detection, or finding the fraud before it is processed and put into the system. In this particular case, the claim is checked and then sent for any further auditing by the automated coding software.

Other text analytic engines not only sort and search claim documents before and after any coding, but also search through call center logs and case management notes. Yet other systems take an incoming claim and run it through different fraud analysis phases to fully investigate the claim. In some situations, the ICD-9-CM codes generated from automated coding are used with pattern recognition software to check for types of fraud[8-11].

Effectiveness of Anti-Fraud Software and Automated Coding

One of the new software tools available to detect fraud is machine learning. An example of this is artificial neural networks (ANNs). ANNs are adaptive in that they don't need a constant update to the software; rather the software learns by constantly analyzing certain pieces of information.

In the case of fraud detection, one of the main uses of this type of system is for pattern recognition. Much like the text analytics in NLP, the data or medical records in ANNs are analyzed for any given claim and provide a statistical estimate that the data will either match or not match desired output. This type of system can be used in all aspects of fraud detection, such as billing anomalies, upcoding, or any other types of falsifications[12].

These types of systems gain the experience to detect fraud by actually training the system with examples of fraudulent cases. Once this is done, the system begins to use the prior knowledge to determine whether a medical claim or data is falsified. These systems can be used for prepayment fraud detection as well as post-payment fraud detection[8-9].

Three mechanisms that help the system deal with prepayment fraud detection include:

- Data profiling
- Advanced analytic models
- Rank scoring

Data profiling works by taking all relevant historical information and condensing it to a file that the program can understand. The incoming claim is compared with the historical

information and data analysis is performed to determine whether the new claim matches the past information or whether it is different in some way. The file is then updated with new information from the current processed claim. The more the data profiling is used the better and more learned the system becomes.

Advanced analytical models that perform pattern recognition are also used in ANN systems. The data is compared to multiple sources of information to eventually try to find patterns in fraud and abuse.

Rank scoring is used to identify which claims have a high fraud risk. A high number on the claim would signify a statistical high fraud risk and a low number would signify a statistical low fraud risk. The rank scoring is checked by managers or staff and they can use the data to examine patterns across providers, settings, diagnoses, and procedures.

Since the 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[8-11].

Anti-Fraud Software and Automated Coding in EHR

The use of electronic health records are expected to facilitate greater automation of the coding process. At the same time, the advent of the EHR presents new problems and issues that must be addressed. No longer will health information be strictly document based, but rather stored in electronic format in a database designed for discrete data element retrieval. With each patient visit, clinical information is updated and processed by the automated coding tools. 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, some software programs (applications) map text as it is generated to potential codes and then prompt the practitioner to select the appropriate code. The code is not automatically assigned and there are limited if any anti-fraud algorithms.

The information can also be run through the anti-fraud software for auditing. Using ANNs, a patient's EHR can easily be analyzed, since all information is now digital and very accessible to incorporating anti-fraud software.

Pattern recognition and text analytics are just some of the ways ANNs can lead to fraud prevention in nationwide interoperable health information systems. Based on current rising fraud cases, it is imperative that the nationwide interoperable health information system have fraud detection capabilities. Advanced analytic modeling can be applied to the claim at hand as well as the entire information system for that specific patient. For example, the model combines the thousand or so profile numbers with the claim information in complex ways, looking at multiple relationships between many pieces of data simultaneously[8]. This powerful pattern recognition capability enables the model to see how all the pieces form a unified picture of the fraud and abuse risk of the claim[8]. A model, such as the one mentioned above, could easily navigate the vast amount of electronic information and check each record.

Without a proper fraud detection system, an automated coding tool will simply do what it has been designed to do. The document will be processed and possibly checked for

inaccuracies at a later time. However, by adding fraud prevention capabilities, processed documents will be checked digitally either before or after the automated coding process.

Other aspects of fraud detection include prepayment fraud detection, or finding the fraud before it is processed and put into the system. In this particular case, the claim is checked and then sent for any further auditing by the automated coding system.

There are other text analytic engines that will not only sort and search claim documents before and after any coding, but that will also search through call center logs and case management notes. Yet other systems will take an incoming claim and run it through different fraud analysis phases to fully investigate the given claim. In some situations, the ICD-9 codes generated from automated coding are used with pattern recognition software to check for types of fraud[8-11].

Automated Coding Capabilities and Data Quality

Auto-coding systems may reduce fraud because they contain coding edits, quality checks, and assistive tools to assure code sets are properly applied to the available clinical data... Theoretically, documentation should drive manual coding but there is always the potential for human error or a lack of quality documentation that results in coding errors. Automated coding systems are available that can generate administrative code sets that can be accessed using an interface with an abstracting or coding evaluation module.

Automated coding systems are not 100% accurate (or even 95% accurate which is the current standard for manual coding)[13]. All systems evaluated in this research require that a coder edit and evaluate the codes before final billing. In some cases, a separate billing system is used that may have additional quality edits, thus completing the revenue cycle. Data generated from multiple systems that may also reside in separate databases could conceivably contain slightly different data.

Two issues arise. First, the education of coders is critical to generation of quality coded data and, second, using multiple systems provides multiple opportunities to evaluate the data and detect potential fraud. There are positive aspects to the use of several different software systems but there are dangers as well. Software packages sometimes have specific edits that pertain to only one step in the coding billing process and, therefore, they are focused on what that specific system is designed to do. For example, an encoder or similar coding software contains edits relative to coding rules and government requirements pertaining to codes. A billing system will have a similar data quality checking mechanism but it will be relative to the elements of the UB92 and specific insurer billing requirements. In terms of fraud and abuse prevention, it is imperative that all systems that retain coding and billing data be consistent and up to date with regulatory requirements. This prevents fraud and abuse and increases data quality to assist in continuous monitoring of appropriate billing and in providing an audit trail.

Automated coding products are available for records associated with facility coding in the inpatient, emergency, ambulatory surgery, radiology, pathology, and other specialties. Many of the existing products are undergoing continuous refinement and applications are growing in capabilities that are likely to extend to more settings than currently applicable. The features of automated coding include the use of three main approaches to translate the health records data into codes: rules-based, statistics- based, and a combination of rules- and statistics-based approaches. Codes may be based on analysis of structured text or

natural language processing. The majority of systems use natural language processing and they can interface with multiple types of healthcare systems including dictation, transcription, billing, health information, encoding, and abstracting systems, in addition to a variety of ancillary hospital systems. There is also text- to- code- mapping software for evaluation and management (CPT) codes for physician offices reporting professional services.

Trends in Healthcare and Automated Coding

Several major trends in healthcare have an impact on the use of automated coding and potential fraud and abuse. As the industry evolves, many products are used in the coding and billing process including dictated reports, automated coding software, and coding assistive tools such as encoders, abstracting systems, and billing systems. Some automated coding systems provide more detailed reference information, including pictures and anatomical references, so that coding can be more accurate. Overall, no automated coding product has an accuracy rate that meets the existing industry standard of 95% that coding professionals are expected to meet.

Multiple interfaces with many types of computer systems are a reality in healthcare. These interfaces may present solutions if they work well or difficulties if they cannot be configured. Potential benefits include the capability for automated coding products to read existing reports, increase data quality, increase productivity, provide for audit trails, and improve the efficiency of services. The potential difficulties are that the systems may be redundant or may not communicate with each other. This can result in duplicate data entry, less efficiency of services, decreased productivity, and frustrated practitioners and staff.

Automated coding systems may complement existing compliance systems by providing more detail to combat fraud. Healthcare organizations are already required to have compliance evaluation mechanisms. The importance of monitoring coded data that is generated by automated coding systems, and the use of it within the context of the EHR, may provide greater focus for compliance evaluation. For example, several systems can analyze aggregate data. This is particularly important if there are patterns of billing that may be viewed as fraudulent.

The potential of the consumer's role may be an important untapped source of information. When patients receive the summary of care, they can help to ensure that the services provided were correctly represented in the billing process. Consumer education could be provided to help the public understand the coding and billing process and to explain how consumers can report potentially fraudulent claims.

Payers are often perceived to approach providers as if assuming they are trying to promulgate fraud, and providers appear to be unclear as to how to accurately submit claims because of varying and inconsistent application of standards by insurers or difficulty understanding code set guidelines. Resolving antagonistic interactions between payers and providers relative to fraud might result in better services, less fraud, and greater efficiency in the healthcare industry

The following list summarizes some important aspects of the healthcare industry that have implications for preventing fraud and abuse.

- Automated coding products are available for many settings of care including many physician specialties.

- Auto-generation of ICD and CPT codes is applicable to both professional fee abstracting for many physician specialties and for facility billing.
- Accuracy rates of initial coding with automated coding software are not 100% so other processes must be undertaken to improve accuracy, including editing by human coders to apply missing data or context affecting code assignments as a final check.
- Once systems are perfected to achieve a "gold standard" accuracy rate productivity gains are expected to increase dramatically as the confidence level for selected procedures may allow the full review process to be reduced to a spot check or audit rather than 100% review
- The use of SNOMED CT as a standard clinical reference terminology is being researched by several automated coding software companies so that there will be ease of interface with the EHR. The use of a standard terminology that underlies the EHR has the potential to further increase the accuracy and efficiency of automated coding because of the use of standardized terms in the electronic documents.
- Data quality edit capability is required.
- ROI and fraud incentives increase acceptance of automated coding.
- Interfaces to abstracting and billing systems need to be considered.
- Multiple software packages are used in the steps of the revenue cycle.
- Final data resides in multiple databases.
- Products deriving codes use rules-based and statistics-based approaches, and a combination.
- The user-friendliness of the software can affect acceptance, coding quality, and productivity.
- Automated coding systems can use both individual case edits and trending of data to evaluate coding quality.
- Education and training of coders has wide variability resulting in variability of coding practices and quality.
- Audit trail capability is an absolute necessity to authenticate documents and any subsequent changes, track the provision of advance beneficiary notices, determine medical necessity, and document bypassing of coding edits and changes to coding and billing assignment.
- Automated coding provides the potential for increased efficiency.
- Antagonism between the provider and payer sides has been identified as problematic and potentially costly.
- Healthcare organizations have existing compliance mechanisms.
- Technology is facilitating consumer involvement in healthcare and more can be done in this area.

Issues Related to Fraud and Abuse

The following list summarizes some important issues related to fraud and abuse.

- User friendliness of software applications can influence the quality of coding and productivity.
- Automated coding systems are capable of using both individual case edits and trending of data to evaluate coding quality.
- Audit trail capability is an absolute necessity to establish authentication of documents and any subsequent changes, to track the provision of advance beneficiary notices, to determine medical necessity, to document bypassing of coding edits, and to track coding, billing and subsequent changes.
- Antagonism between the provider and payer sides has been identified as problematic and potentially costly.
- Healthcare organizations have existing compliance mechanisms which can be applied to fraud detection and prevention.
- Technology has facilitated increasing consumer involvement in the healthcare industry.
- There is unintended incentive for fraud because healthcare organizations and software developers need to prove a return on investment for the coding products. This issue must be considered in fraud prevention activities.

Ease of use and whether or not automated coding systems require a change in clinical workflow are important considerations. The structured text approach, for example, requires that practitioners learn how to develop reports based on the use of pre-developed templates whereas the NLP approach requires less of a workflow change. Similar to the barriers to implementation of the EHR, the physician's propensity for using the system is critical for success. It appears that the less behavior change required by the practitioner, the greater the use of the systems. In fact, some NLP systems are transparent to the physician users. For example, if transcribed reports are analyzed, as are radiology and pathology reports, the coding software generates a code without any additional effort by the practitioner.

Coding software processes information contained in the document rather than requiring structured text to generate the segments that result in a particular code. The tradeoff appears to be consistency. When structured text is used, the mapping is straightforward. NLP is more flexible in generating codes, but it can be less consistent because the software vendor must develop new rules as language changes and the computer must "learn" those rules. For example, one problem with structured text is how to prevent a physician or facility from entering information about a procedure that was not actually performed (or reporting a higher level of evaluation and management service than was actually rendered) to obtain a higher reimbursement. With NLP, automated coding software only codes something that is present in the document text. However, effective structured text systems can prove advantageous in that they limit or more strictly guide coding choices

Some systems have individual edit and analysis capabilities as well as aggregate data evaluation. This allows the individual record to be coded with input from references based in health plan or government program rules, AHA Coding Clinic, and AMA CPT Assistant as well as other authoritative references. It also allows trending of data to detect patterns of reporting that are significantly different from the expected pattern based on industry trends. Some systems also have audit trails that record who generated and modified the document being coded as well as who edited the codes and who bypassed suggested anti-fraud/data quality edits.

Products are ROI-driven, for both the purchaser and the software vendor. This creates an incentive for increasing the revenue that the systems generate to justify the cost of the new system, and can create pressure to further tailor systems in ways that may not be consistent with practice standards.

Automated coding enables many functions (such as multiple tracking, reporting, and administration) to be automated that can help practitioners and their staffs to be more efficient. The majority of practitioners interviewed stated that automated coding enabled better analysis of administrative data and created a smooth workflow.

In terms of coding and billing operations, the automation of job tasks can help assign coding to those with the appropriate expertise to edit the codes. For example, the system can direct interventional radiology cases to a coder who is certified in that area. It can also create queues that list the cases assigned to a particular coder. Some of the products also provide tools that increase efficiency of the coding staff by revealing the number of documents that contain a particular diagnosis and by providing color coded anatomical and document references. The level of efficiency depends in large part on the capability of disparate healthcare software systems to be interfaced so that data can be interchanged.

Results

Vendor Information Matrices

Three different vendor matrices were developed from the product forms received from the vendors. The first matrix (Appendix E) describes the types of anti-fraud software and automated coding products that vendors provide across different healthcare settings. Eight of the 13 vendors have some type of automated coding product. Seven of the vendor systems have some type of anti-fraud capability, six have automated coding with NLP, six have automated coding with structured text, and five have some other type of automated coding product. Four of the vendor systems have both anti-fraud software and automated coding products.

The second vendor matrix (Appendix F) focuses on the cost of these systems. It can be seen that prices tend to vary across vendors and products but that most are based on a per record or per claim basis.

The third vendor matrix (Appendix G) describes the use of coding optimization products and also the use of other coding tools. It can be seen that 7 of the 13 vendors have coding optimization software, none have bar codes, 5 have pick or look up lists, 2 have automated super bills, 7 have logic or rules-based encoders, 3 have groupers, 3 have imaged coding applications, 7 have remote coding applications, and 4 have hard coding via charge master tables.

Coding Process Flowcharts

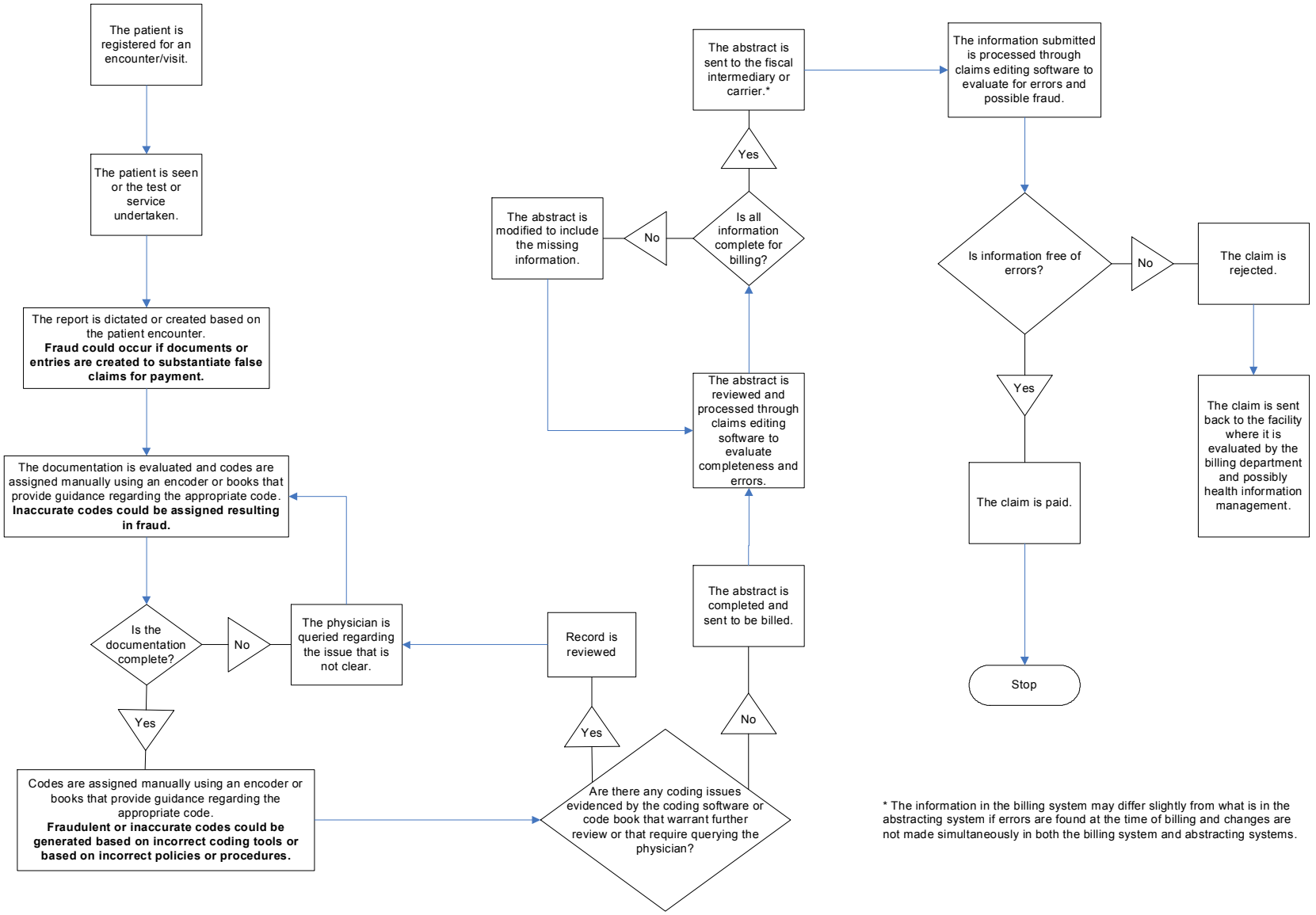
Three different flowcharts were developed to describe the progression of data flow when using automated coding.

Scenario 1 describes the basic flow of data without automated coding.

Scenario 2 describes the basic flow of data with auto-coding via NLP, combined with electronic documents or an EHR, and Quality Assessment Procedures.

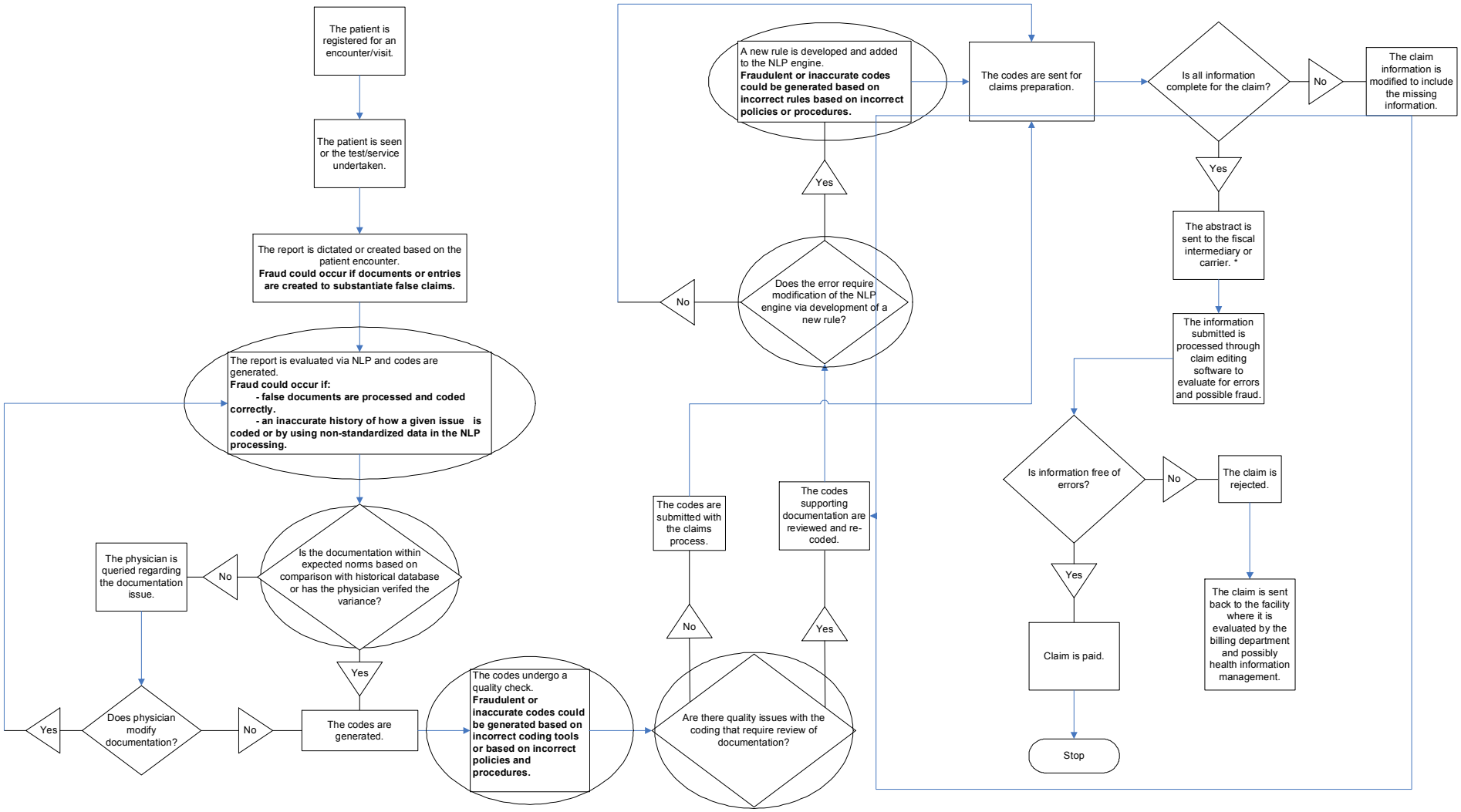
Scenario 3 describes the ideal flow of data using all appropriate systems: the basic flow of data with auto-coding via NLP, combined with electronic documents or an EHR, and Quality Assessment Procedures, quality Assessment procedures, continuous auditing and trend analysis prior to submission to fiscal intermediary or the insurance carrier.

Scenario 1: Basic Flow of Data without Automated Coding



* The information in the billing system may differ slightly from what is in the abstracting system if errors are found at the time of billing and changes are not made simultaneously in both the billing system and abstracting systems.

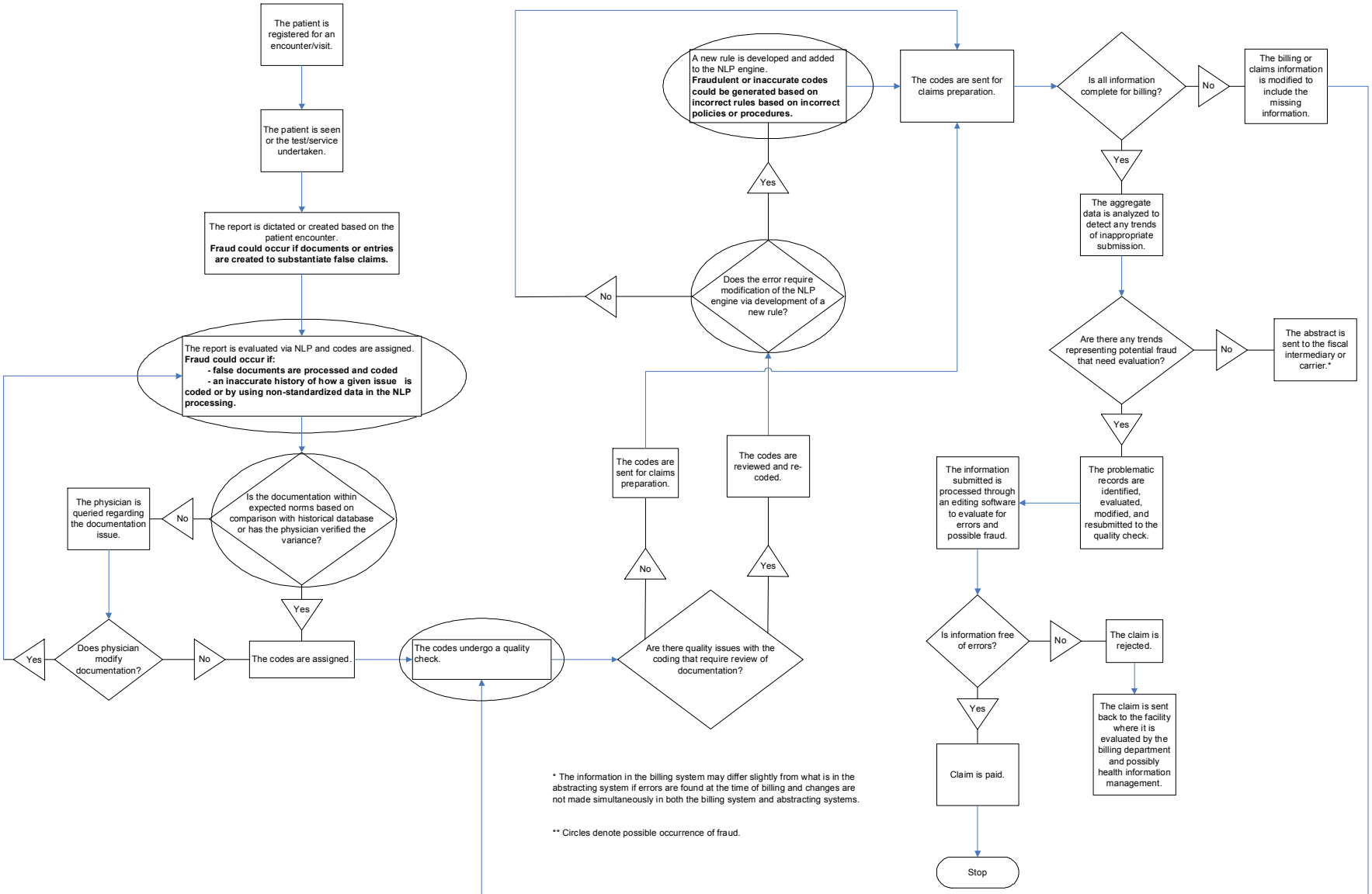
Scenario 2: Basic Flow of Data with Auto-coding via Natural Language Processing Combined with Electronic Documents or an EHR and Quality Assessment Procedures



* The information in the billing system may differ slightly from what is in the abstracting system if errors are found at the time of billing and changes are not made simultaneously in both the billing system and abstracting systems.

** Circles denote possible occurrence of fraud.

Scenario 3: Basic Flow of Data with Auto-coding via Natural Language Processing Quality Assessment Procedures Combined with Electronic Documents or an EHR , Continuous Auditing and Trend Analysis Prior to Submission to Fiscal Intermediary or Carrier



Automated Coding Impact Table

The Automated Coding Impact Table (Appendix H) summarizes the impact of automated coding tools on coding and billing accuracy. This table shows that automated coding with NLP or structured text with a combination of rules-based and statistics-based approaches has the potential for the highest level of accuracy of coded data.

A system of rules-based and statistics-based approaches, coupled with anti-fraud software, provides the strongest combination for fraud prevention and detection. Coding and billing edits can be used concurrently with claim development. Use of anti-fraud software that includes prepayment fraud detection at a higher level should also be used and should include data profiling, advanced analytic models, and rank scoring. A combination of these systems is ideal since it enables the system to continuously learn and become more intelligent over time and therefore, better able to detect fraud.

Weak Links

The Weak Links table (Appendix I) summarizes weak links in fraud and abuse software, user education, and compliance practices.

As the table shows, there will still be incidents of fraud with automated coding software and anti-fraud software is only worthwhile when the potential for fraud related issues is addressed. Also, as anti-fraud software becomes more technologically advanced, such as with the use of ANNs, those involved in fraud detection and prosecution will need to be continuously educated to stay current with the latest advances in technology.

Rules and standards may be inconsistently applied when audits are conducted because software edits are bypassed or the software is not used.

None of the automated coding software systems are 100% accurate and coding quality in many of the systems has not been assessed in actual practice.

Education is needed regarding EHRs, automated coding technology, and anti-fraud software and how all of the systems work together. It is extremely important to educate law enforcement professionals as well as compliance staff within healthcare facilities about how all of the automated systems will work together. Beneficiaries should continue to be educated about fraud potential. Those newly entering medical practice need to understand how the language they use can change codes, and reminder systems should be incorporated so physicians know what to include in their documentation. More consistent education of the coder is needed to include how clinical data interfaces with financial data and the revenue cycle.

Compliance practices will improve if there is less tension between the payer and provider communities. Aggregate data evaluation capabilities that a payer can access or query regarding a claim's accuracy are needed, as well as incentives for payers to decrease fraud. Even a well-thought-out compliance plan does not guarantee consistent implementation. The bigger picture may be to control whether a test is necessary, ordered appropriately, or completed as ordered rather than whether the software is effective in generating the correct code.

Anti-Fraud Model

The Anti-Fraud Model (Appendix J) summarizes features, processes, and staffing to facilitate efficient and effective anti-fraud applications according to persons interviewed for this report and our search of existing literature.

Features

The ideal model to combat healthcare fraud includes:

- Automated coding software with NLP with a combination of rules-based and statistics-based approaches.
- A standardized system of data (based on a representative sample of standardized claims information) for statistical reference.
- Private development of software for code assignment should be bolstered by the establishment of national criteria or certification procedures for software.
- Development of a standard template for evaluation and management coding
- Anti-fraud software that incorporates ANNs and predictive modeling to determine where the potential for fraud lies.
- Audit Trails

Processes

A prepayment fraud detection system should incorporate the following features:

- Data profiling
- Advanced analytic models
- Rank scoring

A standardized method to derive the statistical aspects of code assignment and data analysis should also be part of the process as well as a post-payment fraud detection system.

Staffing

The following list presents staffing considerations for implementing anti-fraud applications:

- Advanced coder analysts with proper education and training are needed to edit and verify all processes for accurate code assignment.
- Information technology (IT) staff is needed to maintain current technology.
- Health information system (HIS) staff is needed to train office staff, physicians, and other clinical professionals.
- The payer, provider, vendor, and consumer need to work together to combat fraud through education and specific incentives. Consumers and providers need to be educated regarding what constitutes fraud, how it impacts them, and the role they can play in helping to detect and prevent it.

Other features of a model to combat healthcare fraud recommended in interviews include:

- Random audits to verify whether or not services were provided and to verify that the coding used corresponded to the level of care provided
- Human review and participation in quality control and fraud detection efforts by trained coders and fraud analysts/investigators.
- System access controls to protect against identity theft and unauthorized access, such as identity authentication, verification, and tracking measures.
- Use of ad hoc interdisciplinary groups or teams to examine indicators of patterns of potential fraud and to collaborate with law enforcement agencies on matters referred for investigation.

With the enforcement of the software features, processes, and staffing described above, the potential for increased healthcare fraud will be diminished, healthcare costs will be decreased, and the quality of healthcare will be improved.

Detailed Recommendations

Recommendations are presented by stakeholder:

Software Developers

Software should be designed that has the following characteristics:

- Use of a combination of statistics-based and rules-based automated coding to determine codes using a standardized database (as opposed to a facility-specific database) to determine initial coding standards. This should then be followed by an evaluation of coding rules by qualified coders assisting to develop associated rules. Codes should be generated only after the text has been developed.
- Audit capabilities for tracking all users of the document and the subsequent coding and billing process.
- Machine learning capabilities through the use of ANNs for data profiling, advanced analytic models, and ranked scoring in fraud and abuse software.
- Continuous monitoring capabilities of aggregate data analysis to detect any potential patterns of abuse with the aim of reducing submission of inaccurate claims.
- Standards based software with fraud alert warnings from the Office of the Inspector General (OIG) areas of focus, authoritative references including ICD-9-CM Official Coding Guidelines for Coding and Reporting the American Hospital Association's *Coding Clinic for ICD-9-CM* the American Medical Associations *CPT Assistant*, and the rules and conventions contained within the coding systems.
- Use of automated coding (NLP or structured text) systems that only code treatment or services that are adequately documented
- Incorporation of access controls to prevent or minimize opportunities for unauthorized persons (whether practitioner or non-practitioner) from entering information about an encounter or healthcare episode that did not occur or for producing false claims for services never rendered.
- Incorporation of identity authentication procedures to verify who accesses EHRs, automated coding software, and medical claim submissions, as well as mechanisms to track this access
- Incorporation of document versioning, to allow identification of the fact that a document has been changed, the date(s) and time of the changes and the ability to access earlier versions.

Software developers should include the following tasks as part of the product development process:

- Work with both payers and providers to determine how to develop products that decrease error and create accurate claims by communicating results of aggregate quality analysis with both groups.
- Utilize coding experts that have appropriate certifications and education in product development and during product support to assure that the code sets and authoritative references are up to date.

- Thoroughly assess the impact of the software on healthcare programs and beneficiaries.

Users

The following list presents recommendations for users:

- Use software that generate codes only after the text is developed.
- Conduct compliance checks to determine the accuracy of coded data and evaluate aggregate data for problematic trends when using automated coding on a continuing basis as part of the routine compliance program.
- Seek products that incorporate the following features and standards: standard coding references and coding edits, simple mechanisms to attain intra-operability with billing and abstracting systems, aggregate data evaluation capabilities in order to detect potential fraud, and audit trails.
- Create accountability within the organization with regard to accuracy of claims and build positive relationships with payers to determine how to improve interactions.
- Establish and enforce standard educational requirements in job descriptions for all coding and billing experts that include minimum education and continuing education requirements.
- Include coding and billing professionals in decisions to purchase and implement software. The impact software has on the accuracy of coding, billing, and the beneficiaries of healthcare programs should also be evaluated so that appropriate modifications to the software, implementation and education can take place to prevent incorrect bills.

Consumers

The following list provides some recommendations relative to healthcare consumers:

- Educate the consumer about the definition and extent of healthcare fraud and the means by which it can occur.
- Provide a report card of compliance to the public regarding accuracy of claims submission.
- Create on-line and other education for consumers regarding fraud and abuse and provide a mechanism for consumers to report potential fraud.
- Simplify the explanation of benefits and billing forms with the aim of demystifying the information. It would be helpful if billing documents could be matched to episodes of healthcare.

Payers

- Work more closely with providers. For example, developing a way to determine whether the current pattern of healthcare claims is within the patient's clinical history can help to prevent false claims.

- Use aggregate data analysis techniques, perform continuous monitoring, and use standard coding conventions and rules. Payers should also seek the guidance of coding professionals in the development of coding policy and in software purchasing decisions.
- Use machine learning capabilities through the use of ANNs for data profiling, advanced analytic models, and ranked scoring in fraud and abuse software.
- Undertake continuous monitoring capabilities of aggregate data analysis to detect any potential patterns of abuse with the aim of reducing submission of inaccurate claims.
- Require the use of standards based software with standard coding references and coding edits from the Office of the Inspector General (OIG) areas of focus, Coding Clinic and CPT Assistant, and CMS and process claims using these standard references.

Government

The federal agencies charged with detecting and prosecuting fraud should consider requiring that the following elements be included in automated coding software:

- A combination of statistics- and rules-based automated coding. So that statistics-based automated coding is fair and accurate, develop a standardized database for software developers to use to determine initial coding standards. Require that statistics-based automated coding be followed by developing coding rules by qualified coders.
- Audit trails for all uses of the document and the subsequent coding and billing process.
- Use of machine learning (ANNs) for data profiling, advanced analytic models, and ranked scoring in fraud and abuse software.
- Compliance checks for all organizations using the systems and incorporate into model compliance plans.
- Aggregate data analysis on a continuous basis (similar to continuous auditing) to detect any potential patterns of abuse with the aim of reducing inaccurate claims.

The following list summarizes governmental action items with respect to automated coding:

- Develop standards for automated coding products and certify only systems that have the following capabilities: standard coding references and coding edits, simple mechanisms to attain intra-operability with billing and abstracting systems, aggregate data evaluation capabilities in order to detect potential fraud, and audit trails.
- Require that only a certified product be used by organizations receiving government healthcare payments
- Support standard educational requirements for all coding professionals that include minimal education and continuing education.
- Provide an incentive for payers and providers to work together in developing standards, including those related to product certification, so that providers understand how to be legitimately reimbursed for services and payers have fewer inaccurate claims submitted to them.
- Require providers to take steps to ensure that they thoroughly assess all new computer systems and software that impact coding, billing, or the generation or transmission of information related to federal healthcare programs or their beneficiaries as part of their compliance program.

- Continuously train law enforcement personnel involved in the investigation of fraud regarding changing technology and its possible uses for fraudulent activities.
- Educate providers regarding what constitutes fraud, the consequences of submitting inaccurate claims, and how fraud and inaccurate claims can be prevented
- Provide ongoing education to law enforcement personnel regarding new developments in technology related to automated coding as the software evolves.

Conclusions

The following conclusions take into consideration the strengths and weaknesses of the current market:

- Computer-assisted coding software should utilize a combination of statistics-based and rules-based automated coding and a standardized national database (as opposed to a facility-specific database) to train the statistics-based engine. Audit trails are essential in all coding and billing software and EHR application to ensure that codes are based on documentation by clinicians. Machine learning such as ANNs should be available for predictive modeling to reveal trends and scores to detect fraud and abuse before it happens.
- Audit trails should be required for all users of the document and the subsequent coding and billing process.
- Machine learning such as ANNs should be used for data profiling, advanced analytic models, and ranked scoring in conjunction with fraud and abuse software.
- Compliance checks for automated coding should be required for all organizations using the systems. Aggregate data analysis should be required on a continuous basis (similar to continuous auditing) to detect any potential patterns of error with the aim of reducing inaccurate claims.
- A specific certifying mechanism should develop standards for products and only those systems that have the following required capabilities should be certified: standard coding references and coding edits, simple mechanisms to attain intra-operability with billing and abstracting systems, aggregate data evaluation capabilities in order to detect potential fraud, and audit trails.
- Organizations receiving Medicare payments and other government payments should be required to use only a certified product.
- Payers and providers need to work together so that providers understand how to be legitimately reimbursed for services and payers have fewer inaccurate claims submitted to them. Payers should also include the coding and billing professionals in their software purchasing and implementation decisions. Accuracy of coding from the payer's perspective should be evaluated to increase the accuracy of the data used for aggregate analysis.
- Standards that include minimum education requirements and continuing education requirements need to be established for all coding professionals.
- Providers should be required to take steps to ensure that they thoroughly assess all new computer systems and software that impact coding, billing, or the generation or transmission of information related to federal healthcare programs or their beneficiaries.

- Financial incentives should be created for compliance rather than the disincentive for fraud.
- A report card of compliance should be provided to the public regarding accuracy of claims submission.
- On-line and other education for consumers regarding fraud and abuse and a mechanism for consumers to report potential fraud should be provided.
- Providers should be educated regarding what constitutes fraud, the consequences of submitting inaccurate claims, and how fraud and inaccurate claims can be prevented.

Limitations of This Research, Future Research, and Next Steps

This research was based on data gathered from selected vendors of automated coding products and a limited number of users. It consisted of Web-based product demonstrations and telephone interviews with vendors, users, and government personnel. Many of the technologies described are quite new and on the cutting edge of healthcare IT and are not yet in widespread use.

Generally, more thorough evaluation is needed regarding how these tools perform in a variety of settings with different types of health records. Particular attention should be directed to the coding features of primary care EHRs that prompt for evaluation and management (E & M) code assignment. It is necessary to develop some agreed upon measures so that these technologies can be evaluated over time.

Short-term research and action plans suggested by this descriptive study are:

- Institute programs to improve national adherence to standard coding guidelines and rules by all stakeholders. This will require education about the consequences of local policy and practice and incentives to drive compliance. Standardization is a necessary prerequisite to improving data quality. It will also make it less costly to develop automated coding solutions, will permit more reliable trending for fraud detection and facilitate adoption of updated code sets.
- Evaluate the use of computer assisted coding technologies in production EHR settings. Compare and contrast the benefits in terms of data integrity, productivity and compliance monitoring for EHRs that feature structured versus unstructured text and those that are based on a reference terminology.
- Create use cases and test databases on which to evaluate the capability of computer assisted coding technologies to assign codes according to standard coding guidelines and rules. Many of these tools are new and not widely used in production settings. Early laboratory-based research could provide useful insight while broad-based field research is not feasible. This will permit assessing how best to certify these technologies in the future.
- Evaluate the potential of automated coding and anti-fraud software used in conjunction with the EHR to relieve coding workforce shortages. Research is needed to better understand this potential and what skills and competencies will be needed by coding experts in the future.

Long-term research questions include:

- What role can the consumer play in fraud prevention?
  - Do barriers such as health literacy prevent consumer participation?
  - What communication methods work best to educate consumers about fraud prevention and reporting?
  - Does the risk of fraud and its impact have any bearing on consumer participation in fraud prevention and reporting?
  - Can patient friendly billing practices increase consumer awareness of and role in helping to identify and prevent fraud?
- What is the statistical impact of automated coding on errors and fraudulent claims?
  - Are there statistically significant differences in error rates based on the type of automated coding method?
  - Is there a statistically significant difference in the number of claim errors in terms of a pre- and post- automated coding usage?
  - What is the best methodology to evaluate error rate of automated coding in various healthcare settings?
  - Is there a statistically significant difference in error rates in diagnostic, procedural, and evaluation and management code assignment?
  - What are the known economic impacts or potential benefits through cost savings and accuracy improvements associated with automated coding software?
- What is the impact of automated coding on various constituents including users, beneficiaries, payers, and law enforcement?
  - What is the impact of automated coding on users?
  - What is the impact of automated coding on beneficiaries and does it affect the interaction with providers positively or negatively?
  - What is the impact of automated coding on payers and law enforcement personnel?
  - What is the prevalence rate of automated coding products across the healthcare industry, including hospitals, outpatient facilities, physician offices, and so forth?
- What is the impact of using ANNs or other machine learning on fraud prevention?
  - Is there one method that is better than another to evaluate and prevent fraud?
  - Can ANNs be integrated easily into existing fraud prevention and compliance activities at the provider level?
  - Is there a statistically significant difference between the fraud prevention capabilities of ANNs vs. the aggregate pattern recognition that some automated coding products already offer?

Appendix A: Government Interview Form

The following questions were used during interviews with members of the federal government:

1. In your view, what are the best processes to prohibit fraud and abuse?
2. What problems do you foresee in relation to fraud and abuse when the Electronic Health Record (EHR) is used?
3. As discussed in an AHIMA practice brief, automated coding was defined as 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. Are you aware of specific facilities or settings that use automated coding systems or automated coding? Have you found patterns of abuse with automated coding?
4. Are you aware of incorrect coding or abuse detected with Natural Language Processing (NLP)? If you are familiar with the approach of the NLP, was it a rules-based approach or data-driven approach? Please describe.
5. Have you found a pattern of abuse with any particular commercial software product that assists in the determination of codes? Examples of products include: bar codes, pick or lookup lists, coding templates or coding protocol, automated super bills, logic or rules-based encoders, groupers, imaged and remote coding applications, hard coding via charge master tables. Please describe any patterns found.
6. According to Managed Healthcare Executive, the most effective anti-fraud and recovery programs include elements of process assessment, both retrospective and prospective technology and investigations and resolutions. Do you agree? Please discuss.
7. With the proliferation of EHRs with embedded reference terminology, such as SNOMED CT, do you envision this to affect fraud and abuse in automated coding systems? Please discuss.
8. What in your view are the weak links in anti-fraud software, education, and compliance practices?
9. What general patterns of abuse have you found by setting (for example, physician office, SNF, hospitals, and so forth) with services that:
    - Were never rendered, either by adding charges to legitimate claims, or by using actual patient names and health insurance information to fabricate claims.
    - Were up coded (second most common).
    - Were a deliberate provision of medically unnecessary services, which include tests, surgeries, and other procedures?
10. Are you aware of programs in which consumers have been educated to alert governmental agencies of fraud? If yes, how has it worked? What is the extent of fraud found by this means and are there any patterns of reporting by setting, diagnosis, or procedure?

Appendix B: Vendor Interview Form

As discussed in AHIMA's practice brief, automated coding is defined as 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. What type of automated coding system do you provide?
2. When was your first installation of the automated coding system? How many installations (users/clients) do you have and in what settings?
3. What is the average installation and training time?
4. Did coder quality change with the use of your automated coding system? Please describe what occurred in terms of coding quality and define how you evaluated coder quality. If coder quality was affected, by what percent was it affected?
5. Do you provide a remote coding application? Has this application improved coding productivity? Please describe what occurred in terms of productivity and by what percent the productivity changed? Please describe the number of outpatient records and inpatient records per hour before and after use of the remote coding application. What was the percent change?
6. How is the automated coding system used with the EHR? What are your thoughts regarding automated coding systems and what will transpire when the EHR is fully implemented?
7. How is the automated coding system used within the coding and billing process? Include the workflow from the coder assigning codes to billing to the payer.
8. What are the anti-fraud features available and how do they link to the automated coding system? Do you have future recommendations for anti-fraud features within automated coding systems?
9. How do you use the coded data in your analytics? Please elaborate on rules based vs. statistics based approach, as well as statistical modeling applications you may be using with the automated coding software.
10. Can you recommend any users of automated coding systems or vendors who are using or developing automated coding applications that we can also interview?
11. What do you believe are the weak links in fraud/abuse software, education, and compliance practices?

Appendix C: User Interview Form

1. What type of automated coding system do you use within your facility?
2. Is the automated coding system natural language processing (NLP) or structured text?
3. What is the approximate cost of the automated coding system (including education and training)?
4. When was the automated coding system developed?
5. How long did it take to implement the automated coding system on-site including education and training?
6. What is the level of accuracy on coding and billing?
7. How is the automated coding system used with the EHR?
8. How is the automated coding system used within the coding and billing process?
9. What are the anti-fraud features available and how do they link to the automated coding system?
10. What do you believe are the “weak links” in fraud/abuse software, education and compliance practices?

Appendix D: Product Information Form

Vendor Name:	
Address:	
E-mail Address:	
Phone:	
Contact Person:	
Title of Respondent:	

Place an X in the appropriate setting for each product listed that your company provides, The following definitions are taken from the *Coding (AHIMA Practice Brief) Glossary:* ED=Emergency Department; SDS=Same Day Surgery; Other OP= Other Outpatient; HH=Home Health, LTC=Long Term Care; ASC=Ambulatory Surgical Center; PO=Physician Office

	ED	SDS	Other Outpatient	X-ray	HH	LTC	Acute	Physician Office	Hospice
Coding Optimization Software									
Anti-Fraud Software - software that provides aggregate data analysis and record-specific audits									
Anti-fraud software									
Automated coding - Software that automatically generates a set of medical codes for review/validation and/or use based upon clinical documentation provided by healthcare practitioners									
Automated coding with NLP									
Automated coding with Structured Text									
Automated Coding Products									
Coding Tools - Tools used by coding professionals in the code assignment process:									
Bar codes									
Pick lists or lookup lists									
Automated super bills									
Logic or rules-based encoders									
Groupers									
Imaged coding applications									
Remote Coding applications									
Hard coding via chargemaster tables									
Automated coding-NLP system									
Automated coding-structured text									
Other automated coding systems									
Maintenance									

Appendix E: Vendor Information

The following vendor information matrix was compiled from the Product Information forms received from vendors. The following definitions are taken from the *Coding (AHIMA Practice Brief) Glossary:* ED=Emergency Department; SDS=Same Day Surgery; Other OP= Other Outpatient; HH=Home Health, LTC=Long Term Care; ASC=Ambulatory Surgical Center; PO=Physician Office

Vendor	1	2	3	4	5	6	7	8	9	10	11	12	13
Settings	ED, SDS, Other OP, Acute	Payer	Other OP, ASC,	ED, SDS, Other OP, X-ray, Acute, ASC, PO, DME, Pharmacy	ED, SDS, Other OP, X-ray, Acute (certain depts.), ASC, PO	Acute, SDS, Other OP, X-ray, HH, LTC, ASC, PO	Payer	ED, SDS, OP, X-ray, PO	ED, SDS, Other OP, Acute, ASC	ASC, PO	ED, X-ray, Acute, PO (Primary Care), Pathology	ED, SDS, Other OP, X-ray, Acute, ASC, PO, Clinics	ED, SDS, Other OP, X-ray, HH, LTC, Acute, ASC, PO, Hospice, DME, Pharm, Other Inpatient
Anti-Fraud Software - software that provides aggregate data analysis and record-specific audits													
Anti-Fraud Software		X (with NLP and structured text)		X	X	X (Acute only)	X					X	X
Automated coding - Software that automatically generates a set of medical codes for review/validation and/or use based upon clinical documentation provided by healthcare practitioners.													
Automated coding with NLP				X (Specific to CDM intent of service only)	X			X			X	X(ED and Adult Medical Clinics Only)	X
Automated coding with structured text			X	X	X					X	X		X
Automated coding products			X	X						X	X		X

1=Innovative Health Solutions; 2=Axonwave Software Inc.; 3=Provation Medical; 4=Craneware; 5=CodeRyte; 6=3M Health Information Systems; 7=McKesson Corporation; 8=KIWI-TEK, LLC; 9=Siemens Medical Solutions Health Services; 10=gMed, Inc. 11=A-Life Medical Inc.; 12=Medquist Inc.,13=Artificial Medical Intelligence Inc.

Appendix F: Product Costs

The following definitions are taken from the *Coding (AHIMA Practice Brief) Glossary:* ED=Emergency Department; SDS=Same Day Surgery; Other OP= Other Outpatient; HH=Home Health, LTC=Long Term Care; ASC=Ambulatory Surgical Center; PO=Physician Office

Vendor	1	2	3	4	5	6	7	8	9	10	11	12	13
Settings	ED, SDS, Other OP, Acute	Payer	Other OP, ASC,	ED, SDS, Other OP, X-ray, Acute, ASC, PO, DME, Pharmacy	ED, SDS, Other OP, X-ray, Acute (certain depts.), ASC, PO	Acute, SDS, Other OP, X-ray, HH, LTC, ASC, PO	Payer	ED, SDS, OP, X-ray, PO	ED, SDS, Other OP, Acute, ASC	ASC, PO	ED, X-ray, Acute, Primary Care, Pathology	ED, SDS, Other OP, X-ray, Acute, ASC, PO, Clinics	ED, SDS, Other OP, X-ray, HH, LTC, Acute, ASC, PO, Hospice, DME, Pharmacy, Other Inpatient
Automated coding-NLP system				(For total Chargemaster Toolkit Solution) Dependent on bed size—roughly $15,000 pa (100 beds)-$41,000 pa (1,000+ beds) plus I&T	Per transaction ASP model; price dependent on specialty			$1.00 per record	N/A		ASP with application specific volume based transaction pricing; Acute: ASP with application specific annual license fee	$1.25 per record (ED, Adult Medical Clinics Only)	$125,000 for hardware; $125,000 for implementation, PO=ASP model priced on a monthly basis; DMD depends on size

Vendor	1	2	3	4	5	6	7	8	9	10	11	12	13
Automated coding structured text system			Varies depending on setting but is close to current expense for transcription	Both the NLP portion and the structured text portion are embedded in product and included with same price Dependent on bed size—roughly $15,000 pa (100 beds)-$41,000 pa (1,000+ beds) plus I&T	Per transaction ASP model; price dependent on specialty				N/A	$12,500/MD	Included in cost of Automated coding-NLP system	$1.25 per record (ED, Adult Medical Clinics Only)	Same
Other automated coding system					Per transaction ASP model; price dependent on specialty	Cost depends on the usage of the institution and is different for every facility	Based on per member per month		N/A	N/A	Included in cost of Automated coding-NLP system	$1.25 per record	Same

Vendor	1	2	3	4	5	6	7	8	9	10	11	12	13
Approximate Maintenance Costs		18%	20% of software	Included in above annuity based pricing	0				N/A	$2,000/year	Transaction and annual license fee include standard maintenance. Special support rates negotiated on case-by-case basis		$50,000 yearly maintenance fee

1=Innovative Health Solutions; 2=Axonwave Software Inc.; 3=Provation Medical; 4=Craneware; 5=CodeRyte; 6=3M Health Information Systems; 7=McKesson Corporation; 8=KIWI-TEK, LLC; 9=Siemens Medical Solutions Health Services; 10=gMed, Inc. 11=A-Life Medical Inc.; 12=Medquist Inc.; 13=Artificial Medical Intelligence Inc.

Appendix G: Coding Optimization Tools

The following definitions are taken from the *Coding (AHIMA Practice Brief) Glossary:* ED=Emergency Department; SDS=Same Day Surgery; Other OP= Other Outpatient; HH=Home Health, LTC=Long Term Care; ASC=Ambulatory Surgical Center; PO=Physician Office

Vendor	1	2	3	4	5	6	7	8	9	10	11	12	13
Settings	ED, SDS, Other OP, Acute	Payer	Other OP, ASC,	ED, SDS, Other OP, X-ray, LTC, Acute, ASC, PO, Hospice, DME,Pharmacy	ED, SDS, Other OP, X-ray, Acute (certain depts.), ASC, PO	Acute, SDS, Other OP, X-ray, HH, LTC, ASC, PO	Payer	ED, SDS, OP, X-ray, PO	ED, SDS, Other OP, Acute, ASC	ASC, PO	ED, X-ray, Acute, Primary Care, Pathology	ED, SDS, Other OP, X-ray, Acute, ASC, PO Clinics	ED, SDS, Other OP, X-ray, HH, LTC, Acute, ASC, PO, Hospice, DME, Pharmacy, Other Inpatient
Coding Optimization Software - Additional prompts that may be added onto existing software to help the coder fully code each case													
Coding Optimization Software			X	X (Except DME)	X	X (Acute Only)	X			X			X
Coding Tools - Tools used for coding professionals in the code assignment process													
Bar Codes													
Pick lists or look up lists			X	X (Except DME)			X			X			X
Automated super bills			X	Future Release						X			
Logic or rules-based encoders	X		X	X (Except DME) (Not specifically a traditional encoder,	X	X				X	X		

Vendor	1	2	3	4	5	6	7	8	9	10	11	12	13
				but has the ability to generate a code based on logic thru different approach — HCPCS only)									
Groupers	X					X(All but X-ray HH, PO)				X			
Imaged Coding Applications									X		X	X	
Remote Coding Applications			X	X (Except DME) (via web tool)	X	X(All but X-ray & HH)			X			X	X
Hard Coding via Charge Master Tables			X	X(Except hospice and LTC)	X (Validate the CDM code in comparison to the MR)					X			

1=Innovative Health Solutions; 2=Axonwave Software Inc.; 3=Provation Medical; 4=Craneware; 5=CodeRyte; 6=3M Health Information Systems; 7=McKesson Corporation; 8=KIWI-TEK, LLC; 9=Siemens Medical Solutions Health Services; 10=gMed, Inc. 11=A-Life Medical Inc.;12=Medquist Inc.; 13=Artificial Medical Intelligence Inc.

Appendix H: Automated Coding Impact

The following table summarizes the impact of automated coding tools on coding and billing accuracy.

Automated Coding Software Tools	Use With EHR	Impact on Coding and Billing Accuracy	Significant Aspects of User Interface	Potential Use to Impact Fraud
Automated coding with NLP	Generates codes following the development of the record of care.	Generates codes based on what is documented in the clinical record so coding is more consistent with documentation.	**Practitioners** Utilizes existing processes and documents. Does not require a significant change in behavior. Can help to improve documentation through education. **Compliance Staff** Associated software capabilities can streamline administrative functions	Codes are generated based on the documentation alone and therefore compliance should increase Aggregate data analysis capabilities can detect claims or patterns of claims that differ from what is expected The same potential exists to fabricate claims as with paper records
Automated coding with structured text	Provides drop-down lists and prompts to generate the record of care.	Generates codes that are tied to structured text and this can increase coding accuracy for CPT codes. Accurate diagnostic codes are difficult to establish using this method unless there are limited clinical scenarios.	**Practitioners** Requires behavior change in terms of how clinical records are developed. **Coding and Billing Staff** Inconsistencies of diagnostic codes can result in increased analysis time **Compliance Staff** Software capabilities can streamline administration functions.	Because of the tie between the structured text and the code assignment for CPT, the coding quality of CPT codes should be high. Because of the range of diagnosis possibilities, the assignment of the correct diagnosis can be more difficult. This can result in an apparent lack of medical necessity unless the codes are edited.
Rules Based	Develops codes based on expert coders.	Because coding rules are the basis for code assignment, this method provides for a high level of accuracy.	**Coding and Billing Staff** Decreases revision of codes.	Results in high levels of accuracy and decreases fraud

Automated Coding Software Tools	Use With EHR	Impact on Coding and Billing Accuracy	Significant Aspects of User Interface	Potential Use to Impact Fraud
Statistics Based	Develops codes based on the history established in the database.	The history of how codes are assigned is the basis for statistical analysis. Therefore, statistics-based methods can be less accurate unless the coding history is based in a standardized or very large database.	**Coding and Billing Staff** Requires revision of codes that are not accurate	Errors can occur if a limited coding history is used. Coding professionals can have varying levels of expertise and therefore bias the statistics approach.
Combination	Develops codes initially based on the statistical analysis of the codes but then requires that expert coders determine rules for code assignment.	Results in the highest level of accuracy of code assignment.	**Coding and Billing Staff** Improves efficiency of coding process and quality of data **Compliance Staff** Results in better compliance through increased accuracy	Results in the highest level of accuracy of coded data
Anti-Fraud Software	Evaluates codes assigned through providing edits and evaluation of data quality. Evaluates aggregate data for patterns of abuse.	Edits assist in developing an accurate claim Aggregate data analysis enables the organization to examine problematic trends or allow payers to identify potential fraud.	**Practitioners** Alerts providers to unusual patterns of documentation or billing. **Coding and Billing Staff** Increases the number of complete and accurate claims **Compliance Staff** Provides alerts and methods to detect potential problems.	Coding and billing edits can be used concurrently with claim development Trends in data can be captured before claims are submitted or paid

Appendix I: Weak Links

The following table summarizes weak links in fraud and abuse software, user education, and compliance practices.

Respondent	Software	Education	Compliance
Federal Government	Incident of fraud with auto-coding software, so it can facilitate fraud Access to the data can be inhibited because of the type of software used. For example, if the coding system is proprietary, it may be hard to get access to the data. Most individuals in crime detection and compliance are not technically savvy so may not be able to deal with the increased technological patterns Antifraud software are turned off, ignored, or the issues detected are not addressed- it is a disincentive if not a deterrent	There is ignorance of automated coding and technology and fraud More education is needed regarding EHRs and how they will work. Further, it is important to know how interoperability will come into play It is important to teach law enforcement professionals and investigators how it will work, the more that investigators understand the more they will be able to apply their knowledge to on-going work in the area of concern Educate beneficiaries about fraud. With more awareness comes help with anti-fraud activities. Provide education about how to report inaccuracies on the Medicare Summary Notices (MSN) and especially the medication report.	Compliance staff will need to understand automated coding, EHR technology and how it impacts fraud More education about EHR is needed so that compliance staff can use appropriate software within the context of interoperability of systems There are mechanisms provided in HIPAA to allow beneficiaries to report fraud (Beneficiary Incentive Program from HIPAA 63FR31123 published June 8, 1998, 42CFR420.400)

Respondent	Software	Education	Compliance
Vendor	Rules may be inconsistent when audits are conducted because software edits are bypassed or software is not used. Difficult to cross-check the accuracy of information from system to system Retrospective and prospective designs, prospective good but difficult to do because of prompt payment laws All vendors require a final evaluation by an experienced or expert coding professional. Coding accuracy of automated coding alone would be problematic. Much of the software is rules based or statistics based; need a combination of both to improve accuracy Loose standards and poor data quality especially for claim elements that aren't tightly linked to payment	Inaccuracy of clinician dictation or documentation but coder should be able to confirm whether information is correct Lack of education of fellows and interns on how language used changes code Lack of education for providers, government personnel, payers and consumers regarding what constitutes fraud and how to detect it Lack of HIM education on how the coded data interfaces with revenue cycle Lack of reminder systems in dictation and auto-coding systems for physicians on what to include in their documentation	Tension between the payer and provider communities Limited administrative power leads to not monitoring medical necessity Not always done correctly or consistently even though they have compliance officer No aggregate data evaluation capabilities a payer can query regarding a claim's accuracy Lack of incentives for insurance payers to pursue and eradicate fraud
User	Inaccuracy of automated coding systems; None 100% accurate. Don't know the weaknesses of the automated coding system Don't know if automated coding software improved coding quality Have capabilities to cut and paste the documentation and need to change content to show what you did that day	Automated coding systems not used correctly by physicians Education and background of coder if lacking can make a huge difference in the quality of coded data Proficient coders more scarce which can also impact cost and quality of coded data Reliance on physicians to code without appropriate coding education background	Inaccuracy of codes of the automated coding product Whether compliance is being done consistently everywhere. No benchmarking standards Control of whether test is necessary, ordered appropriately, completed as ordered is much bigger underlying issue than whether software is effective.

Appendix J: Anti-Fraud Model

The following table summarizes features, processes and staffing for the ideal anti-fraud system or model.

Features	Processes	Staffing
Automated coding with NLP with a rules-based and statistics-based combination	Prepayment fraud detection using data profiling, advanced analytic models, and rank scoring.	Advanced coder analyst to edit and check all processes for accurate code assignment IT staff for maintenance of current technology and HIS staff to train office staff and physicians
A standardized system of data (based on a representative sample of standardized claims information) for statistical reference. Private development of software for code assignment facilitated by the federal government who may also establish criteria or certification procedures for software.	Prepayment fraud detection by using a standardized method to derive the statistical aspects of code assignment and data analysis.	Payer, provider, vendor all working together to combat fraud through education and incentives
ANN and predictive modeling to determine where potential for fraud lies	Post-payment fraud detection	Consumer involvement and education
Audit Trails		

Glossary

ABN

Advanced beneficiary notice.

ANN

Artificial neural network.

Auto-coding

The use of software that automatically generates a set of medical codes for review, validation, and use based upon clinical documentation provided by healthcare practitioners.

Computer-assisted coding (CAC)

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. This definition is from the document called *Delving into Computer-assisted Coding (AHIMA Practice Brief).*

CPT

Common procedural terminology.

Free text

Alphanumeric data that is unstructured, typically in narrative form. Unstructured data is not processed uniquely by the computer without the application of natural language processing tools. Free text provides the benefit of expressivity and flexibility. However, information that is recorded as free text is significantly more difficult to use for data analysis, aggregation, and comparison. (See also **structured data**.)

Data Quality

Data that is valid and reliable.

EHR

Electronic health record (see also **EMR**).

EMR

Electronic medical record (see also **EHR**).

Encoder

A tool used to automate the coding process that is similar to using a code book to assign codes. Encoders are computer software programs that usually prompt the coder to evaluate documentation and coding rules during the process of assigning a code. The use of encoders is generally thought to decrease variability of code assignment and increase accuracy.

HCPCS

Health Care Common Procedure Coding System.

ICD

International Classification of Diseases.

LMRP

Local Medical Review Policy.

NLP

Natural Language Processing.

OIG

Office of the Inspector General.

ROI

Return on Investment.

Statistical NLP

A group of techniques relying on mathematical statistics and used in natural language processing, for example, to find the most likely lexical categories or parses for a sentence. Often, the techniques are based on frequency information collected by analyzing large groups of sentences in a single language, to find out, for example, how many times a particular word ("dog," perhaps) has been used with a particular part of speech. The sentences in the group have usually been tagged in some way (sometimes manually) so that the information about the part of speech, each time each word is used, is known. Statistical NLP may also be referred to as Boolean NLP.

Structured data

Documentation of discrete data using controlled vocabulary rather than narrative text.

Structured input

A form of data entry that captures data in a structured manner (point-and-click fields, pull-down menus, structured templates, macros, and so forth).

Users

Those who use the vendor products.

Vendors

Companies who provide software related to coding and billing.

Footnote References

1. National Health Care Anti-Fraud Association, “Health Care Fraud, A Serious and Costly Reality for All Americans”, www.nhcaa.org/pdf/all_about_hcf.pdf Accessed June 14, 2005.
2. Centers for Medicare and Medicaid, “HIPAA Part 4: Protecting Medicare from Fraud and Abuse”, www.cms.hhs.gov/medlearn/billing Accessed June 14, 2005
3. AHIMA Coding Practice Team. "Internet Resources for Accurate Coding and Reimbursement Practices." (AHIMA Practice Brief) *Journal of AHIMA* 75, no.7 (July-August 2004): 48A-G.
4. AHIMA Position Statement on Consistency of Healthcare Diagnostic and Procedure Coding, 2002
5. AHIMA e-HIM™ Work Group on Computer-Assisted Coding. "Delving into Computer-assisted Coding. Appendix G: Glossary of Terms" *Journal of AHIMA* 75, no.10 (Nov-Dec 2004): web extra.
6. AHIMA e-HIM™ Work Group on Computer-Assisted Coding. "Delving into Computer-assisted Coding. Appendix C: Advantages and Disadvantages of CAC Technology" *Journal of AHIMA* 75, no.10 (Nov-Dec 2004): web extra.
7. AHIMA e-HIM™ Work Group on Computer-Assisted Coding. "Delving into Computer-assisted Coding. Appendix E: Summary of Use Cases" *Journal of AHIMA* 75, no.10 (Nov-Dec 2004): web extra.
8. Fair Isaac. “Automated Exception Management for Efficient Claims Processing.” A Fair Isaac White Paper. January 2003
9. Fair Isaac. “Prepayment Fraud and Abuse Detection”. A Fair Isaac White Paper. January 2003.
10. Popowich, F. “Using Text Mining and Natural Language Processing for Health Care Claims Processing” Axonwave:To appear in A.Kao and S.Poteet (eds.), SIGKDD Explorations, Special Issue on Text Mining and Natural Language Processing, 2005
11. Popowich, F., “Use of Text Analytics and Taxonomies for Fraud and Abuse Detection in Medical Insurance Claims” Axonwave Software Inc., 2004, Burnaby, BC, Canada
12. Sordo, Margarita. “Introduction to Neural Networks in Healthcare”. October 2002. [Online: www.openclinical.org/docs/int/neuralnetworks011.pdf]
13. Dunn, Rose; "Putting Productivity Plans to Work: Sample Coding Incentive Plan 2" Journal of AHIMA, October 2001, 72/9, 96-100

Bibliography

AHIMA. “Advocating for Quality Documentation and Adherence to Official Coding Guidelines." *Journal of AHIMA* 69, no. 1 (1998), insert before p.49.

AHIMA. “After Automated Coding, What?" *Journal of AHIMA* 70, no. 7 (1999): 43.

AHIMA Coding Practice Team. "Internet Resources for Accurate Coding and Reimbursement Practices." (AHIMA Practice Brief) *Journal of AHIMA* 75, no.7 (July-August 2004): 48A-G.

AHIMA e-HIMTM Work Group on Computer-Assisted Coding. "Delving into Computer-assisted Coding" (AHIMA Practice Brief). *Journal of AHIMA* 75, no.10 (Nov-Dec 2004): 48A-H (with web extras).

AHIMA e-HIMTM Work Group on Computer-Assisted Coding. "Delving into Computer-assisted Coding. Appendix A: Primer on NLP for Medical Coding" *Journal of AHIMA* 75, no.10 (Nov-Dec 2004): web extra.

AHIMA e-HIMTM Work Group on Computer-Assisted Coding. "Delving into Computer-assisted Coding. Appendix B: Continuum of Linguistic Competence" *Journal of AHIMA* 75, no.10 (Nov-Dec 2004): web extra.

AHIMA e-HIMTM Work Group on Computer-Assisted Coding. "Delving into Computer-assisted Coding. Appendix C: Advantages and Disadvantages of CAC Technology" *Journal of AHIMA* 75, no.10 (Nov-Dec 2004): web extra.

AHIMA e-HIMTM Work Group on Computer-Assisted Coding. "Delving into Computer-assisted Coding. Appendix D: Potential Uses of Structured Code Output" *Journal of AHIMA* 75, no.10 (Nov-Dec 2004): web extra.

AHIMA e-HIMTM Work Group on Computer-Assisted Coding. "Delving into Computer-assisted Coding. Appendix E: Summary of Use Cases" *Journal of AHIMA* 75, no.10 (Nov-Dec 2004): web extra.

AHIMA e-HIMTM Work Group on Computer-Assisted Coding. "Delving into Computer-assisted Coding. Appendix F: Resources" *Journal of AHIMA* 75, no.10 (Nov-Dec 2004): web extra.

AHIMA e-HIMTM Work Group on Computer-Assisted Coding. "Delving into Computer-assisted Coding. Appendix G: Glossary of Terms" *Journal of AHIMA* 75, no.10 (Nov-Dec 2004): web extra.

AHIMA Position Statement on Consistency of Healthcare Diagnostic and Procedure Coding, 2002

Anonymous. “Healthcare fraud and abuse remains a costly challenge.” *Management Healthcare Executive* 14, no. 10 (2004): 38.

Averill, R. "Honest Mistake or Fraud? Meeting the Coding Compliance Challenge." *Journal of AHIMA* 70, no.5 (1999): 16-21.

Becker, D., Kessler, D, McClellan, M. “Detecting Medicare Abuse.” *Journal Health Economics* 24, no. 1 (2005): 189-210.

Beinborn, J. "Automated Coding: the Next Step." *Journal of AHIMA* 70, no.7 (1999): 38-43.

Biafore, S. "Real-time fraud detection. Systems with predictive analytics can catch fraud and abuse before they happen." *Healthcare Informatics* 20, no. 10 (2003): 82.

Bolin, J N, Clark, LS. "Avoiding charges of fraud and abuse: Developing and implementing an effective compliance program." *Journal of Nursing Administration* 34, no. 12 (2004): 546-550.

Borca, G.J. "Technology curtails health care fraud." *Management Care* 10, no. 4 (2001): 50-53.

Bowman, S. "Room for Improvement." *Journal of AHIMA* 76, no.4 (April 2005): 18, 20-21.

Bowman, J, and Stanfill, M "Physicians Cast Wary Eye at Computer-Assisted Coding." *Journal of AHIMA* 75, no.8 (September 2004): 76-77.

Curtin, L.; Simpson, R.L. "Preventing fraud." *Journal Health Management Technology* 22, no. 10 (2001): 66.

Cummins, R. "Preparing for Automated Coding." Source: *2004 IFHRO Congress & AHIMA Convention Proceedings,* October 2004.

Department of Justice. "Legal Considerations in Designing and Implementing Electronic Processes: A Guide for Federal Agencies." November 2000. http://www.cybercrime.gov/eprocess.htm

Diamond, F. "Prevent crime from paying: Flight managed care fraud." *Journal Managed Care* 11, no. 10 (2002): 20-24.

Dorn, C. "Data Mining Technology Helps Insurers Detect Health Care Fraud." *Journal of National Underwriter: Life and Health* 108, no. 39 (2004): 34.

Fair Isaac. "Automated Exception Management for Efficient Claims Processing." A Fair Isaac White Paper. January 2003

Fair Isaac. "Prepayment Fraud and Abuse Detection." A Fair Isaac White Paper. January 2003.

Fenton, S. "An Introduction to the Unified Medical Language System." *Journal of AHIMA* 76, no.5 (May 2005): 60-62.

Fenton, S. "Clinical Vocabularies: Essential to the Future of Health Information Management." *IFHRO Congress & AHIMA Convention Proceedings*, October 2004.

Frieden, J. "Fraud squads target suspect claims." *Journal of Business Health* 9, no. 4 (1991):
20-24.

Frieden, J. "Health care fraud detection enters the information age." *Journal of Business Health* 10, no. 7 (1992): 29-30.

Garcia, James L. "Using technology to fight fraud." *Journal of Health Management Technology* 23, no. 1 (2002): 32-34.

Gill, R. and Hecht, R. “Audit of Automated Code Assignment in Israeli Hospital’s Computerized Medical Information Systems.” *IFHRO Congress & AHIMA Convention Proceedings,* October 2004

Goedert, J. "Software Helps Fight Fraud and Abuse Problem.” *Journal of Health Data Mining* 11, no. 6 (2003).

Hagland, M. "Revolution in Progress: How Technology is Reshaping the Coding World." *Journal of AHIMA* 73, no.7 (2002): 32-35.

Haron, D.L. "Waste and neglect: Fraud and abuse in the health care industry.” *Journal of Health Care Supervison* 16, no. 4 (1998): 61-67.

Johns, M. "A Crystal Ball for Coding." *Journal of AHIMA* 71, no.1 (2000): 26-33.

Journal Health Data Management. “Readers’ perspectives. Commercial and governmental health care payers should make the use of information technology to detect and prevent fraud and abuse a top priority.” *Health Data Management* 11, no. 6 (2003): 104.

Kaldal, K. "Benchmarks as the Signposts on the Fraud Case Trail." *Journal of AHIMA* 69, no.10 (1998): 44-45.

King, E. "In Pursuit of Healthcare Fraud." *Journal of AHIMA* 70, no.1 (1999): 22-28.

King, E. "Playing a Part: The FBI's Role in Healthcare Fraud Investigations." *Journal of AHIMA* 69, no. 1 (1998): 43.

Lobb, R.Dixon, M. “Readers’ perspective. Is the role of consultants increasing or decreasing in importance in health care information technology?” *Journal of Health Data Mining* 9, no.6 (2001): 72.

Morris, L. “Hard-line approach to compliance pays off.” *Journal of AHIMA* 72, no.5 (2001): 18-20.

Morris, WC., Heinze, D., Warner, H., et al. "Assessing the Accuracy of an Automated Coding System in Emergency Medicine." *Proceedings of the 2000 AMIA Annual Symposium,* 595–99.

National Health Care Anti-Fraud Association. “Health Care Fraud: A Serious Costly Reality For All Americans.” April 2005.

Nelson, L. "Stepping Into Continuous Audit." *The Internal Auditor* 61, no.2 (2004): 27.

Office of Inspector General. “OIG Supplemental Program Guidance for Hospitals.” Federal Register. January 2005: 4858-4876.

Popowich, F. "Use of Text Analytics and Taxonomies for Fraud and Abuse Detection in Medical Insurance Claims.” [Online: www.cscsi.org/home/CSCSI/Members/swig/swig04papers/popowich-swig.pdf].

Prophet-Bowman, S. "Observation Services Present Compliance Challenges: Know These Complex, Changing Requirements to Avoid Risk." *Journal of AHIMA* 74, no.3 (2003): 60-65.

Prophet, S. "Coding Compliance: Practical Strategies for Success." *Journal of AHIMA* 69, no.1 (1998): 50-61.

Prophet, S. "Corporate Compliance: A Look at the Hot Issues and Current Trends.". *AHIMA Convention Proceedings, October 2001*

Prophet, S. "GAO Evaluates Implementation of False Claims Act Guidance." *Journal of AHIMA* 70, no. 10 (1999): 19-20.

Prophet, S. "Is the FCA Effectively Fighting Healthcare Fraud?." *Journal of AHIMA* 73, no.5 (2002): 20-23.

Prophet, S. "Fraud and Abuse: What You Can Do." *Journal of AHIMA* 69, no. 1 (1998): 68-70.

Prophet, S. "Fraud and Abuse Implications for the HIM Professional." *Journal of AHIMA* 68, no.4 (1997): 52-56.

Prophet, S. "Roundtable takes on compliance challenges." *Journal of AHIMA* 70, no. 7 (1999): 20-26.

Scala-Foley, M; Bryant, N. "Medicare fraud and abuse." *Journal Issue Brief Center Medicare Education* 5, no. 6 (2004): 1-6.

Schnitzer, G L. and Stanfill, M. "Outwit, Outlast, Outcode: Surviving in the Autocoding Era." *Journal of AHIMA* 72, no.9 (2001): 102-104.

Schnitzer, G.L. "Natural Language Processing: a Coding Professional's Perspective." *Journal of AHIMA* 71, no.9 (2000): 95-98.

Sordo, M. "Introduction to Neural Networks in Healthcare." October 2002. [Online: www.openclinical.org/docs/int/neuralnetworks011.pdf].

Statement Of Bruce Merlin Fried, Director. Center For Health Plans And Providers On: Combating Medicare Waste, Fraud, And Abuse Through Information Technologies. Before The House Committee On Commerce Subcommittee On Oversight And Investigations. Hearing. 1997.

Taylor, M. "High-tech sleuthing. New computer tools give government tighter handle on hard-to-track healthcare fraud." *Journal Modern Healthcare* 30, no. 50 (2000): 46-48.

Taylor, M. "Fraud detectors. Payers and health insurers are turning to powerful new software to help identify questionable healthcare claims in real time." *Journal Modern Healthcare* 32, no. 12 (2002): 32-34.

TITLE 31. Money and Finance , Subtitle III. Financial Management , Chapter 37. Claims, Subchapter III. Claims Against the United States Government. *U.S. Code* (January 2, 2001)

U.S. Centers for Medicare & Medicaid Services. "Clinton Administration Releases Grants To Fight Fraud And Abuse." 1997. < http://www.cms.hhs.gov/media/press/release.asp?Counter=30>

U.S. DHHS, OIG. "Medicare and state health care programs; fraud and abuse: OIG civil money penalties under the Medicare prescription drug discount card program. Final rule." *Federal Register* 69, no. 239 (2004): 74451-74453.

< http://a257.g.akamaitech.net/7/257/2422/06jun20041800/edocket.access.gpo.gov/2004/pdf/04-27341.pdf>

U.S. DHHS, OIG. "Health care fraud and abuse data collection program: technical revisions to Healthcare Integrity and Protection Data Bank data collection activities. Final rule." *Federal Register* 69, no. 182 (2004): 56364-56367.

< http://a257.g.akamaitech.net/7/257/2422/06jun20041800/edocket.access.gpo.gov/2004/pdf/04-21204.pdf>

Warner, H.Jr., "Can Natural Language Processing Aid Outpatient Coders?" *Journal of AHIMA* 71, no.8 (2000): 78-81.

Warner, H. Jr., "Will Natural Language Processing Help Coders Any Time Soon?" Proceedings of the 2001 AHIMA Convention. Available online at http://library.ahima.org/xpedio/groups/secure/documents/ahima/pub_bok2_001414.html

Weiss, H.P. "Feds unveil new data collection program to fight fraud and waste." *Journal Balance* 2, no. 1 (1998): 5.

Wegener, J and Woodman, A J. "Critical Success Factors in Establishing an Electronic Health Record. The Experience of St. Michael's Hospital, Canada." Source: 2004 IFHRO Congress & AHIMA Convention Proceedings, October 2004

Carus, W., Dictaphone Corporation, and D'Amato,C. "Auto-Coding: HIM Professionals' Role in the Future." *AHIMA's 75th Anniversary National Convention and Exhibit Proceedings,* October 2003