2000

An Intelligent Data Mining System to Detect Health Care Fraud

Guisseppi A. Forgionne
University of Maryland - Baltimore County

Aryya Gangopadhyay
University of Maryland - Baltimore County

Monica Adya
Marquette University, monica.adya@marquette.edu

Chapter VII

An Intelligent Data Mining System to Detect Healthcare Fraud

Guisseppi A. Forgionne
Aryya Gangopadhyay
Monica Adya
University of Maryland Baltimore County, USA

INTRODUCTION

There are various forms of fraud in the health care industry. This fraud has a substantial financial impact on the cost of providing healthcare. Money wasted on fraud will be unavailable for the diagnosis and treatment of legitimate illnesses. The rising costs of and the potential adverse affects on quality healthcare have encouraged organizations to institute measures for detecting fraud and intercepting erroneous payments.

Current fraud detection approaches are largely reactive in nature. Fraud occurs, and various schemes are used to detect this fraud afterwards. Corrective action then is instituted to alleviate the consequences. This chapter presents a proactive approach to detection based on artificial intelligence methodology. In particular, we propose the use of data mining and classification rules to determine the existence or non-existence of fraud patterns in the available data.

The chapter begins with an overview of the types of healthcare fraud. Next, there is a brief discussion of issues with the current fraud detection approaches. The chapter then develops information technology based approaches and illustrates how these technologies can

improve current practice. Finally, there is a summary of the major findings and the implications for healthcare practice.

BACKGROUND

Fraud in healthcare transactions refers to knowingly and willfully offering, paying, soliciting, or receiving remuneration to induce business that healthcare programs will reimburse. Healthcare fraud can result from internal corruption, bogus claims, unnecessary health care treatments, and unwarranted solicitation. As in any commercial enterprise, unscrupulous provider or payer employees can misappropriate healthcare payments for personal purposes. Providers can also issue claims for treatments that were never, or only partially, rendered. Corrupt healthcare providers also can induce patients to undergo unnecessary, or even unwanted, treatments so as to inflate charges to the payers. In addition, unethical providers can willfully solicit business from unprincipled, or unsuspecting, patients for the sole purpose of generating billable procedures and treatments.

According to a 1993 survey by the Health Insurance Association of America of private insurers' healthcare fraud investigations, the majority of healthcare fraud activity is associated with diagnosis (43%) and billing services (34%). In Medicare, the most common forms of fraud include billing for services not furnished, misrepresenting the diagnosis to justify payment, falsifying certificates of medical necessity, plans of treatment and medical records to justify payment, and soliciting, offering, or receiving a kickback (Health Care Financing Administration, 1999).

Early cases of healthcare fraud have applied to gross issues such as kickbacks, bribes, and other fairly transparent schemes. Increasingly, however, the Office of the Inspector General has demonstrated a willingness to pursue cases that are in the gray area and courts have tended to interpret antifraud statues more broadly so as to make criminal prosecution more likely (Steiner, 1993). For instance, waiving a patient's co-payment when billing third-party payers and not disclosing the practice to the insurance carrier has been deemed as fraud and resulted in prosecution (Tomes, 1993).

Fraud has a substantial financial impact on the cost of providing healthcare. Medicaid fraud, alone, costs over $30 billion each year in the United States (Korcok, 1997). According to CIGNA HealthCare

and Insurance groups, the healthcare industry is losing an estimated $80 to $100 billion to fraudulent claims and false billing practices (CIGNA, 1999). Investigators have shown that fraud is found in all segments of the healthcare system, including medical practice, drugs, X-rays, and pathology tests, among others.

The timely detection and prevention of fraud will not only provide significant cost savings to insurance companies but will also reduce the rising cost of healthcare. Money wasted on fraud will be unavailable for the diagnosis and treatment of legitimate illnesses. In the process, research monies may be reduced and critical research may be delayed. Ineffective and cost inefficient treatments may continue. Administrative effort may be diverted to fraud detection instead of being concentrated on the effective management of healthcare practice. As a consequence, patient care may suffer and healthcare costs may continue to soar.

MAIN THRUST OF THE CHAPTER

There are several issues, controversies, and problems associated with fraud detection. An analysis of these issues recommends a solution based on artificial intelligence techniques.

Issues, Controversies, and Problems

In the past, claim fraud has been identified through complaints made, among others, by disgruntled healthcare competitors, beneficiaries and recipients, and present or former employees of providers. A significant volume of false claims, however, still go undetected. Consequently, fraud is still rampant in the healthcare system. The rising costs of, and the potential adverse effects on quality health care, have encouraged organizations to institute measures for detecting fraud and intercepting erroneous payments, especially through electronic means.

Due to the documentation typically required by payers, all forms of healthcare fraud will leave a paper, or electronic, trail that can serve as the basis for detection. However, the transactions useful for fraud detection will generally be buried in the documentation. Furthermore, these transactions may be from disparate sources and in diversified formats. Often, the needed transactions are also discarded as a normal part of transmitting claims from providers to payers.

Another major barrier to fraud detection is the reactive nature of the current approaches. For the most part, detection relies on: (a) complaints made by disgruntled interested parties, (b) random examinations by payers of provider submitted records, and (c) occasional detailed studies by public and private oversight agencies (Tomes 1993). Since such methods tend to be relatively narrow in scope, few fraud cases will be detected in this manner. Even in the identified cases, detection will be time consuming, costly, and difficult to correct.

Solutions and Recommendations

With the increasing number of healthcare transactions and persecution of situations with such uncertainty, it is possible to increase the chances of detecting fraud through the use of information technology. Such technology can be utilized to develop a proactive and effective healthcare fraud detection strategy based on data warehousing, data mining, artificial intelligence, and decision support systems.

Data needed to support the identification of fraud routinely flow, often electronically, between healthcare providers and payers as medical transactions. By filtering and focusing the transactions, warehousing the focused data, and creating tailor-made data marts for the appropriate recipients, requisite information can be made available for significant data mining analyses (Abraham and Roddick, 1998; Davidson, Henrickson, Johnson, Myers, and Wylie, 1999). Artificial intelligence then can be used to help providers and payers detect the underlying fraudulent patterns in the data and, with the aid of additional information technology, form effective proactive correction strategies (Burn-Thornton and Edenbrandt, 1998; Hornung, Deddens, and Roscoe, 1998; Makino, Suda, Ono, and Ibaraki, 1999).

Data Warehousing

Data warehousing involves the physical separation of day-to-day operational healthcare data from decision support systems. Benefits of data warehousing include clean and consistent organization-wide data, protection of transactional and operational systems from user's query and report requirements, and effective updating and maintenance of applications. The more significant purpose of the data warehouse is to support multidimensional analyses of both historical and current data.

A multidimensional model is developed using the MOLAP (multidimensional on-line analytical processing) design. Several data cubes are populated with historical and current data. An example of a three-dimensional data cube consists of patient demographics, time, and procedure code as the dimensions, and the payment as the measure. The actual analysis could require dimensionality reduction, such as a time-series analysis of payment records for patients that underwent a given treatment. In this case only two dimensions of the data cube are investigated. Such an analysis could be required to establish a historical pattern of the amount of payments made for a given medical procedure, sudden changes of which may cause an alarm for further investigations. Average values of payment amounts for medical procedures over a given data set can be used as a normative value to trigger any significant variations in current payment amounts. Other examples of multidimensional analyses include pivoting or cross tabulating measures against dimensions, dicing the cube to study a subpopulation of the data collected over a period of time, and rollup or drill down along dimensions to study any changes that might have taken place along individual dimensions.

Data Mining and Classification Rules

Data mining is an emerging technique that combines artificial intelligence (AI) algorithms and relational databases to discover patterns with or without the use of traditional statistical methods (Borok, 1997). It typically employs complex software algorithms to identify patterns in large databases and data warehouses. Data mining can facilitate information analysis using either a top-down or a bottom-up approach (Limb and Meggs, 1995). While the bottom-up approach analyzes the raw data in an attempt to discover the hidden trends and groups, top-down data mining tests a specific hypothesis.

Effective data mining relies on an effective and representative data warehouse. By definition, data mining is a pattern discovery process that relies on large volumes of data to infer meaningful patterns and relationships between data items. Once the data is "mined" from the warehouse and patterns are cataloged, the patterns themselves can be converted into a set of rules (Borok, 1997). These rules that explain healthcare behavior will be coded into a rule-base and be used for analyzing individual instances.

Classification rules deal with identifying a class of regularities in data (Adam, Dogramaci, Gangopadhyay and Yesha,1998; Ramakrishnan 1997). A classification rule is an expression $(l_1 \leq X_1 \leq U_1)$ Ÿ $(l_2 \leq X_2 \leq U_2)$ … $(l_k \leq X_k \leq U_k)$ Ø $(l_y \leq X \leq U_y)$, where $X_1 … X_k$ are attributes used to predict the value of Y, and $l_1 … l_k$, $U_1 … U_k$ are the lower and upper bounds of the corresponding attribute values, respectively. As an example, in detecting healthcare fraud, a classification rule would be X Ø $(Y \leq l_y)$ or $(U_y \leq Y)$, where X is a surgical procedure and (l_y, U_y) is the prescriptive range of values for the payments made (Y).

A classification rule is said to have a support *s* if the percentage of all cases satisfying the conditions specified in the rule equals or exceeds the support. In other words, *s* is the ratio to the total number of cases where both X and Y values are within the specified ranges. The confidence *c* of a classification rule is defined as the probability that, for all cases where the value of X falls within its specified range, the value of Y will also be within the range specified for Y. In other words, **c** is the ratio of cases where the values of X and Y are within their respective specified ranges, to the total number of cases where only the X values are within the specified range. Both support and confidence can be user or system specified as percentages or ratios.

If the support for a certain rule is low, it indicates that the number of cases is not large enough to make any conclusive inference. In that case, no further analysis is done with the current data set. If the support is large but the confidence is low, the rule is rejected. If both the support and confidence exceed the values specified by the user (or system) then the rule is accepted. Such a case would trigger a flag for a potential fraud and recommend further investigation, which is done by isolating the cases that triggered the flag.

Illustrative Example

Take the instance of determining physician charges for a surgical procedure. Charges for this procedure may vary somewhat by, among other things, physician, location of the practicing facility, and the regulations of the insurance provider. It is challenging, therefore, to identify an acceptable and representative range of charges using traditional statistical techniques. This requires understanding the physicians' practice procedures, determining the practice patterns implicit in the data, and possibly identifying practice patterns over the past few weeks.

Data mining can discover such patterns in the historical data. More importantly, it can uncover atypical patterns of practice within a group. For instance, mining on a large sample of nationwide data may identify that for a simple dental procedure, physicians charge a fee of $45.00 to $60.00 in the state of Maryland. If there is a sufficient number of cases in the data warehouse that support the correlation between the procedure and the range of charges, then the support and confidence in this rule will be high. Otherwise the rule will be rejected and will not be included in the rule-base. If the rule is accepted, a new case regarding this procedure can now be compared against the rule and can trigger a fraud alert if the charges deviate significantly from those specified in the rule.

In another instance, data mining may support the analysis and understanding of temporary conditions which may be incorrectly triggered as a fraud alert. Suppose the classification rules above indicate an increase in the incidence of emergency hospitalizations than in other regions around the area. This deviation can set up a trigger whereby further analysis may reveal the presence of a high-risk construction facility for the next two years. This factor will allow the healthcare providers to prepare for the situation both during and after the construction activity and possibly aid in the prevention of emergency situations at this facility. Similar analysis can be used for chronic conditions such as breast or lung cancer in specific regions.

FRAUD DETECTION SYSTEM

In the next few sections, we suggest the development of a decision support system to support the identification of fraud in healthcare transactions. The architecture for this system is proposed in Figure 1. As this figure shows, the system interactively processes inputs into the outputs desired by healthcare users.

Inputs

The fraud detection system has a data base that captures and stores historical, and industry standard, data on healthcare providers, claims, and payments. These data are extracted from the data warehouse that captures the relevant transactions from the providers to the payers, and vice versa.

Provider information includes the name, address, ID, and other demographics. Claims information includes the patient ID, procedure code, charge,

billing dates, and other financial statistics. Payment information includes the patient and provider IDs, deductibles, co-payments, covered remuneration, and relevant payment dates.

There is also a model base that contains classification rules and artificial intelligence algorithms. The classification rules would establish lower and upper limits, supports, and confidence levels for each covered procedure from historical data and industry standards. These rules would be derived through the data mining tool, and the classification algorithm would determine the support and confidence of the classification rules.

Processing

The health official (health plan administrator, auditor, or other staff assistant) uses computer technology to perform the fraud detection analyses and evaluations. Computer hardware includes an IBM-compatible Pentium-based microcomputer with 16MB of RAM, a color graphics display, and a printer compatible with the microcomputer. Software includes the SAS information delivery system running through the Microsoft Windows operating system. This configuration was selected because it offered a more consistent, less time-consuming, less costly, and more flexible development and implementation environment than the available alternatives.

Users initiate the processing by pointing and clicking with the computer's mouse on screen-displayed objects. The system responds by automatically organizing the collected data, structuring (estimating and operationalizing) the classification rules, and simulating fraud performance. Results are displayed on the preprogrammed forms desired by health officials. Execution is realized in a completely interactive manner that makes the processing relatively transparent to the user.

As indicated by the top feedback loop in Figure 1, organized data, structured classification rules, and fraud performance reports created during the system's analyses and evaluations can be captured and stored as inputs for future processing. These captured inputs are stored as additional or revised fields and records, thereby updating the data and model bases dynamically.

The user executes the functions with mouse-controlled point-and-click operations on attractive visual displays that make the computer processing virtually invisible (transparent) to the user.

Figure 1: Fraud Detection Conceptual System Architecture
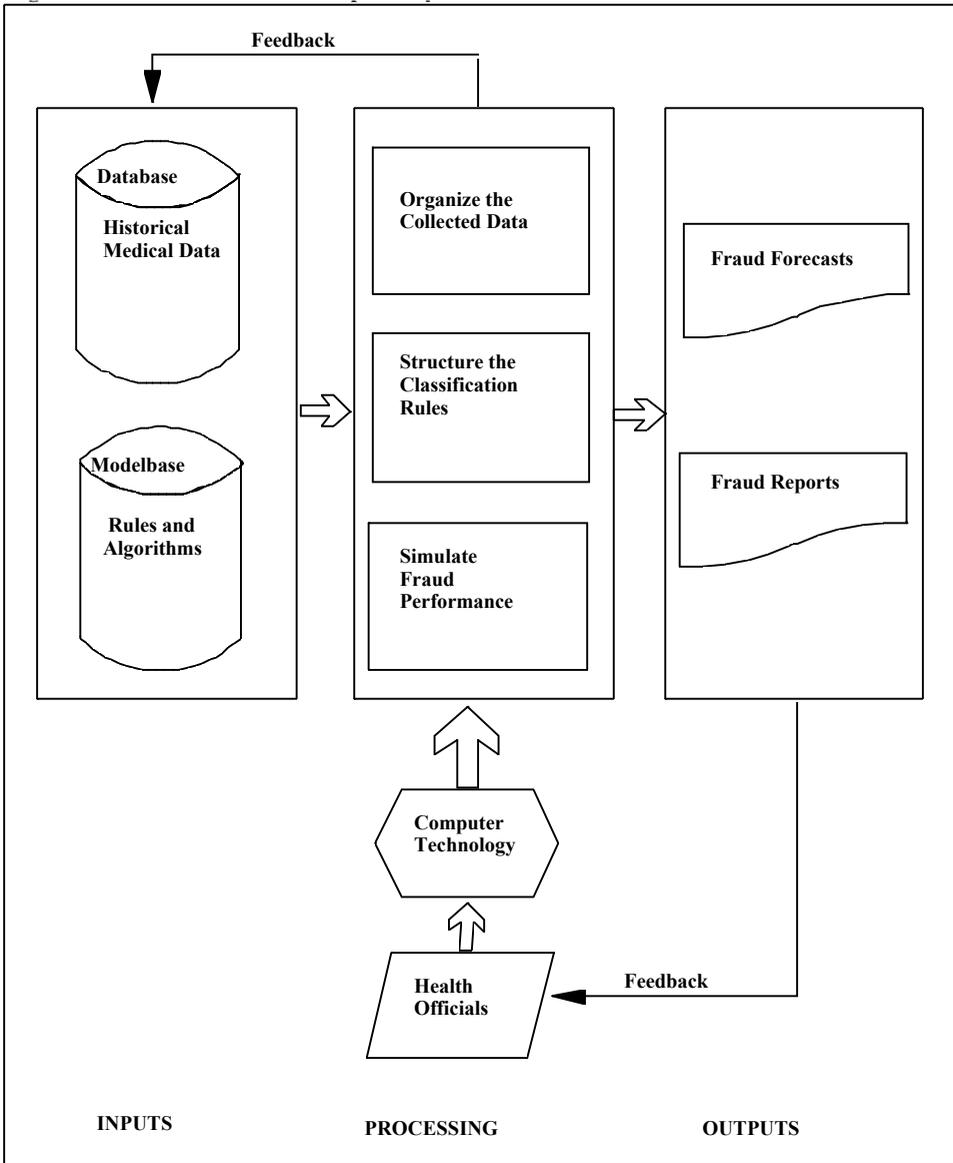
Outputs

The above procedures generate visual displays of the outputs desired by health officials. Outputs include fraud forecasts and reports. These reports are in the form of tables and graphs. Each table displays the forecasted payment value relative to its lower and upper limits for a specified medical procedure.

The corresponding graph highlights deviations outside the limits and allows the user to drill down to the supporting detail (which includes the provider, any extenuating circumstances, and other relevant information). The user has the option of printing or saving the reports.

As indicated by the bottom feedback loop in Figure 1, the user can utilize the outputs to guide further processing before exiting the system. Typically, the feedback will involve sensitivity analyses in which the user modifies support and confidence levels, upper and lower limits, or other pertinent factors and observes the effects on fraud performance.

System Session

There is a graphic icon on the Windows desktop. By double clicking this icon, the user accesses the fraud detection system. Once in the system, the user performs the fraud detection analyses and evaluations by navigating with point-and-click operations through the displays overviewed in Figure 2.

The Welcome display (shown in Figure 3) enables the user to access an embedded executive information system (EIS) shown in Figure 4. Once in the EIS, the user can interactively access the data warehouse, by selecting the database management system (DBMS) button, or go directly to DSS reports by selecting the REPORTS button. Selecting the DBMS button will enable

Figure 2: Display Relationships
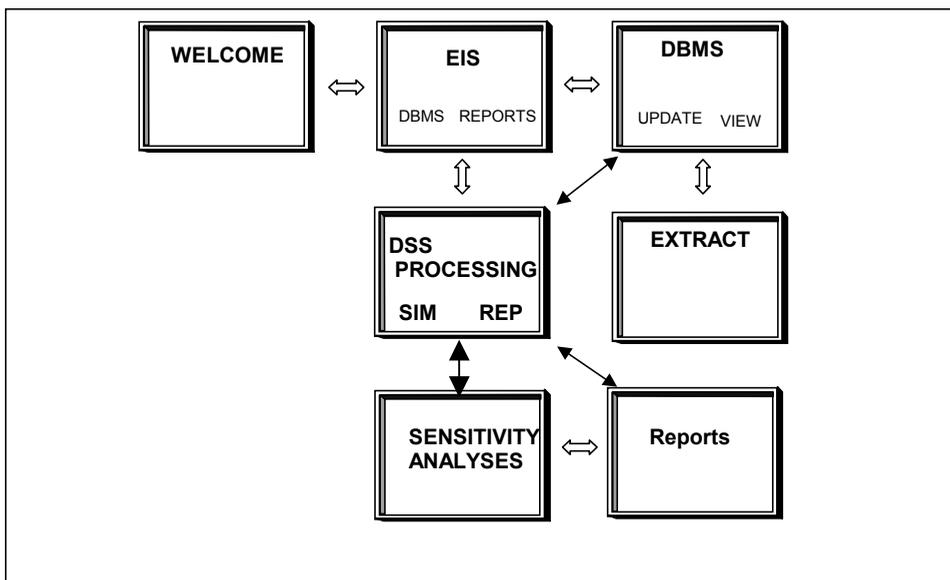
the user to UPDATE the data warehouse and VIEW the contents of the existing or updated warehouse.

In the DBMS, the user can UPDATE the data warehouse and VIEW the contents of the existing or updated warehouse, as shown in Figure 5. Selecting the UPDATE button will place the user in the EXTRACT screen shown in Figure 6. Once there, the user will interactively select the data source for the updating operation from the predefined list. The selection reads data from the specified source, reformats the data (if necessary), and updates the data warehouse values.
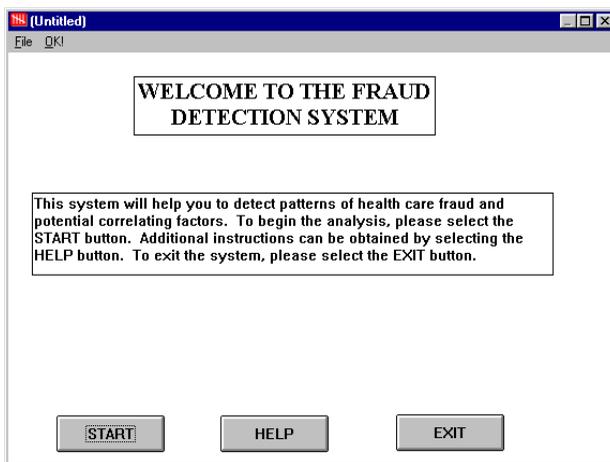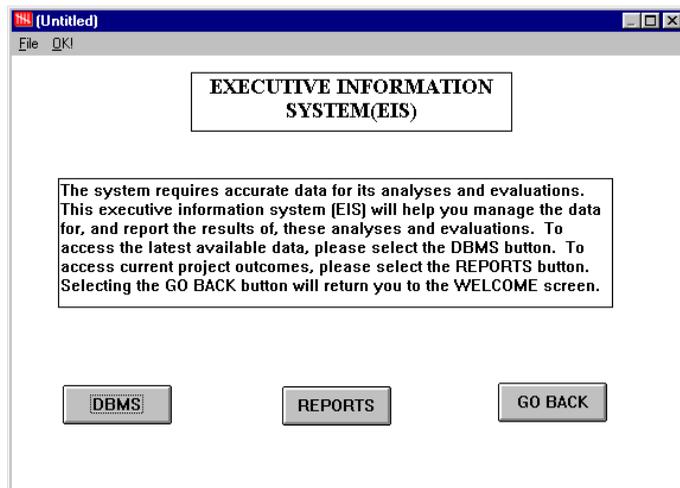
Figure 3: Welcome Screen

Figure 4: EIS Screen

Selecting the VIEW button from the DBMS will access a display that prompts users for the desired information (shown in Figure 7). These selections will form the pertinent Structured Query Language (SQL) call to the data warehouse and generate the desired custom report.

Selecting the REPORTS button from the EIS display will run the fraud detection analysis with the updated or existing data and bring the user to the DSS PROCESSING screen shown in Figure 8. Once there, the user can simulate fraud performance by selecting the simulate (SIM) button. The decision support system will generate the required DSS database from the data warehouse, operationalize the appropriate models, and perform the

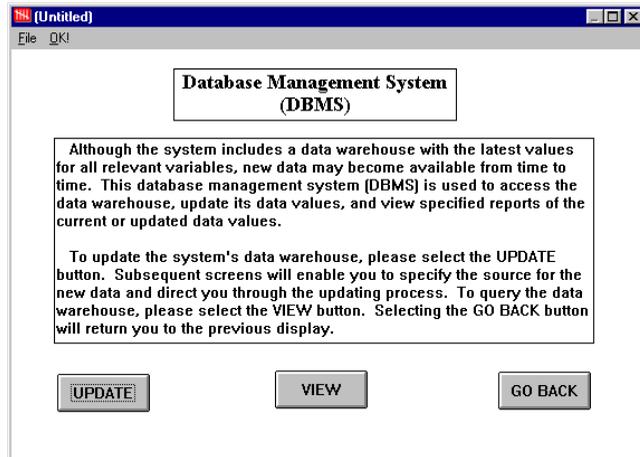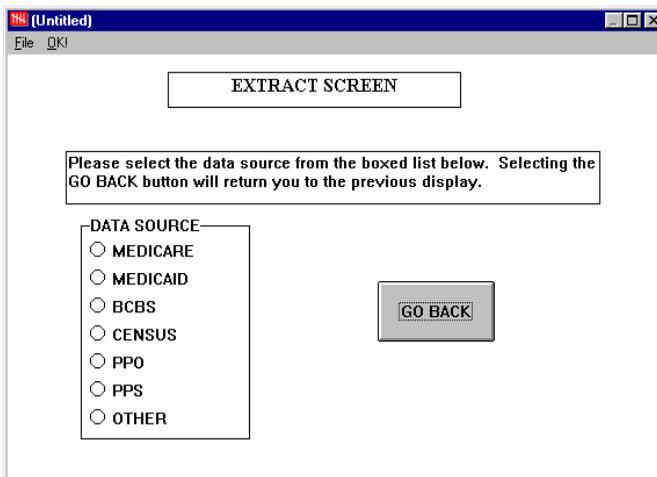*Figure 5: Database Management System Screen*

Figure 6: Extract Screen

needed analyses and evaluations. It is here that data mining techniques will be used to identify new patterns in the updated databases. In deeper level screens, the user will be supported with features that allow the development of classification rules from patterns that hold consistently on data samples. These rules can then be used on incoming transactions to proactively identify fraudulent activities. A report (REP) button selection will display the results in the desired predefined format on the Reports screen (shown in Figure 9).

From the output screen, the user can perform sensitivity analyses on the
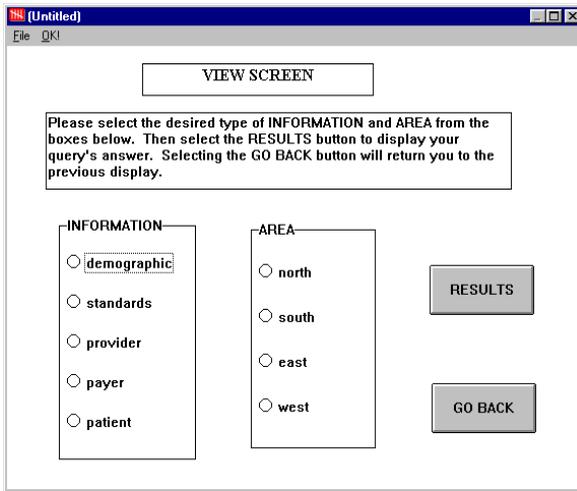
Figure 7: View Screen
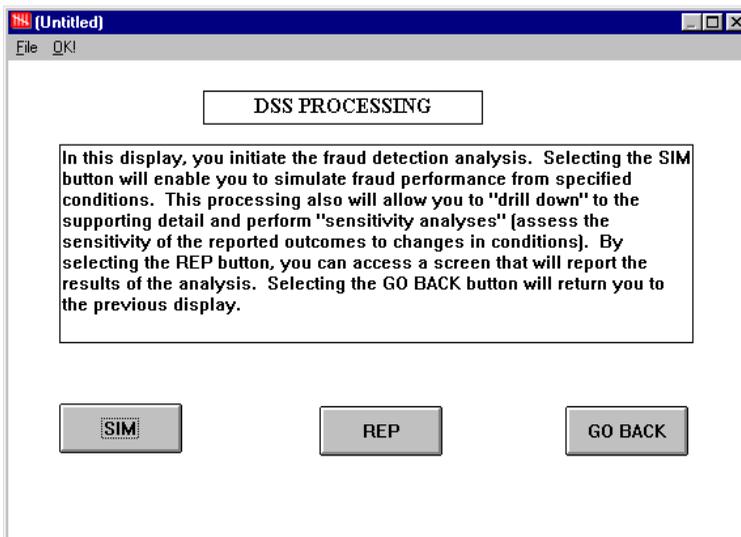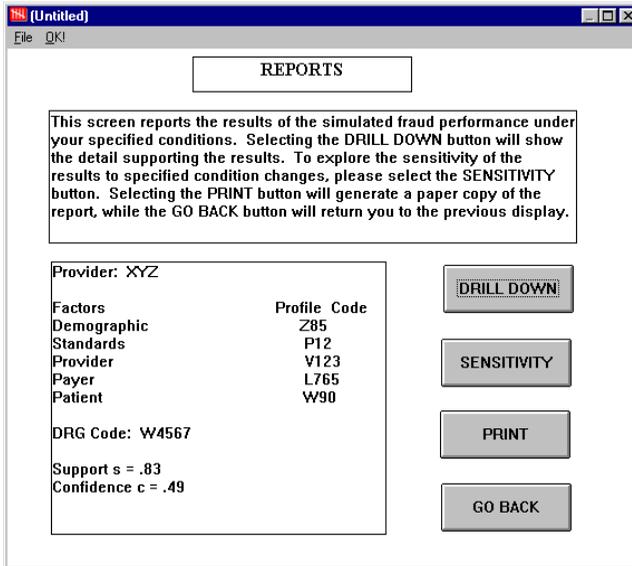

Figure 8: DSS Processing Screen

Figure 9: Reports Screen

results. By making the desired selection from the predefined "What If" list, the user can experiment with changes in: (a) provider characteristics, (b) key local factors, (c) patient demographics, and (d) financing alternatives. Results from the what-if analyses are displayed on the SENSITIVITY ANALYSES screen. Such experimentation can continue in sequence, or the user can generate an entirely new experiment.

Acting as an electronic counselor, the decision support system sequentially guides the user through an effective fraud detection analysis and evaluation. System operations, which are performed in an intuitive, timely (typical five-minute-session) and error-free fashion, liberate the user to focus on the creative aspects of fraud detection and correction.

FUTURE TRENDS

Web-based electronic commerce is an emerging trend that can benefit the healthcare community and the nation in a variety of ways. Such commerce enables the health care organization to be proactive rather than reactive. Transactions can be captured as they are generated, thereby allowing healthcare organizations to compare actual and expected patient outcomes. Such comparisons can help predict, among other things: (1) patient problems, (2) required healthcare interventions, (3) time required for implementation of healthcare

services, (4) accessibility of healthcare services, (5) quality of healthcare services, and (6) cost of healthcare services.

The fraud detection system is conceived as a Web-based technology. A Web site will be established to collect the pertinent data from the various sources. Geographic data would be obtained from state and local government base map files, U.S. Postal Service ZIP Code files, U.S. Geological Survey hydrology data files, and U. S. Bureau of the Census TIGER files. General population characteristics would be obtained by census blocks from the U.S. Bureau of the Census, while health-related demographic data would be acquired from the U.S. Health Care Financing Administration and the National Health and Nutrition Examination Surveys. Health outcome and care data would be obtained from state-specific public health data files. Environmental data would be acquired from state and federal survey data files on water quality, pollutant, toxic waste, ambient air and source emission, air quality, radiation, powerline, and chemical usage and waste generation.

Utilizing the electronic commerce concept, data suppliers will access the system's Web site and select appropriate screen icons (Gerull and Wientzen, 1997). These selections will automatically obtain the data from the supply source and transfer the elements to a data warehouse (Tsvetovatyy, Gini, and Wiecckowski, 1997).

The system's EIS will extract the pertinent data, capture the extractions in user-oriented data marts, and make the marted data available for ad hoc queries by users. Ad hoc queries can be made at the users' sites in an easy-to-use, convenient, and interactive manner, utilizing the Web-based fraud detection system. Results will be displayed in formats anticipated by the requesting parties.

In effect, then, the proposed system provides a vehicle to utilize the emerging Web-based electronic commerce for proactive fraud detection and correction. Without the system, fraud detection involves a very complex process that requires extensive training for provider and payer analysts. By decreasing the volume of documentation, by simplifying the educational process, and by simplifying and automating much of the detection process, the proposed system can be expected to save the medical community millions of dollars per year in fraud detection and correction costs.

With the embedded EIS, an analyst can interactively conduct the initial phase of fraud detection at a computer terminal in a matter of minutes at a nominal expense. Next, the EIS can be used to access pertinent data, mine

patterns from the accessed data, and relate the pattern variables with other correlates. The DSS then can be used to develop an explanatory model and use the model to simulate fraud performance under selected conditions. The fraud detection system can be developed and implemented for a small fraction of the potential cost savings.

From a diagnosis perspective, the manual search for fraud detection patterns is a tedious process that often results in inaccurate, incomplete, and redundant data. Such data problems can leave fraud inadequately detected and corrected. With the proposed system, the user identifies all data relevant to the fraud detection process, and the system provides a mechanism that facilitates data entry while reducing errors and eliminating redundant inputs. Reports from the system also offer focused guidance that can be used to help the user perform fraud searches, detections, and corrections.

Challenges

Realizing the strategic potential will present significant challenges to the traditional healthcare organization. Tasks, events, and processes must be redesigned and reengineered to accommodate the concurrent electronic commerce. Clinicians and administrators must be convinced that the electronic commerce will be personally as well as organizationally beneficial, and they must agree to participate in the effort. Finally, the organizational changes will compel substantial informational technology support.

The organization can have several stand-alone systems to provide the decision analyses and evaluations (Tan, 1995; Tan and Sheps, 1998). Integrating the stand-alone functions, however, can enhance the quality and efficiency of the segmented support, create synergistic effects, and augment decision-making performance and value (Forgionne and Kohli, 1996).

When implemented fully, the innovation will alter the work design for, and supervision of, fraud detection and correction. Requisite operations and computations will be simplified, automated, and made error-free. Training requirements will be reduced to a minimum. Processing efficiency will be dramatically increased. User-inspired creative fraud detection experimentation will be facilitated and nurtured. Management learning will be promoted. Knowledge capture will be expedited.

In short, the fraud detection system's usage would substantially reshape the organizational culture. Faced with significant time pres-

sures and limited staff, healthcare leadership may be reluctant to take on this burden at the present time. In addition, public health officials have developed and cultivated strong and enduring relationships with practitioners and vendors. These practitioners and vendors also have important contacts and allies within the government agencies that oversee healthcare programs. For these reasons, it may be politically wise for public health officials to preserve these practitioner and vendor relationships.

Future Research Opportunities

There are a number of future research opportunities presented by the fraud detection system. To ensure that the information system accurately replicates the inputs, the final version of the system should be tested against Web-collected data from existing institutions. In the testing, warehoused data should be compared against actual values. Statistical tests should be conducted on the estimated models. There should be evaluations of user satisfaction with: (a) the speed, relevance, and quality of ad hoc query results; (b) the system interface; (c) model appropriateness; and (d) the quality of the system explanation. Simulations should be statistically tested for accuracy, and confidence intervals should be established for the results. Tests should also be conducted on the system's ability to improve the decision-making maturity of the user.

Enhancements can be made to the fraud detection architecture. Machine learning techniques can be developed to improve the intelligent modeling, database management, and user interface operations of the system. Communication links can be created to more effectively disseminate system results to affected parties.

The fraud detection system concept can also be adapted for a variety of adjunct healthcare applications. Similar systems can be applied to the diagnosis and treatment of cancer, mental disorders, infectious diseases, and additional illnesses. Effectiveness studies can be done to measure the economic, management, and health impacts of the additional applications.

CONCLUSIONS

The fraud detection system presented in this paper is a combination of data warehousing, data mining, artificial intelligence, and decision support system technology. This system offers the healthcare official a tool that will support a proactive strategy of health care fraud

detection. The system's use can reduce the time and cost needed to detect healthcare fraud, and the system can substantially lower the public and private expenses associated with such fraud.

The fraud detection system delivers the information and knowledge needed to support fraud detection in a comprehensive, integrated, and continuous fashion. The comprehensive, integrated, and continuous support from the system should yield more decision value than the non-synthesized and partial support offered by any single autonomous system. Improvements should be observed in both the outcomes from, and the process of, strategic claims and other electronic commerce decision making (Lederer, Merchandani, and Sims, 1997). Outcome improvements can include advancements in the level of the users' decision-making maturity and gains in organization performance (Whinston, Stahl, and Choi, 1997). Process improvements can involve enhancements in the users' ability to perform the phases and steps of decision making.

To achieve the potential benefits, healthcare officials will have to meet significant challenges. First, a data warehouse must be established to capture the relevant transactions. In particular, there must be continuous user-involvement including careful upfront examination of business requirements and identification of quality and standards. The warehouse must be iteratively developed to deliver increasing value to the organization. Second, to support effective data mining, data marts must be formed to filter and focus the data for fraud detection. A strategy must be formulated for developing the tool. Once again, because of their domain knowledge, users must play a central role in such development. Thirdly, appropriate data mining techniques should be made available to the user and more importantly, validation routines will need to be built into the system to support effective validation of data mining outcomes. Finally, users must be convinced about the efficacy of the fraud detection system and trained in the use of the proactive technology.

Regardless of the proposed system's legacy, the application offers useful lessons for Web-based healthcare decision technology systems' development and management. The system is effectively delivering to the user, in a virtual manner, embedded statistical, medical, and information systems expertise specifically focused on the health care problem. Any single human technical specialist typically will not: (a) be proficient with, or even aware of, all pertinent tools, or (b) possess sufficient domain knowledge to fully understand the medical situa-

tion, propose trials, or interpret outcomes. While practitioners will have the domain knowledge, they usually will not have the technical expertise to effectively develop and implement relevant technology.

The proposed effort suggests that system design, development, and implementation should be a team effort. In addition, the team should be composed of the affected practitioners, information system personnel, and technological specialists proficient with the tools needed to address the healthcare problem.

Fraud detection is inherently a semi-structured (or even ill structured) problem. When initially confronted with such situations, analysts have a partial understanding of the problem elements and relationships. Typically, their understanding evolves as they acquire more information, knowledge, and wisdom about the problem. The fraud detection system is designed to support such decision making.

Relying on the information center, or other traditional information system organization, to design and develop a Web-based fraud detection system will likely be ineffective. These types of organizations typically are staffed by personnel with general skills, limited technological expertise, and restricted problem-specific knowledge. Development and implementation will follow a prescribed pattern designed to provide standard solutions to relatively well-understood and well-structured problems.

A hybrid project-technology organization may work well for Web-based fraud detection system design, development, and implementation in a healthcare environment. The organization would be virtual rather than physical. A project team would be established and administered by the practicing healthcare professional. Team technology specialists would be drawn from within and outside the organization to match the expertise needed for the specific project. Telecommuting and distributed collaborative work would be allowed and possibly encouraged.

REFERENCES

Abraham, T., & Roddick, J. F. (1998). Opportunities for knowledge discovery in spatio-temporal information systems. *Australian Journal of Information Systems*, 5(2), 3-12.

Adam, N. R., Dogramaci, O., Gangopadhyay, A., & Yesha Y. (1998). *Electronic Commerce: Technical, Business and Legal Issues.* New Jersey: Prentice-Hall.

Adam, N. R. & Gangopadhyay, A. (1997). *Database Issues in Geographic Information Systems*. Boston/Dordrecht/London, Kluwer Academic Publishers.

Borok, L. S. (1997). Data mining: Sophisticated forms of managed care modeling through artificial intelligence. *Journal of Health Care Finance*. 23(3), 20-36.

Burn-Thornton, K. E., & Edenbrandt, L. (1998). Myrocardial infarction—Pinpointing the key indicators in the 12-lead ECG using data mining. *Computers and Biomedical Research*. 31(4), 293-303.

Chen, R. (1996). Exploratory analysis as a sequel to suspected increased rate of cancer in a small residential or workplace community. *Statistics in Medicine,*15, 807-816.

CIGNA (1999). CIGNA HealthCare and Insurance Groups Web-site at http://www.insurance.ibm.com/insur/cigna.htm.

Davidson, G. S., Hendickson, B., Johnson, D. K., Meyers, C. E., & Wylie, B. N. (1999). Knowledge mining with VxInsight: Discovery through interaction. *Journal of Intelligent Information Systems: Integrating Artificial Intelligence and Database Technologies*. 11(3), 259-285.

Fischer, M. M. and Nijkamp, P (eds.) (1993). *Geographic Information Systems, Spatial Modeling, and Policy Evaluation*. New York: Springer-Verlag.

Forgionne, G. A. and Kohli, R. (1996). HMSS: A management support system for concurrent hospital decision making. *Decision Support Systems*. 16, 209-223.

Grimson, R. C. and Oden, N. (1996). Disease clusters in structured environments. *Statistics in Medicine* 15, 851-871.

Geographic Information System for the Long Island Breast Cancer Study Project (LIBCSP). National Cancer Institute's Electronic RFP Number NO2-PC-85074-39. Bethesda: National Cancer Institute, 1998.

Gerull, D. B. and Wientzen, R. (1997). Electronic commerce: The future of image delivery. *International Journal of Geographical Information Systems*. 7(7) 38-51.

Heath Care Financing Administration. (1999). Medicare fraud Web-site at http://www.hcfa.gov/medicare/fraud.

Hornung, R. W., Deddens, J. A., & Roscoe, R. J. (1998). Modifiers of lung cancer risk in uranium miners from the Colorado Plateau. *Health Physics*. 74(1), 12-21.

Huxhold, W. E. (1991). *An Introduction to Urban Geographic Information Systems*. Oxford: Oxford University Press.

Kalakota, R. and Whinston, A. B. (1997). *Electronic Commerce: A Manager's Guide.* Reading, Massachusetts: Addison-Wesley.

Keegan, A. J. and Baldwin, B. (1992). EIS: A better way to view hospital trends. *Healthcare Financial Management,* 46(11), 58-64.

Korcok, M. (1997). Medicare, Medicaid fraud: A billion-dollar art form in the US. *Canadian Medical Association Journal*. 156 (8), 1195-1197.

Laden, F., Spiegelman, D., and Neas, L. M. (1997). Geographic variation in breast cancer incidence rates in a cohort of U. S. women. *Journal of the National Cancer Institute* 89, 1373-1378.

Lederer, A. L., Merchandani, D. F., and Sims, K. (1997). The link between information strategy and electronic commerce. *Journal of Organizational Computing and Electronic Commerce*. 7(1), 17-25.

Limb, P.R., and Meggs, G. J. (1995). Data mining -tools and techniques. *British Telecom Technology Journal*. 12(4), 32-41.

Makino, K., Suda, T., Ono, H., & Ibaraki, T. (1999). Data analysis by positive decision trees. *IEICE Transactions on Information and Systems*. E82-D(1), 76-88.

Oden, N., Jacquez, G., and Grimson, R. (1996). Realistic power simulations compare point- and area-based disease cluster tests. *Statistics in Medicine* 15, 783-806.

Ramakrishnan, R. (1997). *Database Management Systems*. Boston: McGraw-Hill.

Regional Variation in Breast Cancer Rates in the U. S. – NIH. National Cancer Institute's Electronic RFA Number CA-98-017. Bethesda: National Cancer Institute, 1998.

Robbins, A. S., Brescianini, S., and Kelsey, J. L. (1997). Regional differences in known risk factors and the higher incidence of breast cancer in San Francisco. *Journal of the National Cancer Institute* 89, 960-965.

Steiner, J. E. (1993). Update: Fraud and abuse Stark laws. *Journal of Health and Hospitals*. 26, 274-275.

Sturgeon, S. R., Schairer, C., and Gail, M. (1995). Geographic variation in mortality rates from breast cancer among white women in the United States. *Journal of the National Cancer Institute* 87, 1846-1853.

Tan, J. K. H. (1995). *Health Management Information Systems*. Gaithersburg, Maryland: Aspen.

Tan, J. K.H., and Sheps, S (eds.)(1998). *Health Decision Support Systems*. Gaithersburg, Maryland: Aspen.

Tomes, J.P. (1993). *Healthcare Fraud, Waste, Abuse, and Safe Harbors: The Complete Legal Guide*. Chicago, Illinois: Probus Publishing Company.

Tsvetovatyy, N., Gini, M., and Wieckowski, Z. (1997). Magma: An agent-based virtual market for electronic commerce. *Applied Artificial Intelligence.* 11(6), 501-509.

Whinston, A. B., Stahl, D. O., and Choi, S. (1997). *The Economics of Electronic Commerce.* Indianapolis, Indiana: Macmillan Technical Publishing.

Workshop on Hormones, Hormone Metabolism, Environment, and Breast Cancer, New Orleans, Louisiana, September 28-29, 1995. *Monographs in Environmental Health Perspectives* supplement 1997, 105(3), 557-688.