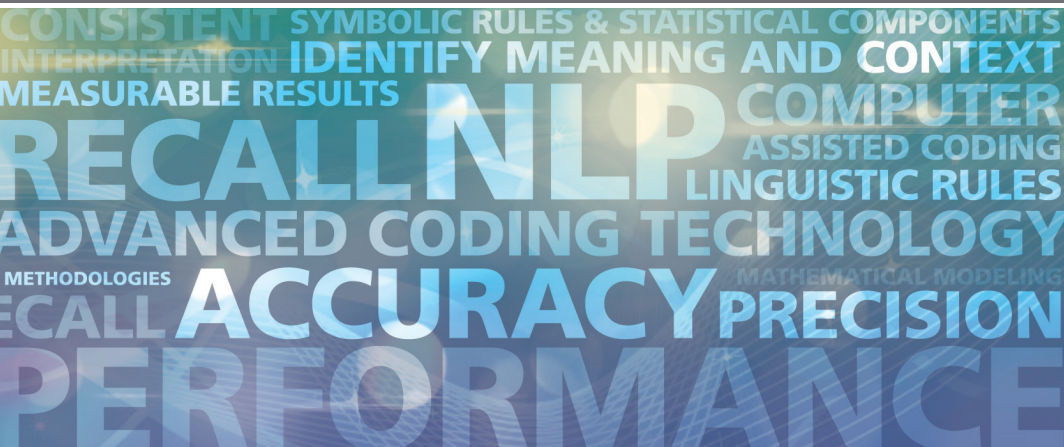# Not all NLP is Created Equal:

CAC Technology Underpinnings that Drive Accuracy, Experience and Overall Revenue Performance

Performance Perspectives

Health care financial leaders and health information management (HIM) professionals have a shared concern: they are looking for the best, most efficient means to safeguard revenue and continue medical coding using the current ICD-9 classification system, while simultaneously preparing for the transition to the ICD-10-CM/PCS code set, ensuring revenue integrity and maintaining coding productivity and accuracy. With the prospect of approximately 155,000 diagnosis and procedure codes available in ICD-10, versus only about 17,000 in ICD-9, these concerns are valid but not insurmountable. Enter computer-assisted coding (CAC) with natural language processing (NLP).

CAC, or the use of computer software that automatically generates a set of medical codes for review, validation and use based upon clinical documentation, has moved beyond the early adopter stage, becoming a valuable tool utilized in hospitals, surgery centers and clinics across the United States. In these health care settings, CAC has enabled improvements in important technical and business performance measures, including improved productivity and accuracy from their coders, a boost in compliance, quicker and more accurate reimbursement, and fewer denied claims. And beneficially, the technology offers traceability of assigned codes to corresponding documentation, which is functionality invaluable when responding to potential audits.

Within CAC applications, the NLP engine provides the enabling technology responsible for automatically reading clinical documentation to identify diagnoses and procedures and then recommend codes to be assigned to clinical cases. Health executives choosing a CAC solution for their organizations need to understand how different methodologies that power NLP engines affect CAC performance. This is critical to maximizing return on investment through realizing immediate, measurable gains in current coding processes, ensuring scalability to ICD-10 and enabling broader use of the technology, for example, in clinical documentation improvement programs and analytics applications.

NLP engines are not all created equal. They are driven by one of five distinct methodologies (see sidebar) for organizing and extracting meaning from clinical documentation. Each dramatically affects performance and accuracy (see the Optum360™ white paper, *Advanced Coding Technology to Advance the Revenue Cycle*).

"Recall" and "precision" (see below for definitions) are industry standards for measuring NLP performance, as calculated by comparing the codes suggested to those selected for billing. NLP engines that deliver higher degrees of recall and precision enable coders to capture all applicable diagnoses and procedure codes on medical records more quickly and accurately, increasing productivity and ensuring correct capture of an organization's case mix index (CMI), a vital determinant of reimbursement rates.

Standard Measures of NLP Accuracy

Precision: Measures the number of accurate results *compared to total results.* Higher rates of precision mean lower false positives.

Recall: Measures the number of accurate results *compared to potential accurate results.* Higher rates of recall mean lower false negatives (or missed codes)

Precision	Recall
Lower false positives	Lower false negatives
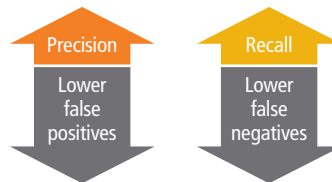
NLP Methodologies

All NLP technologies available today for CAC use one of five methodologies:

1. Medical Dictionary Matching: Words are mapped to medical terminology

2. Pattern Matching: Word patterns describe a diagnosis

3. Statistical: Pre-coded documents train and evolve algorithms

4. Symbolic Rules: Identify language using linguistic rules and symbols

5. **Optum™ LifeCode® NLP**
   Symbolic Rules & Statistical Components: Linguistic rules + mathematical modeling identify meaning and context

Patented LifeCode technology integrates sophisticated linguistic analysis with a massive knowledge base of over 10 million medical facts and leverages deep historical data for consistent interpretation of clinical content.

The Optum Enterprise CAC Platform

Optum™ LifeCode® NLP — the only patented NLP technology on the market today — has been selected by hundreds of health care organizations across the United States, and serves as the engine for all Optum CAC solutions. These include Optum Enterprise CAC, which provides in-depth coding power for both inpatient and outpatient diagnoses and procedures within a hospital setting, as well as Optum CAC Professional, which is optimized for radiology, emergency, pathology and interventional radiology departments and multi-specialty practices.

LifeCode combines the strengths of symbolic rules with statistical components, recognizing the precise clinical detail within medical records, while at the same time lending the flexibility to adapt to variation in syntax and document structure. This sophisticated technology allows LifeCode to identify key clinical facts (including facts that are difficult for coding staff to find in the extended documentation of complex cases) and apply coding guidelines through proprietary rules and algorithms to derive correct coding recommendations. It does this through an integration of linguistic analysis with a knowledge base of more than 10 million medical facts, giving coders the assurance that LifeCode has deep historical data for consistent interpretation of clinical content.

Enterprise CAC presents coders with diagnosis and procedure codes that are more complete and accurate — based on its high degree of recall and precision — transforming their role into an expert auditor and reviewer of coded results. Importantly, NLP engines unable to decipher and understand the high-definition details within clinical documentation in their context will be unable to support the rigorous demands of ICD-10 meaningfully.

While some CAC technologies offer inpatient and outpatient versions powered by completely different types of NLP engines, the Enterprise CAC platform utilizes LifeCode across all settings and for all organizations using it. This ensures that updates instantly apply to all users, reducing IT costs and ensuring compliance and consistency, and eliminates the need for coders to train and work with different CAC technologies.

Understanding NLP Performance

A 2013 study by the American Health Information Management Association (AHIMA) evaluated the ability of CAC technology to support HIM professionals, and improve results over time. In the research, a CAC system, working alone, processed 25 inpatient cases upon implementation. The system then processed the same 25 inpatient cases after six months of use. While preliminary findings presented to the ICD-10-CM/PCS and CAC Summit showed performance of the system improved modestly in recall (from 51 percent to 66 percent) and precision (from 47 percent to 52 percent), the results reveal key inequities between the system tested to more experienced and broadly used applications such as Enterprise CAC with LifeCode NLP.

For example, health care organizations that have implemented Enterprise CAC achieve higher recall (84 percent) and precision (80 percent) accuracy in their very first month of use.


Optum Patents

Vector processing: mathematical model for isolating, comparing and assigning different facts from clinical documentation to build a contextual framework

Mere-parsing: method for determining meaning from free text, including single phrases, within a sentence and from a combination of related phrases. This exclusive capability is essential to supporting ICD-10

Auditing the coding and abstracting of documents: method and design used by Optum CAC for computer-assisted auditing of coded medical documents.

Multi-magnitudinal vectors with resolution based on source vector features: detailed natural language processing method for semantic analysis known as vector processing — a way of comparing two sets of words to determine the similarity of the underlying concepts. This unique invention permits the consideration of multiple dimensions (e.g., identical words, word proximity, word frequency, semantic weight, syntactic role, etc.) between two vectors.

Inpatient Diagnosis Code Accuracy After 1 Month

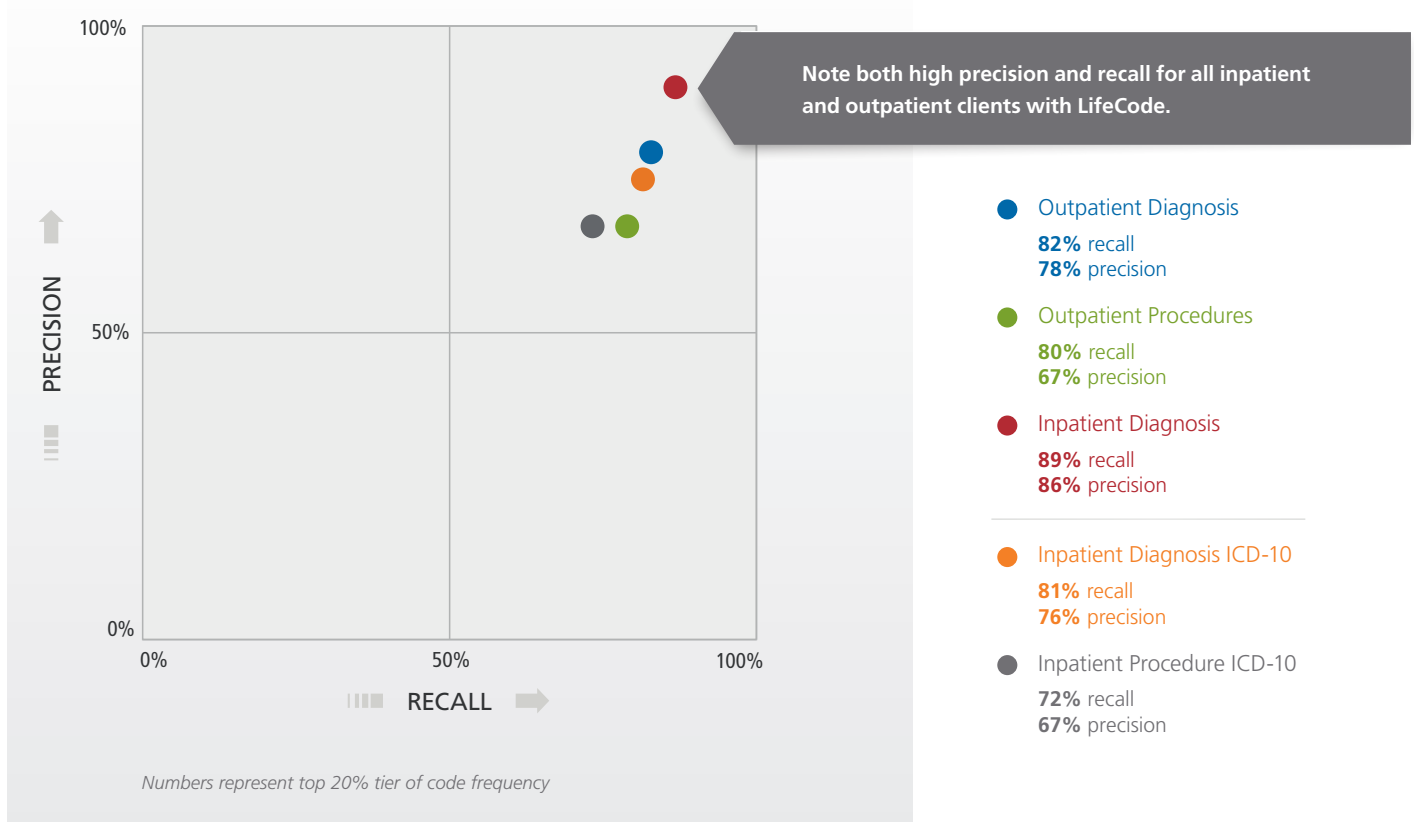| RECALL	PRECISION
84%	**80%**
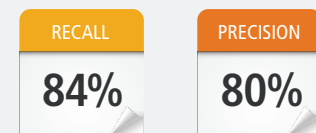
20 clients | over 100,000 codes | Q2 - Q3, 2014

These numbers represent top 60% tier of code frequency

These high rates of accuracy are typical across all Enterprise CAC clients. In a recent analysis comparative study presented to the ICD-10-CM/PCS and CAC Summit, Optum360 examined LifeCode's performance for two months during the third quarter of 2014. This data represents over 150 hospital sites processing both inpatient and outpatient cases, and covers the most frequent codes suggested by LifeCode NLP. For inpatient diagnosis coding, this level represented 650,000 code suggestions, and for outpatient coding, 900,000 code suggestions in both diagnosis and surgical procedure coding. The results for inpatient were 89 percent recall and 86 percent precision. The results for outpatient coding were 82 percent recall and 78 percent precision for diagnoses and 80 percent recall and 67 percent precision for surgical procedures. These outpatient procedure results are for the more complex surgical and interventional procedures, and do not include simple diagnostic and lab procedures, which tend to inflate recall and precision statistics.

Table 1. Inpatient and Outpatient Accuracy

Note both high precision and recall for all inpatient and outpatient clients with LifeCode.

Outpatient Diagnosis
82% recall
78% precision

Outpatient Procedures
80% recall
67% precision

Inpatient Diagnosis
89% recall
86% precision

Inpatient Diagnosis ICD-10
81% recall
76% precision

Inpatient Procedure ICD-10
72% recall
67% precision

Numbers represent top 20% tier of code frequency

The advantage of LifeCode NLP has carried over to ICD-10 results. Measuring results for 3,000 inpatient cases coded by ICD-10 early adopter clients, ICD-10-CM diagnosis coding had 81 percent recall and 76 percent precision. For the more complex ICD-10-PCS procedure codes, results showed 72 percent recall and 67 percent precision. Similar to the ICD-9 results, this was for the most frequent code suggestions.

NLP Experience Matters

The reason for the difference is experience. Since pioneering the computer-assisted coding field in 1999, LifeCode has processed more than 500 million transactions. Since 2008, LifeCode has processed 4.1 million hospital inpatient cases, with a steady rise in number of inpatient cases processed per month to over 300,000 in October 2014 (see Table 2). During that same time period, LifeCode processed 43.2 million outpatient cases, with the most recent number of outpatient documents processed at over 3.5 million per month (see Table 3).
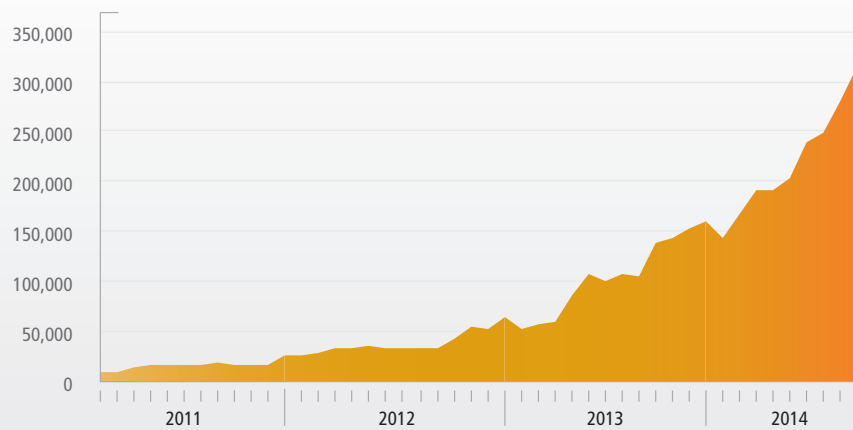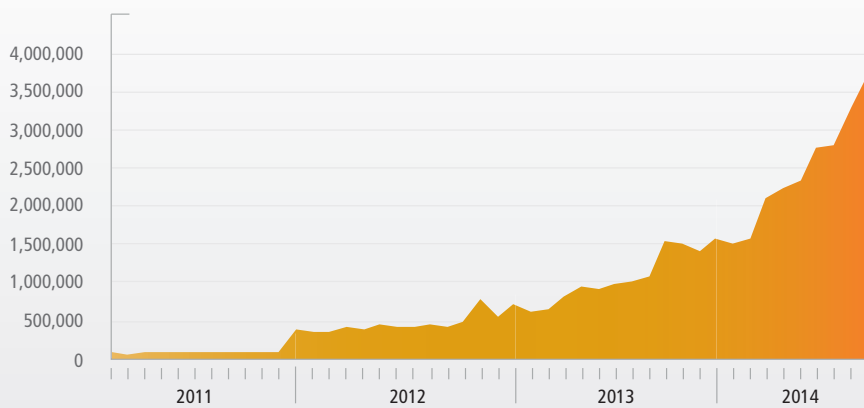
Table 2. # of Inpatient Cases Processed per Month

Table 3. # of Outpatient Cases Processed per Month

Precision and Recall — Key to Measurable Results

These statistics demonstrate how LifeCode's depth of transaction experience drives peak precision and recall accuracy. As a result, the technology, when coupled with the skill of a professional coder, ensures complete, accurate capture and coding of patient care, and dramatically enhances coding productivity and business performance metrics such as discharged-not-final-billed (DNFB) days and more. This translates into true return-on-investment in HIM operations and beyond. For example, LifeCode reads the chart and assigns codes, eliminating time-consuming tasks previously done by coding staff such as personally reading the extensive clinical records and documentation in search of the correct code. This way, the coder can focus on validating codes to help prevent denied claims, a scenario that is estimated by the U.S. Department of Health and Human Services to potentially double from a current average of 3 percent to a level of 6 to 10 percent with the rollout of ICD-10.

With the expansion of the number of codes from which to choose in the ICD-10 environment, NLP's precision becomes even more valuable. With more than 155,000 possible codes in ICD-10, coders will be required to find codes based on highly granular elements (i.e., laterality, severity, acuity, exact body part affected). LifeCode's exclusive, patented mere-parsing capabilities make it uniquely capable to differentiate these features within medical documentation, driving to the highest level of specificity in coding and reducing false positives — or those codes that might be nearly correct, but not as precise as necessary.

CAC and NLP: The Future

Looking beyond the current challenges of coding to ICD-9, use of NLP technology will play a key role supporting ICD-10 and solving provider-to-provider collaboration and outcomes-based reimbursement challenges, facilitated by growing use of electronic medical records (EMRs) and the emergence of new fee-for-value models. Here, selecting the right NLP engine will help health care organizations convert clinical data into other critical health care information including:

- Pairing NLP with health quality data for reporting to the Centers for Medicare & Medicaid Services (CMS) and the Joint Commission
- Using NLP to study clinical documentation against evidence-based medical practices for immediate care improvement
- Providing clinical analytics to identify potential gaps in care and/or clinical documentation

Optum360 has developed applications that use LifeCode technology to automate clinical documentation improvement (CDI) programs. With challenges similar to those faced by coders, including limited time to completely read and comprehend complex medical cases, the NLP technology automatically reviews clinical documentation and identifies likely discrepancies, helping CDI specialists prioritize their work.

LifeCode NLP technology scales easily to the specificity of ICD-10 today, and currently enables dual coding with full logic support. Our solutions are ICD-10 ready, and can help organizations and providers successfully transition to the new code set, supporting both coding and CDI workflow, efficiency and accuracy.

Performance In Practice:
Mission Health

Empowered to perform concurrent coding using Enterprise CAC, Mission Health achieved a 2.4 percent increase in CMI in less than a year. The organization also saw significant improvements in DRG reimbursement and discharged not final coded (DNFC) cases. Read the Mission Health case study on our website.

Additional White Papers

Not all NLP is Created Equal: CAC Technology Underpinnings that Drive Accuracy, Experience and Overall Revenue Performance is part of a series of white papers from Optum360 that explore the impact of advanced computer-assisted coding applications on improving performance for health care organizations. Additional white papers, case studies and videos can be found under Client Results at **optum360.com/EnterpriseCAC.**

Learn how Optum360 is leveraging its NLP technology to automate and support clinical documentation improvement at **optum360.com/CDI3D.**

To learn more about CAC, NLP and Optum360 solutions, call **866-223-4730** or email **optum360@optum.com.** Find out more at optum360.com/EnterpriseCAC.

Authors:

Mark Morsch
Vice President of Technology, Optum360

Chris Martin
Vice President of Client Experience, Optum360

U.S. Patent Nos. 6,915,254; 7,908,552; 8,682,823; 8,731,954; and other Patents Pending

---

1   AHIMA e-HIM™ 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).

2   Liddy, E.D. *"Natural Language Processing."* Encyclopedia of Library Information Science, 2nd Ed. New York City. Marcel Decker, Inc., 2003

3   Department of Health and Human Services. *"HIPAA Administrative Simplification: Modification to Medical Data Code Set Standards to Adopt ICD-10-CM and ICD-10-PCS; Final Rule."* Federal Register 74:11 (16 January 2009) p. 3346. http://edocket.access.gpo.gov/2009/pdf/E9-743.pdf

OPTUM360°™   **optum360.com**