# DATA STANDARDIZATION AND INTEROPERABILITY IN THE DEPARTMENT OF DEFENSE

February 4, 2004

**Executive Summary**

The Department of Defense (DoD) has used a common application, Composite Health Care System (CHCS), throughout all 101 DoD Military Treatment Facilities (MTFs). However, the master files used to encode patient data in CHCS are not identical across MTFs. The encoded data is thus not interoperable from one MTF to another. The DoD selected the 3M Healthcare Data Dictionary (HDD) to accomplish data standardization in the next-generation system, CHCSII. Patient data from all MTFs will be encoded to the same Numerical Concept IDentifiers (NCIDs) from the HDD for storage in the enterprise-wide Clinical Data Repository (CDR).

External standard codes, such as Logical Observation Identifier Names and Codes (LOINC), are mapped to the NCIDs. To support external interoperability, the HDD can translate the DoD data from NCIDs to an external standard code, for communication with organizations outside of the DoD. There are many complex, interrelated issues that must be addressed for the implementation and use of vocabularies. The 3M HDD and the DoD data standardization approach have been designed to overcome these problems.

Practical Issues Related to Using Standard Vocabularies

Significant effort and a thorough understanding are required to implement vocabularies in an Electronic Health Record (EHR). Certain vocabulary practices can put the encoded patient data at risk. One of these is code reuse – the code for a concept is reassigned to a different concept; the code has thus changed its meaning. If the code has been used to encode (identify) the data stored in the patient record, the data will be interpreted incorrectly. Code reuse is a common problem for many standard coding schemes, notably International Classification of Diseases, 9th Edition, Clinical Modification ( ICD9CM), Common Procedural Terminology (CPT) and National Drug Code (NDC). Softwares that provide code sets usually supply just the latest version, with no backward compatibility. In contrast, the 3M HDD insulates the stored patient data from perturbations in the reference coding systems. Concepts in the HDD never change their meanings. External standard codes are mapped to NCIDs in the HDD. Thus, as external codes are reused, the mapping in the HDD will change accordingly to the appropriate NCIDs. Patient data encoded with NCIDs will never be misinterpreted.

A similar problem occurs when standard vocabularies retire or delete codes, and is also handled by the HDD. Often, an external standard code is no longer in use. In the HDD, the code can be "moved" to an inactive context while remaining mapped to the same NCID. The NCID is never deleted, thus encoded patient data will always be interpretable. At other times, a standard vocabulary may remove a code when it is discovered to be a duplicated concept. In the HDD, if two NCIDs are found to be duplicates, one of the pair is inactivated and superceded by the other. Both NCIDs remain in the HDD – linked. For population-based data use, this "superceded by" linkage will allow the query to be structured correctly to retrieve data stored as either NCID.

A standard vocabulary, particularly one that is still growing, may not provide all the codes that correspond to the entire set of data in use right now. In certain scenarios, mandating the use of only the codes from a standard coding scheme may be acceptable (e.g. ICD9CM for reimbursement). In other situations, particularly where clinical care or workflow is concerned, it is important to capture the data accurately. For instance, many laboratory observations will never receive a LOINC code. Some are used for internal system processes, others serve as text placeholders for send-out laboratory results or have attributes that are not compliant with LOINC definitions or rules. Nevertheless, all of them have an NCID in the HDD, so that there is no delay in using the laboratory results at the DoD MTF. With each LOINC release, the HDD is updated by assigning (mapping) the new LOINC codes to their corresponding, existing NCIDs. There is no need to do any update or transformation to the data, encoded with the same NCIDs, in the CDR. Lacking the HDD, if a non-LOINC code is stored temporarily, it will need to be changed to the new LOINC code. This could require significant effort if there is a large amount of data, if the laboratory result is used commonly, or if there is a long interval before the LOINC code is assigned.

A related issue is the coordination of local extensions – codes added at a facility when the appropriate equivalent is not yet found in the standard vocabulary. Local extensions provide concepts needed by the facility for a different granularity or compositional structure, or to support applications or processes. Many of these concepts may never be added to the standards. A common example is the medications compounded at the facility. While local extensions are critical for efficient workflow and data capture at the facility, the lack of coordination is the biggest reason why the DoD has data interoperability issues. The DoD is solving the problem by using the 3M HDD to coordinate the local extensions at all 101 MTFs. The result is a DoD Enterprise Reference Terminology (ERT) that incorporates standard vocabularies and local extensions.

Another related issue is the need for additional translation efforts if data communication is requested in a standard different from that used to encode the stored data. Many domains in health care currently lack a consensus standard vocabulary. When the standard is named or developed, or if a different standard is selected in the future, the current stored codes may need a significant translation effort. In contrast, new or additional standards are mapped to existing NCIDs as part of the HDD's development and update. Encoded with NCIDs, the DoD data can be translated by the HDD to a requested external standard with no additional effort.

Last but not least, historical patient data needs to be preserved. Legacy patient data in CHCS has been encoded with the master file codes at each facility. Replacing the master files with a set of standard codes, without mapping to the legacy codes, would result in the loss of a large amount of valuable patient data to computable clinical or administrative use. CHCS has been in use throughout all DoD MTFs for nearly a decade. The knowledge gained from analyzing the population data can greatly benefit care delivery and outcome, not only to DoD patients, but also to health care at large. This is why the DoD has elected to map all master files codes, historical and current, to the 3M HDD, so that the patient data can be interoperable across all MTFs and across time.

The 3M Mapping Methodology

The 3M mapping processes have been validated and refined through years of operation at commercial and DoD facilities as well as academic and professional scrutiny. Mapping quality is managed through a continuously evaluated quality assurance system. Mapping accuracy (correctness) and consistency (inter-mapper and inter-site consensus) are established through an expert review and quality assurance process and supported by automated, self-learning mapping tools and a Mapping Quality database. Staffing, training, procedures, automated tools and documentation are all in place. Project management and coordination provide for an efficient site implementation, and lessons learned allow for continuous quality improvement. Economy of scale has been used to the DoD's advantage. Ongoing updates from standard terminologies as well as local additions and changes from the sites are managed. Medical Informatics expertise is leveraged for continuous growth and improvement of the HDD.

Additional Advantages Gained by Using the 3M HDD

3M is actively involved in developing standards and is recognized as the leading expert in the implementation and use of standard terminologies. Research and development in lexical services have created a variety of tools and techniques, and 3M continues to explore advances in technology, e.g. in the area of natural language processing.

The 3M HDD includes a comprehensive vocabulary that provides all needed code sets for our customers' health care transactions. The HDD vocabulary sets are built from standard terminologies whenever available. When not, the HDD develops them, using reference sources and 3M Medical Informatics expertise. All system and workflow information from our customers are obtained and incorporated. Customer validation is solicited and customer data is completely tested before go-live. When a standard vocabulary for a domain becomes available, in addition to mapping the standard codes to the existing HDD NCIDs, new concepts and terms from the standard will be added to enrich the HDD.

A rich knowledge base provides domains and relationships to support applications such as picklists, alerts and population queries, and to provide a robust and comprehensive reference model to help in understanding the complex interrelationships of medical knowledge. Because all local DoD extensions have been incorporated, the knowledge base is comprehensive for DoD's needs. the DoD can use the HDD to analyze and understand its own data use and variances, to support guidelines and processes for standardization and improvement. Without the HDD, An organization must start with the vocabulary and hierarchy from the standard terminology. Significant enhancement and customization will be needed.

The HDD also provides a properly structured information model to serve as the template for accurate, comprehensive and meaningful record storage. Domains in the model are provided by the knowledge base, and the vocabulary concepts instantiate the model for data storage. The three technical components, information model, vocabulary and knowledge base, have been implemented in the 3M HDD software product, to work with other applications such as the CDR, Clinical Workstation, decision support, and interfaces. The HDD has been live in real world use at multiple health care facilities over the past few years. Operational implementation, maintenance, and support are in place. In contrast, standard vocabularies are terminologies, not applications, and will require significant work to be incorporated into software systems. The HDD is designed to be flexible and extensible, and to stay current with evolving standards. The HDD can subsume the other ERTs and ensure their interoperability with DoD's ERT.

Data Sharing

The DoD and the Veteran's Administration (VA) need to exchange and share clinical data to provide services seamlessly to both DoD and VA patients. We have been told that an experiment has found no match between the items in the VA and DoD databases for a test sample of 100 NCIDs. Not having been given any additional information, we are unable to remark on the experiment design, but would offer some general comments.

Most reference terminologies are large, in order to provide concepts of sufficient granularity – one concept may differ from another only in a particular detail or attribute. Site-to-site subset variation is common, and a lack of overlap is particularly likely if two subsets are disparate in size, or if they differ in the granularity or compositional structure of concepts. Comparing item to item for exact matching only is one-dimensional and does not make use of the vocabulary and knowledge base information for data interoperability. To provide insight, we compared the race concepts used by the VA and the DoD. We also examined and compared the Primary Drug master files between two of the VA's facilities, and to the DoD metadata in the HDD. The details are shown in the paper, but a summary is provided here.

For the race domain, the DoD uses 44 concepts, while the VA uses 16 concepts. There are 11 concepts in common. If the VA sends a race code to the DoD, for 69% of the VA concepts there will be a direct match to a DoD race. For the remainder, the HDD knowledge base can be used to create relationships to support the DoD's business logic for interoperability. For instance, when the VA sends the code for "Asian", which is not used by the DoD, the DoD might decide that it will serve the same function as the DoD concept "Asian or Pacific Islander". Because the VA race data set is less granular, the majority (75%) of DoD race codes do not have a direct correspondence. The HDD can also be used by the VA to achieve data interoperability. For instance, when the DoD sends the codes for "American Indian/Nat Alaskan, Hispanic" or for "American Indian/Nat Alaskan, Non-Hispanic", which are not used by the VA, the HDD knowledge base could be used to relate either of those codes to "American Indian or Alaskan," the race concept used by the VA. This kind of scenario will be seen repeatedly in medical data, and is one the HDD is built to manage.

The Primary Drug files from VA's Martinsburg and St. Louis facilities were reviewed for overlap. Of the 1980 NCIDs mapped for Martinsburg and 2169 NCIDs mapped for St. Louis, 570 are in common (29% for Martinsburg, 26% for St. Louis). The low overlap is primarily due to a mixture of different compositional structures in the files – ingredient (single or multiple) only, or ingredient with any combination of strength, form and route. For instance, St. Louis has "Acetaminophen" and "Acetaminophen 650mg" whereas Martinsburg has "Acetaminophen Liquid" and "Acetaminophen 500mg Tablet". They both have "Acetaminophen/Oxycodone". However, when HDD ingredient relationships are used to retrieve the ingredient, comparing ingredients increased the match to about two-thirds between the two sites. The remainder comprised largely of non-drug (e.g. "Lotion") or workflow processing items (e.g. "Do Not Use"), and site variation. This case study again illustrates the power of the HDD. Because all DoD formulary items have appropriate ingredient relationships in the HDD, there is a complete overlap for ingredients with the VA.

Conclusion

The problems surrounding data interpretation, interoperability, and interrelationships are not new. The 3M HDD product and implementation strategies for mapping and maintenance are designed to meet our customers' needs. The data standardization approach pioneered by the 3M HDD has been validated by successful operational use at commercial and DoD facilities. 3M is committed to help the DoD achieve its mission as a health care provider and as a national leader in the adoption of information standards for the enhancement of patient safety and quality of care.

**What is Data Interoperability?**

The Department of Defense (DoD) has invested significant effort in its Electronic Health Record (EHR). A common application, Composite Health Care System (CHCS), has been in use throughout all DoD Military Treatment Facilities (MTFs) for nearly a decade. Each MTF has a set of master files used to encode patient data in CHCS. The content of each master file is not identical across facilities (e.g. "1" may mean "Acetaminophen" at an MTF but "Aspirin" at another). The data is thus not encoded the same way from one DoD MTF to another.

The DoD requires data interoperability within its organization. It also needs to exchange clinical data with the Veterans Administration (VA) in order to provide services in a seamless fashion to both DoD and VA patients. Interoperability means that the data encoded at one site is interpretable at another site, as if the data were encoded there. Interoperability allows data to be used in applications regardless of origin, and to be aggregated and compared across location and time. One means to achieve interoperability is data standardization.



What is Data Standardization?

Data standardization refers to the use of the same set of codes to encode data throughout a system. As an example, for the domain of "sex", one may decide always to code the sex of male as "1", female as "2", and unknown as "3". The domain of "sex", consisting of three members, "male", "female" and "unknown", forms a vocabulary, albeit a very simple one. If all data about sex is coded consistently according to this vocabulary, the data should always be understandable and usable for analysis across time and location.

The execution of the data standardization principle is less simple when the domain is large and complex (e.g. medications), and when the system in question is large (e.g. the DoD). Management, business and even political issues can complicate the implementation. Also, while data standardization may be achieved within an organization, the data may still not be interpretable by external systems. To solve this problem, standard vocabulary codes are promoted for use in data exchange among organizations.

What is a Standard Vocabulary?

A standard vocabulary or coding scheme is one that has wide industry acceptance or use. Standards are obtained from a variety of efforts. The federal government has purchased a national license for the Systematized NOmenclature of MEDicine Clinical Terminology (SNOMED CT). The National Library of Medicine (NLM) maintains the Unified Medical Language System (UMLS), with the latest focus being RxNorm, a reference terminology for clinical drugs. Standards are developed by consensus industry effort, such as version 3 of Health Level 7 (HL7). The set of standards to be used by government agencies will be named by the Consolidated Health Informatics (CHI) initiative – the first vocabulary standard selected is Logical Observation Name Identifiers and Codes (LOINC). Examples of other vocabularies that are considered standards for billing are the International Classification of Diseases, 9th Edition, Clinical Modification ( ICD9CM) and the Common Procedural Terminology (CPT). The National Drug Code (NDC) may also be considered a standard for use within the pharmacy industry.

**The DoD Approach to Data Interoperability**

Four years ago, the DoD began the transition to the next-generation system, CHCSII. The DoD selected the 3M Healthcare Data Dictionary (HDD) to accomplish data interoperability in CHCSII and with the health care industry at large. The content of the HDD is built with industry-standard vocabularies. Concepts in the HDD are each identified by a Numerical Concept IDentifier (NCID). NCIDs are used for enterprise-wide encoding of clinical data for storage in the CHCSII Clinical Data Repository (CDR).

To achieve data standardization within the DoD, the CHCS master files are mapped to the HDD. When content is loaded into a CHCS master file (e.g. a list of insurance companies), a unique Internal Entry Number (IEN) is automatically generated as the item identifier (primary key). The identical item from the same file will receive different IENs in different MTFs, unless the same IEN is assigned by a rare coincidence. For interoperability, the different IENs for the same item need to be mapped to a single standardizing concept. To date, a total of 3 million CHCS items from the five domain areas of Demographics/Encounters, Laboratory, Microbiology, Pharmacy and Radiology/Text Reports have been mapped for all 101 MTFs. Because NCIDs, instead of IENs, are used to encode CHCS data in the CHCSII CDR, previously isolated islands of legacy data can now be interoperable across all MTFs, and across time.

Standard vocabularies are incorporated in the HDD. External standard codes, such as LOINC, are mapped to NCIDs. Through mapping, the HDD can translate between one standard and another, between legacy systems, and between a legacy system and a standard. To exchange data with external systems, such as the VA, the HDD can translate the DoD data from NCIDs to the requested external standard code (e.g. LOINC).

**Practical Issues Related to Using Standard Vocabularies**

There are many complex technical and management issues related to data standardization. It is unrealistic to assume that a vocabulary can be operationalized in an EHR without significant effort. Some vocabulary practices (e.g. code reuse) can put the encoded patient data at risk. The 3M HDD was developed to bear the burden of vocabulary implementation and to insulate the health care enterprise from these risks.

Shift in Meaning of a Standard Code

If the code from an external standard vocabulary is used to encode (identify) the data stored in the patient record, then, if the meaning of the code changes over time, the data will be interpreted incorrectly. For instance, NDC 00074433501 was given to the drug product "Liposyn (Fat Emulsions), 10%, Intravenous Solution, Intravenous, Abbott Hospital, 200ml Bag" until July 2002, when it was reassigned to "Paclitaxel (Paclitaxel, Semi-Synthetic), 6mg/ml, Vial, Injection, Abbott Hospital, 5ml Vial". If the NDC is stored as the patient's medication data, then after July 2002, the patient will mistakenly be thought to have been given Paclitaxel. This misinterpretation of the data has obvious negative effect on population based data use. If the patient is still under current treatment, the data error can potentially affect clinical care.

| NDC Code | Date | Drug Product |
|---|---|---|
| 00074433501 | Before July 2002 | Liposyn (Fat Emulsions), 10%, Intravenous Solution, Intravenous, Abbott Hospital, 200ml Bag |
| 00074433501 | After July 2002 | Paclitaxel (Paclitaxel, Semi-Synthetic), 6mg/ml, Vial, Injection, Abbott Hospital, 5ml Vial |

**Table: Code Reuse – The same code is used to identify two different concepts. Relying exclusively on standard code sets can lead to interpretation errors.**

Code reuse is a common problem for many standard coding schemes, notably ICD9CM and CPT, in addition to NDC. Softwares that provide code sets usually provide just the latest version, with no backward compatibility. Often, a vocabulary makes what it considers an "adjustment", e.g. when LOINC changes the specimen attribute of its laboratory results from Serum and Plasma, separately, to Serum/Plasma. However, the resultant implications to the data are often non-trivial.

The 3M HDD insulates the stored patient data from perturbations in the reference coding systems. Concepts created in the HDD never change their meanings, thus, patient data encoded with NCIDs will never be misinterpreted. Since external standard codes are mapped to NCIDs, as external codes are reused, the mapping in the HDD will change accordingly. In the above example, the first drug product has NCID 3000493238 (Liposyn). The second drug product has NCID 3000536480 (Paclitaxel). The NDC of 00074433501 was mapped to NCID 3000493238 (Liposyn) in the Active NDC Context until July 2002, when it was moved out of the Active NDC Context into the Inactive NDC Context. At the same time, the NDC of 00074433501 was added (mapped) to NCID 3000536480 (Paclitaxel) in the Active NDC Context. The patient data is stored in the CDR as NCIDs, not NDCs; so the information is always correctly stored and interpreted as Liposyn or Paclitaxel regardless of the NDC reuse. In addition, through the HDD one can trace the "path" of the NDC and find out what a drug product's NDC is at a point in time. This is particularly important for accurate data interoperability and exchange. Otherwise, two communicating institutions might falsely assume the data to be the same because the code is the same – as in the above example – when they are not, because of code reuse.

| Date | NCID | Drug Product | Active NDC Context | Inactive NDC Context |
|---|---|---|---|---|
| Before July 2002 | 3000493238 | Liposyn (Fat Emulsions), 10%, Intravenous Solution, Intravenous, Abbott Hospital, 200ml Bag | 00074433501 | |
| After July 2002 | 3000493238 | Liposyn (Fat Emulsions), 10%, Intravenous Solution, Intravenous, Abbott Hospital, 200ml Bag | | 00074433501 |
| After July 2002 | 3000536480 | Paclitaxel (Paclitaxel, Semi-Synthetic), 6mg/ml, Vial, Injection, Abbott Hospital, 5ml Vial | 00074433501 | |

**Table: The HDD accounts for shifts in the meaning within standard code sets. Different NCIDs are given to distinct concepts, preserving their unique meaning despite having the same code.**



Removal of Standard Codes

Frequently, standard vocabularies may retire or delete codes. If patient data is stored using the removed code, it will no longer be interpretable. For instance, CPT code 0002T, "endovascular repair of infrarenal abdominal aortic aneurysm or dissection; aorto-uni-iliac or aorto-unifemoral prosthesis", was effective January 1, 2002, and was terminated from use after December 31, 2003. Versions of CPT after this date will not contain code 0002T. If previous versions of CPT are not maintained in a master file or data dictionary, patient data stored as code 0002T prior to January 1, 2004 will not be interpretable.

In the HDD, an NCID, once created, is never deleted. In the above example, the CPT code of 0002T is mapped to NCID 14780136, in the CPT Code Context. When the code of 0002T is no longer needed, while remaining mapped to NCID 14780136, it will be "moved" to the Inactive CPT Code Context. Because it is the NCID that is used to encode data and not the CPT code, the patient data will never be lost to use.

Another instance of code deletion occurs when redundancy is discovered in a coding scheme. Since a standard vocabulary serves as a reference, one of a pair of duplicates is usually removed from use. In the HDD, if two NCIDs are found to be duplicates of one another, then one of the pair is inactivated and superceded by the other. Both NCIDs remain in the HDD – linked – as shown in the following table. For population-based data use, such as queries, the "superceded by" mechanism of the HDD will allow the query to be structured correctly to retrieve data stored as either NCID.

| NCID | CID (Concept IDentifier) | Status | SupercededBy |
|---|---|---|---|
| 154309 | StrepPneumoniae | Active | |
| 11547 | Pneumococcus | Inactive | 154309 |

**Table: The HDD inactivates and supercedes duplicates without deletion of NCIDs.**

Lack of Comprehensive Standard Codes

A standard vocabulary may not provide all the codes that correspond to the entire set of data in use right now. This is particularly true of standard vocabularies that are built via voluntary submission from participating organizations over time (e.g. LOINC). In certain use cases, mandating the use of only the codes from a standard coding scheme may be acceptable (e.g. ICD9CM for reimbursement). In other situations, particularly where clinical care or workflow is concerned, it is important to capture the data accurately according to what really occurred.

The LOINC database for laboratory results was started with the master files from seven US laboratories. It was first released in April 1995 with approximately 6,500 codes, and has since grown through submission from laboratories, hospitals, and other organizations, notably 3M. The latest release in October 2003 contains approximately 20,000 laboratory codes and roughly 14,000 clinical observation codes. LOINC is released periodically. In 2003 it was released in May and October; in 2002, January, February, August and September; in 2001, January and July; in 2000, February and June. If LOINC codes are to be used to code data directly to be stored in the CDR, is the clinician restricted to ordering only those laboratory tests that currently have associated LOINC codes? This could have a significant impact upon clinical practice and workflow, or worse, lead to either imprecise information or data gaps in the CDR.

This is why the LOINC committee recommends that LOINC codes should be recorded "as attributes of existing test/observation master files" for use in the appropriate message segments to communicate among systems. Many laboratory observations will never receive a LOINC code. Some are used for internal system processes (e.g. "DoD DNA Samples" – NCID 14614770). Others serve as text placeholders for send-out laboratory results (e.g. "Cystic Fibrosis DNA Test" – NCID 26133). Yet others have attributes that are not compliant with LOINC definitions or rules (e.g. "ABO Group, Serum or Plasma, Qualitative" – NCID 44979). All three examples are characteristic of LOINC use by the DoD. The last one, for instance, is used by over a third of the MTFs. The situation where laboratory observations do not have a LOINC code is not unique to the DoD. The latter two NCIDs above are used by commercial health care organizations as well.

As a result of mapping laboratory results for 16 health care institutions and 101 DoD MTFs, the 3M HDD has developed a comprehensive domain of over 42,000 laboratory results. The size of the laboratory file from the DoD MTFs ranged from about 1,000 to over 18,000 rows of data, averaging about 5,000 rows. These approximately 500,000 rows from all 101 MTFs have been mapped to just fewer than 20,000 NCIDs. Over half of these DoD laboratory results do not currently have a LOINC code. 3M is preparing them for submission to LOINC. With each LOINC release the HDD is updated by assigning (mapping) the new LOINC codes to their corresponding, existing NCIDs. There is, therefore, no delay in using the laboratory result at the DoD MTF or any interruption to workflow. Since it is the same NCID that is stored whether there is a LOINC code or not, there is no need to do any update or transformation to the data in the CDR.

Lack of Coordination for Local Extensions

Local extensions are codes added at a health care facility to encode data when the appropriate, equivalent code is not yet found in the standard vocabulary. Many of these may never be added to standard vocabularies. Local extensions provide concepts needed by the facility for a different granularity or compositional structure. For example, a standard for race may provide "Asian/Pacific Islander", whereas a facility may wish to differentiate between them. Local extensions are common for medications compounded at the facility. There may also be local codes created to support applications or processes. Local extensions are critical for efficient workflow and data capture at the facility. However, it is the biggest reason why the DoD has data interoperability problems in CHCS.

Local extensions should be coordinated within an enterprise for two reasons. The first is to support clinical functions by capturing complete, accurate and appropriately detailed data, encoded with enterprise-wide codes. The data is thus standardized within the enterprise, even though the code is not found in the external standard vocabulary. Second, coordinating local extensions within an enterprise supports administrative or workflow functions. System-wide processes can then be implemented for efficiency and cost savings. Decision support and population queries can be applied to all enterprise data.

The 3M HDD is used to coordinate the local extensions at all DoD MTFs. For instance, the domain of Pharmacy Items (NCID 1202118) contains all medications – generics, trade name products, and locally compounded drugs. Applications and processes access the same domain. The necessary updates are maintained by the HDD. The result is a DoD ERT that incorporates standard vocabularies and local extensions. To restrict the organization to standard codes only will not serve the needs of the facilities.

Additional Effort of Translation

Many domains in health care currently lack a consensus standard vocabulary. The CHI initiative has named only one vocabulary standard to date, LOINC, for laboratory result names. Other domains in a laboratory result message also require interoperable data. Examples include specimen, units, and coded result values (e.g. "Staphylococcus aureus" or "Resistant"). Effort remains to name other standards, such as SNOMED CT, for use in particular domains. For many domains, the vocabulary set will require development. Therefore, there is a risk of needing to translate from a stored set of codes to another standard, which would require a considerable effort. In contrast, encoding data to HDD NCIDs enables the DoD to stay abreast effortlessly with the evolution in vocabulary standards. New or additional standards will be mapped to the same set of NCIDs as part of the HDD's development and update, thereby ensuring data interoperability.

Loss of Historical Patient Data

CHCS has been in use throughout all DoD MTFs for nearly a decade. The DoD thus has a large collection of very valuable patient data. The knowledge gained from analyzing the population data can greatly benefit care delivery and outcome, not only to DoD patients, but also to health care at large.

The legacy patient data in CHCS has been encoded with the master file codes at each MTF. Replacing the master files at each MTF with a set of standard codes, without mapping to the legacy codes, would result in the loss of this historical patient data to computable clinical or administrative use. Essentially, the system will behave as if patient data collection is only starting now. Paper records or text printouts will be needed for past medical data (e.g. for a follow up visit). This has very serious implications for patient care and population management for the DoD, which has worldwide MTFs and deployment. It is critical to preserve the historical patient data in CHCS for continuity of patient care, quality of care delivery, population health management and outcomes research.

For these reasons, the DoD has elected to map all master files codes to the 3M HDD. Historical codes as well as currently active master file codes have been mapped, so that legacy patient data can be interoperable across all MTFs and across time.

**The 3M Mapping Methodology**

The 3M mapping processes have been validated and refined through years of operation at commercial and DoD facilities as well as academic and professional scrutiny. The key features of the 3M mapping provided to the DoD are highlighted below.

Economy of Scale

Centralized management of mapping provides an economy of scale for the DoD. If each of the DoD facilities were to perform its own mapping for the master files, it would be faced with a significant effort burden. The labor is repeated at each facility, with no access to the lessons learned or experience gained from the work done at previous sites.

Mapping Expertise and Experience

The 3M HDD mapping processes have been validated and refined through years of operation and scrutiny via publications and presentations in professional and academic arenas. Staffing, training, procedures, automated tools and documentation are all in place. Project management and coordination provide for an efficient implementation of the HDD at each site. Lessons learned allow for continuous quality improvement.

Mapping Maintenance and Site Synchronization

The 3M HDD is an implemented, live product at commercial and DoD facilities. Processes and automated tools are in place for updates and sustainment. Ongoing updates from standard terminologies as well as local additions and changes from the sites are managed.

Mapping Accuracy and Consistency

Accuracy refers to the correctness of the mapping. Expert review and quality assurance are part of the mapping procedure. The 3M HDD mapping methodology ensures accuracy through integrity rules. Factors that would affect a mapping judgment are defined and documented in operating procedures. Examples of the factors are granularity, composition, and specificity attributes (e.g. place or person name associated with an item). The automated mapping tools also incorporate the factors and rules. In addition, a self-learning function is included in the mapping tools. An expert mapper will review those items that the tool was not able to map automatically. The mapping decisions and results are fed into the self-learning function. The tool thus improves its automated mapping with every run. A Mapping Quality Assurance database further ensures mapping accuracy.

Mapping consistency refers to the inter-mapper and inter-site consensus/agreement of the mapping across all facilities. Inter-mapper consensus means that, independently, different mappers will map the same item to a single concept. Inter-site consensus means that the same item from different facilities, being worked on at different times, is mapped to a single concept. It is important to make sure that the data variation from site to site is due to local extensions and not inconsistent mapping. The 3M HDD and its mapping methodology were developed to manage this issue. Mapping results for the facilities are aggregated and compared in an expert review. This centralized mapping management for the entire organization is critical.

**What Additional Advantages are gained by Using the 3M HDD?**

Comprehensive and Applicable Vocabulary

Many domains do not currently have an industry-accepted standard vocabulary. Examples include units, dosing frequency, and medication administration instructions. In recognition of this problem, version 3 of HL7 has focused on terminology development. The needed vocabularies are being created but are far from complete. 3M actively participates in these terminology development efforts by contributing both vocabulary content and expertise. Contributed areas include routes of administration, units and the dental lexicon.

To support the health care transactions of our customers, the HDD provides vocabulary sets for all required domains. The HDD vocabularies are built from standards whenever available. When not, the HDD develops them, using reference sources and 3M Medical Informatics expertise. 3M participates actively in standards activities to ensure that the HDD stays on the leading edge. Most importantly, all system and workflow information from our customers are obtained and incorporated. Customer validation is actively solicited. Customer data is completely tested before go-live. The HDD ensures that all concepts and terms needed by our customers are provided.

When a standard vocabulary for a domain becomes available, the standard codes will be mapped to the existing HDD NCIDs. There is no interruption to customer workflow and no disturbance to the stored data encoded with NCIDs. New concepts and terms from the standard will be added as well, enriching the HDD. Thus, the HDD provides a comprehensive, standard-compliant vocabulary that is ensured to apply fully to our customers' data.

Rich Knowledge Base

The 3M HDD includes a knowledge base that contains a rich semantic network of multiple hierarchies and relationships. The knowledge base provides an enterprise reference model that supports the DoD in understanding the complex interrelationships of medical knowledge. Applications and processes use the domains and relationships in the knowledge base for a wide variety of clinical and administrative functions. Domains are built to provide picklists for care documentation (e.g. problem list) or care delivery (e.g. laboratory test ordering). Drug class and ingredient hierarchy is used for medication allergy or interaction alerts. Component relationships link orders and results. Other relationships can be used to provide any required information, such as cause and effect, suggested treatment or medication, or cost. Population queries are made efficient and maintenance-free by using the domains of the HDD to access the member NCIDs in the patient database.

Because the master codes from all DoD MTFs have been mapped into the HDD for the domain areas described earlier, the knowledge base is comprehensive for DoD's needs. For example, a penicillin allergy would be triggered for any DoD compounded medication that contains one of the penicillins. This is because the DoD local medication has been added to the drug class and ingredient hierarchy in the HDD knowledge base. Without an HDD, an organization must start with the hierarchy from the standard vocabulary. Many coding schemes do not provide a rich set of explicit relationships in computable form. Our experience indicates that significant enhancements and customizations will be needed. Otherwise, the trade-off would be restriction of use, or incompleteness or inaccuracy in the use of data.

The flexibility of the knowledge base allows the HDD to create relationships to enhance current knowledge. For instance, an application may need the drug class of Sulfonylurea subdivided into first and second generations. The HDD can create the new subdomains and relationships to be inserted into the present drug hierarchy. Similarly, an application may need to access all variations of mammography reports. A domain can be created to group them. While a standard vocabulary may provide a "starter set" of relationships for its area of application, the HDD can enhance its usefulness by creating relationships "across" domain areas. An example would be the antibiotic susceptibility results for microbiology cultures from LOINC. The HDD has added a "Uses Drug" relationship from the laboratory result NCID to the medication ingredient NCID from the drug hierarchy. The relationship can be used to suggest the appropriate antibiotic for the patient in response to a positive susceptibility result.

| Lab Result NCID | Lab Result | LOINC Code | Drug NCID | Antibiotic |
|---|---|---|---|---|
| 4026 | Ampicillin, Susceptibility, Point in Time, Isolate, Quantitative, Minimum Bactericidal Concentration Measured | 27-3 | 3000253893 | Ampicillin |
| 4027 | Ampicillin, Susceptibility, Point in Time, Isolate, Semi-Quantitative or Quantitative, Minimum Inhibitory Concentration | 28-1 | 3000253893 | Ampicillin |
| 4028 | Ampicillin, Susceptibility, Point in Time, Isolate, Semiquantitative, Agar Diffusion (Bacterial Sensitivity (Kirby-Bauer)) | 29-9 | 3000253893 | Ampicillin |
| 4522 | Vancomycin, Susceptibility, Point in Time, Isolate, Quantitative, Minimum Bactericidal Concentration Measured | 523-1 | 3000253987 | Vancomycin |
| 4523 | Vancomycin, Susceptibility, Point in Time, Isolate, Quantitative, Minimum Inhibitory Concentration | 524-9 | 3000253987 | Vancomycin |
| 4524 | Vancomycin, Susceptibility, Point in Time, Isolate, Semiquantitative, Agar Diffusion(Bacterial Sensitivity(Kirby-Bauer)) | 525-6 | 3000253987 | Vancomycin |

**Table: Examples of the "Uses Drug" Relationship Between a Microbiology Susceptibility Result and the Antibiotic Tested**

Properly Structured Information Model

An information model functions as a template to store data about an episode or instance of care. It represents the knowledge about a particular topic that is used to form an accurate, comprehensive and meaningful record. For instance, an allergy information model would store all relevant and appropriate data regarding the patient's allergies and adverse reactions. The structure would capture the information correctly, keeping in mind the use for the data, such as alerts and population studies.

The 3M HDD provides an integrated information model for the CDR. The information model is tightly bound to the HDD knowledge base and vocabulary. Domains in the model are provided by the knowledge base, and the vocabulary concepts instantiate the model for data storage. The information model ensures the consistent, accurate and complete storage of DoD patient data in the CDR. It provides a context for the patient data. For instance, a patient may be **on** insulin as a treatment for Diabetes or be **allergic** to insulin. There is only one unique insulin concept in the HDD (NCID 3000250167), because the insulin one may be using and the insulin one may be allergic to is the same concept (entity) where the definition of the drug is concerned. However, the clinical implications are obviously different. Thus, the CDR stores medications and allergies in two different information models, providing the context.

**PharmacyOrder** ::= SET {

| | | |
|---|---|---|
| drugs | [0] | **Drugs**, |
| routeInfos | [1] | RouteInfos OPTIONAL, |
| providerAdminInstructions | [2] | ProviderAdminInstructions OPTIONAL, |
| deliverToLocation | [3] | DeliverToLocation OPTIONAL, |
| substitutionStatus | [4] | CodedWOSform (SubstitutionStatus) OPTIONAL, |
| dispensed | [5] | QuantityUnits OPTIONAL, |
| numberOfRefills | [6] | NumberOfRefills OPTIONAL, |
| prescriptionNumber | [7] | PrescriptionNumber OPTIONAL, |
| numberOfRefillsRem | [8] | NumberOfRefillsRem OPTIONAL, |
| numberOfRefillsDis | [9] | NumberOfRefillsDis OPTIONAL, |
| dtMostRecentDoseDis | [10] | DtMostRecentDoseDis OPTIONAL, |
| needsHumanReview | [11] | CodedWOSform (NeedsHumanReview) OPTIONAL, |
| pharmacySpecialIns | [12] | PharmSpecialInstructInfos OPTIONAL, |
| ivVolumeRate | [13] | IvRate OPTIONAL, |
| ivSpecialRate | [14] | IvSpecialRate OPTIONAL, |
| codedComments | [15] | CodedComments (PharmOrderCodedComment) OPTIONAL, |
| prescriptionType | [16] | LCodedWOSformAtt (PrescriptionType) OPTIONAL, |
| comments | [17] | Comments (PharmOrderCodedComment) OPTIONAL, |
| sortIndicator | [18] | LCodedWOSformAtt (RxSortIndicator) OPTIONAL, |
| results | [19] | ClinicalObservations (WITH TYPES {OrderObservationObs * })} |

**Drug** ::= SET {

| | | |
|---|---|---|
| rXComponentType | [0] | CodedWOSform (RXComponentType) OPTIONAL, |
| **componentCode** | [1] | LCodedWOSform (**PharmacyItems**), |
| doseUnits | [2] | DoseUnitInfo OPTIONAL, |
| giveDosageForm | [3] | LCodedWOSform (GiveDosageForm) OPTIONAL, |
| totalDailyDose | [4] | QuantityUnits (WCS {value, units} ) OPTIONAL, |
| ingredients | [5] | **Ingredients** OPTIONAL, **--Where insulin as medication is stored** |
| codedComments | [6] | CodedComments (DrugCodedComment) OPTIONAL, |
| textComments | [7] | TextComments OPTIONAL, |
| setId | [8] | SetId OPTIONAL } |

**AllergyInfo** ::= SET {

| | | |
|---|---|---|
| AllergyType | [0] | LCodedWOSform (AllergyType) OPTIONAL, |
| AllergyId | [1] | AllergyId OPTIONAL, |
| CompositeSubstance | [2] | LCodedWOSform (CompositeSubstance), **--Where insulin as allergy is stored** |
| AllergenIds | [3] | L-AllergenIds, |
| Severity | [4] | LCodedWOSform (AllergySeverity) OPTIONAL, |
| Reactions | [5] | L-Reactions OPTIONAL, |
| IdentificationDate | [6] | IdentificationDate OPTIONAL, |
| SourceOfInfo | [7] | LCodedWOSform (SourceOfInfo) OPTIONAL, |
| CodedComments | [8] | CodedComments (AllergyCodedComments) OPTIONAL, |
| TextComments | [9] | TextComments OPTIONAL } |

Working Product, Operational Processes and Growth

Standard vocabularies are terminologies, not applications. To use standard vocabularies to encode patient data, operational implementation in a software product is required. The 3M HDD is an application that is fully operational and implemented. It has been live in real world use at multiple health care facilities over the past few years. HDD software services work with other applications such as the CDR, Clinical Workstation, decision support, and interfaces.

Updates handled by the HDD ensure that the DoD will maintain compliance and currency with vocabulary standards. Processes for information modeling and for developing and maintaining the vocabulary and knowledge base are established and validated. The HDD is designed to be flexible and extensible, and to stay current with evolving standards. The HDD can subsume other ERTs and ensure their interoperability with DoD's ERT.

Medical Informatics and Lexical Services

3M is actively involved in developing standards with organizations such as LOINC and HL7. We are recognized in the industry as the leading expert in the implementation and use of standard vocabularies. Research and development in the area of lexical services have created a variety of tools and techniques. The methods pioneered by the 3M HDD has led to advances in the field of terminology research. 3M continues to explore advances in technology, e.g. in the area of natural language processing. We will support our customers with our Medical Informatics expertise and experience.

**Data Sharing Between DoD and VA**

The DoD and the VA need to exchange and share clinical data to provide services seamlessly to both DoD and VA patients. We have been told that an experiment has found no match between the items in the VA and DoD databases for a test sample of 100 NCIDs. Not having been given any additional information, we are unable to comment on the experiment design. A recommendation can be offered regarding the sample size. For large domains, a sample size of 100 is inadequate. There are over 42,000 laboratory results, nearly 50,000 generic drugs, and over 145,000 active NDCs. Experience has shown that site-to-site metadata variation is common. Most reference terminologies need to be large in order to provide concepts of sufficient granularity. One concept may differ from another only in a particular detail or attribute. The codes would be different because the concepts are not identical but related. Even if usage is restricted to a limited, non-extensible list of standard codes, the subset used by each site will differ even within the enterprise. As a demonstration, we will compare the race concepts used by the VA and the DoD. We will also examine and compare the Primary Drug master files between two of the VA's facilities, and to the DoD metadata in the HDD.

Because of the HDD, the DoD will be able to receive external data and operate upon it within CHCSII. The standard codes for the external data are translated into NCIDs by the HDD, as described before. Then the HDD knowledge base is used to obtain the component or attribute information presented by the external data. In contrast, a standard vocabulary may or may not provide the semantic network to allow identification of the relationship. Even if it does, for true understanding of the data, a software process needs to implement the relationship traverse and lookup. The HDD knowledge base has been developed to meet this need. Next, the information is stored appropriately in the CDR according to the information model. Decision support and other applications will operate as usual. We will illustrate the operation with pharmacy data.

As an example, NDCs may be used to exchange medication data. An NDC is assigned to a trade name product. While the NDC is a unique number (apart from the issue of code reuse), it does not provide clinically relevant uniqueness to the medication itself. For instance, there are over 1,200 non-obsolete NDCs for Acetaminophen 500mg tablets. A few examples are:

- NDC 00045012400 – Tylenol Extra-Strength (Acetaminophen), 500mg, Tablet, Oral, McNeil Cons., 125 ea. Bottle
- NDC 00045012410 – Tylenol Extra-Strength (Acetaminophen), 500mg, Tablet, Oral, McNeil Cons., 100 ea. Bottle
- NDC 00573034040 – Anacin Aspirin Free (Acetaminophen), 500mg, Tablet, Oral, Whitehall-Rob., 100 ea. Bottle

In the HDD, all 1,200 NDCs are related to the component generic medication "Acetaminophen, 500mg, Tablets, Oral" (NCID 3000261886). All of these NDCs also have an ingredient relationship to "Acetaminophen" (NCID 3000252137). Traversing the relationship trees will retrieve the appropriate NCIDs for data store according to the information model (PharmacyOrder and Drug, shown earlier). The medication NCID will be stored in Drug.componentCode, while the ingredient NCIDs for that medication will be stored in Drug.ingredients. As stated earlier, all medications belong to the domain PharmacyItems. The NCIDs will be used by the appropriate applications, e.g. the ingredient in drug allergy alerts (stored in the AllergyInfo information model, also shown earlier). Without the HDD, the 1,200 NDCs appear as if they are all different concepts, and the receiving system will need significant effort to manage the data appropriately.

Comparing VA and DoD Metadata for Race

When two data sets are widely disparate in size, there will be a lack of overlap. An example is seen in the race concepts used by the DoD and the VA.

| Number of VA Race Concepts | Number of DoD Race Concepts | Race Concepts in Common | Race Concepts in VA only | Race Concepts in DOD only |
|---|---|---|---|---|
| 16 | 44 | 11 | 5 (31%) | 33 (75%) |

**Table: Comparing the number of Race concepts used by the DoD and the VA**

| Race Concept | DoD | VA |
|---|---|---|
| American Indian | | X |
| **American Indian or Alaskan** | | X |
| American Indian Unknown | X | |
| **American Indian/Nat Alaskan, Hispanic** | X | |
| **American Indian/Nat Alaskan, Non-Hispanic** | X | |
| **Asian** | | X |
| **Asian or Pacific Islander** | X | X |
| Asian or Pacific Islander, Hispanic | X | |
| Asian or Pacific Islander, Non-Hispanic | X | |
| Black | X | X |
| Black Hispanic | X | X |
| Black, Non-Hispanic | X | X |
| Caucasian | X | X |
| Declined to Answer | | X |
| Filipino | X | |
| Filipino Hispanic | X | |
| Filipino Non-Hispanic | X | |
| Filipino Unknown | X | |
| Hispanic | X | X |
| Hispanic Native American | X | |
| Hispanic/Latino | X | X |
| Latin American | X | |
| Middle Eastern | X | |
| Mongoloid | X | |

| Concept | DoD | VA |
|---|---|---|
| Mulatto | X | |
| Mutually Defined | X | |
| Native Hawaiian | X | |
| Native Hawaiian or other Pacific Islander | | X |
| Not Applicable | X | |
| Not available | X | |
| Not Provided | X | |
| Oriental | X | |
| Other | X | |
| Other American Indian Tribe | X | |
| Other or Unknown | X | |
| Other, Hispanic | X | |
| Other, Non-Hispanic | X | |
| Scandinavian | X | |
| South East Asian | X | |
| Unknown | X | X |
| Unknown Asian | X | |
| Unknown Black | X | |
| Unknown, Hispanic | X | |
| Unknown, Non-Hispanic | X | |
| Western Hemisphere Indian (Red) | X | |
| White | X | X |
| White Hispanic | X | X |
| White, Non-Hispanic | X | X |
| White, unknown | X | |

**Table: Comparing the Race concepts used by the DoD and the VA**

As can be seen, the VA data set and the DoD data set differ significantly in size (16 for VA, 44 for DoD). While the majority, 69%, of VA races can be found in the DoD data set, only 25% of the DoD races can be found in the VA data set. If the VA sends a race code to the DoD, there will be a direct match to a DoD race for 69% of the concepts. For the remainder, the HDD knowledge base can be used to create relationships to support the DoD's business logic for interoperability. For instance, the DoD might decide that when the VA sends the code for "Asian", for DoD's purpose it will serve the same function as "Asian or Pacific Islander". Note that relationship is used instead of concept synonymy, because the two concepts are, in fact, not identical. However, they can be related. This example is illustrated in the table above by the lower, red arrow.

For the VA, because its data set is less granular, the majority (75%) of DoD race codes do not have a direct correspondence. The HDD can also be used by the VA to achieve data interoperability. For instance, when the DoD sends the codes for "American Indian/Nat Alaskan, Hispanic" or for "American Indian/Nat Alaskan, Non-Hispanic," the HDD knowledge base could be used to relate either of those codes to "American Indian or Alaskan," the race concept used by the VA. This example is illustrated in the table above by the upper, blue arrow. This kind of scenario will be seen repeatedly in medical data, and it is one the HDD is built to manage.

VA Primary Drugs (File 50.3)

We will examine the Primary Drug files provided to 3M at the beginning of 2003 to review the overlap between the two VA master files.

| Martinsburg NCIDs | St. Louis NCIDs | NCIDs in Common | NCIDs in Martinsburg only | NCIDs in St. Louis only |
|---|---|---|---|---|
| 1980 | 2169 | 570 | 1410 (71%) | 1599 (74%) |

**Table: Comparing the Primary Drug concepts from Martinsburg and St. Louis**

The VA Primary Drug file contains a mix of single drug ingredients (Captopril), combination ingredients (Acetaminophen/Oxycodone), ingredients with forms (Collagenase Ointment), ingredients with routes (Cortisone Injection), ingredients with strength (Dextrose 5% Lactated Ringer's), ingredients with strength and form (Clotrimazole 1% Cream), etc. All the above are examples of concepts that the two VA sites have in common.

As mentioned earlier, two concepts can differ from one another only in a single attribute or component detail. For example, "Acetaminophen", "Acetaminophen Tablet", "Acetaminophen 500mg" and "Acetaminophen 500mg Tablet" would be four different concepts assigned four different NCIDs. It is absolutely critical to distinguish them as different concepts for clinical correctness. It is also necessary to relate them via the same ingredient (Acetaminophen), the same strength (500mg), or the same form (Tablet), where appropriate. We do not agree with the VA approach of stating that "Hydroxyzine Pamoate" is a synonym for "Hydroxyzine Hydrochloride Tablet" or that "Pilopine HS Gel 4%," "Pilocarpine 6% Ophthalmic Solution," and "Pilocar 2% Dropperettes" are synonyms for "Pilocarpine 1% Ophthalmic Drops".

| NCIDs in Martinsburg only | NCIDs in St. Louis only |
|---|---|
| Acetaminophen Liquid | Acetaminophen |
| Acetaminophen 500mg Tablet | Acetaminophen 650mg |
| Acyclovir Capsule | Acyclovir |
| Acyclovir Tablet | Acyclovir 5% |
| Acyclovir IV | Acyclovir IVPB |

**Table: Examples of non-overlapping Primary Drug concepts from Martinsburg and St. Louis**

Given the content of the VA Primary Drug file, the lack of overlap for the mapped NCIDs between the two sites is expected. This is where the 3M HDD can be used for data interoperability. As described earlier, the HDD maintains a comprehensive, up to date knowledge base. The relationships can be used to obtain the ingredient NCIDs and other relevant information from the external data. The component information – ingredient, strength, form, route, the drug class to which the ingredient belongs, etc. – is what is used in data applications for such processes as comparison, aggregation, and decision support. To look at only the single NCID identifying the external data is an oversimplification. More importantly, it is inadequate for obtaining all the information provided by the data.

As an exercise, just looking at the single ingredients of the non-overlapping items, there is now almost a two-third overlap in the ingredient NCIDs. The remaining items do not appear in both files for a variety of reasons:

- Some may simply not be used by the site (e.g. "Lidocaine")
- Others are not drugs (e.g. "Lotion")
- Some are workflow processing items, such as "Do Not Use" and "Temporary Drug"
- Others are trade names for which the ingredient cannot be assured (e.g. "Advantage")
- Many are meaningless (e.g. "ZZ1")

Once a VA Primary Drug item has a relationship to an HDD ingredient NCID, there should be complete interoperability with DoD data because the DoD formulary items mapped in the HDD have the appropriate relationships to the same set of ingredients. Thus, there would be complete overlap between the ingredient NCIDs for the VA and the ingredient NCIDs for the DoD.

**Conclusion**

The problems surrounding data interpretation, interoperability and interrelationship are not new. The 3M HDD product and implementation strategies for mapping and maintenance evolved from meeting our customers' needs. 3M is recognized as the leading expert in the implementation and use of standard vocabularies. The data standardization approach pioneered by the 3M HDD has been validated by successful operational use at commercial and DoD facilities. The lessons learned have been widely shared with the industry through Medical Informatics publications, presentations, and tutorials. Our vision is to be the following:

- A universally useful implementation of informatics principles and standardized vocabularies
- A working product that is an integral part of a CDR, supporting applications such as data warehouse, order entry, results review, etc.
- A partner to help our customers make use of their patient data to improve quality of care, delivery, outcomes, costs, and competitiveness

We are committed to help the DoD achieve its mission in health care.