



# **Description of the MHS Health Level 7 Microbiology Laboratory for Public Health Surveillance**

Technical Document NMCPHC-EDC-TD-1-2013

By Uzo Chukwuma and Gosia Nowak  
EpiData Center Department  
October 2012

Approved for public release. Distribution is unlimited.

The views expressed in this document are those of the author(s) and do not necessarily reflect the official policy or position of the Department of the Navy, Department of Defense, nor the U.S. Government

# Report Documentation Page

Form Approved  
OMB No. 0704-0188

Public reporting burden for the collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to a penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.

1. REPORT DATE <b>10 OCT 2012</b>		2. REPORT TYPE <b>Technical Document</b>		3. DATES COVERED <b>00-05-2004 to 00-10-2012</b>	
4. TITLE AND SUBTITLE <b>Description of the MHS Health Level 7 Microbiology Laboratory for Public Health Surveillance</b>				5a. CONTRACT NUMBER	
				5b. GRANT NUMBER	
				5c. PROGRAM ELEMENT NUMBER	
6. AUTHOR(S) <b>Uzo Chukwuma; Gosia Nowak</b>				5d. PROJECT NUMBER	
				5e. TASK NUMBER	
				5f. WORK UNIT NUMBER	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) <b>Navy and Marine Corps Public Health Center, EpiData Center Department, 620 John Paul Jones Circle, Suite 1100, Portsmouth, VA, 23708-2103</b>				8. PERFORMING ORGANIZATION REPORT NUMBER <b>NMCPHC-EDC-TD-1-2013</b>	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) <b>Navy and Marine Corps Public Health Center, EpiData Center Department, 620 John Paul Jones Circle, Suite 1100, Portsmouth, VA, 23708-2103</b>				10. SPONSOR/MONITOR'S ACRONYM(S) <b>NMCPHC</b>	
				11. SPONSOR/MONITOR'S REPORT NUMBER(S) <b>NMCPHC-EDC-TD-1-2013</b>	
12. DISTRIBUTION/AVAILABILITY STATEMENT <b>Approved for public release; distribution unlimited</b>					
13. SUPPLEMENTARY NOTES					
14. ABSTRACT <b>The EpiData Center Department (EDC) at the Navy and Marine Corps Public Health Center (NMCPHC) evaluated the Health Level 7 (HL7) data source for its usefulness in health surveillance activities. This technical document provides a history of the HL7 microbiology database and its contents, explains the creation of microbiology records, describes the pathway of data from healthcare provider to the EDC, provides a detailed descriptions of all variables within the database, and assesses the database's strengths and limitations. Given an understanding of the strengths and limitations of the data, HL7 microbiology data have proven to be a valuable source of health information for surveillance purposes. The data allow the creation of a timeline of events corresponding to a specific disease occurrence. Furthermore, data are received in a timely fashion, allowing for near-real-time surveillance of diseases.</b>					
15. SUBJECT TERMS <b>Health Level 7 (HL7), Microbiology, Surveillance</b>					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			
<b>unclassified</b>	<b>unclassified</b>	<b>unclassified</b>	<b>1</b>	<b>33</b>	

REPORT DOCUMENTATION PAGE			Form Approved OMB No. 0704-0188		
Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing this collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.					
1. REPORT DATE (DD-MM-YYYY) 10-10-2012		2. REPORT TYPE Technical Document		3. DATES COVERED (From - To) May 2004-October 2012	
4. TITLE AND SUBTITLE Description of the MHS Health Level 7 Microbiology Laboratory for Public Health Surveillance			5a. CONTRACT NUMBER		
			5b. GRANT NUMBER		
			5c. PROGRAM ELEMENT NUMBER		
6. AUTHOR(S) Uzo Chukwuma, Gosia Nowak			5d. PROJECT NUMBER		
			5e. TASK NUMBER		
			5f. WORK UNIT NUMBER		
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Navy and Marine Corps Public Health Center EpiData Center Department 620 John Paul Jones Circle, Suite 1100 Portsmouth, VA 23708			8. PERFORMING ORGANIZATION REPORT NUMBER NMCPHC-EDC-TD-1-2013		
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) Navy and Marine Corps Public Health Center EpiData Center Department 620 John Paul Jones Circle, Suite 1100 Portsmouth, VA 23708			10. SPONSOR/MONITOR'S ACRONYM(S) NMCPHC		
			11. SPONSOR/MONITOR'S REPORT NUMBER(S) NMCPHC-EDC-TD-1-2013		
12. DISTRIBUTION / AVAILABILITY STATEMENT Approved for Public Release. Distribution is not limited.					
13. SUPPLEMENTARY NOTES					
14. ABSTRACT The EpiData Center Department (EDC) at the Navy and Marine Corps Public Health Center(NMCPHC) evaluated the Health Level 7 (HL7) data source for its usefulness in health surveillance activities. This technical document provides a history of the HL7 microbiology database and its contents, explains the creation of microbiology records, describes the pathway of data from healthcare provider to the EDC, provides a detailed descriptions of all variables within the database, and assesses the database's strengths and limitations. Given an understanding of the strengths and limitations of the data, HL7 microbiology data have proven to be a valuable source of health information for surveillance purposes. The data allow the creation of a timeline of events corresponding to a specific disease occurrence. Furthermore, data are received in a timely fashion, allowing for near-real-time surveillance of diseases.					
15. SUBJECT TERMS Health Level 7 (HL7), Microbiology, Surveillance					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			UU
U	U	U			19b. TELEPHONE NUMBER (include area code) 757-953-0706

Standard Form 298 (Rev. 8-98)  
Prescribed by ANSI Std. Z39.18



## Abstract

The EpiData Center Department (EDC) at the Navy and Marine Corps Public Health Center (NMCPHC) evaluated the Health Level 7 (HL7) data source for its usefulness in health surveillance activities. This technical document provides a history of the HL7 microbiology database and its contents, explains the creation of microbiology records, describes the pathway of data from healthcare provider to the EDC, provides a detailed descriptions of all variables within the database, and assesses the database's strengths and limitations. Given an understanding of the strengths and limitations of the data, HL7 microbiology data have proven to be a valuable source of health information for surveillance purposes. The data allow the creation of a timeline of events corresponding to a specific disease occurrence. Furthermore, data are received in a timely fashion, allowing for near-real-time surveillance of diseases.



## Table of Contents

List of Figures and Tables .....	v
Executive Summary.....	1
Introduction .....	2
Public Health Surveillance Applications .....	2
Methods and Procedures.....	4
Data Structure and Analysis.....	6
Structure.....	6
Table 1 Example of HL7 microbiology record sorted by MSG ID and SET ID .....	6
Analysis.....	7
Key Fields for Public Health Surveillance .....	8
Duplicates.....	8
Unique Patient/Specimen .....	8
Test Definition .....	8
Laboratory Test Result .....	9
Date/Time References.....	10
Table 2. Timeframes in HL7 data.....	11
Strengths .....	12
Timeliness.....	12
Completeness.....	12
Organization and Structure .....	12
Public Health Applications .....	12
Limitations.....	14
Complexity.....	14
Completeness.....	14
Generalizability.....	14
Inconsistency.....	14



Data Interpretations.....	15
All Data Fields (Variables) .....	16
Automatically Populated Fields.....	16
Variable Descriptions .....	17
Abbreviations .....	26
Definitions.....	27
Department of Defense (DOD).....	27
Department of the Navy (DON) .....	27
Defense Medical Information System (DMIS).....	27



## List of Figures and Tables

Table 1. Example of HL7 microbiology record sorted by MSG ID and SET ID .....	6
Table 2. Timeframes in HL7 data .....	11
Figure 1. HL7 laboratory data flow process .....	4
Figure 2. Timeline of dates in HL7 microbiology .....	11



## Executive Summary

The EpiData Center Department (EDC) at the Navy and Marine Corps Public Health Center (NMCPHC) was funded by the Department of Defense (DOD) Global Emerging Infections Surveillance and Response System (GEIS) with the evaluation of the Health Level 7 (HL7) data source for its usefulness in health surveillance activities. This technical document provides a history of the HL7 microbiology database and its contents, explains the creation of microbiology records, describes the pathway of data from healthcare provider to the EDC, provides a detailed descriptions of all variables within the database, and assesses the database's strengths and limitations.

The HL7 microbiology database is used extensively by the EDC for a variety of tasks, including daily case finding of reportable diseases, identification of antibiotic resistant organisms, preparation of health reports, and responding to congressional requests for disease burden. Disease burden evaluations have included, among others, respiratory infections (e.g., pandemic influenza, pertussis), skin and soft tissue infections (e.g., methicillin resistant *Staphylococcus aureus*) and gastrointestinal infections (e.g., salmonellosis, norovirus). Positive microbiology results can be matched with outpatient or inpatient encounter records to identify whether laboratory tests correlate with encounters, which may help with case validation and confirmation.

These data are limited such that records from purchased care, shipboard facilities, battalion aid stations, or in-theater facilities are not available. Additionally, microbiology testing results only show the organism(s) that were identified, not what the test was intended for. Cases where a physician chooses to treat presumptively without laboratory confirmation will not be captured. Clinical practice with regards to culturing varies between providers and facilities. Examples of situations where cultures may not be performed include confirmatory tests for patients with influenza-like illness symptoms, or patients with superficial infections who are treated presumptively.

HL7 microbiology data are a valuable source of health information for surveillance purposes. The data allow the creation of a timeline of events corresponding to a specific disease occurrence. Data are received in a timely fashion, allowing for near-real-time surveillance of diseases.



## Introduction

The EDC at the NMCPHC evaluated the HL7 data source for its usefulness in health surveillance activities at the request of the DODGEIS. The HL7 data source includes records from anatomic pathology, chemistry, microbiology, pharmacy, and radiology. Laboratory results (microbiology and chemistry) were identified as the most useful type of HL7 data for improving military health surveillance activities. Therefore, extensive work was done to examine the laboratory databases, determine completeness and reliability, identify areas for improvement as needed for surveillance, and establish methods for the surveillance of specific conditions. This technical document describes the data fields in HL7 microbiology, extent of completion of these fields, modifications made to the HL7 data flow and processing schema, data cleaning rules, and other comments regarding surveillance activities.

When HL7 data were first received by the EDC, a significant amount of work was devoted to ensuring messages were parsed and organized properly. Sample extracts for review were received from Defense Health Services System (DHSS) from September 2003 to April 2004. Initial sample extracts showed data were sparse. Conversations with personnel at local military treatment facilities (MTFs) and analysis of particular fields in the sample files revealed several observations. First, not all microbiology results for a given culture were seen as expected in the sample extracts. Further investigation and discussions led to comparison of local MTF results, the original HL7 message for those results, and the records in the DHSS staging database. Based on these comparisons, DHSS reconfigured the HL7 process such that microbiology data were not lost when messages were parsed and submitted.

The second issue identified during the initial review process involved fields that identified which MTFs requested and performed laboratory test orders. Frequencies of these fields showed many MTFs were not represented in the data. As this information was passed onto DHSS, the Military Health System (MHS) Helpdesk was contacted to remedy this situation. Within several months, missing MTFs began appearing in the HL7 data. In addition, DHSS began to monitor incoming message traffic by the Composite Healthcare System (CHCS) sending facility. By May 2004, all major MTFs and most clinics were represented in the HL7 data extracts.

## Public Health Surveillance Applications

The EDC has used the HL7 microbiology database to support Department of the Navy (DON) and DOD preventive medicine activities since 2005. Examples of support include case findings of particular diseases (e.g., malaria, meningococcal meningitis, or influenza) and identification of antibiotic resistance patterns. The data are used for analyses in support of collaboration with MTFs for local infection control activities.

Epidemiologic analysis of these data focuses on defining trends of illness by reviewing laboratory test orders and results. These results may be linked to other databases for a more comprehensive description of a disease event. For example, positive laboratory results are matched to International Classification of Diseases, 9<sup>th</sup> Revision, Clinical Modification (ICD-9-CM) codes in encounter records to identify how specific laboratory tests relate to clinical encounters and comorbid conditions. Prescription data may also play a significant role in



conjunction with laboratory results data. Many variables in a database disease surveillance model depend on clinical practice.

Applications of these data are not limited to traditional medical event surveillance. These data fill a significant gap in the DOD's ability to track and describe antibiotic resistance throughout the MHS. The MHS has found it necessary, but almost impossible, to currently track patients with antibiotic-resistant infections across MTFs as they are transferred from one MTF to another. The HL7 data stream can fill these gaps as all information is collected centrally, without the need to access local CHCS data.

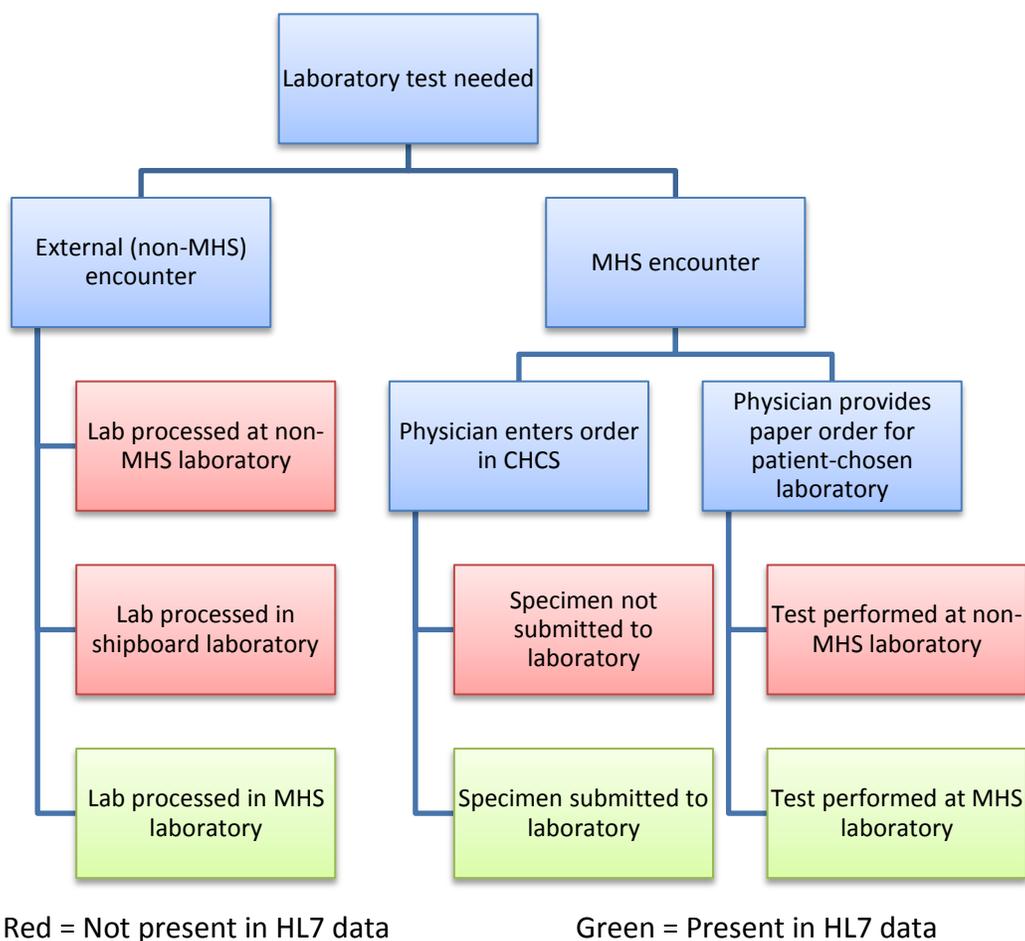
For surveillance purposes, a laboratory result often yields information useful for identifying epidemiologic characteristics such as the timeframe of potential illness, pathogenic agent, patient demographics, and geographic location. Reviewing these characteristics may help describe a disease cluster or other issues of concern to preventive medicine, infection control, and patient safety.



## Methods and Procedures

[Figure 1](#), provides an overview of the flow of information involved in the generation of HL7 laboratory data. The process originates at point of care. When a patient is seen by a healthcare provider in a MTF, the provider can order a laboratory test via local CHCS. Generally, the provider sees a list of possible tests to order in CHCS. This list is generated by the MTF laboratory's Management Information Department (MID) based on laboratory tests that can be performed at the facility or outside laboratories that have contracts with the facility. Laboratory orders performed by non-MHS laboratories should be entered into the CHCS when a patient provides the laboratory order slip or when the supporting laboratory communicates results with the requesting MTF. The consistency of this practice is not well-known.

Figure 1. HL7 laboratory data flow process



There are circumstances where laboratory tests may not be ordered directly from the provider. For example, the laboratory can initiate another test for a sample based on results of a previous test necessitating confirmation of the disease; this practice depends on standard protocols of the local laboratory. Depending on the patient's status, specimens could be drawn within the hospital ward and sent to the laboratory or could be drawn directly at the laboratory. Most hospitals can do microscopic readings in-house for tests such as malaria and acid-fast bacilli (AFB) smears for tuberculosis. However, MTFs may also outsource culture tests for some conditions. Initial culture growth may be performed at the MTF and then sent out for typing if an organism grows. If a test is outsourced, MTFs are required to enter the results into CHCS for clinical evaluation by the provider.

Laboratory results are mapped to their appropriate database based on a value that defines the type of laboratory setting that performs the test (chemistry, microbiology, pathology, etc). Once labeled, the information is sent to the CHCS database where it is stored based on this value.

Laboratory results are certified by a laboratory technician in CHCS before a script based on system triggers can generate an HL7 message. The HL7 message is archived and batched with other HL7 messages on the local CHCS host. At least once a day, these HL7 messages are forwarded to the central CHCS server. Once receipt is verified by the central server, HL7 messages at the local host are deleted. These records are then retrieved by DHSS and parsed into a database design four times a day.

Extracts are retrieved by the EDC using a secure connection to the DHSS feed node. Flat file extracts of the raw, parsed data are received from DHSS on a daily basis. The data are cleaned, including the elimination of true duplicates.



## Data Structure and Analysis

### Structure

The EDC receives HL7 microbiology data in a pipe-delimited flat file from DHSS. [Table 1](#), Appendix C, shows the general structure of microbiology data with one entry per line. An individual test can have multiple entries or records as part of a series. For example, names in the TEST ORDERED field are not always disease-specific because of the nature of the tests. While physicians may suspect a specific viral or bacterial etiology, they order a general test for culture (e.g., respiratory or wound culture). As a result, an individual sample can have multiple entries for each test performed.

**Table 1 Example of HL7 microbiology record sorted by MSG ID and SET ID**

MSG ID	ACCESSION NUMBER	SET ID	TEST NAME	TEST RESULT
999-253	123-45-7	1	Flu Cult	
999-253	123-45-7	2	Bact	Neg A
999-253	123-45-7	3	Bact	Pos A
999-253	123-45-7	4	Inf A	
562-235	123-45-7	1	Influenza A&B	
562-235	123-45-7	2	Bact	Neg A
562-235	123-45-7	3	Bact	Neg A
562-235	123-45-7	4	Virus Cult	Neg A
562-235	123-45-7	5	Haemophilus Influenzae	
562-235	123-45-7	6	Bact	No Growth
562-235	123-45-7	7	Bact	Moderate Growth
562-235	123-45-7	8	Virus Cult	Haemophilus Influenza-Beta Lactamase Positive

**Summary of MSG ID 999-253 (Purple):**

- SET ID 1 identifies the test as an Influenza Culture
- SET ID 2 shows no growth of Influenza A at first time interval in the test result field
- SET ID 3 shows growth of Influenza A at second time interval in the test result field
- SET ID 4 identifies the agent that grew as Influenza A in the TEST NAME field

**Summary of MSG ID 562-235 (Green):**

- SET ID 1 identifies the test for Influenza A and B
- SET ID 2-4 show no growth at first three time intervals in the test result field
- SET ID 5 identifies the second test for the same specimen as *H. influenzae*
- SET ID 6-7 show growth progression at first two time intervals in the test result field
- SET ID 8 identifies growth of *H. influenzae* in the test result field



Multiple entries may also be observed for culture results. Culture results are typically provided at 24 and 48 hour intervals until the final growth is read, usually at the end of a 72 hour growth period, although this period may vary depending on disease. Separate records are created for results at each time interval. All of these records will be certified in CHCS and are visible in the HL7 data. These records should be read in chronological order to understand the progress of bacterial or viral growth.

The two examples in [Table 1](#) should be considered differently since one (MSG ID 999-253) indicates a positive test for a particular species, while the other (MSG ID 562-235) describes growth of a culture over time. Methods have been established in the EDC to distinguish between these instances. Records can be grouped together by **ACCESSION NUMBER**, **SPECIMEN SOURCE**, and **SET ID** to group laboratory results for an individual sample.

Results may be corrected after final results are certified. Corrected results are visible in the HL7 microbiology data as separate lines for each corrected result. The hierarchy for selecting a final line entry for analysis purposes is as follows: C (Corrected), F (Final), P (Pending).

## Analysis

Before analysis of a particular outcome, it is critical to have an understanding of the disease of interest, its symptomatic course, local provider practices with regard to treatment and testing, available tests, and local laboratory testing or outsourcing procedures to determine how the HL7 microbiology data should be used with respect to the outcome of interest. It is also important to note that test names and results are not standardized in CHCS or the HL7 microbiology data. Multiple variations of test names and results may be associated with a particular disease, depending on the testing method and recording practices. Misspellings, abbreviations, variations in spacing, and periods or other punctuation within text should all be considered when conducting a query of particular tests or results in the HL7 microbiology database. Queries may be done across multiple fields to search for tests and results for a particular outcome. All query results should be thoroughly analyzed throughout the process to assess data completeness and accuracy and to ensure data are classified correctly.

Test orders that are not traditionally considered microbiology tests may be found in HL7 microbiology data for particular MTFs, depending on how the particular laboratory is set up at the local level. MTFs may accession some specimens through the microbiology section of the laboratory, even though the order requests a chemistry test. Therefore, when searching for a particular agent or species, all laboratory tables should be searched to ensure full capture of test orders and results.



## Key Fields for Public Health Surveillance

Data for calendar year 2011 were pulled for all DOD medical beneficiaries in the HL7 microbiology database. All data were reviewed and analyzed in order to modify the datasets to more accurately address the disease surveillance needs of the EDC. Methods for identifying duplicate and unique records were established.

### Duplicates

*True duplicates* are records in which all fields have identical values. If true duplicates exist, one record is kept and the duplicates are eliminated by EDC database administrators prior to providing data to the analyst. Each record that remains after removing true duplicates is considered a unique record; there is at least one variable value different than all other records in the database.

### Unique Patient/Specimen

*Unique patients* are identified through a combination of SPONSOR ID (sponsor Social Security number [SSN]) and family member prefix (FMP). This unique identifier can be used to track individual patients through all HL7 microbiology records. There is a variable called PATIENT ID (patient SSN); however, the field is not consistently populated and is unreliable as a way to identify patients within or across databases. It is possible for individuals to have two different SPONSOR IDs in the database over time. For example, if the child of a sponsor becomes active duty, then that child will have his/her own SSN as the SPONSOR ID instead of the parent's SSN. Unique patients can have multiple laboratory orders in the HL7 chemistry data.

A patient will often have multiple samples taken at one time or over a period of time. Each sample, or specimen, would have a different ACCESSION NUMBER, even if many of the other fields are the same. A *unique specimen* is defined using ACCESSION NUMBER, SPONSOR ID, and FMP.

The use of ACCESSION NUMBER and TEST NAME, in combination with the SPONSOR ID and FMP, is the most accurate way to identify a *unique test order* for both chemistry and microbiology databases. Since samples may be continuously tested throughout the day, there can be several records with the same ACCESSION NUMBER. By using the combination for unique test order, the analyst eliminates duplicate records per each patient's ordered tests. Additionally, a person could have multiple samples taken, each with a different ACCESSION NUMBER. The unique test order will determine which test was ordered for a particular specimen per patient. The most recent CERTIFY DATE and CERTIFY TIME for a unique test order generally correspond to the results of the test.

### Test Definition

The TEST ORDERED and the TEST NAME fields are used to identify the laboratory test performed. Both fields may contain non-standardized values, so additional fields and searches of free text may be needed to identify a test. Inaccurate reporting could occur if the analyst does



not know all parameters of the test prior to analyzing the data. Using influenza as an example, a laboratory test could be defined taxa-specific (e.g., order, family, subfamily, genus, or species), or based on test type (e.g., antibody staining test or convalescent test). Each test could have a different specimen source (e.g., serum or nasopharyngeal wash), result type (e.g., numeric or positive/negative), and timeframe which determines each testing method. The outcome of interest determines how the data are organized in the microbiology dataset. It is important to understand the dataset as it relates to the specific project being conducted. For many diseases and pathogens, the observations as described above hold true. For others, the observations (where test results are recorded, how test results are recorded, how to determine what test was ordered) may be vastly different. For example, influenza tests are recorded differently than tests for other pathogens such as *Acinetobacter baumannii* and *Staphylococcus aureus*.

## Laboratory Test Result

Due to the structure of the laboratory data, results could be identified across multiple variables and records ([Table 1](#)). For example, for cultures, results are recorded at time intervals to describe growth patterns which leads to multiple lines for each test ordered. The TEST RESULT field could show growth per the number of colony forming units (CFU), positive or negative growth or results, laboratory technician's comments or methods, or null values. Particularly in the HL7 microbiology data, the order of these lines of data is important in interpreting a laboratory result. The SET ID field signifies the line number of an HL7 message, and records should be sorted accordingly for proper analysis.

The actual test result could be seen in either the TEST NAME or TEST RESULT field ([Table 1](#)). Due to the nature of the test and recording methods, analysis of the HL7 microbiology data will vary when determining case definition for each disease. For records with a result of "Positive," the TEST NAME will typically identify the positive agent's name.

The HL7 microbiology data are not optimally arranged to identify a pathogen and its associated antibiotic susceptibility tests. If a message is sorted by MSG ID and SET ID, a culture result is generally organized as follows:

1. The test results of the laboratory pathogen in terms of growth after 24 and 48 hours. This may appear over several rows of data.
2. The final test result determining either no growth or the name of the pathogen. This may appear over several rows of data.
3. The resulting antibiotic susceptibility pattern of the identified pathogen with one antibiotic result per row of data.
4. If present, the name of a second pathogen.
5. The antibiotic susceptibility results for the second pathogen.

Information for each of these segments is recorded differently. Data recorded in the TEST NAME and TEST RESULT fields are different depending on the segment of the message. If several organisms grow from one specimen, the name of the second pathogen usually appears after the antibiotic susceptibility results of the first pathogen and it is followed by its own antibiotic susceptibility results. This pattern continues for multiple organisms identified from



one specimen. Many pathogens are not tested for antibiotic susceptibility; in these cases, only the pathogen name appears in the record.

The structure of a HL7 microbiology message allows the system to place the name of an organism (if culture positive) in the TEST NAME field. In these records, the TEST RESULT field will be blank or contain extraneous characters (% , & , etc). Methods can be developed to identify pathogens of interest. In [Table 1](#), there were two tests ordered for one specimen. The first SET ID for each test labels the test performed in the TEST NAME field and does not indicate a result. The subsequent SET IDs below the first SET ID are text fields showing what pending results laboratory technicians have entered during interim growth periods (for example, 24 hour and 48 hour read). An ACCESSION NUMBER can have numerous TEST ORDERED associated with it, corresponding to the multiple tests performed per specimen.

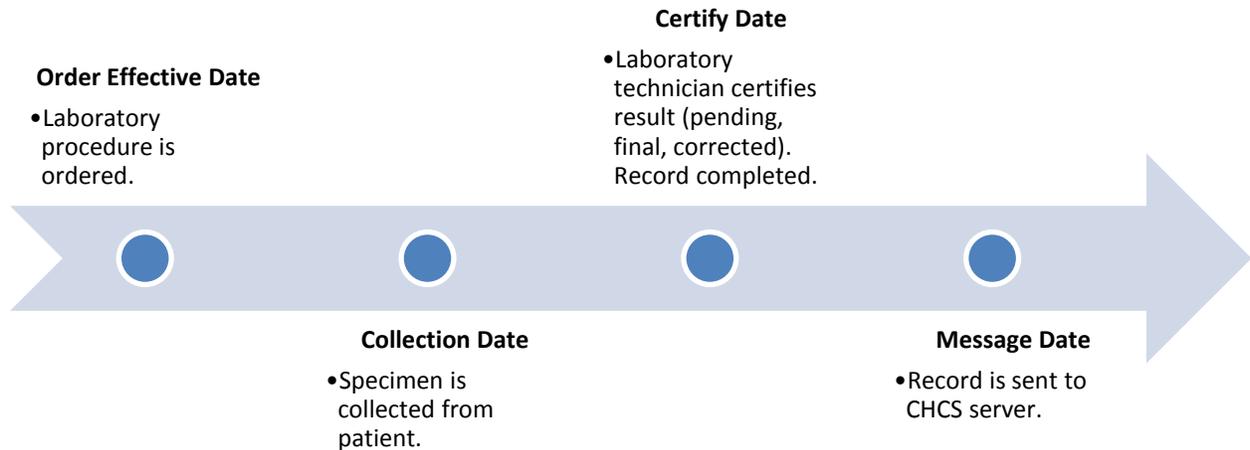
Results of quality assurance tests are also present in the HL7 microbiology dataset. Such records may include text strings such as \*QA\*, \*QC\*, \*LIO\*, \*CAP\*, or \*INTEROP\* in the TEST ORDERED, TEST NAME, TEST RESULT, or CLINICAL COMMENTS fields. Other indicators of quality assurance testing include “ZZZ” appended to the TEST NAME value. Records of laboratory quality assurance tests are not typically included in analyses for public health surveillance.

### Date/Time References

There are five date and time fields within the HL7 microbiology dataset. Historically, the MSG DATE has been used as part of the criteria to pull data from CHCS. The MSG DATE and MSG TIME fields indicate when the data were sent to the central server. Data for epidemiologic projects primarily use the ORDER EFFECTIVE DATE to capture the timeframe when the patient is likely to be symptomatically ill, thus prompting the order of a laboratory test to support the diagnosis. ORDER EFFECTIVE DATE is the date the provider makes the laboratory order effective. It is possible for a provider to enter a laboratory test order to be effective for a future time frame, which may be common practice for chronic or inpatient illnesses. It is also possible that the patient may wait several days before going to the laboratory to have a specimen collected. Therefore, the applicability of using ORDER EFFECTIVE DATE may vary. An alternative date that could be used for analysis is COLLECTION DATE. [Figure 2](#) provides a timeline of dates associated with laboratory tests.



Figure 2. Timeline of dates in HL7 microbiology



[Table 2](#) reflects methods used to determine timeframes within the HL7 microbiology database. Each date/time pair references one action towards the completion of the laboratory test. A time span was created when comparing one date/time variable against another date/time variable.

**Table 2. Timeframes in HL7 data**

Date/Time Variable Range	Interpretation
Message – Order_Effective	Determines timeframe from when test was ordered to when an HL7 message was generated at the local CHCS host
Order_Effective – Date_Of_Birth	Age of member at date of test order
Collection – Date of Birth	Age of member at date when specimen was extracted
Certify – Order_Effective	Timeframe from test ordered to results
Certify – Collection	Timeframe from when sample was taken to results



## Strengths

### Timeliness

Several date and time fields are included in the HL7 microbiology data, including MSG DATE/TIME, ORDER EFFECTIVE DATE/TIME, REQUESTED DATE/TIME, COLLECTION DATE/TIME, and CERTIFY DATE/TIME. To assess the timeliness of the data, the ORDER EFFECTIVE DATE (date the order was placed into CHCS by the provider) was compared to the MSG DATE (date the HL7 message was generated by CHCS) to estimate the time between patient encounter and receipt of data at DHSS. On average, it takes about two days for an HL7 message to be generated and received by the EDC. This indicates that the timeliness of reporting is within acceptable ranges for DON and DOD surveillance activities. Future analysis and assessment may further define lag times in relation to particular MTFs, testing, or outcomes of interest.

### Completeness

Systematic limitations, such as missing records due to atypical data processing, can be addressed through the MHS helpdesk. HL7 data are required to be sent to the central CHCS server by MTFs in support of Veterans Administration (VA) and other data sharing initiatives. Improvements on completeness and reliability of the data, therefore, are vital in the MHS process. Isolates can be tracked MHS-wide from a central location, which is critical for local MTFs as it is difficult to retrieve pathogen data from the local CHCS.

Cultures generally provide a high sensitivity and specificity for laboratory testing. Disease and case burden, therefore, can be determined for many pathogens of interest. It should be noted that culture results may yield pathogens not truly responsible for a patient's illness. For example, bacteria isolated from a wound culture may be colonizers and not the infectious agent. Finally, antibiotic resistance and susceptibility results can be reviewed for those pathogens where laboratories have run these tests.

### Organization and Structure

The fields within HL7 microbiology that are of highest importance for disease surveillance are generally complete and contain expected values. When identifying pathogens of interest, multiple considerations must be taken into account to ensure capture of all test results. The EDC has developed several algorithms to restructure the HL7 microbiology data such that rapid analysis of diseases, pathogens, and antibiotic resistance patterns can be completed.

### Public Health Applications

The HL7 microbiology database provides the ability to establish a disease timeframe based on actual laboratory results. As technology advances, laboratory tests have improved specificity and sensitivity compared to previous methods. Although a laboratory result alone does not indicate a diagnosis, it can be used in combination with clinical symptoms or conditions to determine a more accurate diagnosis. Existing surveillance methods suffer from severe and



systematic underreporting and undercounting. Understanding of the limitations of the HL7 microbiology data has led to more accurate estimation of case burden for several diseases of interest.



## Limitations

### Complexity

The layout of the HL7 microbiology database requires familiarity of the test methodology and parameters in order to conduct proper analysis. For example, a single test can have multiple lines due to multiple tests within a panel, or results at specific time intervals as seen with cultures. Methods need to be applied to ensure the final results are captured, and that each test is analyzed appropriately.

### Completeness

Incomplete demographic information can limit the ability to describe disease burden demographically. Complementary databases, such as personnel records, may be used to supplement HL7 microbiology data for more complete demographic information.

Since MTFs need to be on CHCS so that records are captured in the HL7 database, medical laboratory procedures conducted aboard ships or at field operations are not captured. Laboratory tests for a service member aboard a ship may be seen in the HL7 databases if 1) the ship-based provider referred member to a shore-based facility, or 2) the severity of illness deems the service member to be transported to a shore-based CHCS facility. Shipboard or forward deployed laboratories may also send specimens to shore-based facilities for testing, which would be captured. Additionally, laboratory tests performed outside the MHS are not captured in CHCS unless results are entered manually by laboratory technicians.

### Generalizability

HL7 microbiology data are generated from the medical laboratory records of military service members, their family members, and other beneficiaries. This population differs from the general US population in many ways, including average age, gender distribution, and having universal access to healthcare. Active duty service members also differ from the general population in terms of overall physical fitness and health status. These differences limit the comparability to the general US population.

### Inconsistency

The format of CHCS includes free text fields (e.g., TEST NAME, TEST RESULT, RESULT NOTES), resulting in variations between entries and potential difficulties in data interpretation. Free text fields limit the ability to easily determine testing type, test results, and reference ranges. Non-standardized naming and resulting conventions hinder the ability to use standardized syntax. Methods have been developed to overcome these barriers and flag results of interest. The coding involved in this process can be extensive and requires regular re-evaluation.

Inconsistencies also exist between MTFs. MTF laboratories determine testing capabilities and control what providers can see on the CHCS laboratory order screen. Also, TEST NAME and TEST RESULT options in CHCS (pull down menus) are controlled at the local CHCS host,



allowing variation in recording of test names and results between each MTF's CHCS. This increases the need for extensive coding to determine case definitions, ensuring comprehensive case capture.

### Data Interpretations

While HL7 microbiology data may be useful to identify laboratory confirmed cases of illness, not all cases will be identified if physicians elect to treat patients presumptively without laboratory confirmation. Clinical practice may vary between providers, relying on symptomatic examination versus diagnostic testing. For example, during the influenza season, providers may not order confirmatory tests for patients who present with influenza-like-illness symptoms, and would opt to treat presumptively instead.



## All Data Fields (Variables)

### Automatically Populated Fields

There are several types of automatically populated fields in the microbiology data that are created at different levels, including the CHCS host, each MTF, and each laboratory work section.

When a facility registers within CHCS, several variables are automatically created for use in identification of the facility per DHSS. These variables include: PERFORMING DMIS ID, PERFORMING DMIS FACILITY NAME, PERFORMING FACILITY SERVICE, REQUESTING DMIS ID, REQUESTING FACILITY NAME, and REQUESTING FACILITY SERVICE.

Each DOD beneficiary is registered in the Defense Eligibility Enrollment Reporting System (DEERS) under the SPONSOR ID, which feeds into CHCS. When a patient presents at a medical facility, the SPONSOR ID is entered and their name is chosen from a drop down list of dependents under that SPONSOR ID. The following patient demographic fields are automatically populated after this selection if they were entered when the patient was registered in DEERS: DATE OF BIRTH, ETHNICITY, FMP, GENDER, MARITAL STATUS, PATIENT CATEGORY, SERVICE, PATIENT ID, RACE, and SPONSOR ID. If these data are not present in the system, a designated unknown value is entered; therefore, records should not have missing values for these fields. Administrative personnel at the MTF do have the ability to edit records at the time of visit.

As records are created, edited and completed, several variables are created by the CHCS clock when a specific action is taken, such as order entry or certification of a test result. These variables include: ORDER EFFECTIVE DATE/TIME, COLLECTION DATE/TIME, REQUESTED DATE/TIME, and CERTIFY DATE/TIME. These can be changed if necessary by the laboratory staff, but is not common practice. MSG DATE, MSG ID, MSG TIME, and MSG SENDING FACILITY are created and assigned when the message (record) is sent to the CHCS server.

The CHCS host or the MTF may automate variable selections prior to a physician selecting a test from the drop-down list or lookup tables. After a physician chooses a test for the TEST ORDERED field, a variation of options for TEST NAME could be selected. Based on the physician's request, the SPECIMEN SOURCE, BODYSITE COLLECTION SAMPLE, MEPRS CODE, CPT CODE DATA, and NO OF CPT CODES are selected for each TEST NAME. These can be changed if necessary by the laboratory staff.

The laboratory results and each reference can be entered by the laboratory staff or automated by the laboratory CHCS coordinator or the specific laboratory equipment used. Specific models of laboratory equipment can be specialized to indicate the normal ranges of values, label abnormalities, and create syntax per each specimen's result. Depending on each laboratory



facility and equipment used, the TEST RESULT, RESULT NOTES, REFERENCE RANGE, ABNORMAL FLAG, SENSITIVE RESULT FLAG, and UNITS OF MEASURE could be automated or manually entered by staff.

## Variable Descriptions

Observations and frequencies below are based on HL7 microbiology data from calendar year 2011. There were 53 variables in the dataset.

### ACCESSION NUMBER

ACCESSION NUMBER is a combination of 1) the collection date in an YYMMDD format, 2) two or three alpha characters, and 3) a numeric value. The numeric value can range from 1 to 99999. The ACCESSION NUMBER is created for each unique biological specimen collected from a patient. Different microbiology tests from the biological specimen can have the same ACCESSION NUMBER. These numbers could be recycled throughout the day; therefore, ACCESSION NUMBER should not be used on its own to identify a unique record. ACCESSION NUMBER may be used to determine tests ordered for a patient in conjunction with the SPONSOR ID, FMP, and TEST ORDERED.

### BODYSITE COLLECTION SAMPLE

The BODYSITE COLLECTION SAMPLE refers to where the specimen was taken from the patient. This field is associated with the SPECIMEN SOURCE to determine specifically where the sample was taken.

### CERTIFY DATE

The CERTIFY DATE is the date when a laboratory technician certified the results into CHCS or certified changes to the results. Unlike the ORDER EFFECTIVE DATE, there may be differences between the values for CERTIFY DATE for each SET ID since there are differences in when tests are performed and results are available. The CERTIFY DATE is formatted YYYYMMDD. The field does not have missing values and contains all valid dates. The value of CERTIFY DATE should be between ORDER EFFECTIVE DATE and MSG DATE.

### CERTIFY TIME

The field represents the time component of the CERTIFY DATE formatted within a 24 hour cycle. The timeframe is from 0001 to 2359, and all times are valid entries.

### CLINICAL COMMENTS

The CLINICAL COMMENTS is a free text field which allows the provider or laboratory technician to add additional information regarding the patient's symptoms, contact phone numbers, specimen media, or instructions on procedures for a test. Quality assurance tests may also be identified using the CLINICAL COMMENTS field. Records containing text strings such as \*QA\*, \*QC\*, \*CAP\*, \*LIO\*, or \*INTEROP\* all indicate quality assurance tests and should be removed from analysis. 93.7% of records had blank values for this field.

### COLLECTION DATE

The COLLECTION DATE is the date when the specimen is taken from the patient. This value



should be between ORDER EFFECTIVE DATE and the CERTIFY DATE. The COLLECTION DATE is formatted YYYYMMDD. All records had valid dates.

### COLLECTION TIME

The field represents the time component of the COLLECTION DATE formatted within a 24 hour cycle. All records had valid times and there were no blank entries.

### CPT CODE DATA

The CPT CODE DATA is an alphanumeric field which identifies a particular laboratory test by the Current Procedural Technology (CPT) code. The CPT code is assigned by the American Medical Association (AMA) and is used to identify medical, surgical, and diagnostic services. It is used to communicate uniform, consistent information about medical services and procedures between physicians, coders, patients, accreditation organizations, and payers for administrative, financial, and analytical purposes.

The CPT CODE DATA variable is formatted #####\#\#\AD. The first group of characters defines the CPT code. For example, for influenza tests, the main codes are 87400 (Influenza A or B, each), 87804 (Influenza), 87276 (Influenza A Virus), 87275 (Influenza B Virus), and 86710 (Antibody, Influenza Virus). The second portion is a modifier code which indicates the accession area and work element. The third portion is a status code. A regional CHCS site maps a CPT code to a particular methodology or technique. CPT codes are assigned at various levels to the CHCS test files when the laboratory sets up the procedure. Tests that do not have a specific CPT code may be given unlisted procedure/service codes defined for the specific types of test (immunology, chemistry, microbiology, hematology, etc.).

### DATE OF BIRTH

The DATE OF BIRTH field is formatted YYYYMMDD. If only the year is known, CHCS enters zeros for the month and day. DATE OF BIRTH is a required field within CHCS. There are no blank values and limited false dates.

### ETHNICITY

ETHNICITY is an alphanumeric field with six possible values: 1 (Hispanic), 2 (Southeast Asian), 3 (Filipino), 4 (Other Asian Pacific Islander), 9 (Other), and Z (Unknown). In 2011, the majority of records indicated Other for ETHNICITY (52.8%), followed by Unknown (40.5%). There were no blank values. The most frequent value other than Unknown or Other was Hispanic (3.8%). These results indicate that ETHNICITY is not consistently reported and may be self-identified, which limits the ability to identify disease trends in minority groups.

### FMP

FMP designates the relationship of the patient to the sponsor. The FMP variable is a numeric value, with two digits. In 2011, 42.5% of records had an FMP of 20 (sponsor) followed by 33.3% with an FMP of 30 (spouse of sponsor). Other possible values include 01-19 (child of sponsor, numbered in age order). No records were missing an FMP value. Unknown entries are labeled as 99.



## GENDER

There are three possible values for GENDER: M (Male), F (Female), and X (Unknown). 63.4% of records in 2011 indicated Female for GENDER, and 36.6% indicated Male. Only 5 records out of nearly 8 million records in 2011 had a value of Unknown. There were no blank values for this field.

## MARITAL STATUS

There are nine possible values for MARITAL STATUS: A (Annulled), D (Divorced), I (Interlocutory Decree), L (Legally Separated), M (Married), N (Never Married), S (Single/Not Married), W (Widow/Widower), and Z (Unknown). 39.0% of records had a MARITAL STATUS of Married, followed by 37.2% of records with Unknown and 18.9% of records with Single/Not Married. There were no blank values.

## MEPRS CODE

The MEPRS (Medical Expense Performance Reporting System) CODE is a four letter code that indicates where within a MTF the patient was seen when the sample was collected. The first letter indicates the general areas: A (Inpatient), B (Outpatient), C (Dental), D (Ancillary), E (Support Services), F (Special Programs), and G (Medical Readiness). The entire MEPRS CODE is used to indicate the specific unit or ward, such as Family Practice, General Surgery, or Pediatrics. The MEPRS CODE field is useful for tracking where people were seen within the MTF, which can affect the interpretation of the data. The majority of records in the HL7 microbiology dataset have a MEPRS CODE that begins with B (Outpatient). The most frequent MEPRS code in 2011 HL7 microbiology data was BIAA (22.6%), indicating that the patient was seen in the Emergency Department.

## MSG DATE

MSG DATE is when records are sent from the local CHCS to the central server. This field is in the format of YYYYMMDD. There are no missing values and all are valid dates, and all dates are either the same date or after the CERTIFY DATE. Some MTFs send messages in batches, therefore, the time or date portions may not correlate to the actual transaction time.

## MSG ID

The MSG ID is an alphanumeric code assigned to each batch of messages based on when the message is sent from CHCS to the central server. The MSG ID is not unique to each record; each batch of messages is assigned one MSG ID. The format of MSG ID varies by MTF and may include numbers, letters, numeric code that identifies the MTF, or the function of the message (i.e. RESCHED-057342).

## MSG SENDING FACILITY

This field is in the format of A####, F####, HP####, or N#####. This field allows analysts to identify and track problems that may arise in the transfer of messages from the MTFs through DHSS to the EDC. There are no blank values.

## MSG TIME

MSG TIME is the time component of the MSG DATE formatted within a 24 hour cycle. This



field has four numeric characters, and there were no blank values.

### NO OF CPT CODES

NO OF CPT CODES is a numeric field which lists the number of CPT codes used for each test performed. Two-thirds of the records had one CPT code used. The number of CPT codes is determined at each regional location. This field is not used within the EDC evaluation process.

### ORDER EFFECTIVE DATE

ORDER EFFECTIVE DATE is the date that the laboratory order enters the CHCS system and indicates when the laboratory test was actually ordered. All entries are valid dates, and there were no blank values.

ORDER EFFECTIVE DATE may be used to approximate when a patient was ill, to analyze the difference in time between when the order was issued and when the sample was collected, and to assess the length of time between the dates the test was ordered and when the data are available for use in the EDC.

### ORDER EFFECTIVE TIME

ORDER EFFECTIVE TIME represents the time component of the ORDER EFFECTIVE DATE formatted within a 24 hour cycle. All times are valid entries.

### ORDER NOTES COMMENTS

ORDER NOTES COMMENTS is a text field which allows the provider to provide notes or comments that accompany the test ordered. All records have blank values in this field.

### ORDER NUMBER

The ORDER NUMBER is a numerical code of eleven digits (xxxxxx-xxxxx) unique to each order, but not unique for each record. The first six numbers are the date. The last five numbers are consecutive per the location. An order can have multiple records that correspond to changes made to the order (e.g., changes in test, cancellations), or refer to multiple parts to a test (e.g., results for influenza A and influenza B). All changes appear as individual records with the same ORDER NUMBER. It is a plausible way to track a patient but not useful for identifying unique records.

### ORDERING PROVIDER

ORDERING PROVIDER indicates the name of the ordering physician. It has three components separated by “,”: Last Name, First Name, Middle Initial. It is structured to facilitate analysis but could be separated if necessary.

### PATIENT CATEGORY CODE

The PATIENT CATEGORY CODE (PATCAT) is an alphanumeric code that indicates the patient’s relationship to the Uniformed Services. The first letter of the code refers to the sponsor’s service branch affiliation and is one of the following values: A (Army), B (National Oceanic and Atmospheric Administration), C (Coast Guard), F (Air Force), K (Other Beneficiaries of the Federal Government), M (Marine Corps), N (Navy), P (US Public Health



Service), and R (NATO Recipient). It is followed by two digits that correspond to the status of the sponsor within the service and the patient's relationship to the sponsor. For example, M11=Marine Corps Active Duty Service Member, A31=Army Retired Active Duty, and N41=Navy Dependent of Active Duty. A complete list of PATCAT codes should be obtained from DOD resources. In 2011, the most frequent PATCAT was A41 (Army Dependent of Active Duty) (20.3%) followed by A11 (Army Active Duty Service Member) (11.7%). Less than 1% of records had missing values for this field.

### **PATIENT ID**

The PATIENT ID is intended to serve as a unique identifier for each patient. The format for PATIENT ID is a nine digit numeric listing. The PATIENT ID is the patient's SSN when available; however, the accuracy of the variable cannot be assured based on the EDC's observations and analyses. In place of PATIENT ID, SPONSOR ID and FMP should be used to identify individual patients. This value was missing in 0.02% of records in 2011. It is important to preserve the entire PATIENT ID when importing the data into SAS or other analysis programs. The PATIENT ID variable needs to be imported as a character field so that leading zeros are not dropped.

### **PERFORMING DMIS FACILITY NAME**

This field is the text translation of the PERFORMING DMIS ID field and is assigned by DHSS. Since the field is a translation of PERFORMING DMIS ID, it will be missing when that field is missing in the record. This field was missing in 11.9% of records in 2011.

### **PERFORMING DMIS ID**

The PERFORMING DMIS ID is a four digit code that identifies the MTF that performed the laboratory test. This code allows for grouping of MTFs based on geographic location, as well the ability to identify parent/child relationships between installations. This field was missing in 11.9% of records in 2011.

### **PERFORMING FACILITY SERVICE**

The PERFORMING FACILITY SERVICE field indicates the branch of service with which the MTF is associated. This value is determined from the DMIS ID code list provided to DHSS by the EDC. It will be missing from a record when the PERFORMING DMIS ID is missing. The possible values are A (Army), F (Air Force), and N (Navy). This field is useful for limiting observations to those specific to a service, allowing for analysis of disease burden in facilities of that particular service. This field was missing in 11.9% of records in 2011.

### **PERFORMING LOCATION FACILITY**

PERFORMING LOCATION FACILITY indicates the name of the MTF where the test was performed. Problems may be encountered if the text was entered incorrectly when the facility was registered in the system or due to inaccurate mapping between DMIS ID and facility name. There were 302 different facilities names listed within the microbiology dataset, with no values missing.



### PERFORMING LOCATION WORK CENTER

The PERFORMING LOCATION WORK CENTER field indicates the work center within the MTF that provided the testing service. This field is a relatively unstructured text field with many possible values. Over 400 performing work centers were identified in 2011. These locations are usually laboratories mapped according to the PERFORMING DMIS ID.

### RACE

There are six possible values for RACE: C (White), M (Asian or Pacific Islander), N (Black), R (American Indian or Alaskan Native), X (Other), and Z (Unknown). No records had blank values for RACE in 2011 and 38.8% of records were categorized as Unknown, followed by White (35.9%) and Other (13.1%). A high frequency of Unknown or Other values limits the ability to use the data to look at diseases or conditions according to race.

### RECORD TYPE

RECORD TYPE identifies the database that the records are from (e.g., microbiology, chemistry). There are no blank values for this field and the field is three characters in length. All HL7 microbiology records have a value of "LMI" (for microbiology laboratory) in this field.

### REQUESTED DATE

The REQUESTED DATE is a date field formatted as YYYYMMDD. This field is not frequently used for data analysis, and does not have missing or invalid values. The timeframe of this value should be between ORDER EFFECTIVE DATE and COLLECTION DATE.

### REQUESTED TIME

The field represents the time component of the REQUESTED DATE formatted within a 24 hour cycle. The timeframe is from 0000 to 2359, and all times are valid entries.

### REQUESTING DMIS FACILITY NAME

This field is the text translation of the DMIS ID provided in the REQUESTING DMIS ID field and is assigned by DHSS, although this may be done inconsistently. This field was missing in 12.0% of records in 2011.

### REQUESTING DMIS ID

The REQUESTING DMIS ID is a four digit code that identifies the MTF that requested the laboratory test. This code allows for grouping of MTFs based on geographic location, as well as the ability to identify parent/child relationships between installations. This field was missing in 3% of records in 2011.

### REQUESTING FACILITY NAME

The REQUESTING FACILITY NAME field indicates the name of the MTF where the order originated. Problems may be encountered if the text was entered incorrectly when the facility was registered in the system. The field allows tracking of orders from origin to where they were performed. Less than 1% of records had blank values for this field.



### REQUESTING FACILITY SERVICE

The REQUESTING FACILITY SERVICE field indicates the branch of service with which the MTF is associated. This value is determined from the DMIS code list provided to DHSS by the EDC. It will be missing from a record when the REQUESTING DMIS ID is missing. The possible values are A (Army), F (Air Force), and N (Navy). This field is useful for limiting observations to those specific to a service, allowing for analysis of disease burden in facilities of that particular service. 12.0% of records were missing this value in 2011.

### REQUESTING WORK CENTER NAME

The REQUESTING WORK CENTER NAME field indicates the ward or clinic within the MTF that requested the laboratory test. This field is an unstructured text field with many possible values. Possible entries include DMIS ID number, clinic wards, service centers, and unknown/other MTF locations.

### RESULT NOTES

RESULT NOTES is a free-text field which allows the laboratory technician to add information about the result, a recommendation for additional testing, or the interpretation of the laboratory result. This field can either be an automated drop down menu with the TEST NAME via CHCS or it can be entered as free text. Nearly all records had missing values for this field.

### RESULT STATUS OBX

RESULT STATUS OBX is a free-text field which shows the status of the test performed. There are three entries which could be used, and always go in the following consecutive order: P (Preliminary), F (Final), and C (Correction). Each SET ID for a test should have a status of F. Records with a status of P or C for a test may also be present, either along with or in the absence of F records. Should a test have numerous RESULT STATUS OBX values, each record will have the same SET ID, TEST NAME, and TEST ORDERED. If a SET ID has multiple test statuses available, the following hierarchy should be used to determine which record to use: C→F→P. An entry is corrected (C) when it is amended due to a change in interpretation, operator error, wrong test ordered, or test was performed under wrong patient. The value of “F” was the majority listed within the dataset, followed by “P”, and then “C.” Less than 1% of records had missing values in this field.

### SENSITIVE RESULT FLAG

The SENSITIVE RESULT FLAG permits the laboratory technicians to record results-dependent codes for classifying the observation in CHCS. Nearly all records had missing values for this field.

### SENSITIVITY

The SENSITIVITY field is a character field which contains one of three values: R (Resistant), I (Intermediate), or S (Susceptible). The value is based on the numeric antibiotic susceptibility result in the TEST RESULT field for the corresponding antibiotic in the TEST NAME field.

### SERVICE

The SERVICE field refers to the service branch of the Sponsor. Possible values are: A (Army), B



(National Oceanic and Atmospheric Administration), C (Coast Guard), F (Air Force), K (Other Beneficiaries of the Federal Government), M (Marine Corps), N (Navy), P (US Public Health Service), and R (NATO Recipient). In 2011, 45.3% of records had a value of A (Army) for SERVICE, followed by F (Air Force) (21.4%) and N (Navy) (18.1%). No records had missing values for this field.

### SET ID

The SET ID is a numeric field which can range from 1-9999. This field is a mandatory field populated by CHCS; therefore, there were no missing values. The numbers show the logical order of arrival of data within an HL7 message. Should an entry have a change in its resulting status (from pending to final, or final to corrected), the SET ID will remain the same for that test entry.

### SPECIMEN SOURCE

The SPECIMEN SOURCE is a text field which indicates where the specimen was taken from on the patient. This field is useful to determine if the proper protocol was used for a laboratory test. Also, the laboratory does not always differentiate between the specimen source and the specimen location. This is seen with entries that indicate tissue of the patient's body, such as Both Eyes, Buttocks, Mouth, Oral Cavity, or Toe. Less than 1% of records had missing values in this field.

### SPONSOR ID

The SPONSOR ID field corresponds to the SSN of the sponsor and is in a 9-digit format with no dashes. The SPONSOR ID is not sufficient to identify a unique patient, but may be used in conjunction with the FMP as a unique patient identifier. Only 9 out of 8 million records in 2011 had blank values for SPONSOR ID. It is important to preserve the entire SPONSOR ID when importing the data into SAS or other analysis programs. The SPONSOR ID variable needs to be imported as a character field so that leading zeros are not dropped.

Not all SSNs are ones given by the Social Security Administration. If the patient does not hold a valid SSN, a pseudo SSN number is created. The pseudo SPONSOR ID may begin with 800 or 900, followed by a date. If the number is already assigned to another patient, it could begin with 801 or 901. Also, records for quality assurance/control tests conducted in the laboratory may use SSN-like identifiers in the SPONSOR ID field. The SPONSOR ID for these procedures may resemble a pseudo-SSN, arbitrary identifiers such as 777777777, or three consecutive zeros. These tests will have labels such as Ztest, QC, Quality Control, PSR, CAP, or Non-human (NH, #).

### TEST NAME

The TEST NAME is a text field showing which test was used per the samples provided. This value is usually selected from a pull-down listing from the TEST ORDERED variable. This field will never have missing values because TEST NAME is automated by the regional CHCS system. The TEST NAME includes entries such as tests to be performed, quality controls, temperature, and status of culture growth. When TEST NAME contains the name of an organism, this indicates that the test had a positive result for that organism. Quality control



tests may also be viewed in this field and will often have “ZZZ” prior to the actual test name.

The variance between test names suggests the fields are automated by a regional CHCS system, not the main location. A test procedure can be specific or general. A test name can refer to the procedure type, such as convalescent, cultures, or antibody testing. Therefore, a test name could reflect what an expected result should be. For example, results for a titer should be numeric, while results for organism-specific tests can be positive or negative or have sub-typing.

### TEST ORDERED

The TEST ORDERED identifies the requested observation, test, or panel. Each regional CHCS location has the autonomy to determine the criteria for each test ordered. Therefore, the TEST ORDERED field can have different grouping of tests per CHCS regions. The TEST ORDERED value is repeated among all records for tests associated with it according to the ORDER NUMBER. A provider can use a pull-down menu to determine the test(s) to be performed on a specimen. This shows all available tests per each test ordered.

### TEST RESULT

TEST RESULT is an alphanumeric field which shows the results of a test. The TEST RESULT field can have positive or negative results, control values, dates, reorders, references to comment fields, growth status, or that the test was not performed due to inadequate results or insufficient quality. In 2011, 27% of records had blank values for TEST RESULT. A blank entry is typically observed when either a result or test identification was seen in the adjacent TEST NAME field. There are a wide variety of values in this field, including misspellings and slang language, indicating that the results are either automative per each regional CHCS location or entered manually. Many of these variations show the same result formatted differently, such as Positive, POSITIVE, POSTIVE, POS, and so forth. Currently, CHCS is in the process of regulating the regional CHCS locations to create one specific text for each TEST NAME outcome. This would limit the variation of TEST RESULT significantly.



## Abbreviations

AMA	American Medical Association
CBC	Complete Blood Count
CFU	Colony Forming Units
CHCS	Composite Healthcare System
CPT	Current Procedural Technology
DHSS	Defense Health Services System
DMIS	Defense Medical Information System
DOD	Department of Defense
DON	Department of the Navy
EDC	EpiData Center
EIA	Enzyme Immunoassay
FMP	Family Member Prefix
GEIS	Global Emerging Infections Surveillance and Response System
HL7	Health Level 7
ICD-9-CM	International Classification of Diseases, 9 <sup>th</sup> Edition, Clinical Modification
MEPRS	Medical Expense Performance Reporting System
MHS	Military Health System
MTF	Military Treatment Facility
NMCPHC	Navy and Marine Corps Public Health Center
SSN	Social Security Number



## Definitions

### Department of Defense (DOD)

A DOD beneficiary is a sponsor, dependent, or civilian seen within any military treatment facility.

### Department of the Navy (DON)

DON may refer to *a*) records pertaining to sponsors and family members of either the Navy or Marine Corps seen within a military treatment facility, or *b*) records pertaining to sponsors and family members of any service seen within a Navy military treatment facility.

### Defense Medical Information System (DMIS)

DMIS IDs are recognized within the DOD as the controlling standard for both medical and military facility identification and cost/workload classification. DMIS IDs are used throughout the MHS and worldwide in both healthcare and non-healthcare systems.

