



IMMUNIZATION INFORMATION SYSTEMS PATIENT-LEVEL DE-DUPLICATION TEST CASE DEVELOPMENT, TESTING & UTILIZATION

National Center for Immunization and Respiratory Disease (NCIRD)
Immunization Information Systems Support Branch (IISB)

Version 1.0
July 2, 2013

1	EXECUTIVE SUMMARY	3
1.1	BACKGROUND.....	3
1.2	INTENDED AUDIENCE	4
1.3	METHODOLOGY	4
1.4	WORK PRODUCTS	5
1.5	SUMMARY OF FINDINGS AND RECOMMENDATIONS.....	5
2	PATIENT DE-DUPLICATION TESTING AND TEST CASE DEVELOPMENT	7
2.1	USE OF 2002 TEST CASE TOOLKIT	7
2.2	RESULTS OF 2002 TEST CASE ANALYSIS.....	8
2.3	FUTURE CONSIDERATIONS.....	15
3	PATIENT DE-DUPLICATION TEST CASE MATRIX.....	15
3.1	INTRODUCTION	15
3.2	ORGANIZATION AND KEY FEATURES	16
3.3	SUPPORT FOR MULTIPLE TEST STRATEGIES	25
	APPENDIX A - PANEL MEMBERSHIP	28
	THE CDC EXPERT PANELISTS	28
	CDC EXPERT REVIEWERS.....	31
	NORTHROP GRUMMAN PUBLIC HEALTH CONTRACTOR PERSONNEL.....	33

1 Executive Summary

1.1 Background

This document represents the Centers for Disease Control and Prevention (CDC) Patient-Level De-Duplication expert panel's final report and recommendations regarding Immunization Information System (IIS) patient level de-duplication testing and test case development. For additional information regarding the full scope of work of the CDC expert panel, please consult the companion publication, the IIS Patient-level De-duplication Best Practices.

Patient-level de-duplication testing is an important area. The strategic employment of test cases can help systematically improve IIS operations, improve the sensitivity and specificity of patient de-duplication processing, reduce the need for manual reviews, and meet the objective for greater automation and data quality. Well-constructed, best-practice-oriented test cases represent a significant aid and resource to the IIS national practice community.

With the expansion and evolution of IIS, public health jurisdictions have been called upon to substantially improve their ability to prevent, detect, and correct duplicate patient records; however, the current tools and roadmaps to accomplish this are limited. The strategic improvement of patient-level de-duplication capabilities requires on-going monitoring and measuring of IIS operations to identify areas requiring improvement, as well as the systematic employment of patient de-duplication testing to establish the functionality required to accurately support business process goals. The purpose of this effort was to provide a uniform resource to support these efforts.

To strategically improve IIS patient de-duplication capabilities, there are a number of considerations: 1) monitoring of IIS operations to identify trends and anomalies; 2) the selection or construction of test cases which will meaningfully exercise the logic of an IIS system around areas targeted for improvement; and 3) the execution of test cases in a manner that will reveal system weaknesses for correction and improvement.

Testing and the development of test cases is the dominate method used to verify and improve the behavior of complex computer systems. Patient de-duplication testing is the process of exercising an IIS in a controlled environment to examine if its behavior complies with the patient identity business requirements of a public health jurisdiction. In short, the purpose of testing is to reveal faults in a system. Testing can only show the presence of faults, not their absence.

The last formal examination of IIS patient-level de-duplication methods, along with the development of patient de-duplication tools, was in 2002, and much has changed since that time. In particular, the types of patient de-duplication problems encountered by IIS have evolved and changes in health information technology have produced an environment conducive to greater data interoperability. The CDC notes that given the importance of IIS in national meaningful use objectives, there is much to be gained from a fresh examination of patient de-duplication test cases and best practices.

1.2 Intended Audience

This document is targeted toward programmatic, technical, and operational experts who are involved in creating or maintaining an IIS. The document intends to bridge the gap between technical and program staff so they can have a mutual understanding of the key issues involved in patient de-duplication testing.

1.3 Methodology

The project established a Patient Data De-duplication Expert Panel. The panel consisted of 14 Subject Matter Experts (SMEs) and Expert Reviewers from the following organizations:

- American Immunization Registry Association (AIRA)
- Indian Health Service (IHS)
- EHR vendors
- IIS programs and vendors
- Academic institutions

The work of the expert panel was performed during the period of August 2011 through March 2013. Work was assigned to one of two roles: 1) SMEs for primary content generation and 2) Expert Reviewers for content and product review.

As an expert panel, the group represented decades of profound expertise in IIS patient de-duplication procedures, tools, methods, and system administration. The membership of the expert panel is detailed in Appendix A.

The expert panel conducted a comprehensive Patient-Level De-Duplication National Practice Assessment (NPA). The NPA examined and reported trends, problems, and approaches currently being taken on a national basis. A peer-reviewed publication is the by-product of this effort. Findings from the NPA are summarized in the best practices document; however, notations and findings regarding testing and test case development are located primarily in this report.

The recommendations for new test cases were supported by the following expert panel activities:

- Review of the relevant literature supporting de-duplication concepts and methods
- Exhaustive examination of the existing de-duplication testing application toolkit developed in 2002
- Review of the National Practice Assessment (NPA) to gauge the needs and methods being used by the IIS national practice community
- Discussion and sharing of presentations on the concepts, tools, and methods in use by their own local jurisdictions
- In-person exercises, group work, and other activities to discuss and examine the various considerations associated with test case development

The expert panel also directly submitted examples of representative, complex, patient de-duplication test cases. These examples appear as illustrations of complex, real-world test cases in the test case matrix and are also discussed in Chapter 3 of this document.

In addition to consensus recommendations for approaches to strengthen patient de-duplication testing and best practices, the expert panel accomplished a significant body of work and developed a number of work product artifacts. The expert panel's definition of consensus did not reflect 100% agreement, but rather "I can live with that and support it."

1.4 Work Products

The work products documented in this report volume include:

- An evaluation of the existing 2002 Patient-level De-duplication Tool Kit
- An examination of the IIS data used for patient de-duplication testing
- An examination of the most typical and challenging patient de-duplication problems
- A review of complex test cases and de-duplication problems where two or more variables may cause identity ambiguities
- Important discussions around the evolution of IIS patient de-duplication testing needs
- Consensus views regarding the construction of new and useful de-duplication test cases
- Observations around the testing "roadmaps" needed to improve IIS patient de-duplication capabilities, reduce the need for manual record reviews, and drive greater collaboration and standardization in the national practice community

In addition to this report, the panel developed a Test Case Matrix that contains the actual evolved test cases and tools for utilizing them.

1.5 Summary of Findings and Recommendations

The CDC IIS patient-level de-duplication expert panel made a number of detailed findings and recommendations to guide the development of an improved set of test cases. Highlights of panel discussions are documented below.

- Modifications to the overall patient de-duplication testing philosophy
 - End-users should be free to pick and choose the test cases that are the most meaningful for improving their individual circumstances.
 - There should not be any requirement to publish or compare test result scores among jurisdictions.
 - Test case de-duplication measures need to go beyond sensitivity and specificity measures in order to gain a more holistic understanding of de-duplication operations and measurements of quality and improvements.
 - The new testing toolkit should be created in an intuitive format for which no formal training is required.

- Determination that a new classification framework is needed for the development of IIS patient de-duplication test cases
 - Patient-level de-duplication test cases can be improved by creating a new eight-category framework to test patient matching and patient de-duplication capabilities. These represent common data processing problems.
 - The rationale for this new framework recognizes that patient-level de-duplication is part of overall data quality management and improvement activities. Revising the mix of test cases to conform to these categories will update them to mirror the nature and proportion of problems actually found in the real world.
 - The framework for looking at test cases needs to go beyond just looking at two fields and should be more approximate to what happens in the real world.

- Specific revisions to test data and test cases
 - Data fields need to be removed and new data fields added.
 - The characteristics of the data need to be modified to reflect current IIS processing circumstances.
 - More test cases are needed.

- Best practice considerations to guide patient de-duplication testing efforts
 - Single, discontinuous efforts exploring IIS patient-level de-duplication and interoperability are not adequate to provide the functionality required to sustain continued improvements in the area of patient-level de-duplication. Accordingly, patient de-duplication testing should be integrated into the continuous, on-going efforts to improve the efficiency and accuracy of automated systems.
 - The expert panel believes that explorations should occur within the IIS national practice community to determine how patient-level de-duplication evolution and approaches can be best accomplished and sustained on an on-going basis by the national practice community. The national direction and language around future IIS patient-level de-duplication needs to be explicitly stated in such a manner that the IIS national vendor community can drive towards more common, uniform solutions, similar to those that are being attained in the area of electronic health records.
 - Specifically, the expert panel believes that patient-level de-duplication IIS support needs to have a more collaborative and interactive web presence to support the needs of the IIS national practice community. It is believed that organizations like the American Immunization Registry Association (AIRA) should play a key role in supporting the national practice community, but it is recognized that significant volunteerism may be required as well as initiative funding in order to assist sustainability.

2 Patient De-duplication Testing and Test Case Development

2.1 Use of 2002 Test Case Toolkit

2.1.1 Approach

A key activity area for the CDC expert panel was providing recommendations to guide the development of new IIS patient-level de-duplication test cases. The CDC expert panel's scope of work specifically included examination of the existing 2002 IIS patient de-duplication scoring application and test cases and recommendations for new test cases featuring updated and expanded data utilization. Additionally, the expert panel was asked to consider specifications for the nature, type, mix, and volume of test cases; recommendations for test case packaging and distribution; recommendations for expanded data de-duplication measures; and guidance regarding future directions for the test case development and support.

Expert panel assignments were structured so that the development of the recommendations for the new test cases was integrated with the discussions of the existing test cases.

A key assumption of the expert panel was that the test cases developed as part of the 2002 De-duplication Testing Toolkit could be retained and built upon.

The panel notes that no test case toolkit or process will be able to exercise all aspects of patient-level de-duplication testing. However, certain types of test cases and testing techniques have been identified within this document that can enable efficiencies.

2.1.2 2002 De-duplication Toolkit Background

In 2002, at the request of the CDC National Immunization Program, a patient de-duplication toolkit containing a test data set was developed for immunization registries to evaluate their de-duplication algorithms. Studies conducted during that time indicated that concern about registry data quality was a key barrier to provider participation and widespread adoption. It was noted that each registry had implemented its own variety of de-duplication algorithm approaches.

This 2002 de-duplication algorithm evaluation toolkit was developed to help all registries assess their own system's ability to prevent or remove duplicate records. The data and procedures in the toolkit helped identify weaknesses in the de-duplication algorithms and thus facilitate improvements. The test data set consisted of test cases representative of known duplicate record problems in real data, based on the information provided by various registry personnel (CDC, 2002). The data and procedures in the evaluation toolkit were designed to help virtually any registry identify weaknesses in their de-duplication algorithms and thus facilitate improvements.

2.2 Results of 2002 Test Case Analysis

The panel developed a matrix of test cases that can be used to perform testing of IIS detection and prevention of duplicate patient records. Features of the new test cases include:

- Expanded fields
- Expanded test case volume
- More realistic sensitivity test cases based on common test case categorizations
- More complex test cases
- Improved measures of sensitivity, specificity, accuracy, precision, and false positive rate
- More effective packaging for distribution and use

2.2.1 Expanded Fields

Fourteen (14) data fields were originally included in the CDC de-duplication test database to evaluate the ability of an IIS to detect patient-level duplicates. Based upon the analysis of the 2002 test data problem set, the expert panel recommended expanding the number of data fields and made other revisions to the test data characteristics to align the data set with current practice conditions.

Data commonality was a consideration. The panel specifically considered the following additional questions:

- Should the data fields be based solely upon core National Vaccine Advisory Committee (NVAC) data or some expanded data set?
- How widespread does the use of certain data elements appear to be in registries?
- How can the new test case data support the testing of basic and more advanced functionality?
- What information from the IIS NPA informs national practice needs?
- How helpful or significant would the inclusion of the data element be either alone or in unison with other data elements to improve de-duplication processes?

The expert panel noted that there is significant variation in both the data that is being maintained as part of each immunization registry as well as variation among the use of data elements for de-duplication purposes.

The use and acceptance of these additional data elements for test purposes was judged to be optional. A panel consensus emerged that test cases should be able to be revised by the end users so that they can be utilized to exercise the capabilities of individual implementations. The expert panel discussed the need for and the use of patient Medicaid number and Social Security Number (SSN). It was noted that using and allowing for these two fields can greatly increase the accuracy of matches. Similarly, allowing for current address can assist in the determination of duplicate records. Accordingly, it was decided that these fields would be included and their use would be at the discretion of each individual IIS site.

The expert panel also noted that there was significant variation in the ability to collect demographic information in IIS. Telephone numbers and other guardian information represent examples of useful data that may not be routinely collected. Mandated limited data sets vary from state to state, and those

organizations not collecting enough information may be unable to be successful in automated de-duplication activities using expanded data sets.

The expert panel additionally considered such things as shifts in demographics that are occurring on a national basis—these shifts affect the data contained in individual fields.

The new Test Case Matrix features a greatly expanded use of test data fields. While the use of these additional data fields is completely optional, the use of additional data fields can sometimes expand the ability to accurately match patient records.

The data fields that are represented in the new test case matrix are illustrated in Table 2.2.1: Test Case Matrix Data Fields, shown below.

Table 2.2.1: Test Case Matrix Data Fields		
Field	Field Description	Characteristics
1	Unique Patient (Client) ID Assigned by Registry	Number between 0 and 999999999
2	Questionable match flag (“Y” indicates this record may be a duplicate and requires human review).	“Y” or “N” (Leave as “N” if flag not used by the registry)
3	Patient Last Name	Text
4	Patient First Name	Text
5	Patient Middle Name	Text
6	Patient Suffix Name	Text
7	Patient Date of Birth	Yyyymmdd
8	Patient Gender	“M” or “F” or “U”
9	Mother Maiden Name	Text
10	Mother Last Name	Text
11	Mother First Name	Text
12	Mother Middle Name	Text
13	Vaccine Code (CVX)	Number
14	Vaccination Date	Yyyymmdd
15	Vaccine Manufacturer (MVX)	Text
16	Medicaid ID	

Table 2.2.1: Test Case Matrix Data Fields		
Field	Field Description	Characteristics
17	Address – Street, City, Zip	
18	Social Security Number	
19	Telephone	
20	Guardian	
21	AKA (alias name)	
22	Patient Organization ID	
23	Medical Record ID	
24	Multiple Births	Y/N
25	Birth Order	1,2,3,4, etc.

2.2.2 Expanded Test Case Volume

The expert panel determined that the volume of test cases should be expanded, and the number of sensitivity and specificity test cases was thereby doubled.

The expert panel considered the questions: 1) How should new clean test cases be generated and 2) Should the new test cases be synthetically created or taken from a sample of real IIS data? One can have SMEs come up with the test cases and data, a purely synthetic approach can be taken, or real-world data can be collected and de-identified which will retain its characteristics. Explorations were made regarding the potential to leverage automated tools to support the generation of test cases, and ultimately this approach was implemented.

It became clear that record pairs, utilized in the 2002 test cases, provided a clear and understandable mechanism for processing. The panel indicated that the use of record pairs was a useful and practical way of conducting tests, in particular where the base records represented “truth” or were considered the “gold standard” against which the test conditions can be compared. Therefore, this feature of the original 2002 test cases was retained. In the 2002 test cases, some of the base records had multiple test records associated with them. In the new test cases, there is a one-to-one ratio of base records to test records.

2.2.3 More Realistic Sensitivity Test Cases Based on Categorization

The categories below illustrate the new framework for characterizing patient de-duplication matching problems. These seven categories represent the most commonly encountered data problems associated with accurate IIS patient matching according to the expert panel.

As illustrated in Table 2.2.3, the inability to correctly match patient can usually be traced to one of the seven “Challenge Categories.” Each challenge category represents a significant data processing problem area.

Table 2.2.3: Test Case Matrix Challenge Category Framework Organization		
Challenge Category	Data Processing Problem Description	Examples
Recording and Typographical Errors	Errors in data entry or data transformation. Insertions, deletions, misspellings, transpositions	Gant vs. Grant. Wian vs. Wain. Ryand vs. Ryan. , DOB reported as 01/02/2012 instead of 01/20/2012
Missing Data	Data is missing from incoming or existing record	No last name. No first name. No DOB. No address, phone, mother’s maiden
Values Changes	Marriage, Divorce, Adoption, Move	True data changes, multiple values reported and/or retained, intentional misinformation, e.g., misrepresentation of DOB for school purposes, reuse of SSN within family or community
Common Values Substitutions	Frequently used substitutions or abbreviations	Nicknames (Robert=Bob), street abbreviations (Ct vs. Court, Southeast vs. SE), suffix = 2 vs. 2 nd vs. II vs. Jr
Special Characters	Hyphenation, punctuation, spaces, or other special characters that may be contained within the data in a field, multiple first- or last-names that may be transposed	O’Malley versus Omalley versus O malley, Martinez-Hernandez vs. Hernandez-Martinez vs. Hernandezmartinez
Extraneous Data	Intermingled data that should appear in a separate fields	Middle name appearing in last name field along with last name. Suffix appearing in last name field (LN=Smith Jr.)
Non-standard Data	Data that does not conform to	Middle name used as first name,

Table 2.2.3: Test Case Matrix Challenge Category Framework Organization		
Challenge	Data Processing Problem	Examples
or Field Use	standards, field definitions, or conventions.	townships interchanged for city, addresses entered as cross streets, placeholder or default values such as first name = "baby", zip code = "99999", last name = "Foster Child", address = "Jail"

Each test case in the Test Case Matrix was categorized accordingly. The expert panel looked at real data from the IIS community to determine the frequency of each type of situation described above. After a number of explorations, the expert panel voted on the approximate percentage of test cases in each category. The test cases were then generated to correspond as much as possible with these percentages. Missing data constitute the greatest percentage, followed by values substitutions, and then recording and typographical errors.

2.2.4 More Complex Sensitivity Test Cases

Complex test case are test cases that represent unique processing situations or that contain two or more "problems" or "ambiguities" that would challenge an automated de-duplication process to make correct decisions.

The expert panel examined the scenarios of the 2002 test data, looking at what scenarios were covered and what scenarios were not covered sufficiently. They concluded that the framework for looking at test cases needs to go beyond just looking at two fields and should approximate what happens in the real world.

During the conduct of the expert panel, participants were asked to submit substantive examples of "real world" complex sensitivity test cases for review and consideration. These test cases were included in the new Test Case Matrix.

2.2.5 Improved Measures

The expert panel determined that the test cases provided too much focus on generating sensitivity and specificity scores and really did not provide adequate information or a road map for improvement.

As shown in Table 2.2.5 below, a conceptual understanding exists of five measures that can universally benefit IIS practice efficiency.

Table 2.2.5: Metrics

Measure	Algorithm
Sensitivity – the ability to correctly determine that records are duplicates.	True Positive Matches divided by (True Positive Matches + False Negative Matches)
Specificity – the ability to correctly determine that records that appear to be duplicates are really separate individuals.	True Negative Matches divided by (True Negative Matches + False Positive Matches)
Accuracy – the overall ability of a system to detect patient duplicates.	True Positive Matches + True Negative Matches divided by (True Positive Matches + False Positive Matches + True Negative Matches + False Negative Matches)
Precision – the ability of a system to discriminate true positive patient matches.	True Positive Matches divided by (True Positive Matches + False Positive Matches)
False Positive Rate – the rate at which records are merged in error.	False Positive Matches divided by (True Negative Matches + False Positive Matches)

The expert panel determined that measures are best used to improve internal jurisdictional IIS operations. There should be no requirement to share measures. Additionally, program funding should in no way be dependent upon reporting patient de-duplication measures.

To understand how to use these measures for internal improvement, it is necessary to understand the vocabulary and concepts behind each measure.

For IIS public health applications, these measures can be expressed in a familiar 2 X 2 table, as illustrated below in Figure 2.2.5.

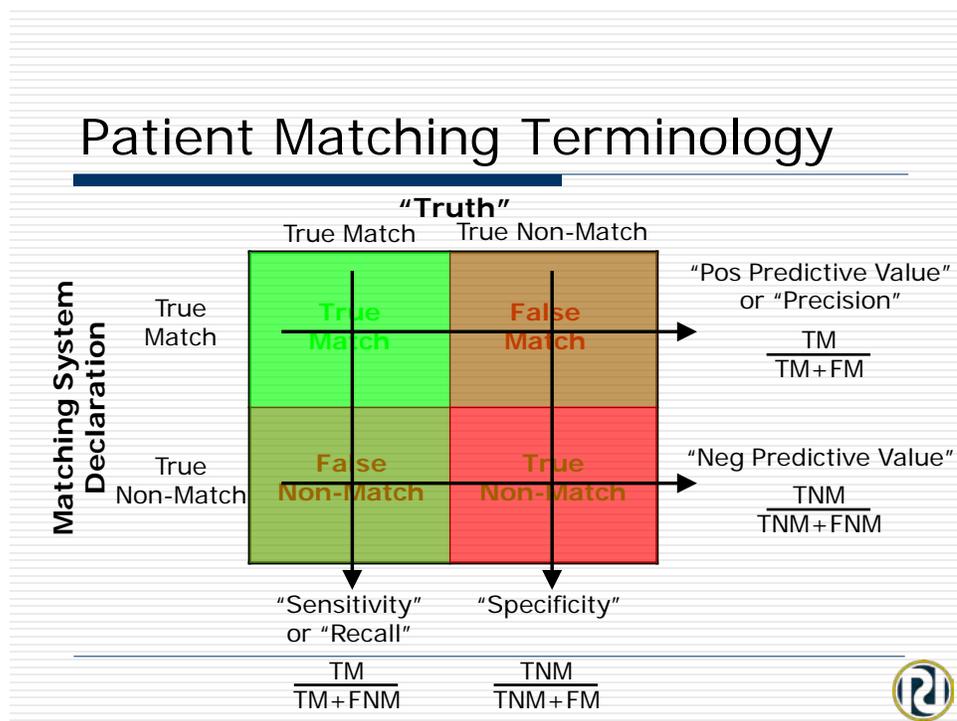


Figure 2.2.5: Measures Expressed in Two by Two Table

The panel considered the question of whether the manual review processes should be part of the scoring. The expert panel determined that this would not be a productive exercise and was out of scope for the project.

2.2.6 Effective Packaging

The new Test Case Matrix is designed to be an intuitive resource with no formal training required. The user interface for the Test Case Matrix is provided through Microsoft Excel. The Test Case Matrix leverages the intuitive aspects of Excel in order to provide its functionality. Because most members of the IIS community have had at least some exposure to Excel, no training and very little orientation is required in order to be immediately productive using it as a resource.

Through the use of Excel, the Test Case Matrix provides an easy-to-use, effective way to strategically manage the various aspects of patient de-duplication testing. As a “toolkit,” the various elements that are provided in the Test Case Matrix can be copied or modified with new and expanded tabs created which more exactly meet the needs of an IIS.

Chapter 3 of this document provides an overview of the Excel spreadsheet that constitutes the Test Case Matrix.

2.3 Future Considerations

The expert panel believes that patient-level de-duplication test case development, testing, and improvements to de-duplication are significant and important areas of ongoing interest to the IIS national practice community. While significant developments are occurring through the production of this report and related work products, there is only so much that can be accomplished within the scope of work of a single expert panel. The needs of the IIS community are not static—mechanisms are needed to foster on-going collaboration. The materials associated with patient-level de-duplication need consistent support and must evolve as experience is gained. Better mechanisms for sharing and collaboration for ongoing patient test case development and discussion are required.

Web-based collaboration tools—beyond a fixed website—need to be established, which will enable efficient leveraging of experience among the national practice community. There are several ways that this can be accomplished. One option to consider is strengthening support to the IIS community through such organizations as the American Immunization Registry Association (AIRA). In connection with this, it is noted that a plan and mechanism for sustainability requires exploration and development. These efforts are beyond the scope of the current expert panel. Specifically, sustained efforts in the areas of electronic health record to immunization information system (EHR-IIS) data interoperability and patient de-duplication are required to advance national practice at jurisdictional levels.

Creation of a short summary “best practice” series of meaningful use publications should be considered as part of the scope for future projects. It is noted that publications produced by the Modeling of Immunization Registry Operations Workgroup (MIROW) contain relevant content and are consistently referenced by IIS practitioners. It is possible that these materials could be produced from efforts within the IIS national practice community.

While specific best practice guidance has been summarized in the various practice-related sections of this report, it is believed that a more formalized testing program can benefit IIS data quality and the strategic use of testing to improve IIS data quality.

The expert panel also indicated that an emerging need would be to “volume test” or “performance test” the amount of time de-duplication software required to perform its functions. With the constant increases in volume and increases in the numbers of external providers submitting patient records, performance will become more critical going into the future.

3 Patient De-duplication Test Case Matrix

3.1 Introduction

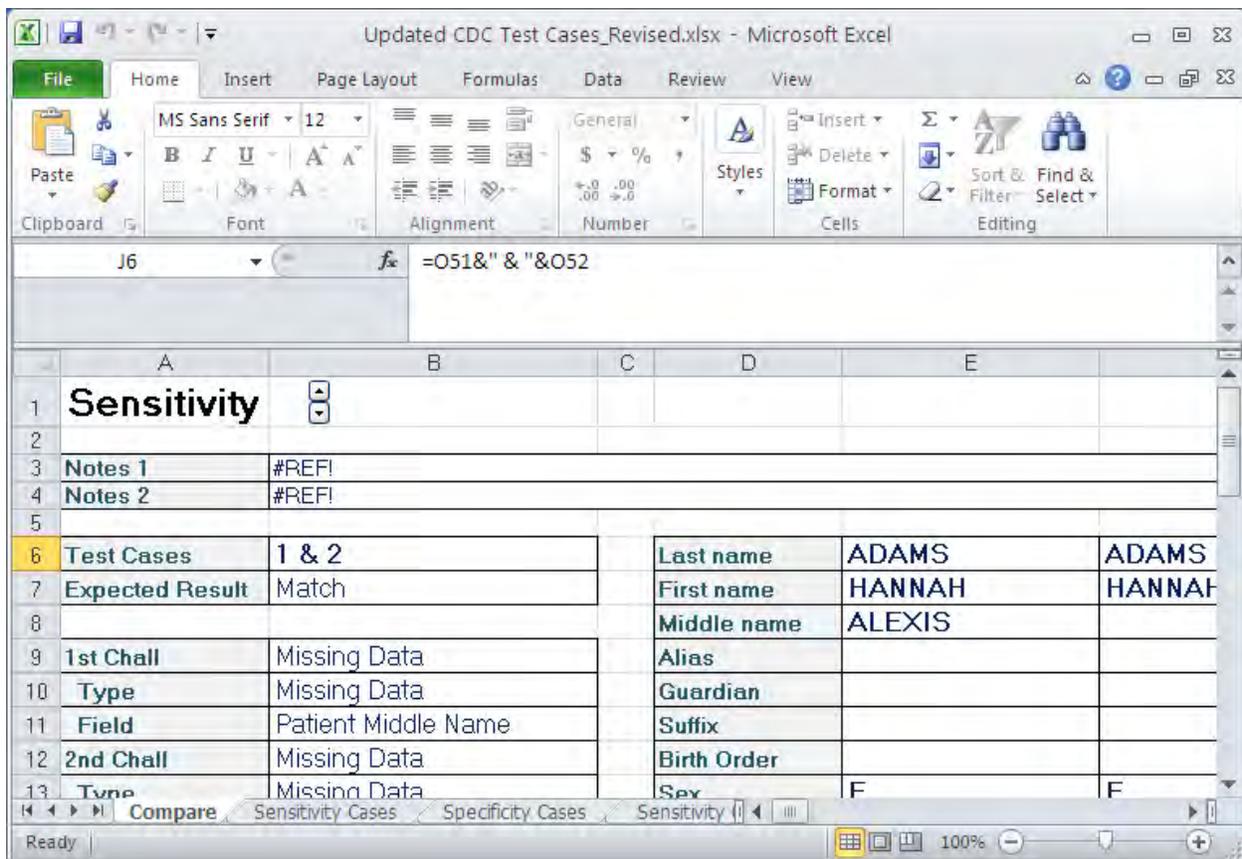
To provide the capability needed to systematically test and improve patient de-duplication capabilities, a new resource has been developed in the form of a comprehensive IIS patient de-duplication Test Case Matrix. The Test Case Matrix provides a logical and intuitive path to improve IIS patient de-duplication processes without requiring extensive technical knowledge.

The Test Case Matrix also supports technically advanced approaches and provides the flexibility needed to meet very complex and/or very demanding patient de-duplication testing needs.

3.2 Organization and Key Features

3.2.1 Compare Tab

Once the Excel Test Case Matrix is launched, it displays the first tab of the Test Case Matrix, “Compare.” The purpose of this tab is to provide a means of easily viewing each test case from a summary perspective. The user can scroll through both the sensitivity and specificity test cases and see all of the data fields in a snapshot view. It clearly details which fields in each test case are missing data, have different data from the base record, or have the same data as the base record. A portion of this screen is illustrated in Figure 3.2.1.



3.2.1: Compare Tab

3.2.2 Sensitivity Cases Tab

The second tab, “Sensitivity Cases,” contains all of the sensitivity test cases. Sensitivity is the ability of an IIS to correctly determine that records containing similar or varying data are actually the same patient. It provides categorization of each test case, columns with which to capture expected and

actual results for each test case, and the data fields associated with each test case, both for its base record as well as its test record. The complex test cases are clearly identified. This portion of this screen is shown in Figure 3.2.2.

	A	B	C	D	E	F	G	H	I
	ID	1st Challenge Category	Type	Field	2nd Challenge Category	Type	Field	3rd Challenge Category	Type
1	1	Missing Data	Missing Data	Patient Middle Name	Missing Data	Missing Data	Mother Middle Name		
2	2	Missing Data	Missing Data	Patient Middle Name	Missing Data	Missing Data	Mother Middle Name		
3	3	Missing Data	Missing Data	Patient Middle Name	Common Values Substitution	Nickname	Mother First Name		
4	4	Missing Data	Missing Data	Patient Middle Name	Common Values Substitution	Nickname	Mother First Name		
5	5	Typo Error	Transposition of Data	Patient DOB	Missing Data	Missing Data	Patient Middle Name		
6	6	Typo Error	Transposition of Data	Patient DOB	Missing Data	Missing Data	Patient Middle Name		
7	7	Values Change	Adoption	Patient First Name	Typo Error	Miskeying	Mother Maiden Name		

Figure 3.2.2: Sensitivity Cases Tab

3.2.2.1 Sensitivity Test Case Categorization and Complex Test Cases

The Test Case Matrix uses the seven challenge categories discussed in Section 2.2.3 to provide a comprehensive array of test conditions that are designed to exercise various aspects of de-duplication processes and logic. These categories represent the data processing problems that affect the establishment of positive patient identification and patient de-duplication. The rationale of this terminology is that the Text Case Matrix contains the specific organization of test cases and data that can be used to systematically challenge any patient-level de-duplication automated process.

Columns B through P provide the categorizations. The Test Case Matrix allows for the categorization of up to five different challenge categories. For each categorization, the challenge category, type, and specific field are indicated.

Figure 3.2.2.1 provides an illustration of a test case (record pair 61 and 62) in which the patient date of birth was mis-keyed and the patient's middle name is missing. It also illustrates a test case (record pair 63 and 64) with the incorrect gender.

	A	B	C	D	E	F	G	H
	ID	1st Challenge Category	Type	Field	2nd Challenge Category	Type	Field	3rd Challenge Category
62	61	Typo Error	Miskeying	Patient DOB	Missing Data	Missing Data	Patient Middle Name	
63	62	Typo Error	Miskeying	Patient DOB	Missing Data	Missing Data	Patient Middle Name	
64	63	Typo Error	Incorrect Gender	Patient Sex				
65	64	Typo Error	Incorrect Gender	Patient Sex				
66	65	Typo Error	Miskeying	Patient DOB				

Figure 3.2.2.1: Example Test Case Categorizations

3.2.2.2 Complex Test Cases

More categorizations indicate increased test case complexity. Column U, “Complex Case,” provides an indicator of whether the test case is complex. This is to assist the user with filtering to show only the complex test cases.

The expert panel provided dozens of examples describing what would constitute a complex test case. The Test Case Matrix displays the complex cases that may be useful for examination within the current patient level de-duplication environment.

The below illustration provides an example of many complex test cases, defined by the number of Challenge Categories.

	ID	1st Challenge Category	Type	Field	2nd Challenge Category	Type	Field	3rd Challenge Category	Type	Field	4th Challenge Category
505	504	Common Values Substitution	Address Abbreviation	Street Name	Common Values Substitution	Address Abbreviation	City	Missing Data	Missing Data	Mother Last Name	Non-Standard Data
506	505	Values Change	Divorce	Mother Last Name	Non-Standard Data	Field Transposition	Patient First Name	Common Values Substitution	Address Abbreviation	Street Name	Common Values Substitution
507	506	Values Change	Divorce	Mother Last Name	Non-Standard Data	Field Transposition	Patient First Name	Common Values Substitution	Address Abbreviation	Street Name	Common Values Substitution
508	507	Non-Standard Data	Field Transposition	Patient Middle Name	Common Values Substitution	Address Abbreviation	Street Name	Extraneous Data	Intermingled Data	Mother First Name	
509	508	Non-Standard Data	Field Transposition	Patient Middle Name	Common Values Substitution	Address Abbreviation	Street Name	Extraneous Data	Intermingled Data	Mother First Name	
510	509	Special Characters	Hyphenation	Patient Last Name	Non-Standard Data	Placeholder Data	Patient First Name	Missing Data	Missing Data	Patient Middle Name	Special Characters
511	510	Special Characters	Hyphenation	Patient Last Name	Non-Standard Data	Placeholder Data	Patient First Name	Missing Data	Missing Data	Patient Middle Name	Special Characters
512	511	Non-Standard Data	Placeholder Data	Patient First Name	Non-Standard Data	Field Transposition	Patient Last Name	Missing Data	Missing Data	Patient Middle Name	Non-Standard Data
513	512	Non-Standard Data	Placeholder Data	Patient First Name	Non-Standard Data	Field Transposition	Patient Last Name	Missing Data	Missing Data	Patient Middle Name	Non-Standard Data
514	513	Missing Data	Missing Data	Mother Last Name	Missing Data	Missing Data	Mother First Name	Typo Error	Miskeying	Street Name	Typo Error
515	514	Missing Data	Missing Data	Mother Last Name	Missing Data	Missing Data	Mother First Name	Typo Error	Miskeying	Street Name	Typo Error

Figure 3.2.2.2: Complex Test Cases with many Challenge Categories

3.2.2.3 Record Pairs

Patient level de-duplication testing is conducted using pairs of records. The Test Case Matrix consists of approximately 668 records pairs for the sensitivity test cases.

The first record of the test case consists of a base record. The base record is considered accurate in all regards and provides the reference data against which a de-duplication test is made. The test record contains variations of data from the base record designed to exercise the de-duplication logic of the IIS system. Column T, “Record Type,” indicates whether a record is a base record or a test record.

Users of the toolkit can take these pairs of records, clone them, modify them, document the modifications, and thereby create additional test cases that are specifically relevant to the needs of their IIS.

Figure 3.2.2.3 illustrates the “Record Type” column and the record pairs.

	T	U	V	W	X	Y	Z	A
	Record Type	Complex Case	LastName	FirstName	MiddleName	Alias (AKA)	Guardian Information	Su
1								
17	Test		ROGERS	STEVEN				
18	Base		HAYAKAWA	ATSUSHI				
19	Test		HAYAKAWA	SUSIE				
20	Base		POWELL	ZACHARY	WILLIAM			
21	Test		POWELL	WILLIAM	ZACHARY			
22	Base		RODRIGUEZ	JOSE				

Figure 3.2.2.3: Record Pairs and the Record Type Column

3.2.2.4 Test Case Filtering

A common approach to the generation of test cases is to formulate the test purposes and then select test cases from the Test Case Matrix to specifically test related functionality.

Sometimes the test purpose can be directly transformed into a simple use of the provided test cases; other times, the test cases may need to be specifically modified to match the needs of the IIS.

The Test Case Matrix is organized to allow either the broad or narrow testing of patient de-duplication capabilities. Testing can be conducted via the data processing challenge categories or by specifically selecting the similarities of data fields and other related criteria.

The user can filter by Category, Type, and Field to best reflect their testing needs.

Figure 3.2.2.4 illustrates the Test Case Matrix being filtered to show only the test cases that involve a typographical error resulting in an incorrect gender.

Updated CDC Test Cases_Revised.xlsx - Microsoft Excel

File Home Insert Page Layout Formulas Data Review View

Clipboard Font Alignment Number Styles Cells Editing

U492

	A	B	C	D	E	F	G	H
	ID	1st Challenge Category	Type	Field	2nd Challenge Category	Type	Field	3rd Challenge Category
1								
65	64	Typo Error	Incorrect Gender	Patient Sex				
80	79	Typo Error	Incorrect Gender	Patient Sex				
81	80	Typo Error	Incorrect Gender	Patient Sex				
220	219	Typo Error	Incorrect Gender	Patient Sex				
221	220	Typo Error	Incorrect Gender	Patient Sex				
1338								

Sensitivity Cases Specificity Cases Sensitivity Challenge C

Ready 6 of 1336 records found 100%

Figure 3.2.2.4: Category Filtering Example

3.2.2.5 Expected Versus Actual Results

As illustrated in Figure 3.2.2.5, each test record pair has been provided a suggested “Expected Result.” For sensitivity test cases, the expected result is always “match.” As a toolkit, the Test Case Matrix permits the recording of “Actual Results” and a “Pass/Fail.” These three columns are Q through S, and in the below illustration, they have been completed for the two test cases shown. Active test cases can be marked for selection and testing; and by using these fields, it is possible to compute various measures of sensitivity, specificity, and accuracy. This data allows for the calculations completed on the Scoring tab.

The screenshot shows an Excel spreadsheet titled "Updated CDC Test Cases_Revised.xlsx". The spreadsheet contains a table with the following data:

	Q	R	S	T	U	V	W	X
1	Expected Result	Actual Result	Pass/Fail	Record Type	Complex Case	LastName	FirstName	MiddleName
2				Base		ADAMS	HANNAH	ALEXIS
3	Match	Match	P	Test	x	ADAMS	HANNAH	
4				Base		FLORES	ASHLEY	ELIZABETH
5	Match	Non-Match	F	Test	x	FLORES	ASHLEY	

Figure 3.2.2.5: Expected Result, Actual Result, and Pass/Fail Columns

A failed test is a failure of the IIS to be able to make the specified determination. If the test is carried out using the cases provided in the Test Case Matrix, then the failure shows that the IIS de-duplication process may need to be systematically examined. When a test fails, its position within the Test Case Matrix provides an exact description of the test conditions, data used, and expected outcome.

3.2.2.6 Optional Data Fields

How testing is accomplished varies among IIS. Additionally, the “test harness” (the software and environment that support the loading of test data, its processes, and results) can vary.

Accordingly, the Test Case Matrix provides resources so they can be adapted to virtually any test harness. Another equally important consideration is the data required for jurisdictional immunization records. The data needed in IIS can vary. The Text Case Matrix accommodates this by providing many optional data fields. These data fields do not necessarily need to be accommodated in the test case record in order to exercise the core functions of the patient de-duplication processes. To accommodate robust testing, values that exceed the minimum or maximum bound of the jurisdictional domains can be chosen. The Test Case Matrix clearly documents the data fields that are relevant to the conditions of the test.

3.2.3 Specificity Cases Tab

The third tab, “Specificity Cases,” contains all of the specificity test cases. It provides columns with which to capture expected and actual results for each test case and the data fields associated with each test case both for its base record as well as its test record. A portion of this screen is illustrated in Figure 3.2.3.

	A	B	C	D	E	F
	ID	Expected Result	Actual Result	Pass/Fail	Record Type	Notes
1						
2	1				Base	Dup has child first name spelled different.
3	2	Non-Match	Non-Match	P	Specificity	Duplicate with different DOB. Different mom.
4	3				Base	Dup has child first name spelled different.
5	4	Non-Match	Match	F	Specificity	Soundex name is alike; DOB is slightly different, mom is d
6	5				Base	Dup has child first name spelled different.
7	6	Non-Match			Specificity	First name like middle; DOB a bit different, mom differe
8	7				Base	Sisters with different first names, different DOB
9	8	Non-Match			Specificity	Sisters with different first names, different DOB
10	9				Base	Dup has child first name spelled different.

3.2.3: Specificity Cases Tab

3.2.4 Sensitivity Challenge Categories Tab

The fourth tab, “Sensitivity Challenge Categories,” illustrates how the different challenge categories discussed in Section 2.2.3 are represented in the sensitivity test cases. The percentage of test cases in each category represents a qualitative consensus view of the approximate distribution of data processing problem types that are occurring within an IIS. The test cases were augmented to reflect the percentages that the expert panel felt best represented reality. This tab is illustrated below in Figure 3.2.4.

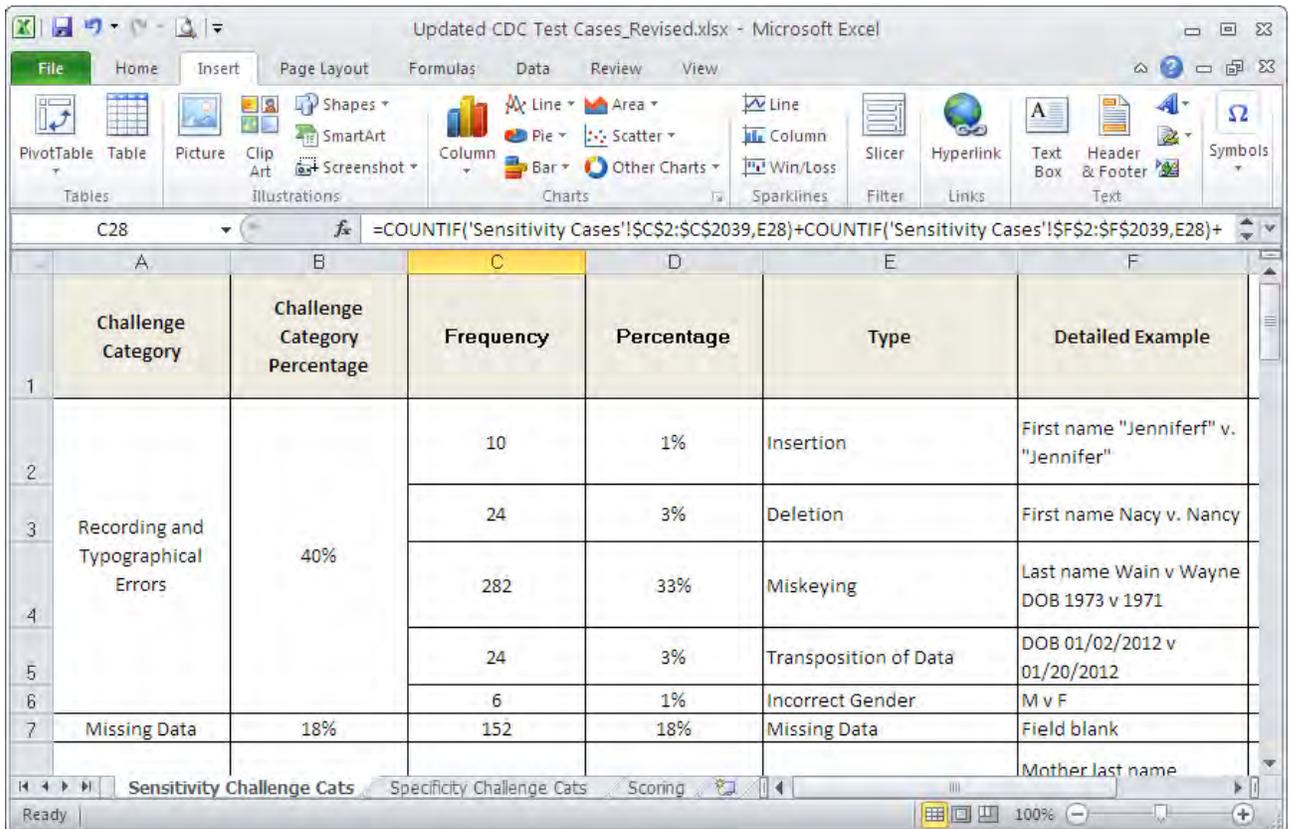


Figure 3.2.4: Sensitivity Challenge Categories Tab

3.2.5 Scoring Tab

The fifth tab, “Scoring,” automatically provides scores based upon the pass/fail information provided in the Sensitivity Cases and Specificity Cases tabs per the execution of the test cases. It does so according to the measures discussed in Section 2.2.5. This screen is illustrated in Figure 3.2.5.

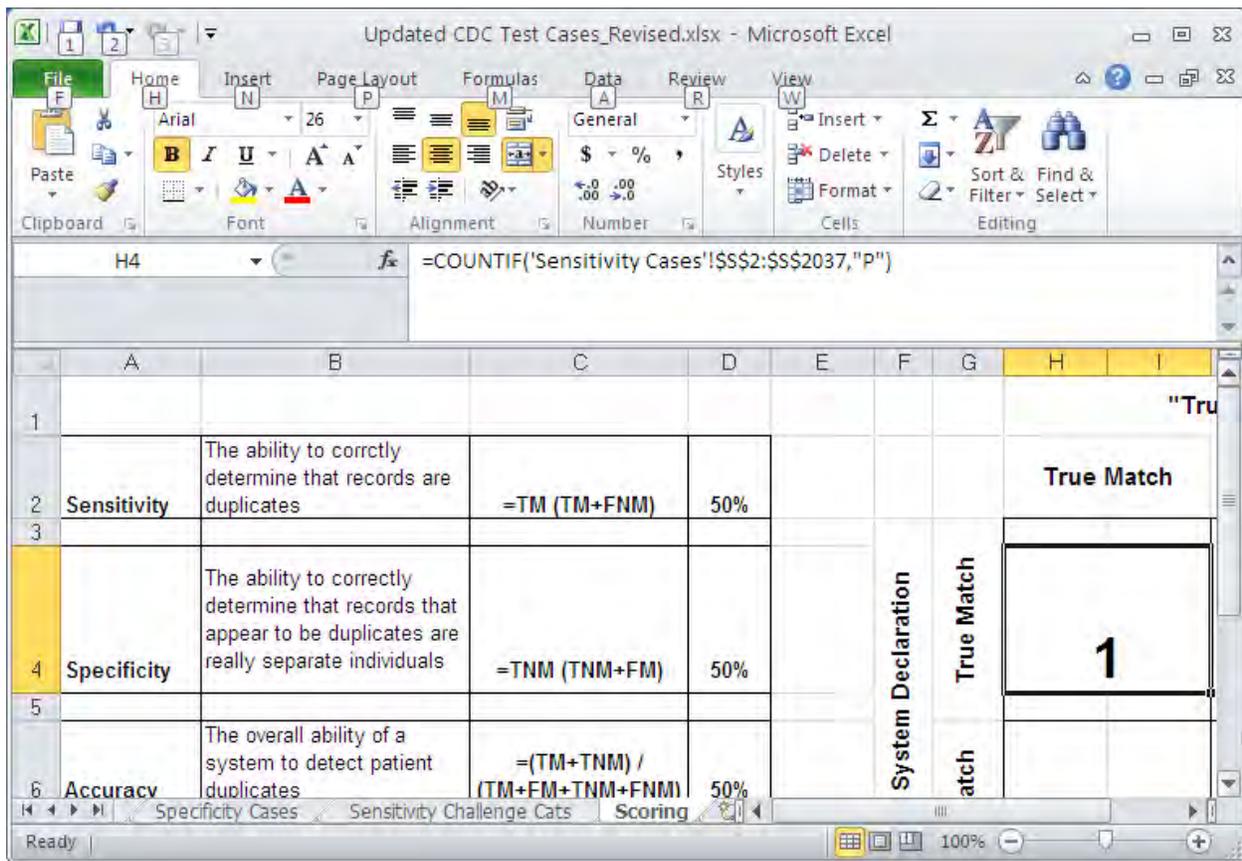


Figure 3.2.5: Scoring Tab

Accurate measurement of test results is an important consideration. However, it is equally important to maintain perspective relative to the meaning of metrics and scores. Test result scores can have significant limitations and should be interpreted with caution. The Test Case Matrix allows the tailored selection of test cases to help improve specific processing aspects of patient de-duplication. Because of the specific selection of test cases, test case scoring can be high or low; therefore, the significance of the score must be maintained in context. Selective testing of functionality may mean that some parts of the IIS de-duplication processes are not exercised at all. No specific tool exists to enable a determination relative to the amount of code patient de-duplication testing exercises.

3.3 Support for Multiple Test Strategies

Developing a patient de-duplication test strategy to systematically improve IIS operations is a complex task. The needs and resources within public health jurisdictions can vary significantly. Also, the hours and resources which can be invested in testing can vary. To support the ability to systematically test areas identified for improvement, an understandable methodology is needed. The methodology must go beyond the simple pass/fail of test cases. The methodology must conform to IIS needs and also suggest a roadmap for improvement in line with the examination of causes and corrections for conflicting and incomplete data and poor data quality. Accordingly, the Test Case Matrix supports multiple types of test strategies.

Testing of IIS system behavior can be categorized into two general types of methodologies: white box

and black box testing. In white box testing, tests are derived from a first-hand knowledge about the structure of the software and from implementation details. In black box testing, test data are derived from general business and functional requirements, without a specific knowledge of the internal program structure. As explained in Table 3.2.6, both white and black box testing approaches are possible.

Table 3.2.6: Test Strategy Support	
White Box	Black Box
<ul style="list-style-type: none"> – A single specific challenge categories can be selected – Test cases can be made more complex and contain multiple challenge categories – Test fields included in the test cases can be tailored – Test field conditions can be specifically identified – A mix of progressively difficult tests can be included – Test conditions can include sensitivity and specificity or a combination of both 	<ul style="list-style-type: none"> – Test cases from single or multiple categories can be selected or all test cases can be selected – Test fields included in the test cases can be tailored – Test conditions can include sensitivity and specificity or a combination of both

As the Test Case Matrix is for general use, the test cases which are presented under each category are based upon a black box testing methodology.

The black box testing methodology delivered in the Test Case Matrix provides a meaningful, standardized categorization of IIS patient de-duplication problems, a comprehensive array of test problems, and a suggested adjudication of the test case outcomes. While a “pass” or “fail” is indicated, the Test Case Matrix methodology acknowledges that circumstances between jurisdictions may affect the interpretation of test result data. Test outcomes which do not conform to expectations should be thoroughly examined. Based upon jurisdictional conditions or differences in data or other local variations, the test itself might not conform to the specification. Accordingly, the methodology recognizes that there may be legitimate needs to vary from the published material. In these cases, the Text Case Matrix provides a helpful resource so that the test case can be changed outside of the system. It is intended that the conditions of the Test Case Matrix should be carefully studied along with the data, as well as the suggested specified outcome.

Test case selection is largely a manual process and can be error prone. With traditional testing approaches, there are no guarantees that the system will be systematically tested. To help overcome this limitation, the Test Case Matrix seeks to contain realistic test cases which exercise possible variations within the seven data processing challenge categories.

The construction of test cases and testing itself is considered by many to be an art as well as a science. Systematic testing of logic in a “black box” format in all cases will require human judgments and excellent coordination between SMEs and technical resources. The methodology and approach of the Test Case Matrix is to provide test cases as models to cover the various conditions of the framework within each category.

For each test category, test criteria are provided which contain synthetically generated test data mimicking real world situations. These test data can be modified to fit the needs of the IIS jurisdiction. The framework provides a set of test cases representing various facets of the problem. The general category should be used to select a set of test cases that together satisfy all the targeted areas for systematic improvement.

Time and currency are common limitations determining the extent to which testing is used. The Test Case Matrix provides values in that it provides test cases that can be immediately exercised by problem category. As patient de-duplication problems can involve several data fields and problem categories, the Text Case Matrix provides complex test cases.

4 Roadmap for Improvements

Table 4 below details possible mitigation strategies by challenge category.

Table 4: Mitigation Strategies Summarized by Challenge Categories	
Challenge Category	Mitigation Strategy
Recording and Typographical Errors	<ul style="list-style-type: none"> – Provide data quality feedback to immunization providers submitting data – Leverage advanced algorithms – Move from deterministic to more probabilistic methods for matching – Prepare incoming data
Missing Data	<ul style="list-style-type: none"> – Publish minimal data standards on website
Values Changes	<ul style="list-style-type: none"> – Leverage high value fields – Leverage secondary patient identifiers
Common Values Substitutions	<ul style="list-style-type: none"> – Provide data quality reports in the way of feedback to immunization providers
Special Characters	<ul style="list-style-type: none"> – Use refined, automated methods to perform data preparation and data standardization
Extraneous Data	<ul style="list-style-type: none"> – Utilized written data entry procedures – Provide data quality reports in the way of feedback to immunization providers
Non-standard Data or Field Use	<ul style="list-style-type: none"> – Utilized written data entry procedures – Provide data quality reports in the way of feedback to immunization providers

Appendix A - Panel Membership

The patient-level de-duplication expert panel gratefully acknowledges and appreciates the support of CDC for constituting this panel and providing attention to the important areas of patient-level de-duplication testing and best practice development.

We applaud the Immunization Information System Support Branch (IISB), National Center for Immunization and Respiratory Diseases (NCIRD), CDC for providing a well-structured, productive, and exciting expert panel experience. Each of the panelists willingly contributed time, talents, passions, and experience. Moreover, new information, new relationships, and new professional connections for developing future practice excellence in patient-level de-duplication and IIS operations were gained by all participants.

The CDC Expert Panelists

Michael Berry is a project manager with HLN Consulting, LLC and has contributed to immunization information systems since 2003. In addition to immunization registries, his work is focused on connecting public health and Health Information Exchange (HIE), standards-based messaging and interoperability architecture, privacy and security in HIE and person-matching technologies for integrated systems. Currently he manages HLN's projects for the Rhode Island Department of Health, and also works as a subject matter expert on the ONC State HIE Cooperative Agreement Technical Assistance Program. berrym@hln.com

Nathan Bunker is a software developer and public health consultant for public and private agencies, focusing specifically on immunization software and data exchange. His work has given him experience with key immunization registry functions, including: immunization recommendation/forecast, HL7 interfacing, data quality analysis, vaccination matching, patient matching, and vaccine barcoding. nathan.bunker@gmail.com

Gerry Bragg, MBA has over 20 years of experience in systems analysis and programming and for the past 15 years, has supported the Michigan Care Improvement Registry (MCIR) as a Senior Systems Developer. He has supported the MCIR system in a variety of capacities, including the development of patient de-duplication/match-merge processes and clinical decision support/immunization forecasting algorithms. Mr. Bragg also specializes in database/SQL performance, scalability, tuning, refactoring, design, technical planning, and configuration management. The system currently supports more than 25,000 users.

Mr. Bragg holds an MBA in Management Information Systems from the University of Minnesota in Minneapolis, Minnesota, and a BA in Accounting from Hillsdale College in Hillsdale, Michigan. He resides with his wife and family in Brighton, Michigan. Gerry.Bragg@altarum.org

Shaun Grannis, MD, MS, FAAFP is a Research Scientist at Regenstrief Institute, Inc. and Associate Professor of Family Medicine, Indiana University (IU) School of Medicine. Dr. Grannis received an Aerospace Engineering degree from the MIT, and underwent post-doctoral training in Medical Informatics/Clinical Research at Regenstrief Institute. He joined IU in 2001. He is a member of World Health Organization (WHO) Collaborating Center for the Design, Application, and Research of Medical Information Systems.

Dr. Grannis completed an analysis of an automated regional electronic laboratory reporting system that revealed substantial increases in the capture rates for diseases of public health. He is project director for an initiative integrating data flows from over 120 hospitals across the state of Indiana for

use in public health disease surveillance. This system has received real-time data from hospitals with more than 2 million transactions per year, and has detected public health outbreaks. As co-chair of the U.S. Health Information Technology Standards Panel (HITSP) Population Health technical work group, he helped lead development of technical Interoperability Specifications.

Dr. Grannis also serves as the Director of the Indiana Center of Excellence in Public Health Informatics, which recognizes that public health practice is driven by a wide variety of data types, data sources, and data management techniques. sgrannis@regenrief.org

Rick Hall, PhD (Physics) has worked on record matching software since 2003, first at ChoiceMaker Technologies, Inc, and later as an independent consultant. He maintains two open source projects, Open Source ChoiceMaker Technology (<http://oscm.sourceforge.net>) and A Data Generator (<http://adatagenerator.sourceforge.net>). His current clients include the New York City Department of Health and Mental Hygiene, the New York State Department of Education, the New South Wales (Australia) Centre for Health Record Linkage, and the Queensland (Australia) Department of Health. rick@rphall.com;

Steve Jarvis is a Data Interface Specialist for the Colorado Department of Public Health and Environment, Immunization Section (CIIS). Since receiving his B.S. degree in Computer Science, he has have been working on various software engineering tasks for the past 25 years, first with group health insurance, then moving into the healthcare field to develop software for clinical management software used by school-based health clinics. In 2003, he was recruited to work with the Colorado IIS and currently manages all aspects of information reported electronically to the CIIS. Steve.Jarvis@state.co.us

Brian Jorgage has worked as a programmer and database developer for over 15 years at various organizations in the Philadelphia area. Over the last several years he has served at the Philadelphia Department of Public Health in support of Philadelphia's city-wide immunization registry. In that capacity, he has processed incoming data files and worked to resolve various data-related issues. His most recent project has been the testing and deployment of a new immunization registry. Brian.Jorgage@phila.gov

Linda Luebchow has been a Data Quality Analyst with the Minnesota Immunization Information Connection (MIIC) for six years, previously holding a similar position with the Minnesota State Registrars' Office, working with the filing of Vital Records. Her day- to-day work is focused on Customer Service and Data Quality, including de-duplication of immunization records. In addition to this she is currently engaged in helping facilities make the switch to HL7 file formats and Real-Time connectivity to qualify for the immunization portion of Meaningful Use.

Linda holds a Bachelor's Degree in Education and is a former Teacher in both Wisconsin and Minnesota. She has two daughters and one grandson (another expected in October) and currently resides with her husband of 35 years in Grand Marais, Minnesota on the North Shore of Lake Superior. linda.luebchow@state.mn.us

Mary Beth Kurilo, MPH, MSW is the Manager of the Oregon ALERT Immunization Information System, or ALERT IIS, and has been with the Oregon State Immunization Program since 2003. Mary Beth worked primarily in health care quality improvement before joining State Public Health. She completed her graduate work in public health and social work in 2001 at the University of Washington, earning a joint MPH/MSW focusing on Health Administration and Maternal and Child

Health topics. She has presented several topics related to immunization information system development and data use at previous National Immunization Conferences, and has co-authored articles in MMWR and other publications. Mary Beth is also the President-Elect of the American Immunization Registry Association (AIRA). mary.beth.kurilo@state.or.us

Christie D. Levy is the Branch Director II (Registry Coordinator) with the Mississippi State Department of Health for the Mississippi Immunization Information eXchange (MIIX) registry system in Jackson, Mississippi. Since 2008, she has assisted in the overall function of the Registry ensuring activities are maintained in adherence to the State of Mississippi, CDC, and HIPAA policies and guidelines. Currently, she works as the project liaison between other programs/agency's project managers and District Administrators to develop strategic plans including marketing and recruitment for the MIIX registry system. She is a member of the American Immunization Registry Association (AIRA) which has a mission to "promote the development and implementation of immunization information systems (IIS) as an important tool in preventing and controlling vaccine preventable disease." christie.levy@msdh.state.ms.us

Megan Meldrum is a Research Scientist (epidemiology) with the New York State Department of Health, Division of Epidemiology, Center for Community Health, Bureau of Immunization, New York State Immunization Information System (NYSIIS). Currently she serves NYSIIS as the data exchange liaison, responsible for creating and maintaining data exchange relationships between NYSIIS and multiple data exchange partners including software vendors, private health care providers, and other state agencies; tests modifications and enhancements pertaining to how NYSIIS handles and stores incoming data; analyzes data quality; and participates in the national IIS community as a subject matter expert. mdm06@health.state.ny.us

Chris Pratt is the Technical Manager for the Utah State Immunization Registry, otherwise known as USIIS. He has been with the USIIS Program since 1996. During this time, he has supported the registry through a broad range of IT roles, including Technical Support, Programmer Analyst, Database Administrator, and Technical Manager. Record matching accuracy factors into the duties of each of these IT roles and each role views the problem of record matching from a different perspective. These experiences have given Chris a diversified and firsthand knowledge of the record matching complexities challenging many patient-centric databases today. cpratt@utah.gov;

Helen Redfield is a software engineer with over 30 years technical experience supporting health care applications for state government. She has spent the last 16 years supporting the Texas Immunization Registry ImmTrac, and is the registry's technical expert on patient matching, data import, immunization forecasting and HL7 data exchange. Ms. Redfield holds a BA from the University of Texas at Austin and currently resides in Corvallis, Oregon with her husband. helen.redfield@dshs.state.tx.us

Bobby J. Sanchez became the Patient Care Training Administrator for Presbyterian Medical Services (PMS) in New Mexico in April, 2012. For the previous 10 years, he worked for the NM Department of Health in Health Promotion and as the New Mexico Statewide Immunization Information System (NMSIIS) Training Coordinator. While working in NMSIIS he created and made available interactive trainings that improved accuracy and 'cleanliness' of the data and increased provider and user participation. He has served on several national expert panels and was a board member of the American Immunization Registry Association (AIRA). At PMS, he is involved in the training for a statewide electronic system that incorporates, electronic health records, electronic patient management (appointments, schedules, billing), accounts receivable, electronic dental

records, behavioral health, e-prescribing, other modules under development and the integration of all these modules throughout the PMS system. bobby_sanchez@pmsnet.org

Cecile Town is a Senior Research Officer assigned from the Centers for Disease Control and Prevention (CDC), National Center for Immunization and Respiratory Diseases, Immunization Services Division to the Indian Health Service (IHS) Immunization Program, IHS Division of Epidemiology and Disease Prevention in Albuquerque, New Mexico. Since 2006, She has served as the IHS Immunization Interface Coordinator, responsible for facilitating and implementing immunization interfaces between RPMS and state IIS. Currently, she works with multidisciplinary groups to ascertain intersystem interoperability; facilitates testing and development of the IHS Immunization Interface Management software (BYIM); actively participate in the national IIS community as a subject matter expert; and provides data exchange support to IHS, tribal, and urban RPMS facilities nationwide. Cecile.Town@ihs.gov

CDC Expert Reviewers

Brandy Altstadter has worked for Scientific Technologies Corporation (STC) for ten years. She has worked in numerous capacities on the immunization registry products, including requirements analysis, development management and managing support. She is currently a Technical Solutions Architect for STC. brandy_altstadter@stchome.com

Noam H. Arzt, PhD, FHIMSS, is president of HLN Consulting, LLC, which has provided HIT services to public health agencies around the country since 1997. Dr. Arzt holds undergraduate, masters and doctoral degrees from the University of Pennsylvania and is active in a number of leading healthcare organizations (HIMSS, PHDSC, AMIA) and standards organizations (HL7, S&I Framework). A frequent speaker at national conferences on healthcare informatics, IIS, and HIE, Dr. Arzt has been supporting the IIS community for nearly 20 years. arzt@hln.com

Tammy Clark, RN, BSN, is the Director of the Mississippi Immunization Program. She is currently responsible for planning, implementing, and evaluating all Immunization programmatic activities, including Vaccines for Children (VFC) and the statewide Mississippi Immunization Information eXchange (MIIX). In addition, she also establishes policies and procedures and for the training/developing program of staff nurses, and provides educational in-services for VFC providers, initiating partnerships with key stakeholders. Tammy.clark@msdh.state.ms.us

Robert R. Grenwelge, Jr. is a Public Health Advisor in the National Center for Immunizations and Respiratory Diseases/Immunization Services Division/Program Operations Branch of the Centers for Disease Control. He has been involved in Immunization Information Systems for over 15 years, beginning during his work as the Administration Manager of the Communicable Diseases Division of the Houston Department of Health and Human Services (HDHHS). While there, Robert was involved in the public and private collaboration to develop and implement the Houston/Harris County Immunization Registry (HHCIR). His association with HHCIR continued after he became a CDC Public Health Advisor assigned to Houston's Immunization Bureau. In 2005, Robert relocated to Wyoming as the CDC Public Health Advisor assignee to the Wyoming Department of Health (WDH) and has been involved in the further development and implementation of the Wyoming Immunization Registry (WyIR). robert.grenwelge@wyo.gov

John Kellgren is the Lead IT Architect for the District of Columbia Department of Health (DC-DOH), and is currently serving as DC-DOH's EHR-IIS Interoperability Enhancement Project Manager. He previously served as the Project Manager for the DC-DOH's NEDSS Project. A software architect for 25+ years, Mr. Kellgren's background includes experience within multiple commercial industries including construction, distribution, and manufacturing. He has worked on software projects for clients such as Anheuser Busch, Avon, Hercules Construction, Marriott, McDonnell Douglas, and Springfield Remanufacturing. In the late 1980s, he built "Just In Case," a copyrighted software application system designed specifically for health industry case workers. John.kellgren@dc.gov

Tammy LeBeau, BS, has served as the Immunization Registry Coordinator for the South Dakota Department of Health (SD-DOH), Immunization Program since the inception of the program in 1995. Providing the program with VFC/AFIX coordination, she was also developed all of the SD-DOH training materials and conducts yearly trainings for the auditors. As a member of the MIROW expert panel, Ms. LeBeau's contributions to South Dakota's Immunization Information System (SDIIS) garnered the win of AIRA's Center of Excellence Award in 2009 (data use) and 2011 (inventory management). Tammy.LeBeau@state.sd.us

Eric Lindskog

Mayo Clinic
Senior Analyst Programmer
Rochester, Minnesota
Lindskog.Eric@Mayo.edu

Thomas Maerz is designer and manager of the Wisconsin Immunization Registry (WIR). An Applications Developer, Computer Electronics Builder and Network Specialist by trade, he has worked with healthcare records and integration with Electronic Medical Record (EMR) systems since 1979 and Vital Records de-duplication of information since 1990. Mr. Maerz has been working with healthcare providers, HMOs, schools and EMR vendors regarding an immunization registry for the state of Wisconsin since 1995. Thomas.Maerz@dhs.wisconsin.gov

Lisa Rasmussen has been the Project Leader for ASIIS, the Arizona State Immunization Information System since July 2007. ASIIS reporting is mandated for administered vaccines for children and any pharmacist-administered vaccines within the state system which currently serves over 5 million patients and 50 million vaccines. Holding a Bachelor's Degree in Public Administration and extensive experience in Public Health related databases and systems, Prior to her career in Immunization, Ms. Rasmussen gained additional experience in Maternal and Child Health programs, such as Newborn Metabolic Screening, Health Start, Child Fatality Review, High Risk Perinatal Programs, and Family Planning. lisa.rasmussen@azdhs.gov

Wendy Scharber, RHIT, CTR is founder and president of Registry Widgets, with more than 25 years of experience cancer registration. She is recognized as an international leader in: electronic reporting automated processing of data, and interoperability between public health and eHealth initiatives. She has created and managed electronic reporting systems, data conversion and electronic processing, and rules-based software support systems for cancer registration and specializes in implementing innovative strategies to meet the needs of public health programs. Ms. Scharber is active in several National eHealth initiatives, serving as a bridge between the public health domain and standard setting organizations and implementation efforts. She has authored two profiles within IHE relating to transmission of cancer data from pathology laboratories and from physician offices to the cancer registry. wendy@registrywidgets.com

Lee Taylor, MB is a medical epidemiologist and public health physician in the Centre for Epidemiology and Evidence at the NSW Ministry of Health, Australia. She has managed record linkage at the Ministry of Health since 1994, was instrumental in the establishment of the Centre for Health Record Linkage (CHeReL) in 2006, and is a member of the CHeReL Management Committee. She has extensive experience in the legal, ethical and privacy issues relating to the collection, use, and disclosure of health data for research and management of health services. She is a past chair of the NSW Department of Health Ethics Committee. Dr. Lee also manages data collections that relate to the health of mothers and babies in NSW, and has been responsible for annual reports on the health of mothers and babies in NSW since 1994.
ltayl@doh.health.nsw.gov.au

Alexandra Ternier, MPH is a City Research Scientist for the New York City Department of Health and Mental Hygiene / Citywide Immunization Registry. Currently a PhD candidate in Epidemiology, Ms. Ternier's area of expertise includes record matching and de-duplication - probabilistic model development and evaluation. aternier@health.nyc.gov

Karen White, MPH, has worked with the Minnesota Department of Health in the areas of Infectious Disease Epidemiology and Immunizations. She joined the Minnesota Immunization Information Connection (MIIC), Minnesota's statewide immunization information system, in 2002 when the system was first deployed. As an epidemiologist she provides analysis of population-based immunization data to the CDC Sentinel Site IIS project and the EHR-IIS Grant and works as a business analyst for development of additional enhancements of MIIC software.
Karen.white@state.mn.us

Northrop Grumman Public Health Contractor Personnel

Frederic Grant, PhD, MPH, MBA, PMP is the Chief Scientist of Northrop Grumman Corporation's Public Health Division—which provides advanced IT and business solutions for government and commercial clients. Dr. Grant is an elected member of the Delta Omega Public Health Honor Society. Additionally, he is a Project Management Professional (PMP) and a Certified Data Processor (CDP). Dr. Grant acts as a senior public health advisor to CDC. He is an experienced strategic planner, informatician, public health SME, and facilitator. He has authored numerous publications and industry reports.

Eric Larson is a Senior Information Architect for Northrop Grumman Corporation and is under contract to the CDC Immunization Information System Support Branch. He is currently the lead technical consultant on four EHR-IIS Interoperability Enhancement Projects involving many subject matter experts in the EHR and IIS community. The projects focus on transport layer, HL7, Patient de-duplication and Clinical Decision Support. Previous to Mr. Larson's current assignment, he spent about 10 years as an implementer helping several statewide immunization programs implement, maintain and improve their IIS.

Lucretia McKenzie, MPH has worked in the field of healthcare technology for over 14 years. During the past 5 years, Mrs. McKenzie has served as a Business Analyst on various projects for the Centers for Disease Control and Prevention as a Northrop Grumman Contractor including supporting software development projects for the Division of HIV/AIDS. Currently, she serves as a business analyst for the Clinical Decision Support for Immunization (CDSi) project that has recently published the Logic Specification for ACIP Recommendations.

Nina Mitchell currently serves as the Quality Assurance Analyst for the NCIRD De-duplication program and as the Lead Quality Assurance Analyst for CIMS Data Message Brokering for 7 years. Early in her career at Northrop Grumman, she was the lead analyst on numerous successful technical projects from the development to implementation. Over the past 5 years, Ms. Mitchell has supported many CDC programs. She interfaces directly with partners from Public Health Labs and federal agencies and oversees their experience within DMB (Data Message Brokering) domain. An expert in development of test cases used in application development and complex problem resolution, Ms. Mitchell is currently working with the expert panel on creating test cases and test messages for the IIS toolkit for the de-duplication program.

She has a strong technical background in SQL, Rhapsody, Selenium Web Browser Automation, IBM Rational Robot, and HL7 (Health Level 7). Nina is currently in the process of obtaining certification in the HIT's Program for EHR.

Stuart Myerburg, JD, is a Health Scientist, Informatics in the Immunization Information Systems Support Branch (IISB) at the Centers for Disease Control and Prevention (CDC). Mr. Myerburg has 15 years of experience working in public health. Before coming to the CDC, he served as an Assistant Director of Information Technology at the Rollins School of Public Health. He now leads the EHR-IIS Interoperability and Clinical Decision Support projects in IISB.

Lindsay Ryan is the Project Coordinator for the EHR-IIS Interoperability Enhancement and Clinical Decision Support Projects for the Centers for Disease Control and Prevention. She has been a contractor for Northrop Grumman for almost 2 years and has over 12 years experience in the field of healthcare that spans across multiple focus areas, including reproductive health and medical education. Her prior experience includes coordinating health policy initiatives through state and legislative agencies for both Florida and Georgia, managing the implementation and monitoring of contracts within state/federal government and universities, performing investigative audits for medical records and clinical research and analysis. Ms. Ryan's business acumen, education, and leadership capabilities have contributed to a number of successful projects within the public and private sectors.

Celia Toles is the Northrop Grumman Contractor Technical Writer & Editor for EHR-IIS Interoperability Enhancement and Clinical Decision Support for Immunization Projects for the Centers for Disease Control and Prevention. Her prior experience in public health includes database management and project coordination in the Office on Smoking and Health and the Division of Cancer Prevention and Control, for which she participated in multiple standardization and reporting projects for the CDC. Her 17 years of experience in health care also includes positions with WebMD, The Emory Clinic, and the Georgia Institute for Lung Cancer Research.

Jennifer Wain has 20 years of experience in project management. She has an extensive resume including employment with Accenture and Unisys and supporting clients such as the Centers for Disease Control and Prevention (CDC), US State Department, Coca-Cola, and AT&T. As a Northrop Grumman project manager and contractor to CDC, Ms. Wain currently leads key immunization-related projects in the areas electronic health record/immunization system interoperability and clinical decision support.

This documentation is available at <http://www.cdc.gov/vaccines/programs/iis/interop-proj/ehr.html#patient>

Centers for Disease Control and Development, June 28, 2013

Last Reviewed Feb 2016