

UCI and You

The SPHERE Institute
HIV/AIDS Bureau (HAB), Health Resources
and Services Administration (HRSA)

November 10, 2008

Presenters from The SPHERE Institute:

Ellie Coombs and Peggy O'Brien-Strain

Contact Information:

Ellie Coombs, ecoombs@sphereinstitute.org, (650) 558-3980, ext 149

Outline

- **What Is the eUCI?**
- **How Does the eUCI Ensure the De-identification of Client Data?**
- **How Was the UCI Selected?**
- **How Will the eUCI Be Incorporated into My Data System?**



2

The “UCI and You” webcast has four short segments. In the first segment, we briefly present the role of the eUCI in the Ryan White Services Reporting System, HRSA/HAB’s new reporting system that involves the submission of client-level data from Ryan White grantees. The second section focuses on privacy and security. We describe how the eUCI de-identifies client data so grantees will comply with both federal and state laws when submitting their Ryan White Service reports. In the third section, we review the UCI selection process, presenting the research methods for evaluating potential UCIs and findings. Finally, we explain how the eUCI can be incorporated into grantees’ and providers’ data systems.

Purpose of the eUCI

- Ensures client privacy and the security of health care information through encryption
- Links client health and service utilization data across providers and over time
 - Each client has his/her own eUCI
 - That eUCI remains constant over time and across providers



3

The eUCI has two main goals. First, because the eUCI contains no personal identifying information, client data within Ryan White Services Reports will be protected and client privacy maintained. Second, through the eUCI, HRSA/HAB can link reports that belong to the same client, but are received at different points in time or from multiple grantees or providers.

Overview of the RSR System eUCI

- The Unique Record Number (URN) will act as the RSR System UCI
 - Commonly used by grantees (e.g. in CAREWare, ARIES)
 - Composed of
 - 1st and 3rd letter of the first name
 - 1st and 3rd letter of the last name
 - Full date of birth (DDMMYY)
 - Gender code
 - “U” if unique within the grantee site, “A”, “B”, or “C”, etc...if not unique
- The UCI will be encrypted with the SHA-1 hashing algorithm



4

The UCI selected for the RSR System is the Unique Record Number, or URN. The URN is used by CAREWare and many other grantee data systems. The URN is composed of the first and third letters of first name, first and third letters of last name, full date of birth, and a code for gender. A final character of the UCI indicates whether any other client within the grantee site has the same first 11 characters of the UCI. If a client does *not* share the first 11 characters with another client, the final UCI character is “U” for unique. If clients *do* share the first 11 characters, the first client receives an “A”, the second client a “B” and so on. Through this process, no two clients within the same grantee site should share the same UCI.

To remove identifying data from the UCI, it will be encrypted with the SHA-1 hashing algorithm. This algorithm, which meets the highest industry standards, is discussed in greater detail in segment 2 of the presentation.

No Identified Information Goes to HAB: eUCIs Are Created at Point of Service

- **Step 1:** A client visits a provider/grantee for services and submits personal information, such as name and date of birth.
- **Step 2:** A program embedded within the provider/grantee data system creates the UCI from the personal information.
- **Step 3:** The program encrypts the UCI, creating the encrypted UCI (eUCI).
- **Step 4:** The eUCI is linked to the client's health and service utilization data, creating the complete Ryan White Services Report.
- **Step 5:** The provider/grantee uploads the client's Ryan White Services Report to the HRSA/HAB server.
- **Step 6:** HRSA/HAB receives RSRs from multiple providers with this same eUCI and merges these reports.



5

A client's eUCI will be constructed at the point of service, either at the site of a sub-contracted provider or the grantee site itself. Imagine a funded client visits a Ryan White provider for health care services. At intake, that client provides personal information such as name and date of birth, which are input into the provider's data system. A program within the system takes pieces of the personal information to create the UCI and encrypt it. The eUCI is then linked to the Ryan White Services Report when it is uploaded to the HRSA/HAB server. When HRSA/HAB receives more than one report with the same eUCI, these reports are merged. With these merged reports, HRSA sees a more complete picture of the client's health and service utilization history.

Outline

- What Is the eUCI?
- **How Does the eUCI Ensure the De-identification of Client Data?**
- How Was the eUCI Selected?
- How Will the eUCI Be Incorporated into My Data System?



6

Now that we have a good understanding of how the eUCI is generated and used within the RSR System, we will go into greater detail about its role in promoting client privacy and the security of Ryan White client data.

The Security of Health-Related Data Has Become Increasingly Important

- The Health Insurance Portability and Accountability Act (HIPAA) (1996) protects patient health-related data
- The Privacy Rule (2003) makes new requirements for the use or transmission of confidential data
- HIPAA security requirements are the minimum allowable; some states have enacted even stricter laws



7

As health-related data is more often being stored and transferred electronically, data security is becoming easier to breach. In 1996, the Health Insurance Portability and Accountability Act, commonly known as HIPAA, was passed to establish protocols for secure data transfer. In 2003, the Privacy Rule made clearer which entities fall under HIPAA's mandate. As HIPAA serves only as the "floor" of client privacy and data security, grantees and providers in many states have to follow even stricter protocols.

Individually Identifiable Health Data is Protected under HIPAA

- HIPAA applies to grantees and their subcontracted providers that:
 - Provide health care services
 - Transmit electronic information on health care services (e.g. for payment purposes)
- Protected data:
 - **Individually identifiable data:** Information alone or in combination with other information that identifies an individual (e.g. name, Social Security Number, phone number)
 - **Health data:** Information related to the physical or mental health of or health care services provided to individuals



8

HIPAA applies to all Ryan White grantees and providers that provide health care services and transmit electronic information about those services.

Individually identifiable health data is protected by the law. Individually identifiable data is information that can be used to identify an individual; examples of this information are name, Social Security Number, and phone number. Some information can be directly used to identify a client, such as Social Security Number, while other data must be accompanied with additional information to do so, such as phone number.

Health data includes all information related to the physical or mental health of an individual; this information is often found in medical records, physician notes or billing information.

eUCIs Protect Health Information by Removing Individually Identifiable Components

Name: John Doe, DOB: February 2, 1964, Male



UCI: JHDE0202641U



eUCI: 8417D5706B0B40E52BA8FE4F95460CB9DC2223AAU



9

Grantees and providers will be HIPAA compliant when they upload Ryan White Services Reports to the HRSA/HAB server because the eUCI removes all individually identifiable information. Imagine a client named John Doe who is born February 2, 1964. The unencrypted UCI in row 2 still has identifiable information, such as full date of birth and the client's initials. Combined with other data in Ryan White Service Report, this information could be used to identify the client. With encryption, however, the UCI is converted into a seemingly random string of letters and numbers, making the eUCI unidentifiable information.

The UCI is Unrecoverable from the eUCI

- SHA-1 Algorithm was selected for the RSR encryption because it is a trap door algorithm



The encryption technique selected for the UCI is the SHA-1 Algorithm. This is a trap door algorithm – meaning that the UCI cannot be deciphered from the eUCI.

The SHA-1 Algorithm Meets the Highest Privacy and Security Standards

- Designed by the National Security Agency
- Is the best established of the existing SHA hash functions
 - Is employed in several widely used security applications and protocols
 - Meets highest federal standards and is required by law for use in certain U.S. Government applications



11

The SHA-1 algorithm was selected because it meets the highest federal standards for privacy and security. Developed by the National Security Agency, the SHA-1 algorithm is commonly used in government and by private institutions deeply concerned with client privacy and data security.

The eUCI Complements Other HAB Security Measures

- **Data Transfer:** All RSR data will be encrypted and transferred securely
- **Data Storage:** Servers will be stored in locked areas; Layers of software protection such as firewalls and antivirus monitors will also be used
- **Data Use:** Data access granted on a need-to-know basis; Good account and password management practices



12

It is important to note the encryption technique is only *part* of HRSA/HAB's security model. All data in the Ryan White Services Reports will be encrypted while being transferred from providers or grantees to HRSA/HAB. Upon arrival at HRSA/HAB, data will be stored in secure servers, protected by locked doors and firewalls. Data access will be granted on a need-to-know basis and protected by strong account and password management.

HRSA/HAB can only ensure the security of data while it is being transferred to and stored at HRSA/HAB. Grantees and providers are responsible for ensuring the security of data, while it is located at their sites. We will be posting additional information on how providers and grantees can protect their data.

Outline

- What Is the eUCI?
- How Does the eUCI Ensure the De-identification of Client Data?
- **How Was the UCI Selected?**
- How Will the eUCI Be Incorporated into My Data Management System?



13

The URN was selected as the RSR System's UCI based on extensive research conducted by HRSA/HAB and its contractor. In this section, we review the UCI selection process.

UCI Selection Criteria

- **Unique**
 - Each Ryan White Program client has his/her own UCI
- **Permanent/Consistent**
 - The UCI remains constant over time and across grantees and providers
- **Feasible**
 - The UCI does not require extensive system changes



14

To act as an effective tool for record linkage, a UCI should satisfy three criteria, which served as the foundation for the evaluation process. First, the identifier should be unique, in that each client has his or her own UCI. The UCI also must remain permanent or consistent over time and across providers so clients don't have more than one UCI. Finally, the UCI's application should not place undue financial and data entry burden on grantees, providers, clients or HRSA/HAB.

Construction of UCIs for Testing

- Identified data elements based on grantee input
 - Already collected and used by providers
 - Easily recalled and reported by clients
 - Permanent over time
- Selected elements: First name, Last name, Date of birth, Gender, Social Security Number (SSN)
- Created 105 test UCIs from segments of selected data elements



15

The research process for choosing the UCI first involved the selection of data elements, or pieces of personal information, that would construct the UCI. First name, last name, date of birth, gender, and Social Security Number were selected as the UCI's potential data elements because they possessed three important characteristics. First, they are already collected by providers, meaning that data intake staff would not have to introduce a new data element into their intake process. Second, they are easily recalled and reported by clients, which would result in fewer reporting errors. Finally, they remain relatively permanent over time, making the UCI more permanent.

105 potential UCIs were created for testing with different segments of these data elements.

UCIs Were Tested on Two Measures of Performance

- **Uniqueness** – Measured by false positive rate
 - Two clients given the same UCI
 - Data source: Marketing database of unique individuals
 - Created test UCIs for each individual in database
 - Any duplicate UCIs were deemed false positives
- **Permanence/Consistency** – Measured by false negative rate
 - One client given two or more UCIs
 - Data source: Ryan White grantee data
 - Created test UCIs for each record at grantee sites
 - Grantees verified when two records, belonging to the same individual, received different UCIs

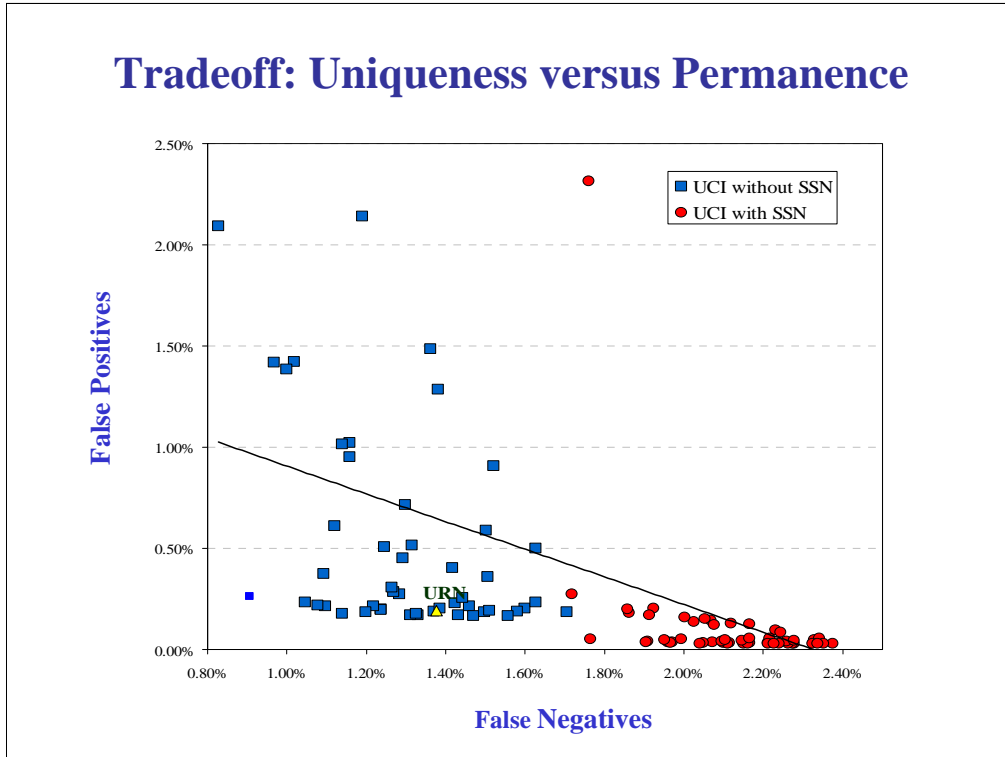


16

The uniqueness of each of the 105 UCIs was evaluated by measuring rates of false positives. A false positive occurs when two different clients receive the same UCI. To calculate the false positive rates, the 105 test UCIs were created for each individual in a marketing database. A false positive was identified when two individuals received the same UCI.

The permanence/consistency of the test UCIs was evaluated by measuring their rates of false negatives. A false negative occurs when a client receives two different UCIs. This may happen due to a recording error, such as a typo in date of birth, or because a client changes his or her last name. False negatives were evaluated using Ryan White grantee data. Grantees that participated in the research process identified false negatives by verifying when records that received different UCIs actually belonged to the same person.

Tradeoff: Uniqueness versus Permanence



A major finding of this research process was the tradeoff between false positives and false negatives. In other words, the UCIs that tended to be more unique were also less permanent. The red circles on the right side of the screen represent the UCIs that contain Social Security Number as a data element. These UCIs had high false negative rates, meaning that Social Security Number was likely to be reported inconsistently. However, these UCIs also had low false positive rates, meaning that they were quite unique.

Findings Do Not Justify System Overhaul

- URN strikes a good balance between uniqueness and permanence
- Minimal differences in total error rates (false negative rate + false positive rate)
- Grantees reluctant to collect new data elements
 - Costs of staff training and system modifications
- Grantees concerned with collecting Social Security Number
 - Barrier to care
- Significant costs of implementing a new UCI



18

Although the URN did not have the lowest total error rate, which was the sum of the false positive rate and the false negative rate, it was selected because it strikes a good balance between uniqueness and permanence. Additionally, it is commonly used by grantees and providers, meaning that they will not have to undergo extensive changes in their data intake and management processes. Grantees were also very concerned with collecting Social Security Number at intake because of issues related to client security and immigration status. Therefore, it was decided that Social Security Number should not be incorporated into the RSR System's UCI.

Some Error Is Expected in the RSR Database

- URN: Total Error Rate = 8.8%
 - False Negative Rate = 3.8%
 - False Positive Rate = 5.0%



19

Despite the URN's strengths, there will still be some level of false negatives and false positives in the RSR database. More than half of errors will stem from false positives; the rest from false negatives. HRSA/HAB acknowledges that the system will not be perfect, but staff believe that any duplication within the RSR System will be better than the current duplication rate.

Outline

- What Is the eUCI?
- How Does the eUCI Ensure the De-identification of Client Data?
- How Was the UCI Selected?
- **How Will the eUCI Be Incorporated into My Data System?**



20

In the final segment of this webcast, we provide grantees with an overview of how to the eUCI will be incorporated into their data systems.

All Ryan White Grantees and Providers Must Use the Same eUCI Algorithm

- The eUCI algorithm will be integrated into major data systems
 - AIRS, ARIES, CAREWare, Casewatch, eCOMPAS, LabTracker, Provide Enterprise, Sage
 - These grantees do not need to do anything!
- Grantees with less common or “home-grown” systems will need to install the eUCI Application by requesting it on the RSR page of the TARGET Center Website at <http://careacttarget.org/rsr.asp>



21

For the same eUCI to be generated for a given client across provider and grantee sites, all providers and grantees must use identical algorithms for creating and encrypting the UCI.

Grantees and providers that use the following major data management systems: AIRS, ARIES, CAREWare, Casewatch, eCOMPAS, LabTracker, Provide Enterprise, and Sage will not need to do anything! HRSA/HAB and its contractor are currently working with these systems developers to integrate the UCI and encryption algorithm along with other RSR requirements.

Grantees and providers that have less common or “home-grown” systems will need to install the eUCI Application, available through the RSR page of the TARGET Center website. The following two slides are related to those grantees that have less common or “home grown” systems, and therefore, will need to install the eUCI Application.

The eUCI Application Performs Three Main Functions

- Creates the UCI from data elements located within the grantee's data system
- Encrypts the UCI to create the eUCI

OR

- Encrypts the UCI, already created by the grantee's data system



22

The eUCI Application performs three main functions depending on the needs of the grantee or provider. If the grantee or provider currently does not use the URN, the eUCI application will call the necessary data elements from the data system, create the UCI and encrypt it. This eUCI will then be exported into the Ryan White Services Report.

If the grantee already creates the URN internally, the eUCI Application will encrypt the pre-created URN.

The eUCI Application Installation Process Depends on Your Data System

- **Most major data systems use the .Net Framework**
 - Use the eUCI Application alone
 - Code must be modified to call the correct fields from your data system to create the UCI
- **Other systems**
 - Create a file containing the necessary data elements for each client
 - Input the file into a second application



23

If your data system is run on .NET, you only need to install one program. The application's code must be modified to call the correct data elements from your system to create the UCI.

If your system does not run on .NET, you need to install a second application. This application coupled with a file that contains your clients' data elements will serve as inputs to the eUCI Application.

Determine Your System's Readiness and Request Support

- Go to the RSR page of the TARGET Center Website to fill out a survey that will allow you to determine how well you are prepared for the Ryan White Services Reporting System
<https://performance.hrsa.gov/hab/CLDXMLSchema/content.aspx>
- Please fill out the survey to receive additional support on incorporating the eUCI into your data system!



24

All grantees should go to the RSR page of the TARGET Center Website and fill out a Ryan White Services Reporting System Readiness survey. If you need assistance with the new reporting system, including the installation of the eUCI Application, a representative will contact you.

END

