BioSense Platform Data Flow

Part 3

Legacy Data Conversion and Data Transition Plan

May 5, 2016

Center for Surveillance, Epidemiology, and Laboratory Services
Division of Health Informatics and Surveillance

Agenda

- Introductions
- BioSense Platform Update
- Recap: ESSENCE Settings
- Legacy Data Conversion
- Recap: Access & Management Center
- Data Transition Plan
- Next Steps

BioSense Platform Update

Michael Coletta, MPH, NSSP Program Manager

BioSense Platform Update

1. PLANNING

- ✓Collaborate to identify critical activities
- ✓Collaborate to develop requirements
- ✓Set up staging environment
- ✓Set up production environment

2. DEVELOPMENT

- ✓ Create baseline Master Facility Tables
- ✓ Document new data flow
- ✓ Document ESSENCE settings
- ✓ **Document legacy data flow**
- ✓ Establish data mart
- ✓ Develop support documents
- ○ Develop Access & Management Center
- ○ User Acceptance Testing

3. TRANSITION

- ○ **Transition 9 sites per month, beginning summer 2016**
- ○ Develop Facility Admin Tool
- ○ Convert legacy data

4. SUNSET

- ○ Sunset BioSense web application

Recap: ESSENCE Settings

Shayne Gallaway, PhD, MPH, Health Scientist

Participant Input

- **Overview**
  - 9 participants (over 4 sessions) assessed BioSense Platform ESSENCE tool settings
  - Initial effort focused on determining critical settings for release
  - More improvements will likely be made before deployment

- **Participants add considerable value**
  - Their feedback is grounded in experience
  - They know how to present data in appropriate and meaningful ways

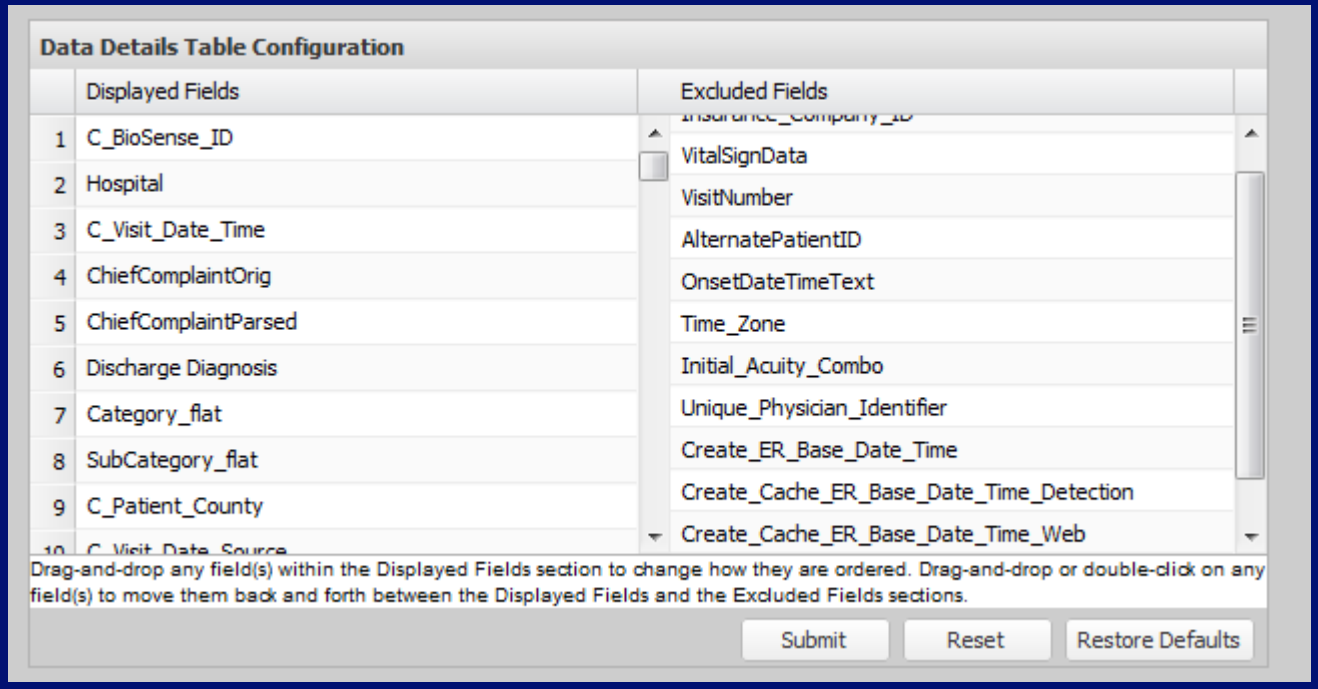
ESSENCE Settings – Local Data View

Participants . . .

- Defined a subset of viewable data elements
  - Default display and order of data elements for "full details" view
  - Decision making affected by ability to download hidden data elements, user type (local vs. state), variable description, and relevance to syndromic surveillance
- Proposed an order for the default subset list by relevance to syndromic surveillance (e.g., prioritize case info, demographics, potential exposures, CCDD type info)
- Discussed rationale for initially hiding specific data elements
- Discussed how individual users can control their view of data (e.g., reveal hidden columns or reorder data elements)

Example: Local Data View

- In ESSENCE, use “Configuration Options” on query results page to control columns seen on screen
- Drag and drop fields into “Displayed” or “Excluded” sections



ESSENCE Settings – National View

Data Elements

- Encounter Date
- Patient Gender
- HHS Region
- Patient Class
- Categorized Age:
  - Standard Groups (5 strata)
  - ILI Reporting Age Groups
  - 10-Year Age Groups
- Syndrome
- Sub-Syndrome
- Disposition

National View

- Restricted, high-level view granted to BioSense Platform ESSENCE users by default. The National View aggregates data by HHS Region.
- Users can only see local details if granted permission.

Who contributed to decision?

- Pilot group discussion (Spring 2015)
- BioSense Governance Group (Summer 2015)
- Extensive communication with site partners (via participants and informally)

ESSENCE Settings – Chief Complaint Query Validation

***NEW* Datasource**

- **Separate Chief Complaint, Discharge Diagnosis, and CCDD information from detailed data**
  - **Goal:** allow users to train queries against the full data set without viewing details
  - Limited Query: Chief Complaint and Discharge Diagnosis
  - Limited View: Chief Complaint, Discharge Diagnosis, Syndrome, and Visit Date (date only)
- **Chief Complaint Query Validation Tool would NOT return individual-level data details such as patient/hospital location or demographics**

ESSENCE Settings – Alert Summaries

Participants . . .

- **Discussed permissions and restrictions associated with controlling regions and facilities displayed in alert summary**
- **Defined appropriate regional and hospital syndrome alert views for local, state, and national users**
  - Summary Alerts – visible to all users
  - Region/Syndrome Alerts – visible to all users and configurable by region
  - Facility/Syndrome Alerts – access controlled by site administrators
  - Spatial Alerts – visible to all users
  - Facility/SubSyndrome Time of Arrival Alerts – access controlled by site administrators
- **Proposed development of a *NEW* alert at the HHS Region level**

ESSENCE Settings – Data Source Names

Participants . . .

- **Found data source labels in ESSENCE Query Tool confusing**
- **Suggested descriptive alternative labels:**

  - Patient Location (Full Details)
  - Patient Location & Visit (Full Details)
  - Patient Location (Limited Details – HHS Region)
  - Patient Location (Limited Details – State & HHS Region)
  - Facility Location (Full Details)
  - Facility Location & Visit (Full Details)
  - Facility Location (Limited Details – HHS Region)
  - Facility Location (Limited Details – State & HHS Region)
  - Chief Complaint Query Validation
  - DoD Data
  - Veterans Affairs Data
  - Weather Data

Legacy Data Conversion

Roseanne English, BS, Analytic Data Mgmt. Team Lead

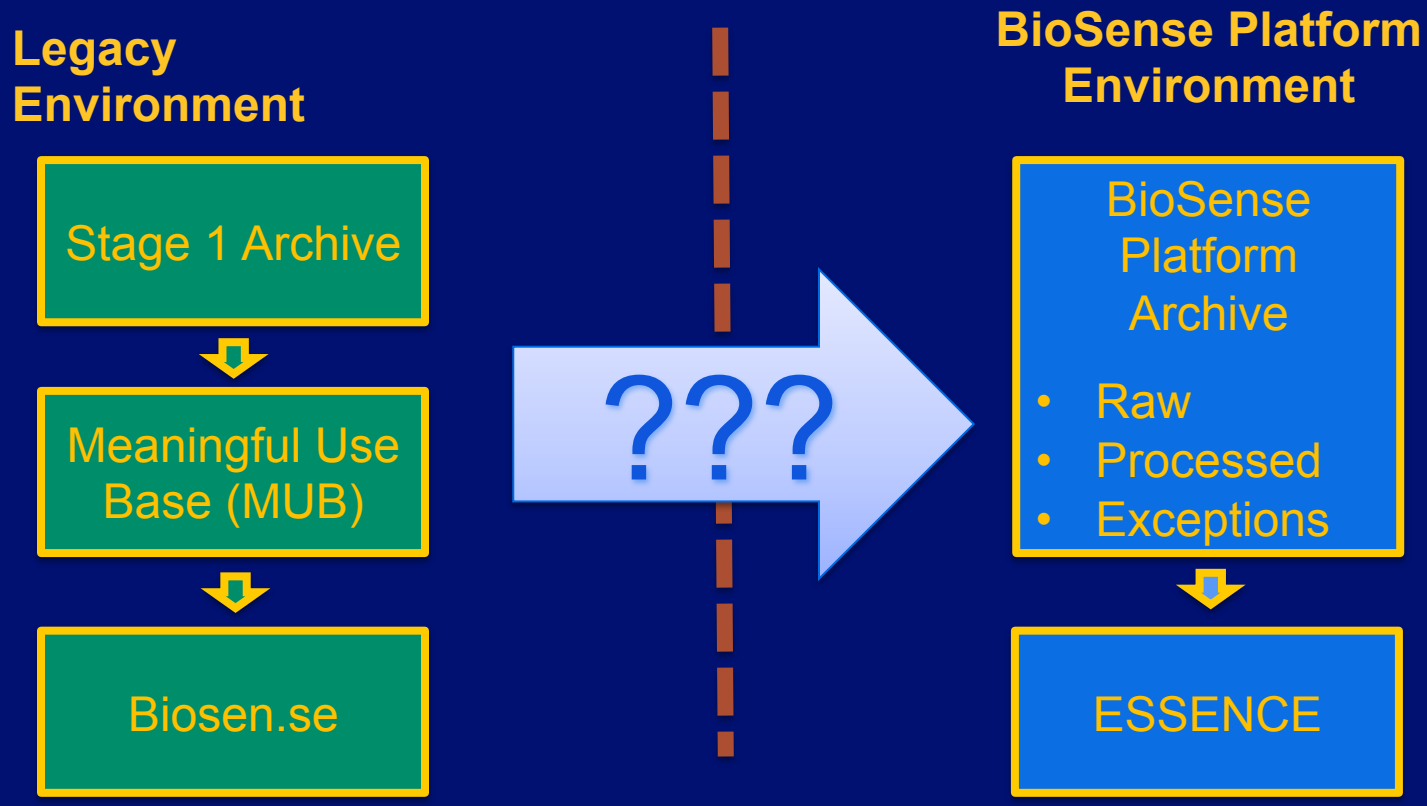
BioSense Platform Data Flow Webinars

Part 1. Data Ingestion into the BioSense Platform

Part 2. Data Ingestion into ESSENCE

Part 3. Migrating Legacy BioSense Data

BioSense Platform Data Structures



Legacy Data Conversion

- **Goal**
  - Convert Stage 1 (“Legacy”) BioSense data into new BioSense Platform data structures
- **Process**
  - Identify Stage 1 fields that map to columns in Archive Processed table
  - Determine processing
    - Map data directly (where possible)
    - Identify areas where processing modifications are required
  - Replace legacy separator (“:” or “:SEP:”) with new separator (“;”)

Check out the protected **NSSP Doc Review** folder on the ISDS Forum to review the proposed mapping between legacy Stage 1 variables and the BioSense Platform Archive Processed table variables.

Legacy Data Conversion Assumptions

- **Records in Stage 1 infrastructure for conversion to each site**
  - Each site will specify the date range for its data conversion
- **During the conversion process…**
  - Records will be associated with a valid Facility ID listed on MFT or Crosswalk
  - Calculated fields will be generated as described for BioSense Platform data flow
- **BioSense Platform Archive won't contain legacy data for…**
  - Archive Raw table
  - Combo fields in Processed table
  - Segment fields in Processed table

How do I know I'm reviewing a legacy record?
Legacy_Flag* = Yes
*new column

Legacy Data Conversion: Facility Information

- **Legacy Processing**
  - Legacy data includes outdated facility information
  - Legacy data does not include facility characteristics like name and location
- **Challenge**
  - Legacy records should be mapped to confirmed MFT information
- **Proposed Solution**
  - Preserve legacy facility information
    - Sending Facility (MSH-4)
    - Treating Facility (EVN-7)
  - Generate calculated facility ID – select the first non-null and valid ID found in the related jurisdiction's MFT
    - EVN-7
    - MSH-4

Associated Archive Variables

- Treating_Facility and Sending_Facility
- C_Facility_ID
- C_MFT_Patient_Class and C_Patient_Class
- C_BioSense_ID and C_Processed_BioSense_ID

Legacy Data Conversion: Site

- **Legacy Processing**
  - Generally, SFTP-routed data is directed to site-specific database
  - PHINMS database contains data for sites that use PHINMS
- **Challenge**
  - Difficult to identify which records are tied to each site when a single legacy database contains data for multiple sites

Legacy Data Conversion: Site (continued)

- **Proposed Solution**
  - Assign site to each legacy record
    - For legacy site-specific databases → determine site from feed name
    - For databases that contain multiple sites
      - Check related site's MFT for the C_Facility_ID
      - Challenge: the same facility ID may appear in multiple sites' MFTs
        - Confirm selection by checking source file names and visit information
        - Consult with sites to confirm results
  - Store converted legacy records in appropriate site's Archive Processed table

Legacy Data Conversion: Date/Time Stamps

Archive Processed Timestamp information	Legacy Conversion Processing Description (proposed)
Arrived_Date_Time – *when the message arrived at the BioSense Platform*	Null
Create_Raw_Date_Time – *when the message was written to the Archive Raw table*	Null
Create_Processed_Date_Time – *when the message was written to the Archive Processed table*	Insert legacy value for **Create_Date_Time** (when the record was written to legacy data table)
Update_Processed_Date_Time – *when the record was last updated in the Archive Processed table*	System timestamp indicating when the legacy record was reprocessed/ converted *note that the legacy field Update_Date_Time is dropped

Legacy Data Conversion: Unique Patient ID

- **Legacy Processing**
  - Select first non-null value across segments containing a patient ID
- **Challenge**
  - Legacy data does not include a medical record number
- **Proposed Solution**
  - Use legacy "first patient ID" as a proxy for medical record

Legacy conversion processing for C_Unique_Patient_ID	Archive processing for C_Unique_Patient_ID
First non-null value from: • PID_First_Patient_ID (PID-3) • PID_2_1_Patient_ID_External (PID-2) • PID_4_1_Alternate_Patient_ID (PID-4)* • PID_18_1_Patient_Account_ID (PID-18)* • PV1_19_1_Patient_Visit_ID (PV1-19)	First non-null value from: • Medical Record Number (PID-3, type = "MR") • Patient ID (PID-2.1) • First Patient ID (from PID-3) • Patient Account Number (PID-18)* • Visit Number (PV1-19)

* Information not stored separately in the BioSense Platform Archive

Legacy Data Conversion: Chief Complaint

- **Legacy Processing**
  - Store Chief Complaint Information from OBX-5
- **Challenge**
  - Legacy processing does not differentiate between chief complaint information sent as a coded value or sent as text
- **Proposed Solution**
  - Preserve legacy information for the Chief Complaint from OBX-5
    - *Legacy* OBX_5_1_Chief_Complaint → *Archive* **Chief_Complaint_Text**
  - Calculate chief complaint as defined for the Archive C_Chief_Complaint
    - First non-null value from:
      - Chief_Complaint_Text
      - Admit_Reason_Description
    - Archive will not include the legacy concatenated chief complaint

Legacy Data Conversion: Admit Reason

- **Legacy Processing**
  - 3 data elements – code, text, alt text
- **Challenge**
  - Map legacy values to new Archive variables
- **Proposed Solution**
  - Leverage new processing to select the first non-null text value

Legacy Stage 1 variable name	Archive Processed variable name
PV2_3_1_Admit_Reason_ID	Admit_Reason_Code
Select first non-null value from: • PV2_3_2_Admit_Reason_Text • PV2_3_5_Admit_Reason_Alt_Text	Admit_Reason_Description
N/A	Admit_Reason_Combo
N/A	Admit_Reason_Segment

Legacy Data Conversion: Visit Date

- **Legacy Processing**
  - Earliest Date Time among available date/time values
- **Challenge**
  - Legacy processing differs from Archive processing
- **Proposed Solution**
  - Generate a calculated visit date/time following Archive processing for legacy data
  - Ingest legacy date/time values but do not preserve "Earliest Date Time"

Legacy Processing

Earliest date among:

- OBX-14 - Date/Time of Observation
- PV1-45 - Discharge Date/Time
- PV1-44 - Admit Date/Time
- PR1-5 - Procedure Date/Time
- PID-29 - Patient Death Date and Time
- EVN-2 - Recorded Date/Time
- MSH-7 Message Date/Time

Archive Processing

First non-null value from:

- Admit Date/Time (PV1-44)
- *Earliest date among:*
  - Discharge Date/Time (PV1-45)
  - Procedure Date/Time (PR1-5)
  - Patient Death Date/Time (PID-29)
  - Recorded Date/Time (EVN-2)
  - Date/Time of Message (MSH-7)

Legacy Data Conversion: Race and Ethnicity

- **Legacy Processing**
  - Concatenate any repeating non-null information for the code or description
- **Challenge**
  - No easy or consistent way to split legacy race and ethnicity data into the columns defined in the Archive
- **Proposed Solution**
  - Set Code and Description columns in the Archive to null
  - Insert legacy data directly into the combo field
  - Include in documentation that the legacy combo field data is not structured like the typical Archive combo field data

PID_10_Patient_Race (*Legacy*)
2054-5:Black
U:Unknown
HISPANIC OR LATINO:H
1~3
D

Legacy Data Conversion: Initial Acuity

- **Legacy Processing**
  - Concatenate any non-null value for OBX-5.1 – OBX-5.5
  - Did not account for null values

OBX_5_1_Acuity_Assessment
1
2
I-Resuscitation
V-Non urgent

- **Challenge**
  - Cannot reliably separate legacy data into distinct components for codes and descriptions
- **Proposed Solution**
  - Set Code and Description columns in the Archive to null
  - Insert legacy data directly into the combo field
  - Include in documentation that the legacy combo data is not structured like the new combo data

Legacy Data Conversion: Procedure Information

- **Legacy Processing**
  - Legacy data includes PR1-3.3, which identifies the type of code present in PR1-3.1 (e.g., I9CP, I10P)
- **Challenge**
  - Archive does not capture this information in a distinct column
  - Information is available in the _Segment column, if it is sent
- **Proposed Solution**
  - Option 1: drop the legacy column and do not include in the Archive
  - Option 2: artificially create a Segment column, noting it will appear differently than typical Segment columns in the Archive

Legacy	Archive Processed
PR1_5_1_Procedure_Date_Time	Procedure_Date_Time
PR1_3_1_Procedure_Code_ID	Procedure_Code
PR1_3_2_Procedure_Code_Text	Procedure_Description
PR1_3_3_Procedure_Code_NS	??

Legacy Data Conversion: Recap

- **Goal is to convert existing Stage 1 ("Legacy") BioSense data into the new BioSense Platform data structures**
  - Improve historic challenges with legacy data
  - Provide clear documentation about processing decisions
- **Plan to start converting legacy data once sites have moved to production data flow**

We need your help!

We're asking 6-9 volunteers to help us explore and finalize the plan to convert legacy data. If you'd like to help, email us at nssp@cdc.gov.

Recap: Access & Management Center

Max Worlund, ICF Project Manager

Access & Management Center (AMC)

- ***NEW* Admin Tool renamed!**
  - **BioSense Platform: Access and Management Center (AMC)**
- **AMC Development Update**
  - First release:
    - User management
    - Data access rules
  - Future releases:
    - Data templates
    - User groups
      - Shared facility
    - Restricted accounts
    - Reports
    - Additional enhancements

What is the AMC?

- Performs user management for the BioSense Platform
- Allows sites to control access to their data in ESSENCE

Additional functionality for facility administration is coming soon.

For details, see the ISDS Webinar on the BioSense Platform Admin Tool.

Access & Management Center: New Functionality

- **Removed concept of "Role"**
  - *NEW* Epidemiologist Flag – indicates if a user account is associated with an epidemiologist (yes/no – optional)
- ***NEW* User Groups**
  - Site administrators can create their own groups of users
    - Public user groups – visible to all site administrators
    - Private user groups –visible ONLY within a site
- ***NEW* Restricted Accounts**
  - Site administrators can "restrict" accounts
  - Restricted accounts are visible ONLY within a Site

Default User Groups (public groups)

- **All System User Group** – all system users have access to National View in ESSENCE
- **All Site X User Group** – one group per site
- **All Site X Epidemiologists User Group** – one group per site

User Group Example

Site	User Group Name	Public/Private?
001	My User Group	Public
001	My Internal Group	Private
001	All Users	Public/Default
001	All My Site Users	Public/Default
001	All My Site Epis	Public/Default

Site 001 Administrator sees…

Data Access Rule: Select User

My User Group
My Internal Group
All Users
All My Site Users
All My Site Epis

Only **your site** can see your site's **private** groups

Site 999 Administrator sees…

Data Access Rule: Select User

My User Group
All Users
All My Site Users
All My Site Epis

Shared Facility Example

- **Scenario**
  - General Hospital is a shared facility between sites X and Y
  - Site X is the designated steward of General Hospital
  - Site Y wants to see and share data from General Hospital in ESSENCE
- **How can the Access & Management Center accomplish this goal?**
  - Site Y creates public User Group called "Site Y: General Hospital Users"
  - Site X creates a *Data Access Rule* that includes
    - Who? – Site Y: General Hospital users
    - What? – All General Hospital data details

How can Site X share data?

- Create additional *Data Access Rules*

How can Site Y share data?

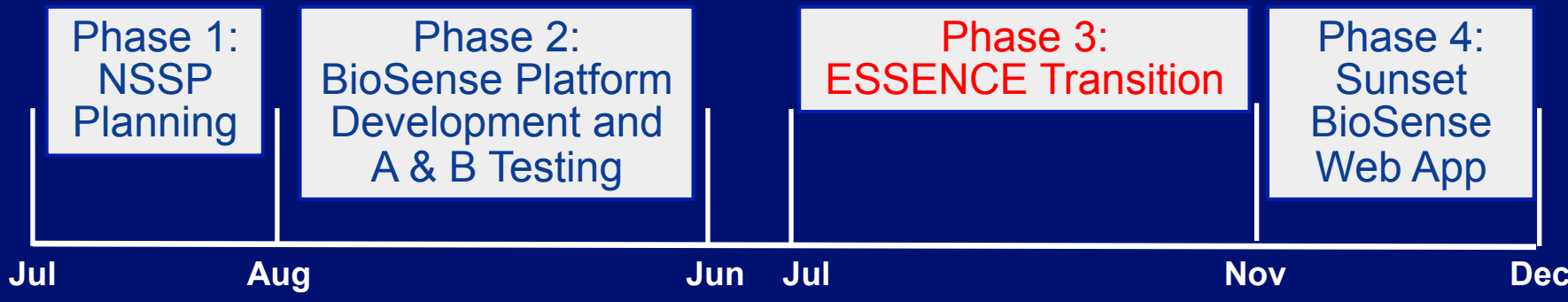
- Add users to the "Site Y: General Hospital Users" public group

Data Transition Plan

Michael Coletta, MPH, NSSP Program Manager

Phase III: Transition to ESSENCE

- **Phase III: July – November 2016**
  - Every 4 weeks, nine (9) sites transition to the new data flow
  - Prior to transition, the MFT for your site must be finalized
- **Transition Plan**
  - 2 webinars will be conducted with each set of sites
    - Transition Plan and Adminer* Orientation
    - Access & Management Center (AMC) and ESSENCE Orientation



* Adminer is the SQL tool that allows you to view the MS SQL data in the BioSense Platform Archive

Transition Plan

- **Purpose**
  - Ensure key functions of tools work as intended
  - Ensure data permissions and access are accurate
  - Provide general feedback on tools and the *BioSense Platform Quick Start Guide*
- **Process**
  - Attend webinar overviews
  - Test applications
    - Adminer
    - AMC
    - ESSENCE
  - Refer to *BioSense Platform Quick Start Guide*
  - Explore utility of applications to meet your individual needs
  - Provide feedback about tools, applications, and documentation

We're here to help!
If you have questions or encounter challenges, use the NSSP Helpdesk to submit a ticket at support.syndromicsurveillance.org

Transition Schedule

Phase III Operations	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52
	22-Jul	29-Jul	5-Aug	12-Aug	19-Aug	26-Aug	2-Sep	9-Sep	16-Sep	23-Sep	30-Sep	7-Oct	14-Oct	21-Oct	28-Oct	4-Nov	11-Nov	18-Nov	25-Nov	2-Dec	9-Dec	16-Dec	23-Dec	30-Dec
	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
	18-Jul	25-Jul	1-Aug	8-Aug	15-Aug	22-Aug	29-Aug	5-Sep	12-Sep	19-Sep	26-Sep	3-Oct	10-Oct	17-Oct	24-Oct	31-Oct	7-Nov	14-Nov	21-Nov	28-Nov	5-Dec	12-Dec	19-Dec	26-Dec
Illinois (includes Cook) / Massachusetts / Kentucky / Arizona / Mississippi / Arkansas / West Virginia / Kansas / Houston, TX																								
Nevada / Utah / New Mexico / Denver Public Health / Riverside, CA / Idaho / North Dakota / Montana /Alaska																								
Stanislaus, CA / Linn County, IA / Santa Clara, CA / Nevada, CA / Florida / Ohio /Pennsylvania / New York / North Carolina																								
Georgia / New York City / New Jersey / Indiana / Tarrant County TX / Missouri / Louisiana / Maryland / Washington																								
Oklahoma / Minnesota / Connecticut / South Carolina / Oregon / Maine / Nebraska / New Hampshire / Rhode Island																								
Boston Public Health Commission / County of Sacramento, CA / District of Columbia / Delaware / San Diego, CA / Hawaii / Vermont / South Dakota / San Mateo, CA																								

**** Confirm your dates with us no later than May 13, 2016 ****

Transition Details: Group 1

Arizona, Arkansas, Houston (TX), Illinois (includes Cook), Kansas, Kentucky, Massachusetts, Mississippi, West Virginia

- Conference Call 1: July 19, 1:00–2:30 pm (ET)
- Conference Call 2: July 26, 1:00–2:30 pm (ET)
- Site Confirmation: August 11

Week 1					Week 2				
18-Jul	19-Jul	20-Jul	21-Jul	22-Jul	25-Jul	26-Jul	27-Jul	28-Jul	29-Jul

Week 3					Week 4				
1-Aug	2-Aug	3-Aug	4-Aug	5-Aug	8-Aug	9-Aug	10-Aug	11-Aug	12-Aug

LEGEND	
Conference Call 1	
Conference Call 2	
Site Confirmation (Email)	

Transition Details: Group 2

Alaska, Denver Public Health, Idaho, Montana, Nevada, New Mexico, North Dakota, Riverside (CA), Utah

- Conference Call 1: August 9, 1:00–2:30 pm (ET)
- Conference Call 2: August 16, 1:00–2:30 pm (ET)
- Site Confirmation: September 1

Week 1					Week 2				
8-Aug	9-Aug	10-Aug	11-Aug	12-Aug	15-Aug	16-Aug	17-Aug	18-Aug	19-Aug

Week 3					Week 4				
22-Aug	23-Aug	24-Aug	25-Aug	26-Aug	29-Aug	30-Aug	31-Aug	1-Sep	2-Sep

LEGEND	
Conference Call 1	
Conference Call 2	
Site Confirmation (Email)	

Transition Details: Group 3

Florida, Linn County (IA), Nevada (CA), New York, North Carolina, Ohio, Pennsylvania, Santa Clara (CA), Stanislaus (CA)

- Conference Call 1: August 30, 1:00–2:30 pm (ET)
- Conference Call 2: September 7, 1:00–2:30 pm (ET)
- Site Confirmation: September 22

Week 1					Week 2				
29-Aug	30-Aug	31-Aug	1-Sep	2-Sep	5-Sep	6-Sep	7-Sep	8-Sep	9-Sep

Week 3					Week 4				
12-Sep	13-Sep	14-Sep	15-Sep	16-Sep	19-Sep	20-Sep	21-Sep	22-Sep	23-Sep

LEGEND	
Conference Call 1	
Conference Call 2	
Site Confirmation (Email)	

Transition Details: Group 4

Georgia, Indiana, Louisiana, Maryland, Missouri, New Jersey, New York City, Tarrant County (TX), Washington

- Conference Call 1: September 20, 1:00–2:30 pm (ET)
- Conference Call 2: September 27, 1:00–2:30 pm (ET)
- Site Confirmation: October 13

Week 1					Week 2				
19-Sep	20-Sep	21-Sep	22-Sep	23-Sep	26-Sep	27-Sep	28-Sep	29-Sep	30-Sep

Week 3					Week 4				
3-Oct	4-Oct	5-Oct	6-Oct	7-Oct	10-Oct	11-Oct	12-Oct	13-Oct	14-Oct

LEGEND	
Conference Call 1	
Conference Call 2	
Site Confirmation (Email)	

Transition Details: Group 5

Connecticut, Maine, Minnesota, Nebraska, New Hampshire, Oklahoma, Oregon, Rhode Island, South Carolina

- Conference Call 1: October 11, 1:00–2:30 pm (ET)
- Conference Call 2: October 18, 1:00–2:30 pm (ET)
- Site Confirmation: November 3

Week 1					Week 2				
10-Oct	11-Oct	12-Oct	13-Oct	14-Oct	17-Oct	18-Oct	19-Oct	20-Oct	21-Oct

Week 3					Week 4				
24-Oct	25-Oct	26-Oct	27-Oct	28-Oct	31-Oct	1-Nov	2-Nov	3-Nov	4-Nov

LEGEND	
Conference Call 1	
Conference Call 2	
Site Confirmation (Email)	

Transition Details: Group 6

Boston Public Health Commission, County of Sacramento (CA), Delaware, District of Columbia, Hawaii, San Diego CA, San Mateo CA, South Dakota, Vermont

- Conference Call 1: November 1, 1:00–2:30 pm (ET)
- Conference Call 2: November 8, 1:00–2:30 pm (ET)
- Site Confirmation: November 22

Week 1					Week 2				
31-Oct	1-Nov	2-Nov	3-Nov	4-Nov	7-Nov	8-Nov	9-Nov	10-Nov	11-Nov

Week 3					Week 4				
14-Nov	15-Nov	16-Nov	17-Nov	18-Nov	21-Nov	22-Nov	23-Nov	24-Nov	25-Nov

LEGEND	
Conference Call 1	
Conference Call 2	
Site Confirmation (Email)	

Next Steps

Michael Coletta, MPH, NSSP Program Manager

Next Steps

- **Review documentation in the protected NSSP Doc Review folder on the ISDS Forum**
- **Volunteer for a follow-up session about the legacy data conversion by May 10, 2016**
  - 6–9 participants needed
- **Review the transition timeline for your site**
  - Let us know of scheduling conflicts by May 13, 2016

Want to sign up or ask a question?

Contact us: **nssp@cdc.gov**

We appreciate your input.

Michael A. Coletta, MPH
Manager, National Syndromic Surveillance Program
CDC/CSELS/DHIS
mcoletta@cdc.gov

For more information, please contact Centers for Disease Control and Prevention

1600 Clifton Road NE, Atlanta, GA 30329-4027
Telephone: 1-800-CDC-INFO (232-4636)/TTY: 1-888-232-6348
Visit: http://www.cdc.gov | Contact CDC at: 1-800-CDC-INFO or http://www.cdc.gov/info