Final Recommendation:

Core Processes and EHR Requirements for Public Health Syndromic Surveillance

International Society for Disease Surveillance (ISDS)

Meaningful Use Workgroup

January 31, 2011

For questions, contact:

Charles Ishikawa, MSPH
Program Manager
Public Health Practice
cishikawa@syndromic.org
617-779-0886

Table of Contents

Table of Figures & Tables

2. The Core Business Processes of Syndromic Surveillance

Business processes describe the means by which organizations accomplish a goal to produce something of value. These processes detail the consecutive tasks or task sets through which value is added. In contrast to use cases, business processes provide a holistic picture for understanding how an information system can add value for its users and interface with other organizational activities to build efficiencies.[1]

The three, core business processes for PHSS is based on a current-state analysis of practice. Informed by widely used and best practices, this model is intended to guide stakeholders in planning, designing or implementing EHR solutions during Meaningful Use, Stage 1. This model may also serve as a basis for system redesign.

This section describes critical, characterizing components of the core business processes, including:

1. Role of PHSS within the three core public health functions
2. Core PHSS business objectives
3. Entities and high level transactions that are key to PHSS
4. Inputs that are required for production

2.1 Goal

Of the three, core public health functions recommended by the Institute of Medicine in, "The Future of Public Health", PHSS is a part of the assessment function. Similar to other surveillance processes (e.g., laboratory-confirmed reportable conditions, or behavioral risk factor surveillance), PHSS systems utilize health data to produce information in support of an overarching public health surveillance goal to:

> "…regularly and systematically collect, assemble, analyze, and make available information on the health of the community, including statistics on health status, community health needs, and epidemiologic and other studies of health problems."[2]

In contrast to other surveillance processes, PHSS is unique due to its use of near "real-time" patient data and statistical tools. PHSS processes enable public health authorities to provide timely assessments of population health that, in conjunction with other core and event-specific activities, assist with determining and assessing the implementation of public health action. This is particularly useful for event detection, situation awareness, and response management. However, it is important to emphasize that PHSS information is one of many artifacts produced by a public health authority's surveillance function.

Given its value to public health surveillance, PHSS is best contextualized within the Common Ground Preparedness Framework (Figure 3). PHSS processes produce information that may trigger a response, alter risk mitigation strategies, or impact the allocation and distribution of resources.

---

[1] Public Health Informatics Institute. (2006). *Taking Care of Business: A collaboration to define local health department business processes.* Decatur, GA: Public Health Informatics Institute

[2] Committee for the Study of the Future of Public Health. *The Future of Public Health: Summary & Recommendations.* (1988) pg. 7
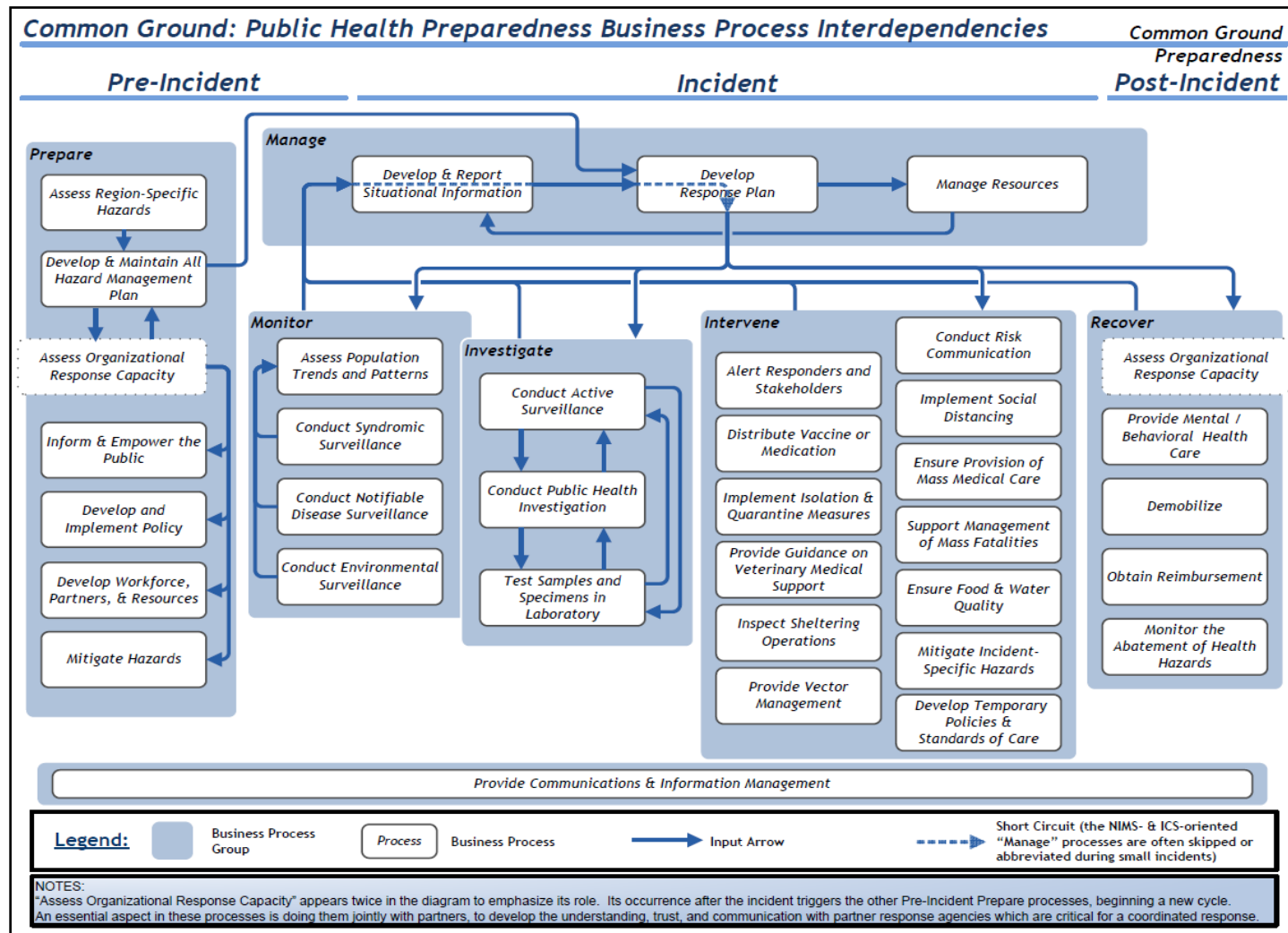
Figure 3: The Common Ground Preparedness Framework was developed through a three-year collaboration of eight state or local health departments, brought together to define public health's business processes related to preparedness. The framework has three phases: Pre-Incident, Incident, and Post-Incident. 33 business processes are contained in six business process groups: Prepare, Monitor, Investigate, Intervene, Recover, and Manage. Syndromic surveillance is located within the Monitor process group. A 34th process involving communications supports all the other processes. Arrows indicate information flow between processes or process groups. (Gibbson, Theodore and Nichole)

2.2 Objectives

Business objectives are statements that reflect what an organization seeks to achieve with its processes. Based on the Workgroup's analysis of PHSS, seven objectives were identified. Each, by way of a PHA's core surveillance function, contributes to ensuring the health and well-being of community and population health through public health interventions and activities.

In conjunction with other core public health activities, PHSS business objectives include the following:

1. Provide ongoing, timely intelligence and data on public health threats or health conditions of interest
2. Support early identification or ruling out of public health threats, conditions of public health importance, or suspected incident(s)
3. Assist in characterizing population groups at greatest risk
4. Assist in assessing the severity and magnitude of possible threat(s) and the effectiveness of control measures
5. Assist with continual evaluation and development of new and improved surveillance practices
6. Keep stakeholder organizations, public health leadership, and the public informed (as appropriate) about conditions of public health importance
7. Support collaborative efforts with health providers, media, first responders, and government decision makers

2.3 Entities & Transactions

A **context diagram** is used to illustrate the participants and the information flows necessary for business. Participants, referred to as entities, are represented in the diagram as circles. Lines between entities represent information flow or transaction. The straight lines have arrows that indicate the direction of the transaction as information is exchanged between entities.

The context diagram reflects the relationships and boundaries that exist between entities of PHSS (Figure 4). While the context diagram shows the relationships between PHSS entities, they do not reflect the sequence or order of transactions, processes or tasks.

The focus for this document is the interaction between Hospital ED and UC health data providers and the PHA or their designee. The health data sent from an ED or UC provider consists of the recommended minimum data set described in Section 4.4.

Transactions between the PHA and the remaining entities may occur at different parts of the business processes. However, it is recommended that an assessment be conducted to detail these interactions when applying this recommendation to any specific practice environment.

Public Health Syndromic Surveillance Context Diagram
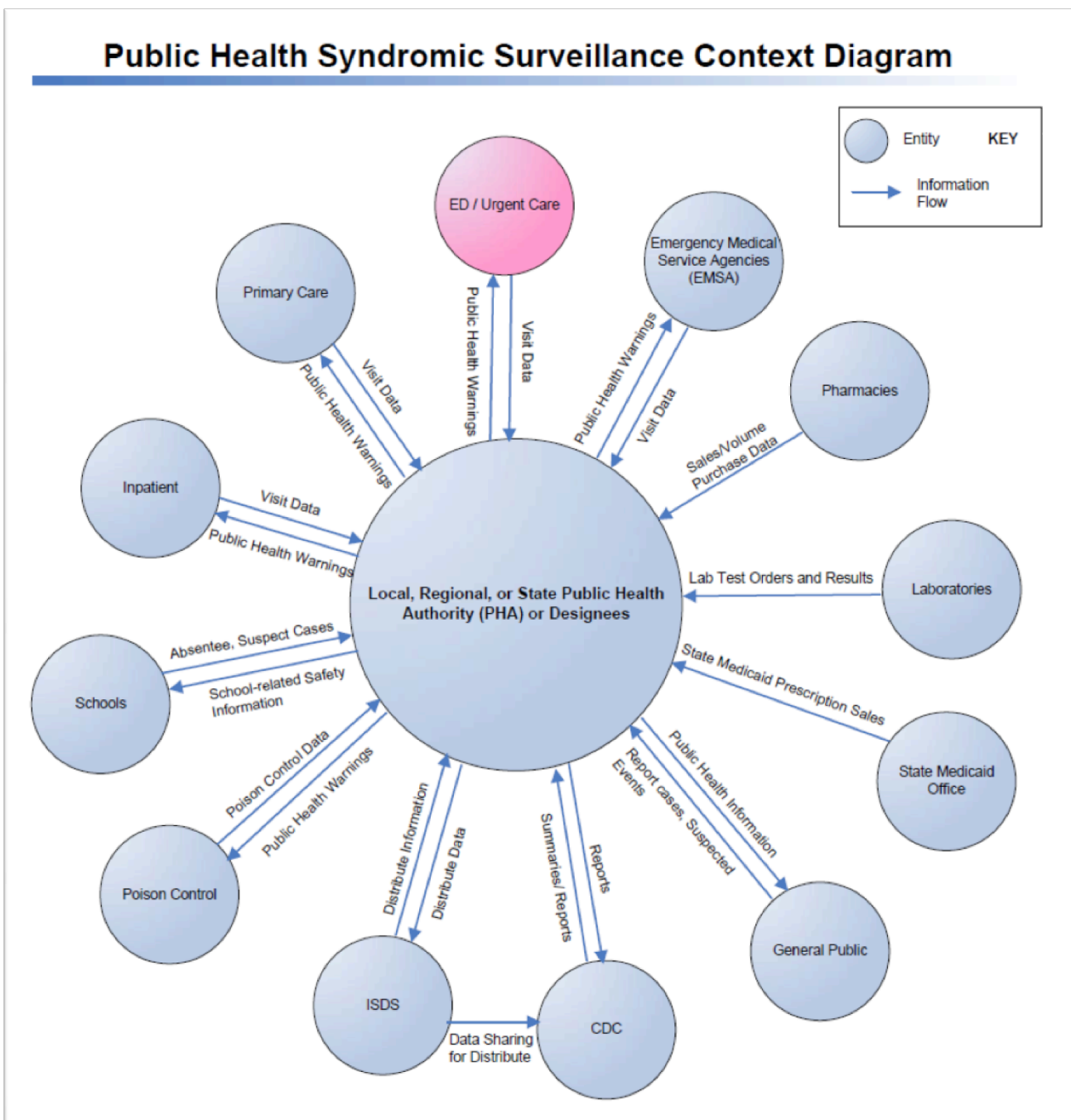
Figure 4 Business context diagram for PHSS: This context diagram illustrates the relationships that exist between entities of PHSS and shows the flow of information required by the core business processes. The focus of this final recommendation is the interaction between ED and UC data and the PHA or their designee. The data sent from the ED or UC provider consists of the recommended minimum data set described in Section 4.

2.4 Inputs

Inputs are information or variables received by a business that affect process outputs. For PHSS, there are seven kinds of inputs that influence the surveillance information produced by the core business processes.

Table 1 summarizes the various inputs. These inputs affect the processes at points that are indicated in the business process task flow diagrams detailed in Section 3.

Input Name	Summary of Input
Health Data	Emergency Department (ED), Urgent care (UC) visits
Health Condition of Interest	Indicators of: Infectious Disease, Environmental Exposures, Injury, Mental Health Conditions, Health Care Utilization, and Exacerbations of Chronic Disease Conditions
Purpose	Event Detection, Response Management, Situation Awareness
Context	Routine, Elevated risk due to an anticipated threat, Elevated risk due to a present threat
Information from other Public Health Activities and Processes	Advice from subject matter experts; Distribute data from neighboring jurisdictions; statistical patterns or aberration characterized by other surveillance systems (e.g. NEDSS); or an astute clinician's case report
Outside Influences	Media reports; weather patterns; air quality measurements; and political factors
Level of Authority	Local, State, Federal, District, Senior Decision-Maker

Table 1: Summary of inputs into the core, PHSS business processes that influence the decision points and pathway of the task flow.

Health Data

Health data from a provider sent to a public health authority are required for PHSS. These data are the observations for the PHSS epidemiological operations and analyses. The scope of this recommendation is *health data on ED and UC visits*.

Health Condition of Interest

As with other public health surveillance processes, population health is assessed for defined health conditions in PHSS. Best practices for monitoring the following conditions are well defined: *Indicators of Infectious Disease, Environmental Exposures, Injury, Exacerbations of Chronic Disease Conditions, and Health Care Utilization*. Practices for monitoring Exacerbations of Mental Health Conditions are rapidly developing.

Purpose

PHSS systems are utilized for three general purposes: *1) Event Detection; 2) Response Management; and 3) Situation Awareness*. When a public health authority engages its PHSS system, the purpose influences the delivery of the surveillance information.

Context

A context is formed by the circumstances within which a public health authority operates a PHSS system. Context is subject to change. There are three contexts within which PHSS systems operate:

- *Routine use*: PHSS systems are routinely used for public health surveillance information. Routine utilization can be daily, weekly, monthly, seasonal, or annual. Examples include: Seasonal influenza-like-illness monitoring, daily review for unusual clusters of illness, or annual ED utilization for injury care.

- *Elevated risk due to an anticipated threat:* There are circumstances where the potential for a population health threat is high. In this context, a threat is anticipated, but not actually present. Examples include: Recurring events such as marathons or local sporting events; and single events, such as larger sporting events (e.g. Super Bowl), or events of national significance, including political conventions and international diplomatic visits.

- *Elevated risk due to a present threat:* Sometimes a threat to population health is present. Examples include: Recurring events, such as heat waves or the response and recovery phases following a natural disaster; and single events, such as communicable disease outbreaks or exposures to hazardous materials.

Information from other Public Health Activities and Processes

PHSS outputs are influenced by the information and participation of parties from other public health activities and processes. Examples include: Advice from subject matter experts; Distribute data from neighboring jurisdictions; statistical patterns or aberration characterized by other surveillance systems (e.g. NEDSS); or an astute clinician's case report.

Outside Influences

Information from outside processes influences PHSS processes. Examples include: Media reports; weather patterns; air quality measurements; and political factors.

Level of Authority

Parties from multiple levels of authority contribute to PHSS. Surveillance analysts, public health investigators, epidemiologists, and senior decision makers from local, regional, state, or federal jurisdictions can be involved.

3. Schematics: Core Business Processes and Task Sets

A **business process** describes a set of activities and tasks that logically group together to accomplish a goal or produce something of value for the benefit of the organization, stakeholder, or public.[3]

In this section, the core business processes of public health syndromic surveillance (Table 2) are described in detail, along with the tasks and decision points that consecutively produce PHSS information. Drawn primary from the perspective of a PHSS analyst and rendered as a generalized model, EHR vendors, hospitals and eligible health professionals should work with PHA's to identify significant idiosyncrasies.

ID	Business Process or Task Set Name
BP 1	Conduct Syndrome-Based Population Health Monitoring
TS 1	Collect and Process Data
TS 2	Characterize, Interpret, and Analyze Data
TS 3	Notify and Engage Partners / Leadership
TS 4	Conduct Reach-back
BP 2	Establish and Maintain Data Sharing Partnerships
BP 3	Conduct Data Quality Assurance

Table 2: List of the Core Public Health Syndromic Surveillance (PHSS) Business Processes

3.1 Conduct Syndrome-Based Population Health Monitoring (BP 1)

Conducting syndrome-based population health monitoring is the core business process of public health syndromic surveillance. Upon identifying a potential public health concern through the characterization, interpretation, and analysis of data, the public health syndromic surveillance unit determines whether to escalate a potential concern, notifies and engages partners and leadership when applicable, assists in determining whether a response is needed, and may assist in the response actions.

Objective of the Business Process: To conduct syndrome-based population health monitoring and assist in the assessment, detection, communication, and response to public health conditions of interest.

Trigger:

For routine monitoring, the trigger is ED / Urgent Care data made available or sent at least every 24 hours.

Other triggers include: knowledge of events considered to be of elevated risk; information from other Public Health activities or outside influences that indicates a possible health condition of interest that warrants monitoring.

---

[3] Public Health Informatics Institute, "Taking Care of Business: A Collaboration to Define Local Health Department Business Processes", Decatur, GA, 2006.

Assumptions of the Business Process: It is assumed that a data sharing partnership has already been established between the data provider and receiver. See Business Process 2 (BP2): *Establish and Maintain Data Sharing Partnerships.*

Input to the Business Process: Health data; Level of authority; Health condition of interest; Purpose; Output of QA business process; Information from other public health activities and processes; Outside influences.

Output of the Business Process: May include: Reports (Routine/SITREP); Documentation of the response or what was seen or not seen in the data; Health alerts; Information to public health leadership/PIO role.

.
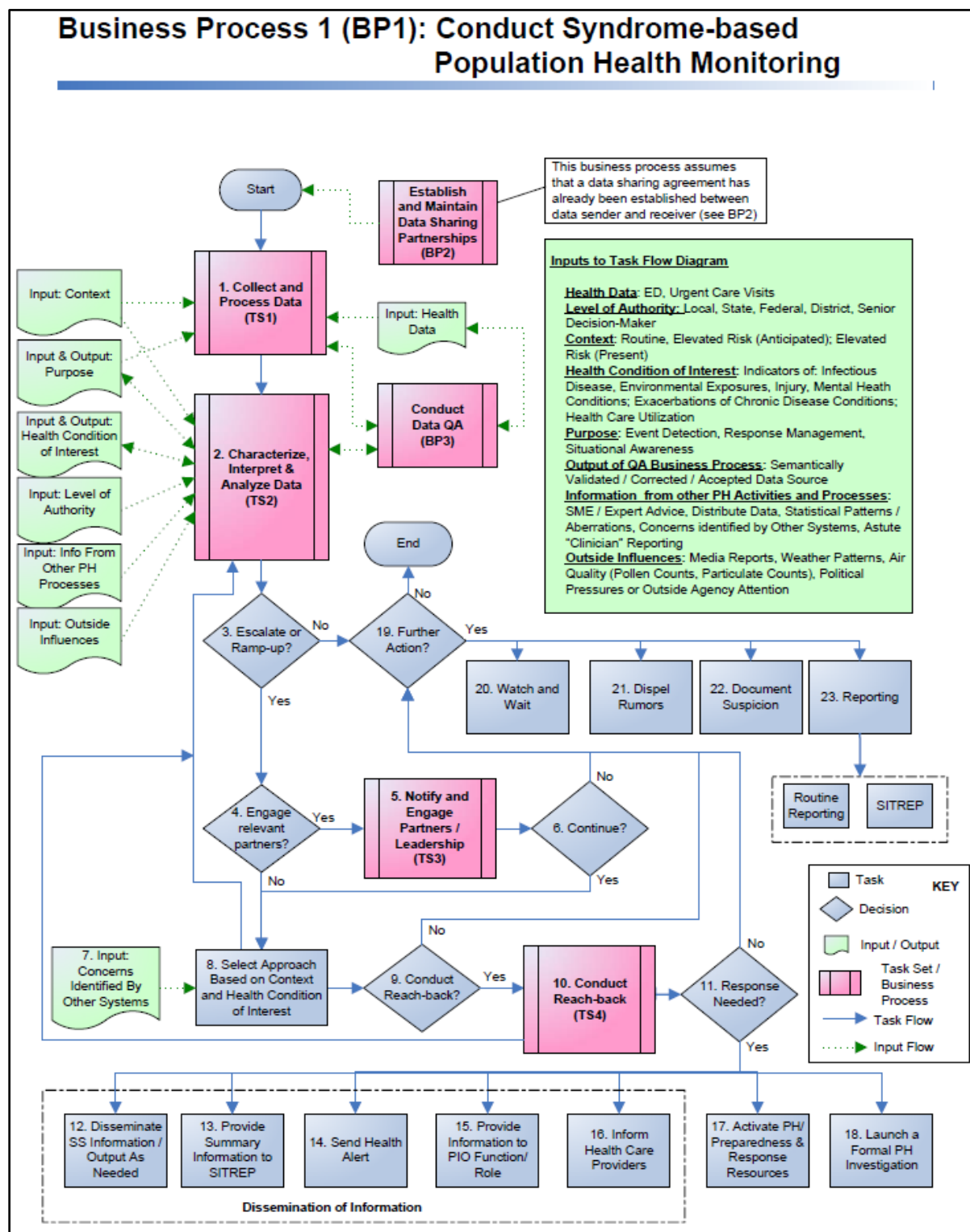
Figure 5: Task Flow Diagram of BP 1 - *Conduct Syndrome-Based Population Health Monitoring:* Monitor and assist in the assessment, detection, communication, and response to public health conditions of interest.

Conduct Syndrome-based Population Health Monitoring (BP 1)

Monitor and assist in the assessment, detection, communication, and response to public health conditions of interest.

All numbers refer to the flow diagram for BP 1 (Figure 7) unless otherwise noted.

BP1-1: Collect and Process Data (TS1)

- The Public Health Authority (PHA) receives ED/Urgent Care Visit health data and engages the task set *Collect and Process Data* (TS1)

- See *Collect and Process Data* (TS1) for a detailed description of this task set.

- The inputs that affect the decisions and pathway within this task set are Health Data; Context; Purpose; and the output of the *Conduct Data QA (BP3)* business process, which is a semantically validated/corrected data source.

- The main output of this task set is the results of the automated statistical algorithms and counts, which include time series and summary counts of flags.

BP1-2: Characterize, Interpret, and Analyze Data (TS2)

- Once the data are collected and processed, the data are characterized and analyzed.

- See *Characterize, Interpret and Analyze* (TS2) for a detailed description of this task set.

- The inputs that affect the decisions and pathway within this task set are Context; Health Condition of Interest; Purpose; Level of Authority; Data from Other Public Health (PH) Processes; Outside Influences, and the output of the *Conduct Data QA (BP3)* business process, which is a semantically validated/corrected data source.

- The output of this task set is conclusions and results of characterization and pattern analysis; Purpose; and Health Condition of Interest.

BP1-3: Escalate or Ramp-up?

- This question is answered by the PHSS analyst who analyzed and characterized the data in TS2-*Characterize, Interpret, and Analyze Data.* Based on the results and conclusion in TS2, the analyst determines whether to escalate or ramp-up any issues based on the specifics of results, the severity of the issue, and potential for spread.

- The specific criteria to determine whether an issue should be escalated vary across jurisdictions and depend greatly on the specifics of the issue.

BP1-4: Engage Relevant Partners?

- If an issue is escalated, the question of whether to engage relevant partners and/or leadership is answered by the PHSS analyst and depends on a variety of inputs, including: variables that quantify the severity and potential spread of the potential concern; the level of authority of the PHA; the context; purpose; health condition of interest; outside influences (e.g. media or political).

  o Example: From the results of analysis, if an analyst suspects a possibility of bacterial meningitis, but the probability is questionable or low, then the analyst may decide to contact the data provider directly to follow-up. However, if the analyst detects a strong possibility of bacterial meningitis from data gathered in one day, the analyst would likely engage leadership.

Conduct Syndrome-based Population Health Monitoring (BP 1)

Monitor and assist in the assessment, detection, communication, and response to public health conditions of interest.

All numbers refer to the flow diagram for BP 1 (Figure 7) unless otherwise noted.

- o Example: Some local jurisdictions may engage data providers more frequently since there are fewer intermediaries between the two parties. At a local jurisdiction, if an analyst sees 'Anthrax' in a chief complaint field, prior to engaging leadership, the analyst may first confirm with the data provider that 'Anthrax' is not referring to the vaccination.
- The criteria for making this decision vary widely across jurisdictions.

BP1-5: Notify and Engage Partners / Leadership (TS3)

- See *Notify and Engage Partners / Leadership* (TS3) for a detailed description of this task set.
- Partners refer to Local, Regional, State Authority, Colleague, Designee, or Federal. Partners include public health and emergency preparedness.
- The output of this task set is a decision for next steps.

BP1-6: Continue?

- Based on the output of the task set *Notify and Engage Partners / Leadership* (TS3), the PHSS analyst / PHA determines whether to take further action or not.

BP1-7: Concern Identified By Other Systems

- There are instances where other systems identify a potential concern and request follow-up by public health syndromic surveillance to find relevant PHSS data.

BP1-8: Select Approach Based on Context and Health Condition of Interest

- The PHSS analyst determines whether additional analysis is needed (TS2) or whether to proceed with conducting reach-back based primarily on the context and health condition of interest.
- Example: If an analyst discovers the possibility of Anthrax, the analyst may decide there is an immediate need to acquire additional information ("breaking the glass", e.g. chart review) rather than further refinement of the data.

BP1-9: Conduct Reach-back?

- The question of whether to conduct reach-back is answered by the PHSS analyst based primarily on the results and analysis thus far, the context, and the health condition of interest.

BP1-10: Conduct Reach-back (TS4)

- See *Conduct Reach-back* (TS4) for a detailed description of this task set.

Conduct Syndrome-based Population Health Monitoring (BP 1)

Monitor and assist in the assessment, detection, communication, and response to public health conditions of interest.

All numbers refer to the flow diagram for BP 1 (Figure 7) unless otherwise noted.

BP1-11: Response Needed?

- This question of whether a response is needed is answered by the PHA and PHSS analyst based on variables that quantify the severity and potential spread of the concern, including: volume; characteristics; geographic spread; tight cluster in time; or tight cluster in time and space.
- Inputs that may influence the decision include: context, health condition of interest, and outside influences.

Response Actions (BP1-12 through BP1-18)

BP1-12 through BP1-18: When a response is initiated, gradient of activities may be activated depending on the purpose, context, level of authority, health condition of interest, and outside influences. The PHA may engage in one or more of these response actions; not all response actions will be necessarily activated. These actions are not conducted in a linear order and one action may influence one or more of the other actions.

Dissemination of Information (BP1-12 through BP1-16)

BP1-12: Dissemination of Information: Disseminate SS Information / Output As Needed

This task is to capture activities related to disseminating PHSS information or output not captured by the other dissemination of information actions (BP1-13 through BP1-16)

BP 1-13: Provide Summary Information to SITREP

Summary information is reported to a situation report (SITREP) in order to establish and maintain situation awareness for the PHA and public health leadership.

BP 1-14: Send Health Alert

Health alerts are sent through Health Alert Networks (HANs), which help disseminate important health information and link local health departments to one another and to other organizations critical for preparedness and response: community first-responders, hospital and private laboratories, state health departments, CDC, and other federal agencies.[4] Health alerts may also be sent through CDC's Epi-X, which provides communication and sharing of preliminary health surveillance information for CDC officials, state and local health departments, poison control centers, and other public health professionals.[5]

---

[4] http://www.bt.cdc.gov/documentsapp/han/han.asp
[5] http://www.cdc.gov/epix/

Conduct Syndrome-based Population Health Monitoring (BP 1)

Monitor and assist in the assessment, detection, communication, and response to public health conditions of interest.

All numbers refer to the flow diagram for BP 1 (Figure 7) unless otherwise noted.

BP1-15: Provide Information to PIO Function/Role

- The PHA provides necessary information to the person fulfilling the Public Information Officer (PIO) role. The PIO role may vary across jurisdictions ranging from a dedicated PIO person/unit in larger PHAs, to public health leadership fulfilling the role.
- The role of the PIO is to be a communication spokesperson for the PHA by providing information to the public and media.

BP1-16: Inform Health Care Providers

The PHA may communicate separately with health care providers to provide additional information or instructions that are relevant to and can be performed by health care providers.

- Example: A PHSS unit detects a cluster that may potentially be a gastrointestinal outbreak within a geographic area. Therefore, the PHA sends out targeted alerts to specific hospitals within the geographic area so that they may be able to identify, evaluate, and provide medical intervention more effectively to their patients and/or those affected persons seeking medical care.

BP1-17: Activate PH / Preparedness and Response Resources

Public health and preparedness and response resources are activated if it is determined that the issue is significant enough to spend resources for intervention. Intervention may be activities related to increasing educational awareness or situation awareness.

BP1-18: Launch a Formal PH Investigation

A formal public health investigation is launched if the issue is severe enough or poses a threat large enough where further investigation is required.

BP1-19: Further Action?

When a potential issue does not require additional escalation, this question is answered by the PHSS analyst to determine if any further action is needed.

Further Actions for Non-Escalated issues (BP1-20 through BP1-23)

BP1-20 through BP1-23: A gradient of responses for non-escalated issues that may still require further action depending on the purpose, context, level of authority, health condition of interest, and outside influences. The PHA may engage in one or more of these actions; not all actions will be necessarily activated. These actions are not conducted in a linear order and one action may influence one or more of the other actions.

BP1-20: Watch and Wait

Watch and wait is to bookmark the issue and to note it in the system and/or documentation that no action is required at the present, but will be "watched" in the event that additional

Conduct Syndrome-based Population Health Monitoring (BP 1)
Monitor and assist in the assessment, detection, communication, and response to public health conditions of interest.
All numbers refer to the flow diagram for BP 1 (Figure 7) unless otherwise noted.

data prompts further analysis and evaluation.

BP1-21: Dispel Rumors

Public health syndromic surveillance analysis and data interpretation are often used to assist with dispelling rumors that an issue thought to be of public health significance is not actually a threat.

BP1-22: Document Suspicion

Whether or not an issue is concluded to be significant, any suspicions along with the data, activities, results, analysis, and interpretation are documented. Documentation is important since an issue that is not considered significant at present may later be found to have significance over a time period.

BP1-23: Reporting

Depending on the purpose and context, reporting may be conducted in order to provide increased situation awareness. Reports may be in the form of routine reporting or a SITREP.

3.1.1 Task Set 1 (TS1): Collect and Process Data

In this task set, the PHA collects data, pre-processes the data to prepare them for automated analysis, and runs statistical algorithms against the data to provide the PHSS analyst an initial set of results to review.

Objective: To collect data from data providers; semantically (BP2) and syntactically validate and process data in order to make the data usable for analysis; calculate counts; and to run automated statistical algorithms on the processed data for the PHSS analyst to review.

Trigger: ED / Urgent Care data sent by the data provider at least every 24 hours.

Input to the Business Process: Health Data; Context; Purpose

Output of the Business Process: Counts and results of the automated statistical algorithms, which include time series and summary counts of flags. Within this business process, aggregated, categorized data sets may be shared with external systems, such as Distribute[6] and BioSense.

---

[6] Distribute is a project of ISDS that is dedicated to real-time data evaluation, allowing for more rapid surveillance and improved decision-making.

Task Set 1 (TS1): Collect and Process Data

Figure 6: Task Flow Diagram for TS1 – *Collect and Process Data*: The PHA collects and pre-processes data to prepare them for automated analysis, and runs statistical algorithms against the data to provide an initial set of results for review.
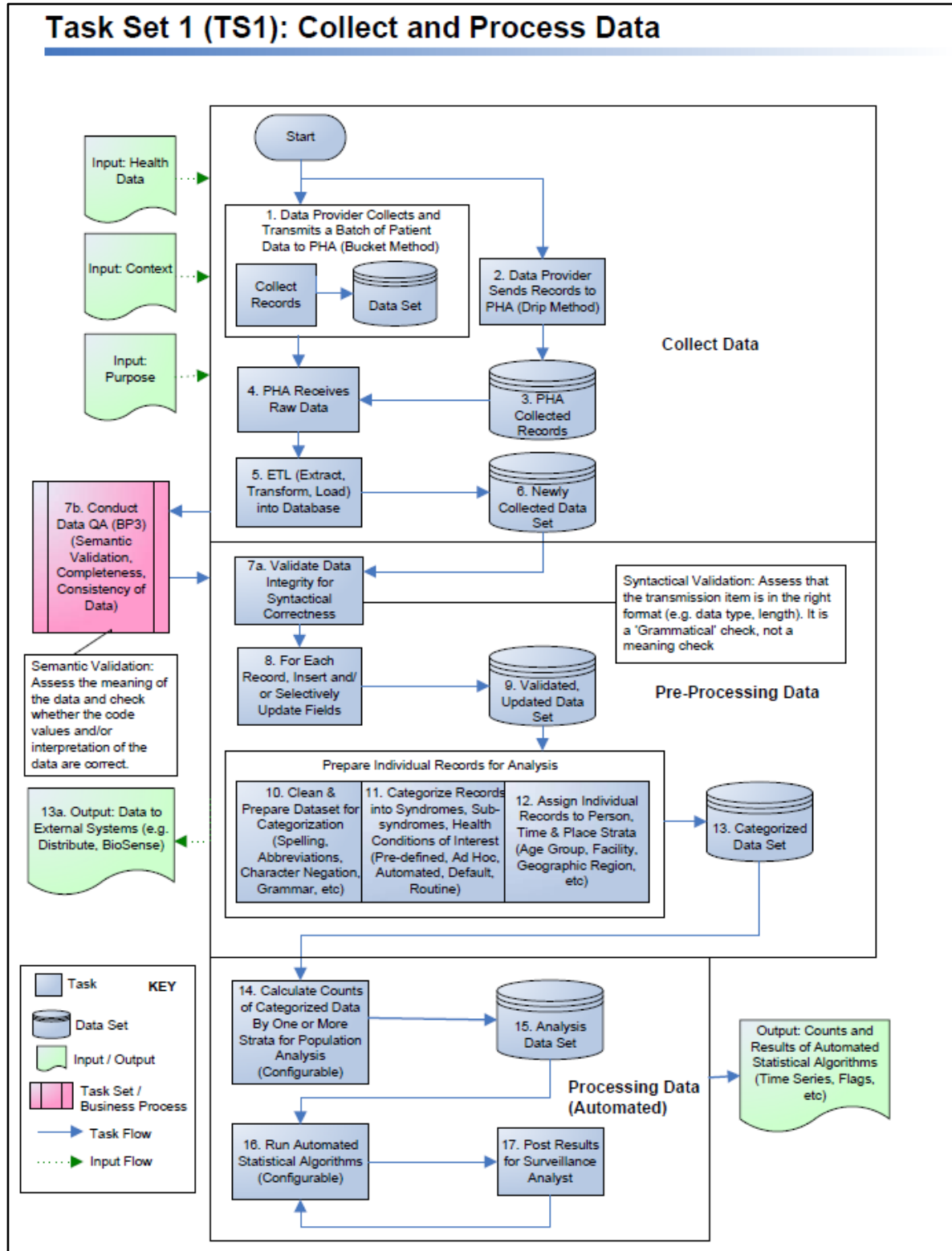
Collect and Process Data (TS 1)

The PHA collects and pre-processes data to prepare them for automated analysis, and runs statistical algorithms against the data to provide an initial set of results for review.

All numbers refer to flow diagram for TS 1 (Figure 8) unless otherwise noted.

Collect Data

#1: **Data Provider Collects and Transmits a Batch of Patient Data to PHA (Bucket Method)**

#2: **Data Provider Transmits Patient Data to PHA (Drip Method)**

#3: **PHA Collected Records**

- There are two identified ways in which data providers transmits data to the Public Health Authority (PHA)

    o Bucket Method: The data provider collects records on their end and transmits a *batch* of records to the PHA within the agreed timeframe (e.g. once every 24 hours). (TS1-1)

    o Drip Method: The data provider transmits individual records without batching (TS1-2) by a more frequent time period (e.g. every 15 minutes) and the PHA collects the records prior to processing data (TS1-3). In this case, the collection/batching of records is conducted on the PHA side.

    o The method of data delivery varies across jurisdictions and is established and agreed upon between the data provider and PHA.

- Filtering - see detailed note at the bottom of this section about filtering of data.

    o It is highly recommended that data providers transmit all available patient data as specified in this provisional recommendation and **do not** filter their data. The effectiveness of public health syndromic surveillance is greatly increased when all records are available for analysis. It is recognized that there are jurisdictions that have data providers transmit select records (e.g. records that indicate a reportable disease condition).

#4: **PHA Receive Raw Data**

- The PHA receives the raw data from the data provider.

#5: **ETL (Extract, Transform, Load) into Database**

- Extract, Transform, and Load (ETL) is an automated process in which an individual record (HL7, comma delimited, etc) is parsed and transformed so that the record can be loaded into the public health syndromic surveillance database.

- The load process may include very basic checks that preserve the integrity of the database, such as verifying that record has a primary key, etc.

#6: **Newly Collected Data Set**

- After the ETL process, the individual records are collected and a data set is created.

- Whether data are received from the data provider via the bucket or drip method, if needed, the PHA will collect data for a specific timeframe (e.g. 24 hours) prior to mov-

Collect and Process Data (TS 1)

The PHA collects and pre-processes data to prepare them for automated analysis, and runs statistical algorithms against the data to provide an initial set of results for review.

All numbers refer to flow diagram for TS 1 (Figure 8) unless otherwise noted.

ing forward. This is done so that the analysis of data is meaningful and complete.

- The collection timeframe will vary by jurisdiction and by individual data provider. For example, if the PHA normally receives 90% of a data provider's records by 10:00am, it may proceed to the next step.

- The output of this step may be used as an input into the *Conduct Data Quality* business process (BP3),

Pre-Processing Data

#7a: Validate Data Integrity for Syntactical Correctness

- For the newly collected data set, the PHA validates data integrity for syntactical correctness.

- *Syntactical correctness* is a "grammatical" check to verify that the record being transmitted is complete and in the correct format, e.g. correct data type and length, no missing fields, etc. The meaning or interpretation of the data is not validated.

- Example: Syntactical validation fails due to a clerk entering a 'pipe' character in the chief complaint field, which is interpreted by the PHSS system as a new field, therefore shifting all of the subsequent fields over.

#7b: Conduct Data QA (BP3)

- In addition to the validation of data for **syntactical** correctness, the PHA also conducts data quality assurance to verify 1) that data are received from all expected data providers, and 2) the data are validated and corrected for semantic correctness, completeness, and consistency (the meaning of the data). See *Conduct Data QA* business process (BP3) for further details on this task flow.

- The output of the *Conduct Data QA* business process is data that are semantically validated, corrected, or accepted.

#8: For Each Record, Insert and/or Selectively Update Fields

- The PHA system will insert and/or selectively update existing records based on syntactical and semantically correct data on a matching identifier.

- There are variations in the type of identifiers that are used across jurisdictions for matching. Unique identifiers include fields such as Visit ID, Patient ID, Medical Record Number (often different from a more anonymous Patient ID) or a combination of identifiers such as Hospital ID and Visit ID.

- The matching identifier helps determine if the record is a new record or an update of an existing record.

- For updates, not all fields will require updating. The criteria of how, when, and what to update varies across jurisdictions. In addition, some jurisdictions may opt to concatenate updated values with original values rather than overwriting previous values.

Collect and Process Data (TS 1)

The PHA collects and pre-processes data to prepare them for automated analysis, and runs statistical algorithms against the data to provide an initial set of results for review.

All numbers refer to flow diagram for TS 1 (Figure 8) unless otherwise noted.

#9: Validated, Updated Data Set

- After inserting new records or updating existing records, the data set is considered a validated, updated, or accepted data set by the PHA.

#10: Prepare Individual Records For Analysis: Clean and Prepare Dataset for Categorization (Spelling, Abbreviations, Character Negation, Grammar, etc).

- This step is part of the process of preparing the individual records for analysis.

- In this step, the PHA cleans and prepares the validated, updated dataset for categorization. The process is done to improve the quality of the data.

- The approaches to processing varies across jurisdictions and can include any of the following:

  o Standardize text; remove punctuation; fix spelling; change abbreviations into whole words; remove CAPS; remove extra spaces; remove extra characters; fix punctuation; conduct character negation; fix grammar, etc.

- This is the pre-processing step prior to text parsing conducted in the next step.

#11: Prepare Individual Records For Analysis: Categorize Records into Syndromes, Sub-syndromes, Health Conditions of Interest (Pre-defined, New, Automated, Default, Routine)

- Once the data are cleaned, the PHA parses the text, categorizes and maps records to different syndromes, sub-syndrome, and health conditions of interest. This step may use natural language processing methods.

- Example: If chief complaint data says "I have nausea and am throwing up." If the words "throwing up" and "nausea" are GI components, they would be mapped to the GI category.

- These categories that are used may be pre-defined, ad hoc, automated, default, or routine. Ad hoc or new categories may be added as a default or routine category.

  o Example: Pre-defined or default categories may be disease conditions, such as varicella or meningitis, or an environmental exposure, such as a heat-related or cold-related disease.

  o Example: An ad hoc category may be developed as a result of a single, recent event with an elevated risk, such as an oil spill.

- The ad hoc queries are important within this step since it allows PHSS to adapt to and further process data based on a particular context or situation that may be new or unique.

  o This reinforces the need for data sent to the PHA for syndromic surveillance to be **unfiltered by the data provider**. It provides the broadest flexibility for surveillance activities in various contexts.

  o This also reinforces the emphasized request by PHSS that the data providers

Collect and Process Data (TS 1)

The PHA collects and pre-processes data to prepare them for automated analysis, and runs statistical algorithms against the data to provide an initial set of results for review.

All numbers refer to flow diagram for TS 1 (Figure 8) unless otherwise noted.

transmit **the chief complaint field as a free text value, with the most descriptive text available**. Free text provides the greatest flexibility in categorizing records into conditions that may not be apparent or match coded values.

#12: **Prepare Individual Records For Analysis: Assign Individual Records to Person, Time, and Place Strata (Age Group, Geographic, etc)**

- After categorizing records into syndromes, sub-syndromes, or health conditions of interest, the PHA assigns or "bins" the records into person, time, and place strata.
    - o Examples of strata: age group; facility or hospital system; geographic region; disposition.
- Ad hoc strata may be created, used, and incorporated as part of the default stratification groups.
- The assigning of records to strata may be saved and stored in the PHSS database, or may be conducted through the system interface as a "report" or "view' of the data without saving.

#13: **Categorized Data Set**

- The output of the Prepare Individual Records for Analysis process results in a categorized data set ready for analysis.

#13a: **Output: Data to External Systems (e.g. Distribute, BioSense)**

- Sometime during the Pre-Processing Tasks, the PHA may send data to external systems, such as Distribute or BioSense.

Processing Data (Automated)

#14: **Calculate Counts of Categorized Data By One or More Strata for Population Analysis (Configurable)**

- The PHA takes the categorized data and has the PHSS system calculate counts of the categorized data by one or more strata for population analysis.
- There is variation in what strata are used across jurisdictions. Strata include: syndromes and sub-syndromes; age group; zip code; hospital; gender; other demographics.
- These counts may or may not be stored in the database.
- Strata are configurable.

#15: **Analysis Data Set**

- After counting by strata, the data set is now available for automated statistical algorithms.

#16: **Run Automated Statistical Algorithms (Configurable)**

- The PHA takes the analysis data set and runs automated statistical algorithms using its

Collect and Process Data (TS 1)

The PHA collects and pre-processes data to prepare them for automated analysis, and runs statistical algorithms against the data to provide an initial set of results for review.

All numbers refer to flow diagram for TS 1 (Figure 8) unless otherwise noted.

syndromic system. Systems, such as ESSENCE[7], allow multiple data sources to be compared.

- With most systems, the PHA may recalibrate their algorithms as needed. Statistical algorithms are automated, but not static, so custom and ad hoc statistical algorithms may be added and configured. In addition, custom or ad hoc algorithms may be added as a default or "automated" algorithm.

- Statistical algorithms may include time series, maps, summary counts of flags, comparison of results to historical data.

- The context and purpose (inputs) may affect the type of statistical algorithms that are run. For example, if there is limited baseline data, the timeframe for any statistical algorithms run may be limited to daily trends.

#17: Post Results for Surveillance Analyst

- Following the running of statistical algorithms on the data set, the PHSS system posts the results for the surveillance analyst.

  o The analyst may include: the Public Health syndromic surveillance coordinator; PH nurse; epidemiologist; analyst.

- After reviewing the data and reviewing any flags, the surveillance analyst may choose to rerun the automated statistical algorithms on a subset of the data set, or on additional data.

---

[7] http://essence.jhuapl.edu/ESSENCE/

Additional Notes and Recommendations Related to the Task Set:

<u>Filtering</u>

It is highly recommended that data providers **do not** filter their data and transmit all available data elements as specified in this recommendation.

The power and effectiveness of PHSS is significantly increased when all records are available to the PHA for epidemiological analysis. Complete, unfiltered data are the basis for a robust and clear picture of community health. Filtering records at a facility level applies a selection process that may bias the results of epidemiological analyses on an inter-facility or inter-regional level, thereby limiting situation awareness and a PHA's ability to assess population health.

3.1.2 Task Set 2 (TS2): Characterize, Interpret, and Analyze Data

This task set is *the central activity* that the public health syndromic surveillance (PHSS) unit conducts. This task set is the "art of syndromic surveillance," where characterization of data and manual analysis and review are conducted in order to understand the meaning and significance of the data.

<u>Objective of the Business Process</u>: To characterize, interpret, and analyze the data to understand the meaning and significance of the data.

<u>Trigger</u>: Completion of Task Set 1 (TS1) *Collect and Process Data*.

<u>Input to the Business Process</u>: Results generated from Task Set 1 (TS1) *Collect and Process Data*; Health Condition of Interest; Purpose; Context; Level of Authority; Data from Other Public Health Processes; Outside Influences.

<u>Output of the Business Process</u>: Conclusions and Results of Characterization and Pattern Analysis; Health Condition of Interest; Purpose
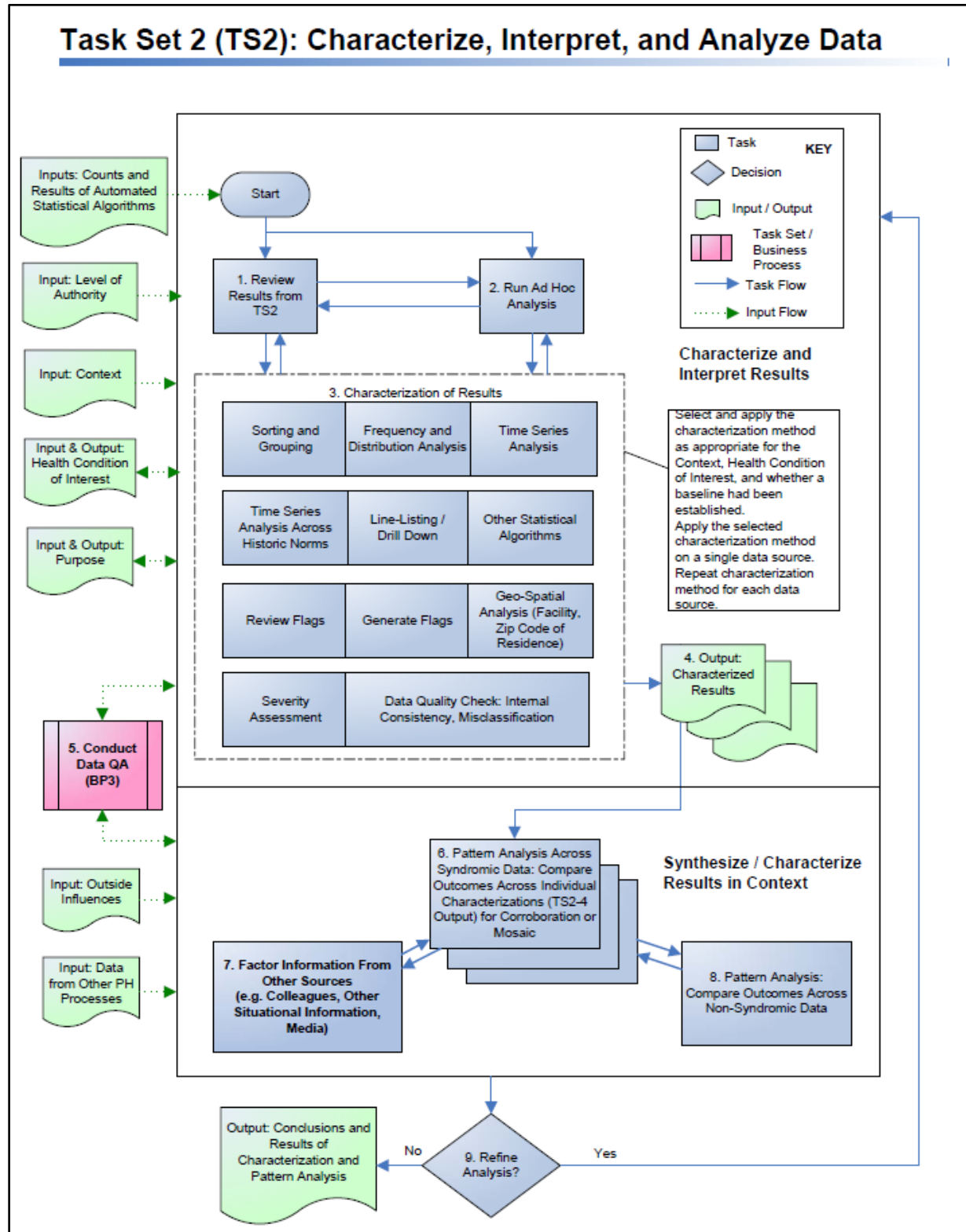
Figure 7: Task Flow Diagram for TS2 – *Characterize, Interpret, and Analyze Data*: This task set is *the central activity* that the public health syndromic surveillance (PHSS) unit conducts. It is the "art of syndromic surveillance," where characterization of data and manual analysis and review are conducted in order to understand the meaning and significance of the data.

Characterize, Interpret, and Analyze Data (TS 2)

This is the "art of syndromic surveillance," where characterization of data and manual analysis and review are conducted in order to understand the meaning and significance of the data.

Numbers refer to the flow diagram from TS 2 (Figure 9) unless otherwise noted.

Characterize and Interpret Results

#1: **Review Results from TS2,**

#2: **Run Ad Hoc Analysis,**

#3: **Characterization of Results**

- The PHSS Analyst manually reviews the result generated from Task Set 1 (TS1) *Collect and Process Data*, runs ad hoc analysis of the data, and uses a variety of methods to characterize the results.

  o Examples of individuals that may fill the PHSS analyst role include, the Public Health syndromic surveillance coordinator, a PH nurse, an epidemiologist, and/or other PHA-designated staff.

- These three tasks, (TS2-1) (TS2-2) (TS2-3), occur in conjunction with one another. The three tasks proceed in a circular flow, where the results of one task may affect a second task, which may then prompt or affect the third task conducted.

  o Different inputs and variables will affect how the three tasks are engaged. This process may be repeated in different ways taking into account different inputs, such as the context or health condition of interest.

  o As the three tasks are conducted, there is an output of increased information about the health condition of interest and purpose. Both of these variables are also inputs into this task set.

- Menu set of characterization methods: The analyst uses any of the following methods, as applicable, for the characterization of data. The list shows a representative list of characterization methods; it does not necessarily represent an exhaustive list. As a menu set, all of the methods are optional and select methods may be applied. The order in which the methods are listed does not imply a linear order:

  o Sorting and Grouping: By zip code; age group; disposition; facility; time of arrival, syndromes, sub-syndrome combinations.

  o Frequency and Distribution Analysis

  o Time Series Analysis: Compare the trend line over time, e.g. looking at the moving average over 7 days.

  o Time Series Analysis Across Historic Norms (Years): Compare current data against data over previous years, e.g. Flu is 5 times higher this year than last year.

  o Line-Listing / Drill Down

  o Other Statistical Algorithms

  o Review Flags:

    ▪ Review flags generated by automated statistical algorithms in TS1: *Col-*

Characterize, Interpret, and Analyze Data (TS 2)

This is the "art of syndromic surveillance," where characterization of data and manual analysis and review are conducted in order to understand the meaning and significance of the data.

Numbers refer to the flow diagram from TS 2 (Figure 9) unless otherwise noted.

lect and Process Data, and

- ▪ Review additional flags generated during the refinement of results within this task set.

- o Generate Flags: Generate additional flags as a result of further refinement and analysis of data within this task set.

- o Geo-Spatial Analysis (Facility, Zip Code of Residence): Maps; geographical representation and/or spatial-oriented flags, e.g. comparison of data from where a patient lives vs. where the patient receives care (which hospital they presented).

- o Severity Assessment

- o Data Quality Check: Internal Consistency, Misclassification

- The characterization methods are applied as appropriate for the Context, Health Condition of Interest, and whether a baseline had been established.

  - o For example, if the context is a single event classified as a present elevated risk of short duration, then a baseline may not exist nor can be established. Thus, the time series analysis across historic norms would not be applicable.

- The analyst applies the selected characterization method on a single data source. The characterization method is repeated for each data source.

#4: Output: Characterized Results

- For each data source, the PHSS analyst analyzes, interprets, and characterizes the results. The output is a series of characterized results for each data source.

#5: Conduct Data QA (BP3)

- Throughout this task set, the PHSS analyst may conduct data quality assurance (BP3) on the data in order to further refine and improve data quality.

- In BP3, the analyst verifies that the data are validated and corrected for semantic correctness, completeness, and consistency (the meaning of the data). See *Conduct Data QA* business process (BP3) for further details on this task flow.

- The output of the *Conduct Data QA* business process is data that are semantically validated or corrected.

Synthesize / Characterize Results in Context

#6: Pattern Analysis Across Syndromic Data: Compare Outcomes Across Individual Characterizations (TS2-4 Output) for Corroboration or Mosaic

- Once each PHSS data source has been analyzed, interpreted, and characterized, the PHSS analyst compares and combines the data sources with each other.

Characterize, Interpret, and Analyze Data (TS 2)

This is the "art of syndromic surveillance," where characterization of data and manual analysis and review are conducted in order to understand the meaning and significance of the data.

Numbers refer to the flow diagram from TS 2 (Figure 9) unless otherwise noted.

- The analyst conducts a pattern analysis across all syndromic data. By comparing and combining data, the analyst assesses patterns across the data sources to see if there is a larger, overall pattern of the data through a mosaic or Gestalt perspective.

- The analyst looks for patterns that raise suspicions or flags.

  o Example: An analyst is monitoring and observes an unexpected pattern (e.g. age distribution) that may be meaningful and warrant additional investigation, but may not have raised a flag through the automated system analysis (TS1) because it does not have statistical significance.

- This task is conducted by the PHSS Analyst and/or PHA-designated staff in conjunction with TS2-6: *Factor Information From Other Sources* and TS2-8: *Pattern Analysis: Compare Outcomes Across Non-Syndromic Data.*

#7: **Factor Information From Other Sources (e.g. Colleagues, Other Situation Information, Media)**

#8: **Pattern Analysis: Compare Outcomes Across Non-Syndromic Data**

- In synthesizing data for interpretation, the PHSS analyst and/or PHA-designated staff analyzes other sources and factors for additional information. Other sources include information within PHSS (e.g. other colleagues, epidemiologists, field staff; Distribute data); within public health (e.g. other situation information, NEDSS); and outside of public health (e.g. weather patterns, media reports, other governmental agencies and offices).

- In using additional data sources, the analyst and/or PHA-designated staff looks for patterns of concern and compares outcomes across non-syndromic data.

- This task is conducted in conjunction with TS1-6: *Pattern Analysis Across Syndromic Data.*

#9: **Refine Analysis?**

- The PHSS analyst and/or PHA-designated staff determines if the results have been refined to the point where the analyst is able to draw meaning out of the results, draw conclusions or interpretations, has excluded all null hypotheses, and/or satisfied all of the important questions.

- If the answer is no, then the analyst reengages the task set and continues the manual analysis, interpretation, and characterization until a meaningful result is attained or null hypothesis is proved.

- If the answer is yes, then the results have been refined enough so that the results have meaning and the analyst is able to contribute to / advise the next steps.

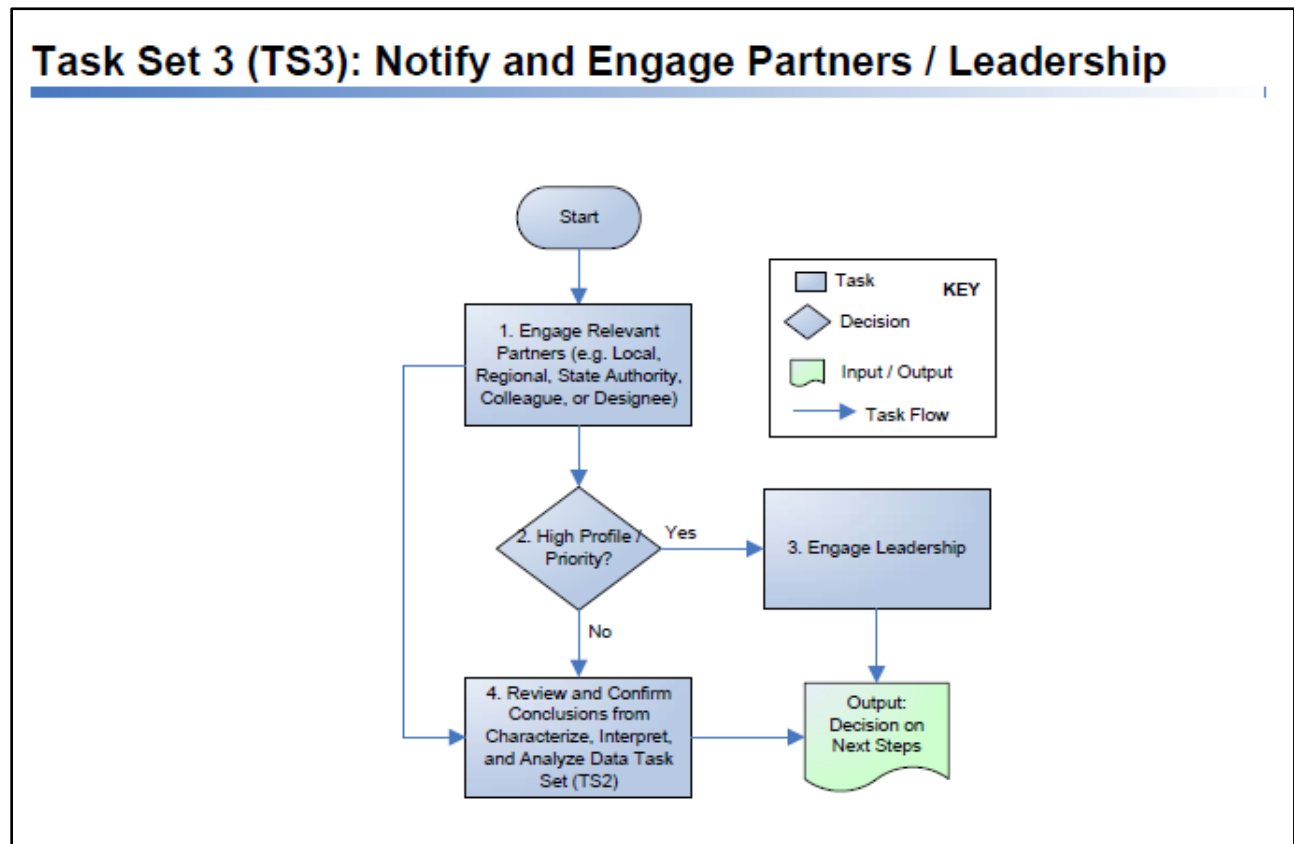### 3.1.3 Task Set 3 (TS3): Notify and Engage Partners / Leadership

When a potential public health concern is identified and escalated, the public health syndromic surveillance unit notifies and engages relevant partners and leadership when applicable, to determine follow-up actions.

Objective of the Business Process: To notify and engage relevant partners and leadership to determine whether to further escalate a potential public health concern.

Trigger: A potential public health concern is escalated and determined that relevant partners and/or leadership is necessary to determine next steps.

Input to the Business Process: Results generated from Task Set 2 (TS2) *Characterize, Interpret, and Analyze Data.*

Output of the Business Process: Decision to further escalate the potential public health concern.

**Figure 8 Task Flow Diagram of Task Set 3 (TS3) – *Notify and Engage Partners/Leadership:* When a potential public health concern is identified and escalated, the public health syndromic surveillance unit notifies and engages relevant partners and leadership when applicable, to determine follow-up actions.

Notify and Engage Partners / Leadership (TS 3)

When a potential public health concern is identified and escalated, the public health syndromic surveillance unit notifies and engages relevant partners and leadership when applicable, to determine follow-up actions.

Numbers refer to the flow diagram from TS 3 (Figure 10) unless otherwise noted.

#1: Engage Relevant Partners (e.g. Local, Regional, State Authority, Colleague, or Designee

- Relevant partners are engaged so that a determination can be made on next steps of whether to continue to escalate the issue or to abort. Partners include local, regional, state authority, colleague or designee. Partners include both public health and preparedness. The partners engaged vary across jurisdictions and level of authority.
- Some jurisdictions may not need to evaluate whether leadership should be engaged at this point. For these jurisdictions, the PHA and relevant partners may review and confirm conclusions from previous tasks and proceed to determining the next steps.
- For jurisdictions that may need to engage leadership, the PHA and relevant partners determine if this issue is high profile / priority.

#2: High Profile / Priority?

This question is answered by the PHSS unit and engaged partners. The issue is categorized as high level priority based on health condition of interest, outside influences, purpose, and context.

#3: Engage Leadership

If the issue is categorized as high priority, then leadership is engaged.

Once leadership is engaged, leadership makes a decision on the next steps.

#4: Review and Confirm Conclusions from Characterize, Interpret, and Analyze Data Task Set (TS2)

If the issue is not high profile/priority, the PHSS unit and relevant engaged partners review and confirm the conclusions from *Characterize, Interpret, and Analyze Data (TS2).*

3.1.4 Task Set 4 (TS4): Conduct Reach-back

Reach-back is the process of acquiring additional data about patient(s) where there is a potential public health concern. Reach-back is conducted when the PHA needs additional information to determine whether a response should be initiated.

Objective of the Business Process: To conduct reach-back for additional patient data to inform the decision of whether to escalate a response for a potential public health concern.

Trigger: Decision to conduct reach-back for additional data.

Input to the Business Process: Health Condition of Interest; Level of Authority; Purpose

Output of the Business Process: Additional information to inform the decision of whether to further escalate a response for a potential public health concern.

Task Set 4 (TS4): Conduct Reach-back
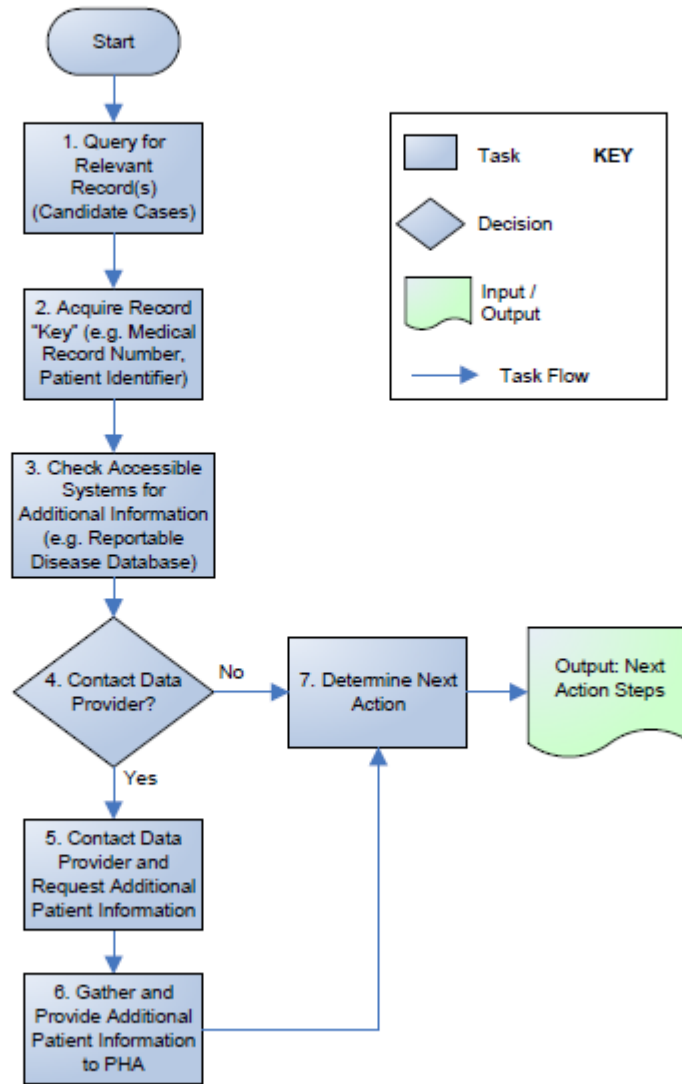
Figure 9: Task Flow Diagram of Task Set 4 (TS4): *Conduct Reach-back:* Reach-back is the process of acquiring additional patient data where there is a potential public health concern. Reach-back is conducted when the PHA needs additional information to determine whether a response should be initiated.

Conduct Reach-back (TS 4)

Reach-back is the process of acquiring additional patient data where there is a potential public health concern. Reach-back is conducted when the PHA needs additional information to determine whether a response should be initiated.

Numbers refer to the flow diagram for TS 4 (Figure 11), unless otherwise noted.

#1: Query for Relevant Record(s) (Candidate Cases)

- The PHSS analyst queries the PHSS database for the relevant record (candidate case)

#2: Acquire Record "Key" (e.g. Medical Record Number, Patient Identifier)

- During the query, the PHSS analyst acquires the unique record. The record "key" that is used varies across jurisdictions. Examples include: medical record number, unique patient identifier, visit identifier, and any combination of identifiers.

#3: Check Accessible Systems for Additional Information (e.g. Reportable Disease Database)

- The PHSS Analyst and/or PHA-designates staff may first check accessible systems, such as the reportable disease database, for additional information. There are cases when these accessible systems may provide enough information to proceed to the next step.

- For example, if the syndromic data indicates 'measles', and the reportable disease database has a record of a person with a matching profile (using key attributes such as age, gender, facility location, etc) seen on the same date, then the analyst may decide to end the reach-back process since the case is already being investigated.

#4: Contact Data Provider?

- Based on the available information discovered through the accessible systems, the PHSS analyst determines whether the data provider should be contacted for additional information.

#5: Contact Data Provider and Request Additional Patient Information

- If the PHSS analyst decides to contact the data provider for additional information, then the PHSS analyst proceeds to contact the data provider and makes a request for additional patient information.

#6: Gather and Provide Additional Patient Information to PHA

- The PHSS analyst gathers any additional patient information from the data provider and provides the results to the PHA.

#7: Determine Next Action

- If the PHSS analyst decides not to contact the data provider, then the analyst will determine the next set of actions.

- Once the PHSS analyst gathers and shares with the PHA any additional patient information from the data provider, the analyst and the PHA determine the next

Conduct Reach-back (TS 4)

Reach-back is the process of acquiring additional patient data where there is a potential public health concern. Reach-back is conducted when the PHA needs additional information to determine whether a response should be initiated.

Numbers refer to the flow diagram for TS 4 (Figure 11), unless otherwise noted.

set of actions.

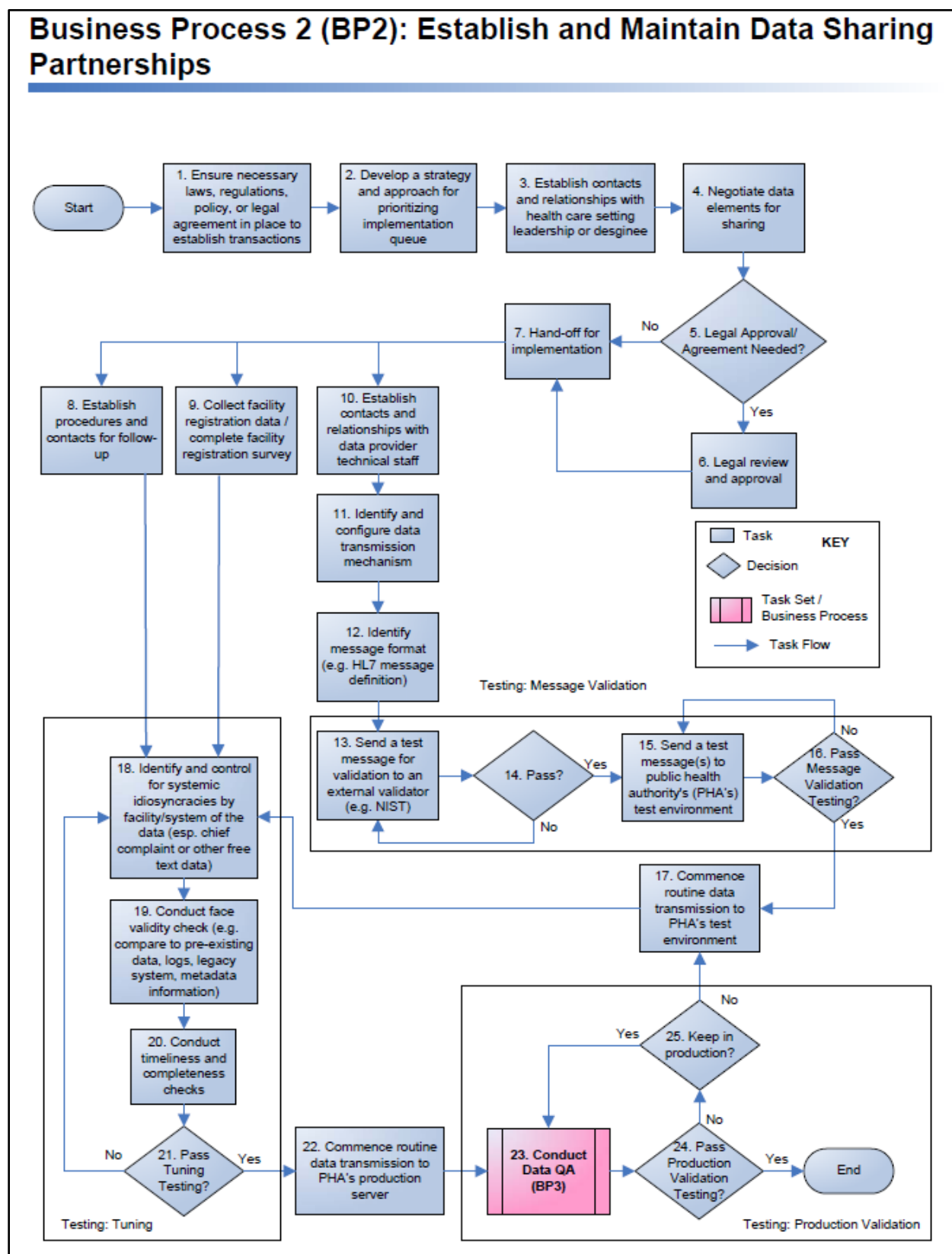## 3.2 Establish and Maintain Data Sharing Partnerships (BP 2)

The core syndromic surveillance activities rely on conclusions drawn from data that are characterized, interpreted, and analyzed. This business process establishes the partnerships between the PHA and data providers so that the PHA may receive data to conduct public health syndromic surveillance.

Objective of the Business Process: To establish and maintain data sharing partnerships for the purpose of obtaining data to conduct PHSS.

Trigger: The public health authority (PHA) identifies a potential data sharing partner.

Input to the Business Process: Context; Health Condition of Interest.

Output of the Business Process: An established data sharing partner.

Figure 10 Task Flow Diagram for BP2: *Establish and Maintain Data Sharing Partnerships:* Data sharing partnerships are established for the purpose of obtaining data to conduct public health syndromic surveillance.

Establish and Maintain Data Sharing Partnerships (BP 2)

Data sharing partnerships are established for the purpose of obtaining data to conduct public health syndromic surveillance.

All numbers refer to the flow diagram for BP 2 (Figure 12) unless otherwise noted.

BP2-1: Ensure necessary laws, regulations, policy, or legal agreements in place to establish transactions

- Prior to establishing any data sharing partnerships, the PHA ensures that all of the necessary laws, regulations, policy, or legal agreements are in place to be able to establish transactions

BP2-2: Develop a strategy and approach for prioritizing implementation queue

- The PHA develops a strategy and approach for prioritizing the implementation of data sharing with data providers. This step is important to identify from the PHA perspective which data providers need to be prioritized in order to increase effectiveness of public health syndromic surveillance.

BP2-3: Establish contacts and relationships with health care setting leadership or designee

- The PHA establishes contacts and relationships with the leadership of the health care settings and/or their designee. Buy-in and support from the healthcare setting leadership is important to strengthen the partnership.

- Both parties discuss and agree on the purpose, roles, responsibilities, and expectations of the data sharing partnership. As part of establishing the partnership, the PHA provides basis and rationale to the data provider to justify the need for the requested data, and what the data provider may expect in return, e.g. how their data will be protected.

- There may or may not be a formal data sharing agreement established between the data provider and the PHA. Data sharing partnerships vary widely across jurisdictions and data providers.

BP2-4: Negotiate data elements for sharing

- Data provider and receiver negotiate the data elements that will be shared and discuss any local or state-specific laws or regulations that may affect data sharing.

BP2-5: Legal Approval / Agreement Needed?

- This question may be answered by the data providers (health care setting) to determine if legal approval / agreement is needed prior to the data sharing partnership being established. Some data sources (e.g. temporary med tents at the Boston Marathon) may or may not need same level of legal agreement.

BP2-6: Legal review and approval

- If required, the data provider conducts a legal review and provides approval for the data sharing partnership.

Establish and Maintain Data Sharing Partnerships (BP 2)

Data sharing partnerships are established for the purpose of obtaining data to conduct public health syndromic surveillance.

All numbers refer to the flow diagram for BP 2 (Figure 12) unless otherwise noted.

BP2-7: Hand-off for implementation

- The data sharing specifications are handed off to staff who will implement the data sharing.

BP2-8: Establish procedures and contacts for follow-up

- The PHA and data provider establish procedures and contacts for follow-up. Follow-up contacts include both: 1) technical contacts for data transmission issues, and 2) clinical contacts in the event that the PHA needs to "break the glass" and request additional data about patients while conducting syndromic surveillance, and 3) administrative contacts for procedural or unresolved technical or clinical issues.

BP2-9: Collect facility registration data/complete facility registration survey

- If used, facility registration metadata are captured for the data provider through a facility registration survey that the data provider completes.

- Many jurisdictions capture and facility data used as metadata in order to streamline data transmissions and prevent the need for facility data elements (such as facility address) to be sent repeatedly. Only the facility identifier would be sent with the data and the identifier would be cross-referenced to a facility registration database maintained by the PHA.

- See Section 4 for additional information about Facility Registration data.

BP2-10: Establish contacts and relationships with data provider technical staff

- The PHA establishes contacts and relationships with the technical staff of the data provider to begin implementing the data sharing specifications.

BP2-11: Identify and configure data transmission mechanism

- The data provider technical staff identifies and configures the data transmission mechanism. The type of data transmission mechanism is determined by the PHA and may include, but is not limited to: NHIN, PHIN-MS, VPN, sFTP, EBXML, HTTPS.

BP2-12: Identify message format (e.g., HL7 message definition)

- The PHA and data provider technical staff work together to identify the message format that will be used for data transmission, e.g. the HL7 message definition.

Testing: Message Validation

BP2-13: Send a test message for validation to an external validator (e.g., NIST)

- The data provider is instructed to transmit a test message for validation, such as to an external validator, such as NIST, to conduct basic message validation and help troubleshoot basic message construction issues.

Establish and Maintain Data Sharing Partnerships (BP 2)

Data sharing partnerships are established for the purpose of obtaining data to conduct public health syndromic surveillance.

All numbers refer to the flow diagram for BP 2 (Figure 12) unless otherwise noted.

BP2-14: Pass?

- The question is answered by the data provider once the test message is tested against the validator.

BP2-15: Transmit a test message(s) to public health authority's (PHA's) test environment

- Once the test message passes the baseline test against the validator, the data provider transmits a test message(s) to the PHA's test environment.

BP2-16: Pass Message Validation Testing?

- This question is answered by the PHA once the data provider transmits a test message to the PHA's test environment.

- If the test does not pass, then the PHA and data provider identify the issue and resolve the problem. The data provider is instructed to retransmit the test message to the PHA's test environment.

Testing: Tuning

BP2-17: Commence routine data transmission to PHA's test environment

- If the message validation testing passes, then the data provider is instructed to commence the testing phase of regular ongoing data transmission to the PHA's test server to determine the steps needed, if any, to fine-tune the data.

BP2-18: Identify and control for systemic idiosyncrasies by facility/system of the data (esp. chief complaint or other free text data)

- Once the testing phase of regular ongoing data transmission is initiated, the PHA begins to identify and determine how to control for systemic idiosyncrasies that are produced by the data provider system.

BP2-19: Conduct face validity check (e.g., compare to pre-existing data, logs, legacy system, metadata information)

- The PHA and the data provider conduct a face validity check of the data, including comparing the data against pre-existing non-HL7 PHSS data, logs, legacy system data, and metadata information

BP2-20: Conduct timeliness and completeness checks

- The PHA and data provider check on the timeliness of data transmissions and the completeness of data being transmitted.

BP2-21: Pass Tuning Testing?

- This question is answered by the PHA once the data are fine-tuned and validated.

- Note that the process of tuning the data and testing data against the test server

Establish and Maintain Data Sharing Partnerships (BP 2)

Data sharing partnerships are established for the purpose of obtaining data to conduct public health syndromic surveillance.

All numbers refer to the flow diagram for BP 2 (Figure 12) unless otherwise noted.

may be an extended period of testing, such as 30-60 days.

- If the data do not pass tuning testing, then the PHA and data provider collaborate to identify and resolve any issues. The data are put through the tuning testing process until the data pass.

BP2-22: Commence routine data transmission to PHA's production server

- If the data pass the tuning testing, the data provider is instructed to commence routine data transmission to the PHA's production server.

Testing: Production Validation

BP2-23: Conduct Data QA (BP3)

- Once data are transmitted to the PHA's production server, the testing of data undergoes the data quality assurance business process as described in BP3: *Conduct Data Quality Assurance.*

BP2-24: Pass Production Validation Testing?

- This question of whether the data transmission passes production validation testing is answered by the PHA.

- If the data transmission passes production validation testing, then this process is completed and the data provider begins reporting its data.

BP2-25: Keep in Production?

- If the data transmission does not pass production validation testing, then the PHA determines whether it can continue to be tested against the production server, or whether the testing must go back to the testing protocol.

3.3 Conduct Data Quality Assurance (BP 3)

The PHA conducts quality assurance to determine whether data are being transmitted from expected data partners and whether the data being transmitted are semantically correct, complete, and consistent with data normally sent by the data provider.

Semantic correctness refers to validation of the meaning of the data and whether the code values and/or interpretation of the data are correct. In this business process, the PHA troubleshoots any errors that prevent data from being transmitted. After semantically evaluating transmitted data, if any "flaws" exist, the PHA will control for the "flaws", accept the data as they are, or collaborate with the data provide to correct the "flaws" if possible. The determination of "flawed" data is from a public health perspective to verify whether the data can be used for PHSS.

Objective of the Business Process: To conduct quality assurance on data being reported for public health syndromic surveillance and to resolve any data reporting issues.

Trigger: Completion of *Collect Data* tasks within Task Set 2 (TS1) *Collect and Process Data*.

Input to the Business Process: Context; Health Condition of Interest; Purpose.

Output of the Business Process: Semantically validated, corrected, or accepted data source(s).
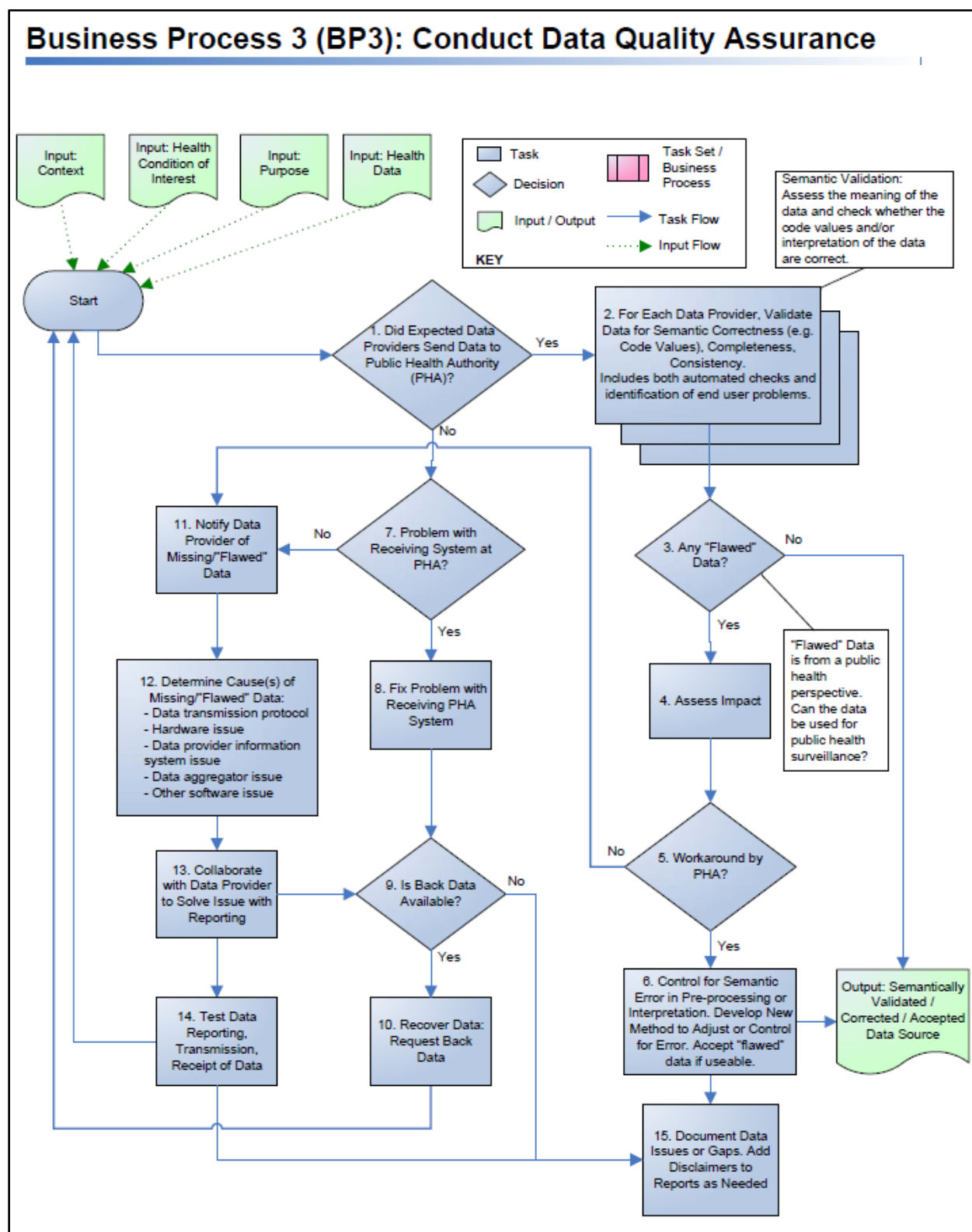
Figure 11: Task Flow Diagram of Business Process 3 (BP3) - *Conduct Data Quality Assurance:* Quality assurance is conducted on the data in order to determine 1) whether data are being transmitted from all of the expected data partners, and 2) whether the data being transmitted are semantically correct, complete, and consistent with data normally sent by the data provider.

Conduct Data Quality Assurance (BP 3)

Quality assurance is conducted on data being reported for PHSS and to resolve any data reporting issues.

All numbers refer to the task flow diagram for BP 3 (Figure 13) unless otherwise noted.

BP3-1: Did Expected Data Providers Transmit Data to Public Health Authority (PHA)?

- Transmission-level check: This question is answered by the Public Health Authority (PHA) to determine whether the PHA has received all expected data from all of their data sharing partners. The expectation is based on the data sharing agreement established between each data provider and the PHA in *Business Process 2 (BP2) – Establish a New Data Sharing Partner*.

BP3-2: For Each Data Source, Validate Data for Semantic Correctness (e.g. Code Values), Completeness, Consistency. Includes both automated checks and identification of end user problems.

- For each data source, the PHA validates the semantic correctness, completeness, and consistency of the data.

- *Semantic correctness* refers to validation of the meaning of the data. The validation checks whether the code values and/or interpretation of the data are correct. Translation errors are corrected on the PHA side.

  o Example of code value validation: The data sent for the field, Patient Class, contains a value "Newborn" that is not a part of the expected code set for the field.

  o Example of interpretation validation: Chief complaint contains "OD", which is translated to "Right Eye" instead of the correct translation of "Overdose".

  o Example: the abbreviation of "ms" for "mother states" is translated into "mental status", resulting in a large number of neurological flags.

- *Completeness* validation refers to the completeness of

  o The data set. Based on the data provider's reporting history, the PHA determines if it appears as though only a portion of the expected entire volume of the data provider's records have been sent.

  o Individual data elements. In addition to completeness, the semantic validity of each or selected data elements is assessed. (e.g., completeness of gender and expected ratio of gender values).

- *Consistency* validation refers to whether the data being sent are inconsistent with previously sent records from the data provider. Example: The gender of a patient is reported as "Male", whereas the previously reported gender of the same patient was "Female".

BP3-3: Any "Flawed" Data?

- This question is answered by the PHA after semantic validation has been completed. The determination of "Flawed" Data is from a public health perspective to verify whether the data can be used for syndromic surveillance. See BP3-2 for the description of semantic validation to determine if data is "flawed".

Conduct Data Quality Assurance (BP 3)

Quality assurance is conducted on data being reported for PHSS and to resolve any data reporting issues.

All numbers refer to the task flow diagram for BP 3 (Figure 13) unless otherwise noted.

BP3-4: Assess Impact

- If the data is categorized as ''flawed" in any way, the PHA assesses the impact or severity of any "flaws".

BP3-5: Workaround by PHA?

- This question is answered by the PHA receiving the data. Based on the assessment of the impact of any "flawed" data, the PHA determines whether it can conduct a workaround to be able to use the data for PHSS.

BP3-6: Control for Semantic Error in Pre-processing or Interpretation. Develop New Method to Adjust or Control for Error. Accept "flawed" data if useable.

- If the PHA determines that it can conduct a workaround for the "flawed" data, or that the data are useful in a "flawed" state, it proceeds to control for the semantic error or accept the data with the known limitations. New methods for adjusting or controlling for any errors are developed as needed.

BP3-7: Problem with Receiving System at PHA?

- This question is answered by the PHA if an expected data provider has not sent their data, in order to determine if the error is caused by the receiving system at the PHA. The PHA conducts an internal investigation of its receiving system to make the determination.

BP3-8: Fix Problem with Receiving PHA System

- If the PHA determines that the error in accepting the data from the data provider is caused by the receiving PHA system, the PHA fixes the problem with its system.

BP3-9: Is Back Data Available?

- The PHA asks the data provider whether back data are available during errors where the data was not received by the PHA.

BP3-10: Recover Data: Request Back Data

- The PHA requests back data from the data provider to recover any data not received.

BP3-11: Notify Data Provider of Missing/"Flawed" Data

- If PHA determines that the missing expected data is not due to the receiving PHA system, then the PHA will notify the data provider of the missing data.

- If PHA receives data from the data provider but assesses that the impact of any "flaws" is too severe or a workaround not possible, then the PHA will notify the data provider with details of the "flawed" data.

BP3-12: Determine Cause(s) of Missing/"Flawed" Data

- The PHA works with the data provider to identify the cause of the error where

Conduct Data Quality Assurance (BP 3)

Quality assurance is conducted on data being reported for PHSS and to resolve any data reporting issues.

All numbers refer to the task flow diagram for BP 3 (Figure 13) unless otherwise noted.

data is not being reported. Causes of the error may include:

- o Data transmission protocol error
- o Hardware issues
- o Data provider information system issues
- o Data aggregator issues
- o Other software issues

BP3-13: Collaborate with Data Provider to Solve Issue with Reporting

- Once the cause of the error is determined, PHA collaborates with the data provider to resolve the problem.

BP3-14: Test Data Reporting, Transmission, Receipt of Data

- Once the data error is determined as being resolved, the PHA and data provider test the data reporting, transmission, and receipt of data.

BP3-15: Document Data Issues or Gaps. Add Disclaimers to Reports as Needed

- PHA will document any issues, workarounds, or gaps in the data, and add disclaimers to reports as needed, when:
  - o PHA controls for any semantic errors in data, which may impact the analysis of the data.
  - o Back data are not recoverable
  - o Any issues or problems are noted that may impact analysis of the data.