PHIN Preparedness

(DRAFT for discussion)

EARLY EVENT DETECTION FUNCTIONAL REQUIREMENTS AND PROCESS FLOWS

Version 0.1 Draft

05/17/2004

VERSION HISTORY

Version #	Implemented By	Revision Date	Approved By	Approval Date	Reason
1	Jennifer Johnson	05/07/2004	<name>	<mm/dd/yy>	Initial draft
2					
3					
4					
5					

TABLE OF CONTENTS

TABLE OF FIGURES

1 INTRODUCTION

This document describes functional requirements and general workflows for systems implemented to manage Early Event Detection. Early Event Detection (EED) is intended to support the earliest possible detection of events that may signal a public health emergency. In addition, EED supports continued detection of subsequent cases of an event after the initial emergency has been recognized. Frequently, the initial detection is made by the astute care provider who recognizes the presentation and informs public health and other involved parties to investigate. EED is not intended to replace, but instead integrate with, and supplement, conventional detection. There is substantial evidence to indicate that the earlier an event is identified, the faster countermeasures can be applied--resulting in a better public health outcome. Events may be caused by acts of nature, accidents (such as chemical spills), or intentional acts including bioterrorism. Once an aberration is detected, EED must support the ability to localize the population and geographic areas affected, and respond quickly and appropriately to reduce morbidity and mortality in the affected population.

Early Event Detection can use substantiated health-related data sources as the basis for analysis, including: diagnostic human data (i.e. lab test results, ambulatory care diagnoses codes, "insurance" codes), environmental, agricultural, as well as potentially pre-diagnostic data (i.e., chief complaint data, lab test orders, over the counter drug sales). The EED infrastructure should be able to support adopting, more experimental data sources (school absenteeism reports, nurse call lines reports) as they are evaluated for effectiveness and value. Systems supporting EED must be able to detect aberrations in data patterns, link and aggregate the data collected, and integrate with systems that manage the response to public health emergencies. Because public health emergencies may occur across jurisdictions, across different areas of business (agriculture, water, veterinary), EED requires the ability to exchange data and support collaboration across jurisdictions and with local, state and national systems involved in public health.

This document provides minimum operational requirements necessary to support an Early Event Detection system and should in no way preclude a system from incorporating additional functionality beyond what this document addresses.

2 REQUIREMENTS

2.1 EARLY EVENT DETECTION FUNCTIONAL REQUIREMENTS

The following requirements describe baseline functionality for any system implemented to support Early Event Detection:

- *Data Sources*: A variety of substantiated data sources should be considered. Data sources may include diagnostic and potentially pre-diagnostic data, and should be selected to support detection of a broad array of public health emergencies.
- *Data Requirements and Linkages*: EED analysis occurs on several levels and therefore must meet specific data characteristics and include traceable linkages.

- *Data Receipt and Storage*: Data collected from multiple sources must be accumulated and standardized to support analysis across data sources and integration with external information systems.
- *Data Analysis*: Established algorithms are applied to aggregated data to detect deviations from normal patterns.
- *Data Visualization and Analytical Reporting*: Analytical results should be supported by visual representation (i.e., maps, graphs, and charts), and pre-defined and ad-hoc reporting at aggregate and detailed levels.
- *Alerts and Communication:* EED should have the ability to issue alerts to partners that need to be notified of a possible or confirmed public health emergency. Alerts should be based on threshold levels that can be configured and defined by the receiver.
- *Data Exchange and System Integration*: EED information must be exchangeable, based on established standards, between systems involved in the detection of, monitoring of, and response to public health emergencies.
- *Consequence Management Operations*: EED systems must support the activities, roles and responsibilities necessary to determine if an aberration is a positive signal or a false alarm and perform the steps to ensure appropriate response. Personnel, roles, and responsibilities necessary to support EED systems must be clearly defined.
- *Security and Availability:* Security (i.e., authentication, authorization and access control) and availability (i.e., disaster recovery within designated time frames and continuity of operations) requirements should be enforced across the EED system.

2.1.1 Data Sources

- The EED system should accept data from multiple established sources, such as claims clearinghouses, hospital systems, clinical laboratories, health plans, integrated delivery systems, or retail pharmacy chains.
- Wherever possible data should be pre-existing and available electronically. Because compliance with manual reporting (among other factors) has historically been poor, it is preferable that a data source not require manual data entry into an EED system.
- Data should be collected frequently enough to support the needs of an EED system (at least daily), but should not introduce latency to the data source during its attempt to provide timely data.
- EED requires the ability to access data that extends beyond the traditional case data reported for notifiable diseases. This additional data should be collected primarily from diagnostic sources. Data may be collected from pre-diagnostic sources as they become available and evaluated as compatible with existing requirements.
    - Diagnostic data collected from sources such as lab results reports, or case reports should be used to support analysis and investigation. Sufficient diagnostic data must be collected from critical care sites to identify outbreaks of established or emerging diseases.

  - Pre-diagnostic data are collected prior to a diagnosis being determined. Examples of these sources include laboratory test orders, CPT codes, over-the-counter drug sales, nurse call lines.
- Examples of other experimental data sources that may be evaluated for EED include school and work absenteeism reports, and 911 calls.
- Data sources that monitor changes in the environment (i.e., BioWatch) should be considered for EED in conjunction with confirmatory data sources. Ideal sources would represent major metropolitan areas where significant populations may likely be targeted.
- Before use in EED systems, data sources must be evaluated for the following characteristics:
  - Data must be available in near real-time in order to minimize the delay between an event and data accumulation.
  - A baseline of one year's data is desirable to train many algorithms for seasonal variations.
  - Data collection and management issues will substantially effect the usefulness of data sources. The quality of collected data must be consistently accurate and maintainable.
  - Data sources should ideally meet accepted coding characteristics to ensure that the data support standardized formats and requirements.
  - Data sources should be representative of the population characteristics for the selected geographic area, and contain enough records to provide appropriate power for statistical analyses.

2.1.2 Data Requirements and Linkages

The following high-level data requirements are necessary to ensure that the data being collected, analyzed, and reported to support EED are clearly defined.

2.1.2.1 Patient Data

- Demographic patient data should be collected to support localizing and characterizing an outbreak, including: age (but preferably not date of birth), gender, and zip code.
- Patient data should be linked to the original data source and also to the supplying data source. For example, the data source may be a data processor such as a claims clearinghouse, but the clearinghouse data come from claims submitted by various hospitals and health plans. Both the clearinghouse and the hospital must be identified in the data and linked to the patient information.

2.1.2.1.1 Patient Event Data Linking

- Each patient event (i.e., scheduled lab test, reported lab result, etc.) should be assigned an unambiguous, identifier which can be used to link back to the original data source as necessary for an appropriate public health investigation. This identifier should not include the patient's name, medical record number, or other identifier that is individually unique to the patient.

- The data source must have the ability to link patient event data to clinical and environmental lab results and provide that linkage to support a public health investigation.
- During the investigation of a possible outbreak, additional information may be requested from the data source to support linking test results, additional demographic and environmental characteristics, and associated data relative to the circumstances of the case to the event. To support treatment as a part of response and contact tracing, this supplementary data may include the patient's identifier if requested by an authorized public health agency in the in the context of an investigation.

2.1.2.2 Organization data

- Organization data should be stored for all organizations participating in EED (including data source providers, laboratories, hospitals, state and local health departments, providers, etc.).
- Organization data must include: organization name, address (including street address, city, state, county, and zip code), primary contact name, phone number, fax number, and organization type.
- Standardized organization data will be defined to enable interoperable systems to report at local, state, and national levels. Moreover, the data will allow for the differences in organization structures (i.e., boroughs, MSAs, counties, departments, etc) from the multiple data sources.
- Organization data standards should be implemented to facilitate the delivery of data from multi-jurisdictional sources to the appropriate recipients.

2.1.3 Data Receipt and Storage

- Data collected from multiple sources must be standardized prior to aggregation in a centralized data store (or data staging area).
- Standardized, aggregated data may be transferred to a data warehouse designed to optimize analysis, extraction and reporting.
- Data store structures used for EED data should be derivable back to the public health logical data model.

2.1.4 Data Analysis

- Analysis should account for expected seasonal fluctuations (i.e., allergies in the spring, flu-like symptoms in the winter).
- The data should provide the ability to drill-down to a more detailed level from aggregated data to support evaluation of possible data anomalies.
- Public health entities (PHEs) need to have the analytic capacity to process data and identify signals of possible outbreaks from large data sets.
- Analysis should utilize established methodologies to reduce false alarms and minimize end-user burden.
- Historical trending and comparative analysis methods should be established to detect aberrations with sensitivity and specificity.

- Aberration detection algorithms must be utilized, compared, tested, and refined based upon a thorough understanding and working knowledge of the interpretation of previous findings.
- Data must support associating possible events to data in other sectors (i.e., agriculture, environment, etc.).

2.1.5 Data Visualization and Analytical Reporting

- Data must be visually represented using geospatial mapping and/or temporal charting.
    - Geospatial mapping should be leveraged to display events in different geographic areas (zip, county, state, region, etc) by data source (i.e., BioWatch cities, states, or MSAs).
    - Temporal charts, such as time series graphs, should be used to visualize how quickly is an event is spreading across geographic borders (zip, county, etc.) and should include environmental factors (i.e., wind directions and speed which affect the spreading of airborne agents).
- An EED system should support the ability to perform a variety of ad-hoc queries for electronic data investigation, including reporting for single or multiple zip code areas, MSA comparisons, or national comparisons.
- Automated reporting tools and pre-defined report templates should be supported to ensure consistency and quality.
- Early Event Detection requires the ability to generate both detailed and aggregated reports.
- Historical trending must provide a baseline against which new outbreaks may be compared.
- An infrastructure must be established to support cross-jurisdictional investigations and may allow PHEs to view trends that extend beyond their jurisdiction.

2.1.6 Alerts and Communications

- EED must be able to initiate alerts to key personnel involved in responding to public health emergencies.
- EED must be able to interact with public health directories to support routing of data to partners based on profile information that includes the participant's name, role, and associated organizations.
- Thresholds should be configurable based upon an outbreak type, level of potential risk to the population, and the existence of external factors that may prompt increased watchfulness.

2.1.7 Data Exchange and System Integration

- Data exchange must be bi-directional and secure.
- Data exchange must support investigations across jurisdictions and require collaboration at multiple levels (i.e., local, state, and national).

- EED messaging should adhere to PHIN messaging standards (HL7) and secure data transport (ebXML as provided by PHIN MS or other transport mechanisms) should be used when exchanging information between organizations and systems.
- Efficient data exchange must be established between the data sources, state and local health departments, and national health partners or data brokers.
    - When a state or local health department warrants further information in order to investigate a public health concern, it must be able to electronically request and receive that information from the data source either directly, or via the national broker.
    - Upon receipt of a request for additional information, a data source must be able to electronically provide that information to the requesting party.
- EED recommends that local jurisdictions associated with a metropolitan area receive all data for the metropolitan area.
- Secure data exchange is required and should include appropriate security and privacy considerations, including data encryption and both destination and source authentication.
- EED should integrate with conventional surveillance systems and corroborate surveillance findings.
- EED should integrate with systems that support Incident Response to identify emergency response team members, assess prophylaxis, training and qualification necessary to respond to the event.
- After possible cases are identified, systems supporting EED must be able to interoperate with outbreak management (OM) systems by providing the data needed to identify affected persons and their exposure levels, as well as to enable case management and contact tracing.
    - The response team(s) should continue surveillance of the outbreak location to evaluate whether the response is effectively containing the event.
    - Linkages to detailed patient information should be leveraged in order to support public health investigations, treatment, and contract tracing as well as to contact those patients for follow-up exams and any necessary treatments.

2.1.8 Consequence Management Operations

- Guidelines must be established for determining whether an event constitutes an outbreak or a false alarm.
- When an event is identified with complete anomaly characterization, it must be immediately communicated to public health partners and other related parties involved in public health investigations via appropriate communication and alerting systems.
- Operational requirements including processes, personnel, and responsibilities must provide clear instruction about supporting, maintaining, and testing EED systems.

- Operational processes must be defined in detail for successful data exchange (bundling, parsing, formatting, etc), data mapping, analysis, visualization, reporting, and alerting of public health events.

2.1.9 System Security and Availability

- The security layer must be managed for authentication, authorization, and access control.
    - Authentication is required to validate that the user is registered to use the system and has signed on with the appropriate user name and password or other identifiable key. Strong authentication mechanisms, such as X.509 certificates or secure token based technology, are required.
    - Authorization levels must be supported to manage access to system functions and data. Authorization levels can include user based, role based and/or context based authorization.
    - Access control rules must be implemented to enforce authorization levels and control user access to the system. For example, access control should allow a jurisdiction to view its own data but should not allow access to data for other jurisdictions.
- Proper access control must be implemented to allow an organization to see information pertinent to its needs but prevent it from seeing restricted data.
- Privacy concerns must be addressed in order to protect the patient and organizations from fraudulent use of their information.
- Continuity of Operations (COOP) and Disaster Recovery plans must be clearly established to ensure the system will not suffer performance, availability, security, and validity failures during emergencies and destructive events.
- Data sources should be monitored for compliance with the data collection, quality, consistency, and integrity standards.
- Interfaces with other systems must be monitored and managed by qualified personnel to ensure the lines of communication remain constantly open and accessible.
- Metrics should be developed to evaluate the reliability of data feeds, latency, performance, and throughput.
- Personnel should be available to help resolve data exchange and connectivity issues.
- PHIN security standards can be viewed using the following link: www.cdc.gov/phin/architecture/automated_data_exchange.htm

3 PROCESS FLOWS

The processes that comprise an EED system are illustrated in Figure 1-1 below:

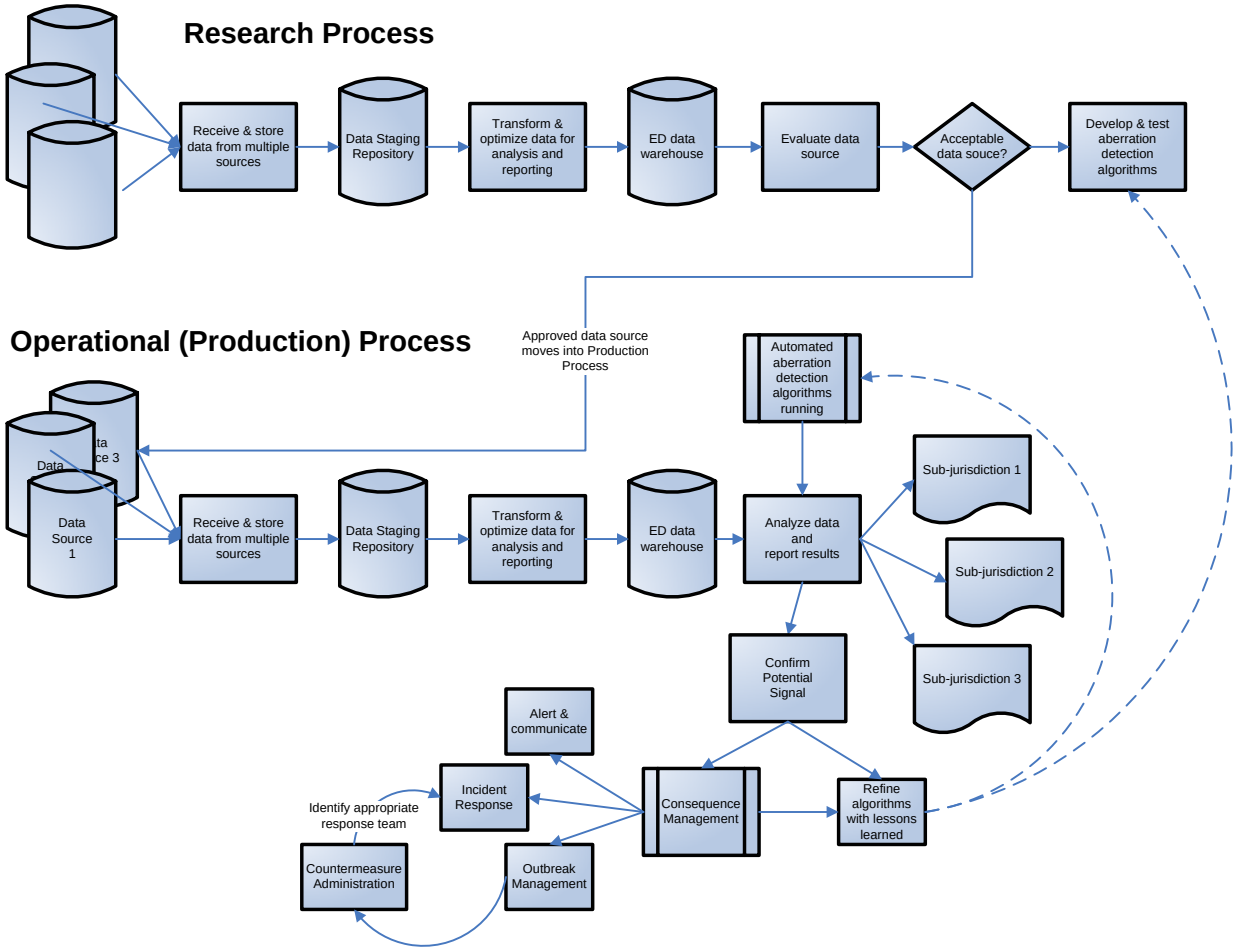

Figure 1-1: Early Event Detection Process Flow

The research process flow runs parallel to the operation process flow in that as potential data sources are identified, evaluated, and approved, they will become a data source used in the live, operational (production process). When production data analysis indicates an event, the consequence management process will initiate, which is set to