

DAILY OPERATIONS

Process: How to prioritize and characterize an anomaly

1) Identify anomalies that meet statistical criteria.

To examine large amounts of data and identify statistical increases that may be of potential public health importance, statistical algorithms are used. Anomalies of statistical significance are then characterized, decisions made regarding potential public health importance, and further actions taken as necessary. Currently, the criterion being used to identify statistical anomalies is that they must have a W2 Rate Recurrence Interval of at least 200 days. Anomalies meeting the recurrence interval threshold criteria are entered into the Anomaly database each morning. Analysts are responsible for certain geographic regions of the country and for characterizing the anomalies identified within these regions. Anomaly locations can be individual facilities, a facility group, or for data from all facilities in the country.

2) Prioritize daily anomaly list.

As the number of facilities increases, the number of statistical anomalies each analyst is responsible for characterizing on a daily basis will greatly increase. It is not possible or practical to evaluate every anomaly. The following criteria provide guidance regarding how to prioritize individual analysts' daily anomaly lists so as to ensure that the most potentially severe or rare anomalies are evaluated on a daily basis.

Prioritization Criteria

- a. Characterize anomalies with a location of an individual facility (not anomalies with a location of "facility group" or "all hospital facilities").
- b. Characterize anomalies with a patient class of Emergency or Inpatient (since these patients are presumably more ill).
- c. Characterize anomalies based on rate (since rate takes into account visit volume). Anomalies can have a statistically significant rate only, or both the count and rate can be statistically significant.
- d. Characterize anomalies with a W2 Rate recurrence interval ≥ 200 days (to identify syndrome counts/rates that are expected to occur only once every six months).

3) Characterize the anomaly.

Once statistical anomalies have been identified and prioritized, they must be characterized to determine if any may be of potential public health importance. The following process and criteria are used to characterize each prioritized anomaly and determine appropriate action.

Criteria

- I. Confirmatory
 - a. Known/confirmed event by state and/or local public health official
 - b. Known/confirmed event by news source
 - c. Known/confirmed event by CDC
- II. Negative

- a. Rate-based anomaly no longer statistically significant due to denominator increase
 - b. Application QA/QC issue
 - c. Holiday effect
 - d. Data feed disruption depressed the expected value
 - e. Coding and/or data entry errors
 - f. Syndrome/sub-syndrome binning issue
- III. Positive
- a. Exceeds previous maximum count and/or rate within last six months
 - b. Exceeds count and/or rate during the same time period the previous year
 - c. Temporal clustering
 - d. Severe illness
- IV. Descriptive
- a. Patient class clustering
 - b. Large proportion of co-morbidities
 - c. Homogeneity among chief complaints and/or diagnoses
 - d. Demographic clustering
 - e. Sub-syndrome distribution change
- V. Future positive criteria (pending additional data being added to the application)
- a. Geographic clustering
 - b. Evidence of anomalies among different syndromic data sources
 - c. Evidence of increases among other surveillance sources
 - d. Positive laboratory results for specific illnesses of concern

4) **Make a decision.**

Anomalies are evaluated using the above criteria. A decision regarding potential public health significance is made using the following rules. Based upon this decision, further actions may be taken.

5) **Determine action.**

All anomalies are entered, documented, and tracked using the Anomaly Database. Analysts should annotate each anomaly based upon the process, criteria, and decision. This allows for analysis of the types of anomalies that are being identified by BioSense, as well as the characteristics, decisions, and actions made based upon the data.

6) **Criteria, Decision, and Action Assignments**

Criteria: Anomaly did not meet prioritization criteria.

Decision: Anomaly not characterized.

Action: Anomaly not characterized.

Criteria: [Any of the negative criteria are met.]

Decision: Anomaly is not likely to have public health significance.

Action: Anomaly characterized but no action taken.

Criteria: [Any of the confirmatory criteria are met.]

Decision: Anomaly has potential public health significance.

Action (one or more of the following): Report submitted to BIC leadership. Report submitted to state and/or local public health officials (DEPR leadership approval required). Report submitted to CDC DEOC (DEPR leadership approval required).

Criteria: [Does not meet any negative criteria and meets one or more positive, confirmatory, or descriptive criteria AND analyst determines that the anomaly is of public health significance.]

Decision: Anomaly has potential public health significance.

Action (one or more of the following): Report submitted to BIC leadership. Report submitted to state and/or local public health officials (DEPR leadership approval required). Report submitted to CDC DEOC (DEPR leadership approval required).

Criteria: [Does not meet any negative criteria and meets one or more positive, confirmatory, or descriptive criteria AND analyst determines that the anomaly is NOT of public health significance.]

Decision: Anomaly is not likely to have public health significance.

Action: Anomaly characterized but no action taken.

Process: How to use the BioSense Application to characterize an anomaly**All Modules**

1. Application QA/QC issue
 - a. Definition
 - i. Negative criteria
 - ii. Anomaly information inconsistent across application modules.
This can include inconsistent counts, rates, and/or recurrence intervals, different patient ID being assigned to the same patient, or potentially other issues that will be added as identified.
 - b. How to use the application
 - i. From the Statistical Anomalies module, click on the anomaly of interest to view the time series for the syndrome, facility, patient class, and data type of interest. Mouse over the star on the time series associated with the anomaly. Is the count, rate, and recurrence interval information the same as what is displayed in the Statistical Anomalies module (and the Anomaly Database)? If there are discrepancies between various application modules, then this criterion applies. From the Time Series module, click on the Patient List link; is there evidence of potential assignment of multiple patient ID's to the same patient? Checking the Patient Detail for these patients will provide additional information for cross-checking, including date of birth, admission and discharge times, etc. If there is evidence of duplication, then this criterion applies.
 - ii. **If this criterion applies, then STOP.**

Chief Complaint and Diagnosis Module

This module is used primarily for application QA/QC as well as to quickly assess data availability for the current day. The counts are checked to determine if data availability is at expected levels. Statistical anomalies for each syndrome are indicated in this module, but the Statistical Anomalies module is the primary starting point for characterizing anomalies.

Select Jurisdiction:

Select the jurisdiction that is appropriate.

Select Options:

Recurrence Interval Threshold = 100 days

Check "Apply these analytic settings to all screens."

Statistical Anomalies Module

This module is used to obtain the details regarding statistical anomalies. Information includes syndrome, data type, patient class, anomaly date, anomaly location, count, recurrence interval for the count, expected count, ratio of observed to expected count, rate, expected rate, and rate recurrence interval.

Select Jurisdiction:

Select the jurisdiction that is appropriate.

Select Options:

Recurrence Interval Threshold = 100 days

Check "Apply these analytic settings to all screens."

Date Range = 2 weeks

End Date = today

Using this module, the following criteria can be evaluated:

2. Rate-based anomaly no longer statistically significant due to denominator increase
 - a. Definition
 - i. Negative criteria
 - ii. The anomaly was statistically significant previously, but as the denominator increased this anomaly was no longer significant. Diagnosis based anomalies may take several days to drop off due to data latency issues.
 - b. How to use the application
 - i. Has the rate-based anomaly from the database disappeared from the Statistical Anomalies module of the application? If there are no other known application QA/QC issues that might be causing this, then this criterion applies.
 - ii. **If this occurs, then STOP.**

Time Series Module

This module is used to examine longer term trends in the data and to compare the anomaly count and rate to historical data. Patient class, data type, event type (syndrome or sub-syndrome), event, and demographic strata (age, race, and gender) can be selected.

Select Data:

Based on anomaly of interest

Select Options:

Recurrence Interval Threshold = 100 days

Check "Apply these analytic settings to all screens."

Left vertical axis = count (7-day expected value optional based on user preferences)

Right vertical axis = rate (7-day expected value optional based on user preferences)

Date Range = 26 weeks

End Date = today

Using this module, the following criteria can be evaluated:

3. Holiday effect
 - a. Definition
 - i. Negative criteria

observed during the same enclosing 4 week period the previous year, then this criterion applies.

8. Patient class clustering
 - a. Definition
 - i. Descriptive criteria
 - ii. Anomalies for that syndrome occur among different patient classes on that day or during the past week at the facility.
 - b. How to use the application
 - i. View the Time Series for the syndrome of interest grouped by data type and stratified by patient class. On the anomaly date or within the past week, are there anomalies for that syndrome occurring among different patient classes at that facility? If yes, then this criterion applies.

Patient List Module

This module is used to examine a line list of patients associated with a particular anomaly. Information presented includes the patient zip code and state, facility name, patient ID, age, gender, patient class, reason for visit or chief complaint or reason for admit, working diagnosis, and final diagnosis. Patterns among a group of patients' clinical data can be quickly assessed.

Using this module, the following criteria can be evaluated:

9. Coding and/or data entry errors
 - a. Definition
 - i. Negative criteria
 - ii. At least 50% of the visits making up the anomaly have coding errors. Applies only to ICD9-based anomalies.
 - b. How to use the application
 - i. Click on the Patient List link associated with the anomaly date in the table beneath the Time Series (or the anomaly date on the actual graph itself). Examine the patient list and determine how many visits are associated with a hypothesized or confirmed coding and/or data entry error. Knowledge regarding coding and/or data entry errors can be postulated by using the application, but can only be confirmed by the source. We might hypothesize that there is a coding issue if we observe that a patient's reason for visit, chief complaint, or reason for admit are completely unrelated to the final diagnosis. For example, we have observed previously at one facility that patients reporting to the ED for pregnancy related complaints were all being assigned a final diagnosis ICD-9 code for cholera. The diagnosis was completely unrelated to other clinical information for these visits and was also a rare diagnosis being used on a frequent basis. Among the visits making up an ICD9-based anomaly, if at least 50% appear to have miscoded

diagnoses, or if there is a confirmed coding and/or data entry error at the source that can be attributed to at least 50% of the visits making up a diagnosis-based anomaly, then this criterion applies.

ii. If this criterion applies, then STOP.

10. Working diagnosis based anomaly negated by final diagnosis

a. Definition

- i. Negative criteria
- ii. At least 50% of the visits making up the anomaly have a final diagnosis that negates the working diagnosis. Applies only to working diagnosis anomalies.

b. How to use the application

- i. Click on the Patient List link associated with the anomaly date in the table beneath the Time Series (or the anomaly date on the actual graph itself). Examine the patient list and compare the working diagnoses to the final diagnoses. If at least 50% of the patients making up the working diagnosis based anomaly have working diagnoses that are unrelated to the final diagnoses, or the final diagnoses negate the working diagnoses, then this criterion applies.

ii. If this criterion applies, then STOP.

11. Syndrome/sub-syndrome binning issue

a. Definition

- i. Negative criteria
- ii. At least 50% of the visits making up the anomaly are mapped to the syndrome due to inaccurate text parsing. Applies only to free-text based (reason for visit, chief complaint, reason for admit, or in some cases working diagnosis) anomalies.

b. How to use the application

- i. Click on the Patient List link associated with the anomaly date in the table beneath the Time Series (or the anomaly date on the actual graph itself). Examine the patient list and determine how many visits are mapped to the syndrome due to incorrect text parsing. If there are questions regarding text parsing or clarification needed regarding whether or not something is an error, please discuss with science leads in this area. Any text parsing issues that are noted should be recorded in the appropriate internal drive location so that they may be included in future improvements to the text parsing masters. An example we have noted is visits with “sore throat” are being mapped to the localized cutaneous lesion syndrome because of the word “sore.” Among the visits making up the anomaly, if at 50% are associated with a text parsing issue that results in these visits being incorrectly assigned to the syndrome, then this criterion applies.

ii. If this criterion applies, then STOP.

12. Large proportion of co-morbidities
 - a. Definition
 - i. Descriptive criteria
 - ii. At least 50% of the visits making up the anomaly have co-morbid conditions. Co-morbid conditions include chronic diseases and history of tobacco or alcohol use. Applies primarily to ICD9-based anomalies.
 - b. How to use the application
 - i. Click on the Patient List link associated with the anomaly date in the table beneath the Time Series (or the anomaly date on the actual graph itself). This criterion applies primarily to ICD9-based anomalies; a patient's co-morbid conditions are often not included in the reason for visit, chief complaint, or reason for admission fields. Examine the patient list and determine how many visits are associated with co-morbidities. Among the visits making up the anomaly, if at least 50% have co-morbid conditions, then this criterion applies.

Patient Map Module

This module is used to examine the geographic distribution of patients associated with a particular anomaly. As of now, the maps provide only an indication of where patients live who are seeking care for a particular condition, which may not give us much more information beyond the hospitals' catchments. The addition of geospatial data analysis capabilities will allow us to assess geographic clustering in the future.

Select Options:

User preference

Describe Module

This module includes distributions for servicing facility, demographic groups (age, race, and gender), and hour of day, Patient County of residence, data type, syndrome, sub-syndrome, patient class, and diagnosis type. The analyst can compare demographic patterns for this anomaly to historical patterns, as well as examine syndrome and sub-syndrome distributions. Subsets and combinations of patient groups can be made by clicking on characteristic sub-groups.

Select Data:

Based on anomaly of interest

Select Options:

To look at distributions for the day of the anomaly, Date Range = 1 day and End Date = Anomaly Date

To compare distributions for the day of the anomaly to the previous 4 weeks, Date Range = 4 weeks and End Date = Date prior to the Anomaly Date

Using this module, the following criteria can be evaluated:

13. Severe illness
 - a. Definition
 - i. Positive criteria
 - ii. At least one visit making up the anomaly maps to the Severe Illness/Death syndrome
 - b. How to use the application
 - i. After the Patient List is reviewed, click on Describe. Make sure that the date for this module is set to the anomaly date. Click on the “Syndrome” thumbnail. If at least one visit is associated with this syndrome, then this criterion applies.
14. Homogeneity among chief complaints and/or diagnoses
 - a. Definition
 - i. Descriptive criteria
 - ii. At least 50% of the visits making up the anomaly map to the same sub-syndrome.
 - b. How to use the application
 - i. After the Patient List is reviewed, click on Describe. Make sure that the date for this module is set to the anomaly date. Click on the “Sub-syndrome” thumbnail. If at least 50% of the visits are associated with the same sub-syndrome, then this criterion applies.
15. Demographic clustering
 - a. Definition
 - i. Descriptive criteria
 - ii. For this anomaly, at least one age or sex stratum is 25% higher than the percentage for that age or sex stratum has been over the 4 week period prior to the anomaly.
 - b. How to use the application
 - i. After the Patient List is reviewed, click on Describe. Make sure that the date for this module is set to the anomaly date. Click on the “Age Groups” and “Gender” thumbnails. Note the age and gender distributions for the anomaly date. Change the date range to 4 weeks and the end date to the day before the anomaly. Click on the “Age Groups” and “Gender” thumbnails. Note the age and gender distributions for the 4 week period prior to the anomaly. If the proportion for at least one age or gender stratum for the anomaly date is 25% higher than the proportion for this age or a gender stratum over the previous 4 weeks, then this criterion applies.
16. Sub-syndrome distribution change
 - a. Definition
 - i. Descriptive criteria

- ii. For this anomaly, the distribution of at least one sub-syndrome is 25% higher than the percentage for that sub-syndrome has been over the 4 week period prior to the anomaly.
- b. How to use the application
 - i. After the Patient List is reviewed, click on Describe. Make sure that the date for this module is set to the anomaly date. Click on the “Sub-syndrome” thumbnail. Note the sub-syndrome distributions for the anomaly date. Change the date range to 4 weeks and the end date to the day before the anomaly. Click on the “Sub-syndrome” thumbnail. Note the sub-syndrome distributions for the 4 week period prior to the anomaly. If the proportion for at least sub-syndrome for the anomaly date is 25% higher than the proportion for this sub-syndrome over the previous 4 weeks, then this criterion applies.

Census Module

This module is used to examine hospital utilization data at the facility and hospital unit levels. Information presented at the hospital level includes occupancy rate, number of admissions, number of discharges, and number of deaths. Hospital unit level data include number of occupied beds and number of available beds. Due to data definition and data validity issues, these data are not used routinely.

Select Jurisdiction:

Select the jurisdiction that is appropriate.

Select Options:

Recurrence Interval Threshold = 100 days

Check “Apply these analytic settings to all screens.”

National Map Module

This module displays data for selected patient classes, data types, syndromes or sub-syndromes, and demographic groups (age, race, and gender) on a national map. Counts for the syndrome or sub-syndrome are mapped based on county of patient residence. The addition of geospatial data analysis capabilities will allow us to assess geographic clustering in the future.

Select Data:

Based on anomaly of interest

Select Options:

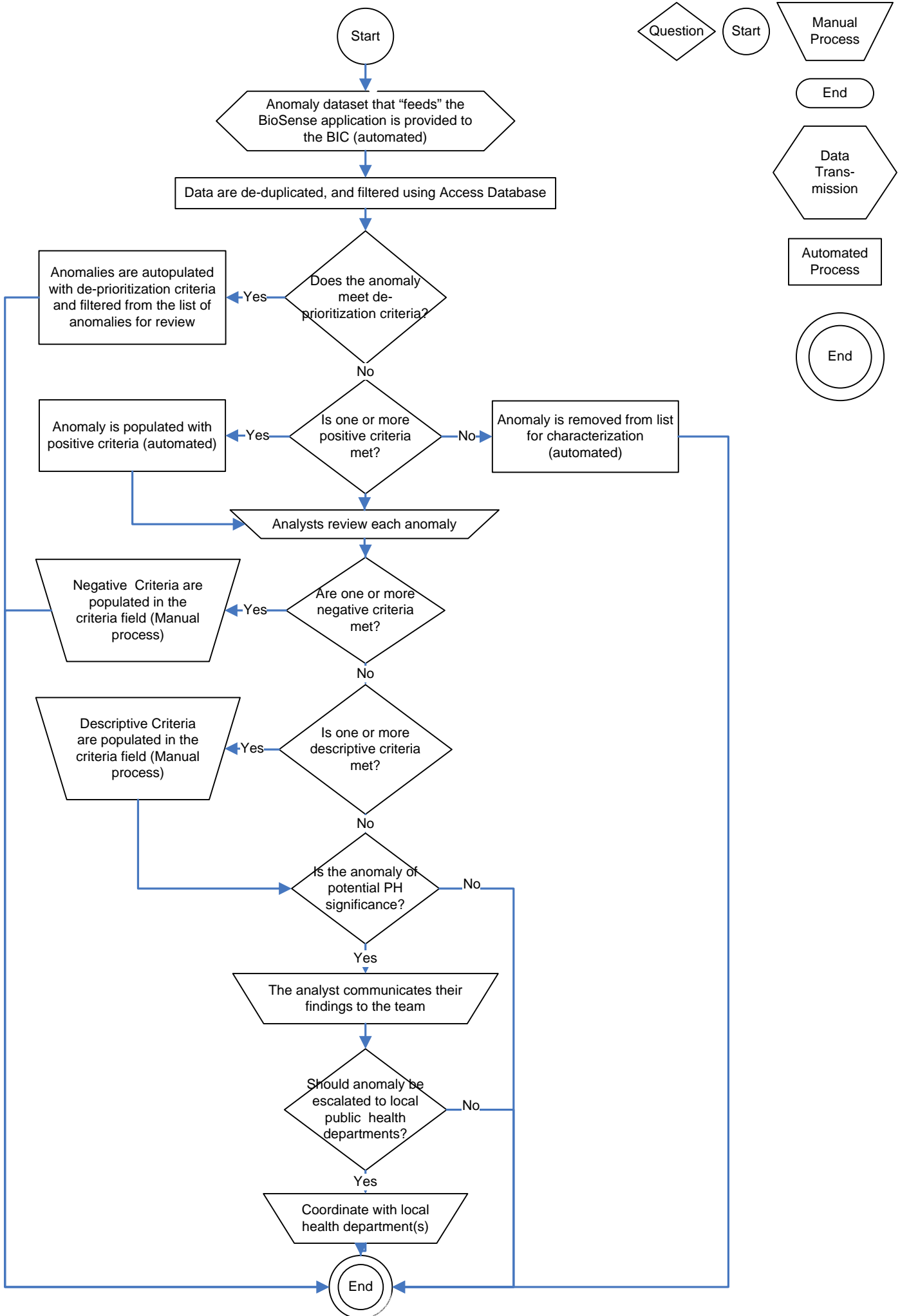
User preference

Anomaly Decision Making Process Flow

The order in which the anomaly characterization criteria should be evaluated is as follows:

1. Confirmatory criteria – Information is gained from outside sources, independent of the BioSense application. These criteria will not apply in most circumstances.
2. Statistical Anomalies – Evaluate negative criteria first, followed by positive criteria.
3. Time Series – Evaluate negative criteria first, followed by positive criteria.
4. Patient List – Evaluate negative criteria first, followed by positive criteria.
5. Describe – Evaluate negative criteria first, followed by positive criteria.

At any point during this process, if a negative criterion is determined to apply to an anomaly, the analyst can stop the process. The negative criterion can be assigned to that anomaly in the Anomaly Database for tracking and analysis purposes.



Start

Anomaly dataset that "feeds" the BioSense application is provided to the BIC (automated)

Data are de-duplicated, and filtered using Access Database

Does the anomaly meet de-prioritization criteria?

Anomalies are autopulated with de-prioritization criteria and filtered from the list of anomalies for review

No

Is one or more positive criteria met?

Anomaly is populated with positive criteria (automated)

Anomaly is removed from list for characterization (automated)

Analysts review each anomaly

Are one or more negative criteria met?

Negative Criteria are populated in the criteria field (Manual process)

No

Is one or more descriptive criteria met?

Descriptive Criteria are populated in the criteria field (Manual process)

No

Is the anomaly of potential PH significance?

The analyst communicates their findings to the team

Should anomaly be escalated to local public health departments?

Coordinate with local health department(s)

End

Question

Start

Manual Process

End

Data Transmission

Automated Process

End