

T-Cube Standard Operating Procedures for Two main Detection Analysis Scenarios *Real-Time Biosurveillance Program*

Nuwan Waidyanatha (LIRNEasia, Sri Lanka)
Artur Dubrawski (Auton Lab, USA)

Purpose: Give the epidemiology units: PDHS-REU (Sri Lanka) and DDHS-IDSP (Tamil Nadu) a set of guidelines for carrying out the operations of the two main detection analysis scenarios: 1) Investigating a suspected or ongoing disease outbreak and 2) routine monitoring of notifiable/communicable (high priority) disease.

Scope: These instructions or standard operating procedures for investigating known disease outbreaks and monitoring high priority (i.e. communicable diseases or notifiable disease) are for the RTBP (Real-Time Biosurveillance Program) users; namely health officials belonging to the Deputy Director of Health Services Integrated Disease Surveillance Program (DDHS-IDSP) and the Provincial Director of Health Services Regional Epidemiology Units (PDHS-REU) staff members. The DDHS-IDSP and PDHS-REU should apply the *Scenario B* for monitoring the 21 disease in Sri Lanka and 11 diseases in Tamil Nadu. *Scenario A* is for investigating any event of ongoing or potential disease outbreak.

Background: Reader should be familiar with the basics of accessing T-Cube Web Interface (TCWI), loading data, manipulating time series, maps, and pivot tables as described in the manual. It is recommended that the user go through the main manual and the exercises described in that manual before engaging in the activities described in this document.

Detection analysis Scenarios

Scenario A: Investigating a suspected or ongoing disease outbreak

Assume you have been notified or have come to learn of increased cases of a particular disease and want to investigate that particular disease, then follow the steps described in this section but do not restrict yourself to just these steps. You may use all other features available in TCWI as well.

Time Series Analysis (Temporal Scan)

After loading the data, click on the **Query Panel** to expand (Figure 1). Click the blue square next to label **Disease**, which will display the list of diseases. Click in the check box (i.e. mark with \checkmark) of the particular disease you wish to investigate (e.g. *Common-cold*). Rename the query name, in the text box located next to label **Name** (e.g. *common-cold*) and click Save. Scroll down to Time Series Analysis display. De-select (i.e. uncheck) **All Data** and **Current Query**. Now you should see the saved query (e.g. *common-cold*) as shown in Figure 2.

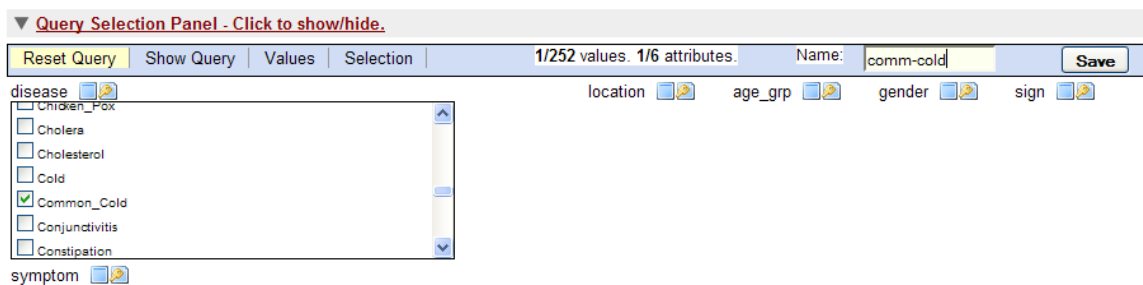


Figure 1: Expanded Query Panel with attributes

The time series plot for the selected disease (e.g. *common-cold*) compared with the time series plot of **All Data** may not show any disparities for you to identify any significant increases or decreases in disease counts. Therefore, it is recommended that you run **Temporal Scan** to find periods in time where there are statistically significant differences of the selected disease (e.g. common-cold).

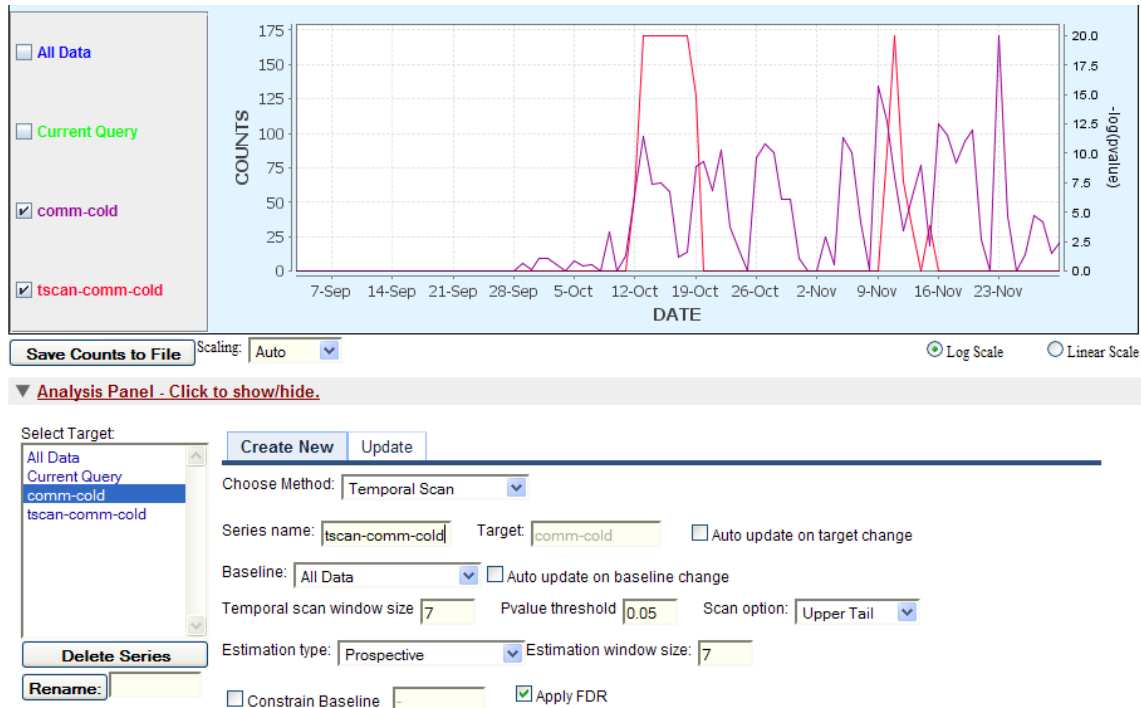


Figure 2: Time Series display showing saved query and temporal scan results with Analysis Panel

First set the parameters in the **Analysis Panel** (see Figure 2). In **Select Target** box click on the saved query, in this case *common-cold*. Immediately, the parameters in the analysis panel will be displayed. Leave all default values as they are but change the following parameters - **Choose Method** = *Temporal Scan*, **Baseline** = *All Data*, **Scan Option** = *Upper Tail*, **Estimation Type** = *Prospective*, **Estimation Window Size** = 7, Apply False Detection Rate (**FDR**) = $\sqrt{\quad}$ (check). Change the **Series Name** to one that you can be identified (e.g. *tscan-comm-cold*) After all parameters are set and the series name is changed, click **Submit** to execute temporal scan. Once the process is complete and graphs are displayed, click on **Log Scale** located at the bottom right corner of the plots.

The Temporal Scan scores are shown by the plot in red with legend label *tscan-comm-cold*. Figure 2 shows two periods: Oct 12-19 and Nov 9-16 with high scores of disease under investigation, in this case *common-cold*. The plot of the disease under investigation (*common-cold*) is in purple.

Map analysis (Spatial Scan)

Please do not click on the Back/Forward/Refresh buttons during using this interface.

Map Display has been initialized

Time Series **Maps** Pivot Tables Poor Performers

Please select the spatial attribute in your dataset.

Dataset 1 location Load Map Clear Map Load Attributes

Figure 3: Loading the map for spatial scan

Navigate in to the Map section, set the **Dataset 1** = *Location*, and click **Load Map** (Figure 3). After several seconds the map will be displayed with blue circles around the locations where data is available.

Scroll down to the **Attribute Selection Panel** and click on the link to expand the panel (see Figure 4). Check (✓) the box associated with the Disease list, which will show all diseases highlighted in blue, indicating all of them to be selected. Scroll through list and click on the disease under investigation (e.g. *common-cold*), which will be the only disease highlighted in blue shown in Figure 4. Click on the **Submit** button at the top of the **Attribute Selection Panel**. Almost, immediately the map and the time series plot will change, reflecting the counts of the submitted query with the disease under investigation; i.e. *common-cold* chosen in our example.

Attribute Selection Panel 1 - Click to show/hide.

Submit Reset

Split By -
Select attribute -

Auto Filter

disease location age_grp gender sign

Cholesterol
Cold
Common Cold
Conjunctivitis
Constipation

symptom

Temporal window size: 7 Max. group size: 6 Run Spatial Scan Clear Scan Results

Min. threshold: 0.0 Submit Threshold

Figure 4: Spatial Scan attribute selection panel

Change the **Temporal Window Size** = 7 and **Max Group Size** = 6 (Figure 4). After the parameters are set, click on **Run Spatial Scan**. The results will comprise the map with global score and time series window with global scores. Figure 5 shows the time series for the chosen example: *common-cold* (in blue) and the score (in red). The date slider has

been moved to Oct/18/2009, which shows a high score of 0.9964 between Oct 12 and Oct 18. Also shows another period Sep 09 to Oct 07 with a high score ranging from 0.700 to 0.8500.

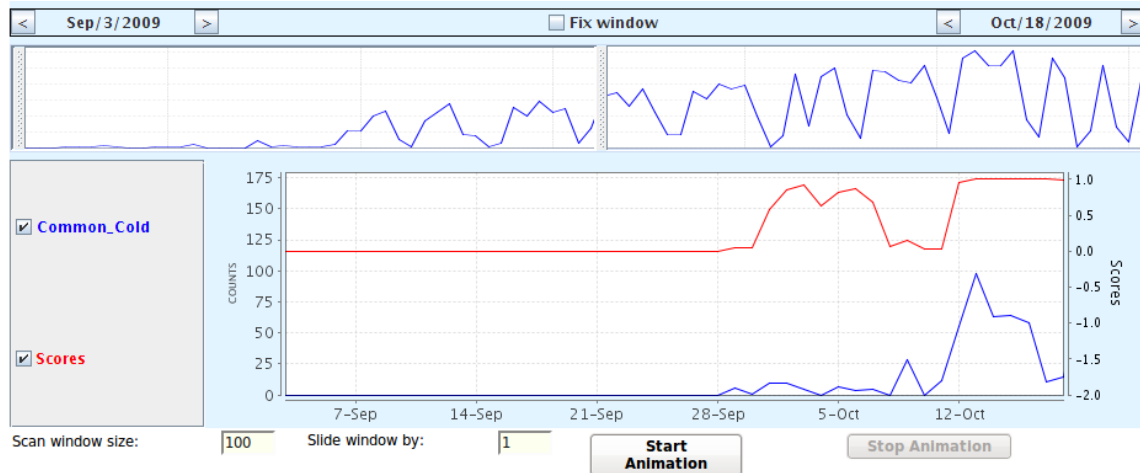


Figure 5: Spatial scan score and time series plot for disease under investigation

Figure 6 shows the corresponding map with the global score of 0.9964. One area (around the Hospital in Katupotha) with a count about 270 (blue circle) and score of 0.833 (orange). Place mouse pointer inside the circle and click to view the name of location and corresponding counts. List will also show the counts of neighboring villages. From the

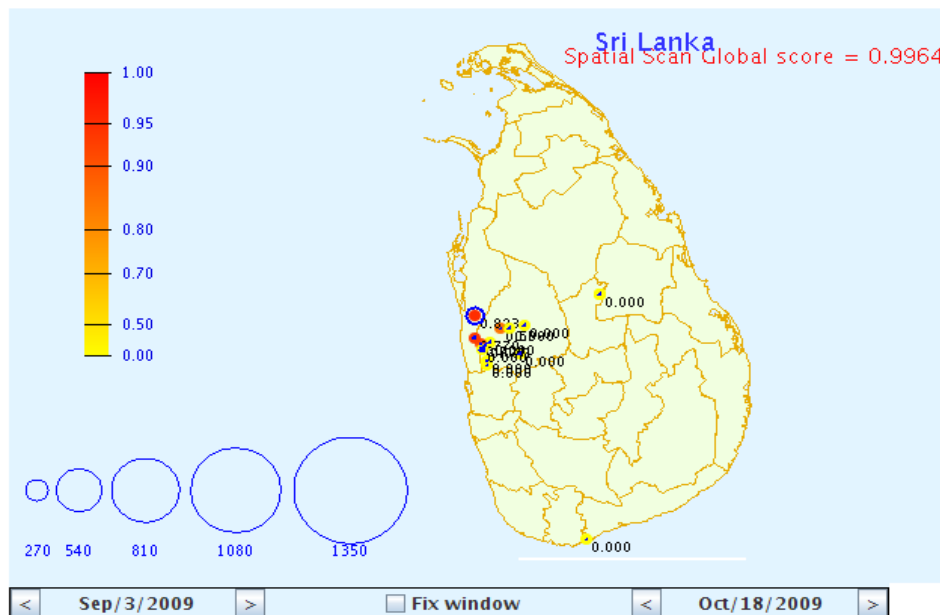


Figure 6: Spatial Scan map with counts and scores of disease under investigation

color scheme we may conclude that the neighboring areas are also showing relatively high counts of the disease under investigation (*common-cold*). By changing the date slider or date arrows the user can see the colors in the respective areas to change proportional to the score.

Scenario A: Monitoring high priority (notifiable/communicable) diseases

Tamil Nadu and Sri Lanka monitor a set of diseases termed as notifiable disease that are currently communicated through the S/P form and H-544 form, respectively. This section illustrates the steps for carrying out the disease surveillance procedure. This procedure should be conducted frequently (daily or weekly).

Time Series Analysis (Massive Screen)

After loading the data, click on the **Massive Screening Panel** to expand (Figure 7). Let all default values be as it is but change the following: **Scan Window Size = 7,14,21,28** **Estimation Type = Prospective**, **Scan Option = Upper Tail**, **Check (✓) Last Day Only**, and **Baseline = All Data**.

The screenshot shows the 'Massive Screening Panel' with the following settings:

- Temporal Scan** (selected) / Change Analysis
- Scan window sizes: 7,14,21,28
- Pvalue threshold: 0.05
- Estimation type: Prospective
- Estimation window size: 7
- Start Date (mm/dd/yyyy): 09/03/2009
- End Date (mm/dd/yyyy): 11/30/2009
- Scan Option: Upper Tail
- Last Day Only
- Constrain Baseline
- Rank Pvalues: 0.05
- Rank Alternate Pvalues: 0.05
- Most significant window only
- Use Hierarchical Attribute
- Use Drill-Down Query
- Baseline: All Data
- Select attributes (upto3) and specific values(optional) to run massive screening. You will be able to select specific values only if the attribute is selected.
 - disease: Denque Fever, Dental Carries, Diabetes, Diabetes mellitus, Diarrhoea
 - location: Ammanamulla, Bogahapitiya, Bopitiya, Deegalla, Dunakadeniya
 - age_grp
 - gender
 - sign
 - symptom
- Run Screening** button

Figure 7: Massive Screening Panel setting of parameters and values

Click the blue square next to label **Disease**, which will display the list of diseases. Click on the particular set of disease you wish to monitor (e.g. *dengue, Diabetes, Diarrhoea*). To select multiple disease hold the **CTRL** button and click on the desired diseases. Similar to Disease, click on the blue square next to label **Location**. This will select all the locations to generate the alerts. Finlay click on **Run Screening**, which will almost immediately generate a list of alerts as shown in Figure 8.

Click once more on the **Massive Screening Panel** to shrink it, then you can see the **Time Series** and **Recent Massive Screening Results** close to each other. Click on the button **Log Scale**. De-select (or uncheck) the legend labels *All Data*, *Current Query* and *MS Baseline*, which will leave you with the *MS Pvalue* and *Results Query* displayed in the **Time Series** window. When you select a alert (or Massive Screening Result) the plot will change.

Figure 8 shows the example of the results *disease = Diarrhoea*, *location = Narammala*, and *Pvalue = 0.02*. We are mostly interested in alerts that have a Pvalue closer to zero; where *Pvalue = 0.02* is considerably high. Alert window (yellow in Figure 8) shows where T-Cube would have generated the alert as early as Nov 10.

The list of Massive Screening results are ordered according to significance with the top most results with Pvalue closer to zero. User should go through the list to find any results that are of interest. Then click on that to view the time series. Thereafter, use the Map to view the spatial clustering of the disease, which is similar to executing the Map Analysis (Spatial Scan) procedure described in the previous section. In this case the disease under investigation will be the disease name listed in the Massive Screening result, which in our example is Diarrhoea.

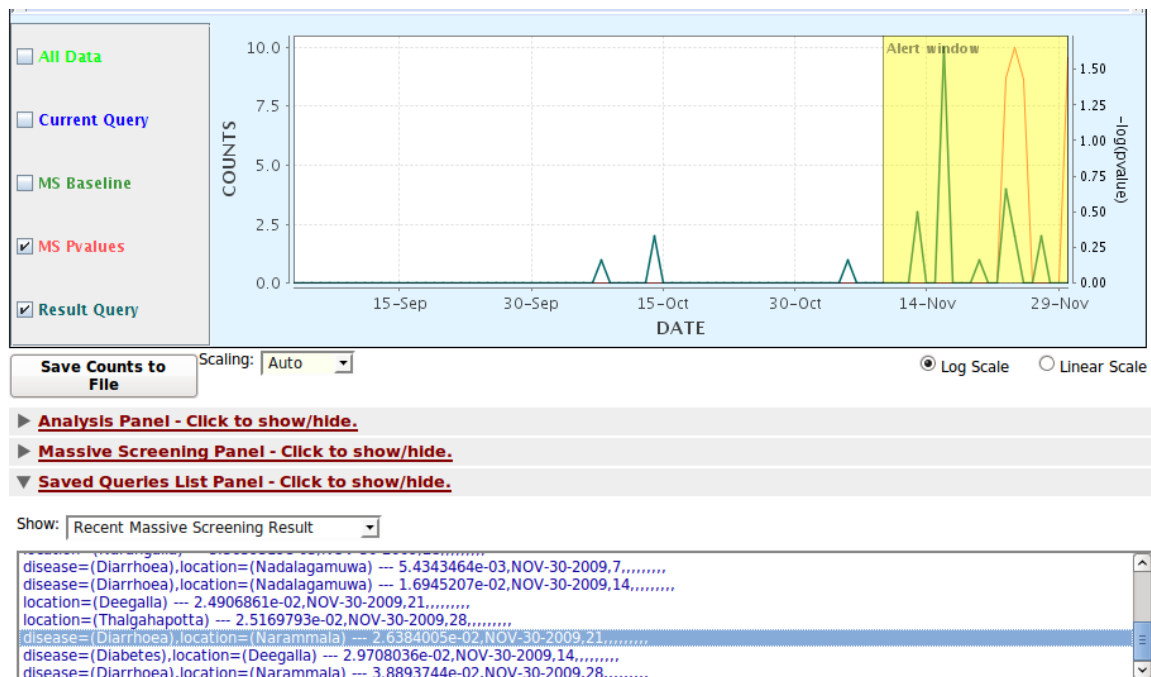


Figure 8: Massive screening results and time series plot for selected result

Pivot Table

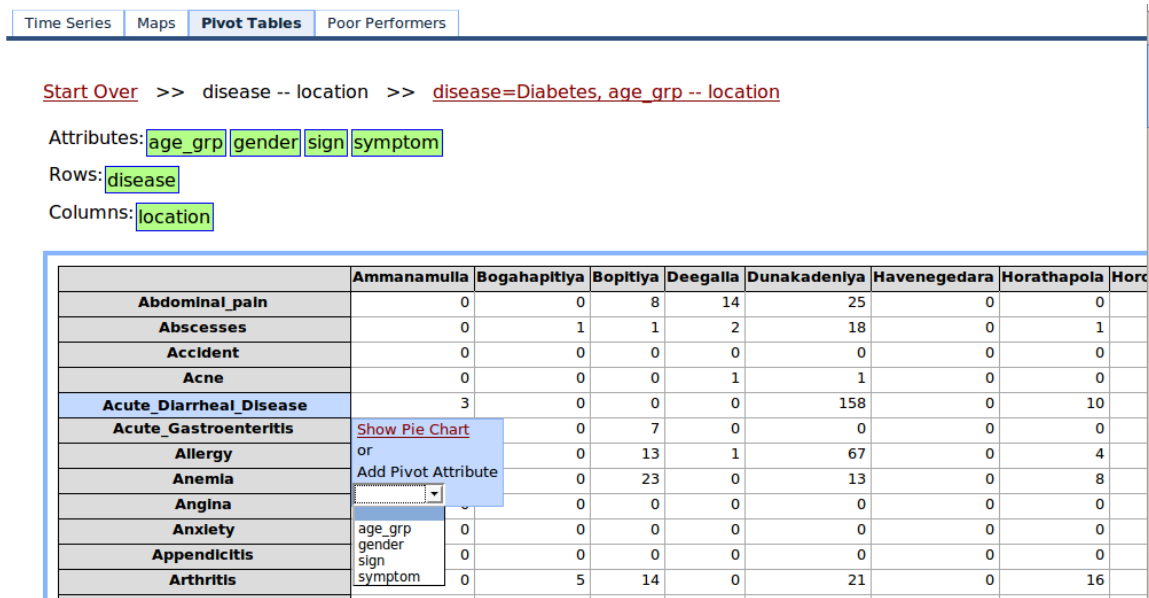


Figure 9: Pivot Table with disease in rows and locations in columns with counts in cells

One can use the pivot table to view counts of the disease for all the locations. First navigate to the **Pivot Table** by clicking the labeled tab. Drag and drop the dimension *Disease* in to the **Rows** and *Location* in to the **Columns**. After a few seconds a table with total counts corresponding to the respective location and disease will be displayed in each of the cells (see Figure 9). Use the **Start Date** and **End Date** controls at the bottom of the table to adjust the dates to your liking.

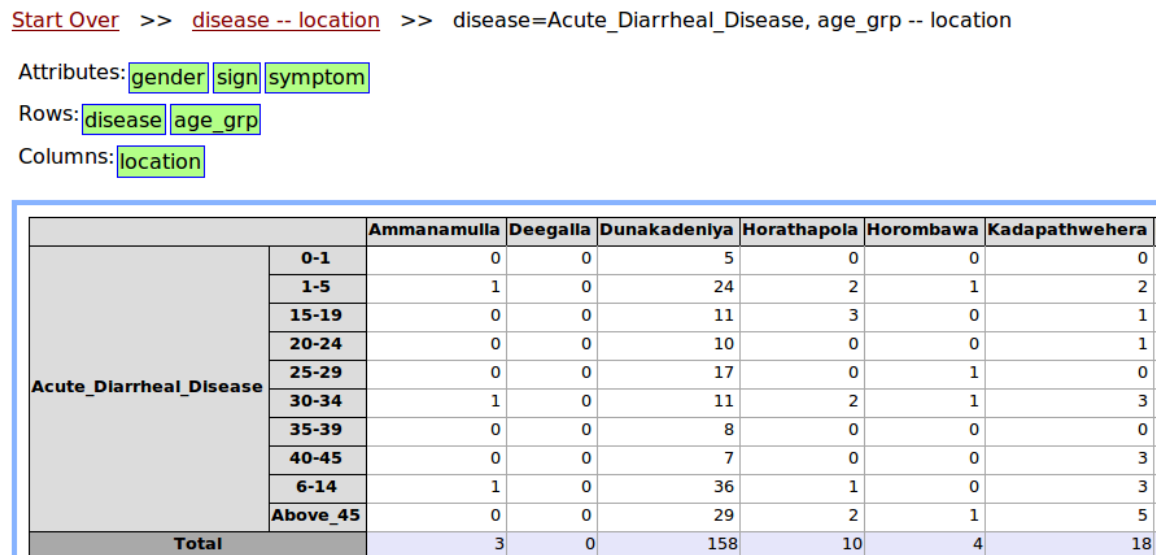


Figure 10: Pivot Table for selected disease split by age group and locations

If the user wants to investigate a single disease slitting it by dimension, click on the disease you wish to investigate (e.g. *Acute_Diarrheal_Disease*), which will show a drop down panel with the list of attributes, as shown in Figure 9. In our example we have selected *age_group* as the dimension, which results in the table shown in Figure 10. To return to the previous table, as in Figure 9, with all diseases and all location click on the link *disease—location* placed on the top of the page, next to link *Start-Over*.