Real-Time Bio-surveillance Pilot Programme in Sri Lanka: Lessons Learned

W.G. C. Sampath MBBS
Researcher
LIRNEasia
Colombo
Sri Lanka
E-mail: chamindusampath@gmail.com

N. Waidyanatha
Senior Researcher
LIRNEasia
Kunming
China.
E-mail: nuwan@lirneasia.net

Vinya Ariyaratne MBBS, MPH, MSc (Community Medicine), MD (Community Medicine), PhD
General Secretary-Sarvodaya
Moratuwa
Sri Lanka
E-mail: vinya.ariyaratne@gmail.com

R.M.S.K Rathnayaka MBBS, MSc (Community Medicine)
North West Provincial Director of Health Services
Kurunegala
Sri Lanka

P.S. Hemachnadra MBBS
Regional Epidemiologist
Kurunegala
Sri Lanka.

E.A.L.C.K. Edirisinghe MBBS, MSc (Community Medicine)
Assistant Director Planning
North West Provincial Director of Health Services
Kurunegala
Sri Lanka.

DOI: 10.4038/sljbmi.v1i3.1774
Abstract

The latter parts of 2007 and early months of 2008 witnessed an alarming number of deaths due to a Leptospirosis outbreak in Sri Lanka\(^1\). An unusual number of patients presenting with symptoms of fever, headache or myalgia concentrated in particular geographic areas (North Central and North Western Province in Sri Lanka) could have signalled the epidemiologists of an abnormal event with the help of a quicker surveillance programme leading to possible implementation of optimal strategies which could possibly have minimized the early deaths and even prevented the progression of the outbreak. The present day paper-based disease surveillance and notification systems in Sri Lanka\(^2\), confined to a set of notifiable diseases, often require 15-30 days to communicate data and for the central Epidemiology Unit to process it. This latency does not allow for timely detection of disease outbreaks and it limits the ability of the health system to effectively respond and mitigate their consequences. Therefore it negatively affects the health status of the work force and productivity of the country. The Real Time Bio-surveillance Program (RTBP) is a pilot study aiming to introduce modern technology to the Health Department of Sri Lanka to complement the existing disease surveillance and notification systems. The processes involve digitizing all clinical health records and analysing them in near real-time to detect unusual events to forewarn health workers before the diseases reach epidemic states. Similar studies have been conducted on bio terrorism surveillance in Winnipeg, Canada\(^3\), pandemic surveillance in Morocco\(^4\) and North Carolina Disease Event Tracking and Epidemiologic Collection Tool (NC DETECT) in North Carolina\(^5\).

The infrastructure of the project RTBP is composed of an interconnected network between health care workers via mHealthSurvey mobile phone application, T-Cube web interface (TCWI) and Sahana Messaging/Alerting Module. Health records from health facilities, namely demographic information, symptoms, suspected and diagnosed diseases are collected through the mHealthSurvey, a mobile phone application\(^6\), that feed in to the TCWI\(^7\), a browser based software tool that detects adverse events; health officials are notified of the adverse events using the Sahana Alerting module that transports via Short Message Service (SMS), Email, and Web\(^10\). Evaluation of the RTBP involves a replication study and parallel cohort study. This pilot study indicates the need for more robust mobile application for data collection with complete ontology, semantics and vocabulary in disease-syndrome information to reduce noise and increase reliability in the datasets. More rigorous capacity building and frequent use is required for health officials to take advantage of the full potential of TCWI. This paper discusses the technologies used in the pilot and the initial findings in relation to usability of the system.

**Keywords:** Bio-surveillance, Epidemiology, Information Communication Technology, m-Health, Disease Outbreak, Event Detection, Alerting, Sri Lanka

*Sri Lanka Journal of Bio-Medical Informatics* 2010;1(3):139-154

DOI: 10.4038/sljbmi.v1i3.1774

**Introduction**

The Real Time Bio-surveillance Program is a pilot project aiming to answer the question “Whether software programs that detect events in health symbolic and categorical data sets and mobile phones that collect health data and receive health alerts are able to predict and prevent disease outbreaks in near-real-time ?.”
The success of the introduced Information and Communication Technologies (ICT) depends not only on the quality of the technology artifacts but also on the actors (i.e. the people and the organizational environment)\(^{(11)}\). Hence, there are three main components that the RTBP researches are investigating: the workability of the technology in the given environments (whether the technology can actually live up to the expectations), understand the set of newly introduced processes that impact the human element (will it aid the healthcare workers with the protocols as it was proven to be the case in the Uganda study)\(^{(12)}\) and the policy implications (are the health workers and epidemiological units ready to accept the changes; i.e. business process improvements or re-engineering).

RTBP provides the ability to detect and monitor a wide variety of health events, involving multiple kinds of diseases, including communicable and non-communicable, as well as reportable and non-reportable ones, following WHO’s general recommendations for disease surveillance systems\(^{(13)}\).

**Research design**

Implementation of RTBP essentially means to make available the right information at the right place, at the right time and in the correct form\(^{(14)}\). In that respect, this section discusses the information flow that completes a cycle (steps 1–8 in Figure 1) where information provided by health workers is processed and resulting decisions are communicated back to the health workers.

Step 1 – Health workers record patient information in various registries such as outpatient registry, inward registry, morbidity report, etc. Once policies are in place for a wider scale deployment, the paper registries can be obsolete and same data can be supplied via electronic means; i.e. skip directly step 2.

Step 2 – Patient complained symptoms, healthcare provider identified signs, and diagnosed disease along with patient’s gender, age, and point of care location (i.e. hospital, clinic, or village name) are entered in to the mHealthSurvey mobile application.

Step 3 – Information is sent to the central database through the Global System for Mobile (GSM) cellular network over the General Packet Radio Service (GPRS) transport layer. In the event the signal is absent, the record is stored in the offline storage in the mobile Record Management System (RMS) until connectivity is established and data is transferred.

Step 4 – Periodically (on the average, once a day), Epidemiologists analyze the information using the T-Cube, time series and spatial scan using web interface software.

Step 5 – If the Epidemiologist detects an adverse event, then a decision is made whether or not to intervene.

Step 6 - Events of interest that require intervention and prevention or are worthy of notifying are disseminated to targeted health workers, in the form of a CAP message, through the Sahana Alerting Module, by the authorized health officials.

Step 7 – A toned down version of the CAP messages that can fit in a Short Messaging Service
(SMS) are transmitted via GSM cellular networks to the health worker mobile phones. The complete CAP message is published on the web for health workers to access through Wireless Application Protocol (WAP).

Step 8 – Based on the received alert message, the health workers, if necessary, activate relevant response plans.

Computer Security is not the main emphasis of the pilot. It is more relevant to study the policy, content, application, and transport layers of the vertical components illustrated in the evaluation matrix (Figure 5). Basic user authentication is in place. We leave it to the implementers of an RTBP ICT system to apply standard global computer security solutions that would protect information, control access, and reducing risks.

**Technologies**

*mHealthSurvey mobile application*

The main menu of the mHealthSurvey comprises: download list, profile, location, offline survey, and health survey, shown in Figure 2 (a). After installing the application the first step is executing the download list function, which will retrieve the lookup values from the database such as the list of disease, sign, symptoms, age-groups, gender names, location types, and health worker types. This is usually a onetime step but the users are encouraged to execute this function from time to time to update the list of disease, signs, and symptoms on
their mobile phones to reflect the changes in the global database. Thereafter, the user must configure the application with their preferences such as their profile, Figure 2(b), and working locations, Figure 2(c). The mHealthSurvey allows for multiple profiles permitting for more than one health worker to share the same mobile phone. After installation and configuration the user is ready to begin sending data through the health survey form (Figure 2(d) and 2(e)). For confidentiality, the project’s initial design had not incorporated the patient’s name the sensitive demographic information, due to ethical clearance reasons, are not offered in the Sri Lanka implementation.

Figure 2: mHealthSurvey Screens (a) main menu (b) profile (c) location (d) patient record page 1 (e) patient record page 2
**Auton Lab T-Cube Web Interface**

T-Cube Web Interface (TCWI) is a generic tool to visualize and manipulate large scale multivariate temporal and spatio-temporal datasets commonly encountered in public health applications\(^{(15)}\). The interface allows the user to execute complex queries quickly and to run various types of statistical tests on the loaded data. Upon uploading the working dataset, the user can manipulate and visualize data through the Time Series, Map, and Pivot Table panels.

The user may choose to apply one of the available statistical modelling and anomaly detection techniques. The list of choices includes moving average, moving sum, cumulative-sum, temporal scan, change scan, linear trend, peak analysis and range analysis. The users can interactively manipulate, navigate, summarize and visualize data at interactive speeds. That supports focused investigations, drill-downs as well as summarizing and reporting operations. The users may choose to simply execute a Massive Screening procedure, which performs an automatic and comprehensive search for anomalous patterns across large number of queries spanning multiple dimensions of data. This function could be invoked interactively by the user, or it could be scheduled to execute periodically to generate a set of alerts. The alerts are sorted according to statistical significance of the corresponding anomalies found in data, and they can be interactively reviewed by the Epidemiologists for the factual confirmation of their practical importance. TCWI supports these efforts by allowing focusing attention on the most surprising patterns in current data and by providing the ability to quickly drill down or roll up the data for further explanations.

The core of the TCWI Massive Screening procedure is the Temporal Scan bi-variate anomaly detection algorithm\(^{(16)}\). It leverages the efficiency of the T-Cube data structure to perform a massive number of tests of statistical hypotheses in order to find the most significantly anomalous patterns in current data. In addition, the RTBP introduced five pre-emptive massive screenings: fever like disease/syndrome, non-communicable diseases, notifiable diseases, and other communicable diseases. The TCWI also implements Multivariate Bayesian Spatial Scan algorithm\(^{(17)}\), that complements the analyses by testing the spatial as well as temporal correlations between health events. This algorithm computes the overall probability of a disease outbreak anywhere in the scope of data selected by the user, separately for each day within that scope. The national score for the current day is reported in the upper right corner of the map display window (Figure 3).
Figure 3 shows an example detection of a potential outbreak of food poisoning. The spatial and temporal distributions of corresponding recorded disease cases are shown in blue. The history of the estimated probability of the food poisoning occurring on a given day anywhere in the nation is depicted with the red line plot. Massive screening automatically identifies the periods of time of abnormally high frequency of cases of food poisoning, relative to cases of other diseases reported nationally. Spatial distribution of probabilities of the food poisoning outbreak computed for separate regions of Sri Lanka for the current day is depicted with filled circles coloured according to the value of the estimated probability. In this example, a central East-West swath of the country seems to be primarily affected by this disease. Figure 3 shows a global score of 0.9795 (i.e. almost a 98% chance) for food poisoning event on August 13th, 2008. This event was later identified as an actual food poisoning outbreak in Kurunegala District. When prompted by the system to the new discoveries, the analysts can further drill into the data by narrowing their filters or selecting various additional modelling techniques to confirm the outbreak.

TCWI also provides computationally efficient, interactive data summarizing capability. Multidimensional data of counts of events (such as the numbers of reported disease cases) can be aggregated into a multi-way matrix view - a pivot table. Multiple attributes can be selected to denote rows and columns of the table by dragging the corresponding attribute names from the attributes list. Once a table is created and automatically filled with values, the user can click on a cell to view the corresponding time series graph, or a pie chart depicting the frequency distribution of the underlying data.
**Sahana Alerting Module**

The RTBP Alerting, Situational-Awareness and Notification Guides are based on the United State (US) Centre for Disease Control’s Public Health Information Network (PHIN) Communication and Alerting Guide (PCA). The PCA Guide has been identified as useful model on which to base the RTBP Guide because it addresses the problem of inter-jurisdictional alerting, provides a comprehensive set of alerting attributes using CAP and Emergency Data Exchange Language (EDXL). RTBP will incorporate CAP as data interchange standards for use in the research in order to serve the primary objective of the project but also to take into account other objectives related to system growth and regional interoperability.

The Messaging/Alerting Module is a Sahana Module that is used for the sending and receiving of messages and/or alerts. The module allows for the generic sending and receiving of messages as Short Messages via the SMS, sending messages as Email, conducting SMS based surveys and sending CAP alerts. The CAP alerting section, which falls under the scope of the RTBP is accessed via the Alert and Templates subsections (Figure 4). The system allows users to create CAP templates and store them in the system. These can then be used when creating CAP messages, which allows the message to be populated based on the relevant template. The template can then be used to pre-populate fields when creating a new alert.

![Figure 4: Browser interface of the Sahana alert generation screens](image)

Figure 4 shows an example of an alert generated to notify a dengue outbreak. Each message
carries a unique identifier, a set of attributes that identifies the source and sender for audits. The scope is set as restricted meaning the message is for those targeted recipients only. Category is naturally set to Health with the event described as an “outbreak”. Priority defines the response actions that should be taken by the receiving health workers or health officials. If the priority was set to “urgent” then recipients may be required to take prompt action; while a “low” priority may mean being vigilant and observe the situation. The description section contains a full synopsis of the alert. Once the attributes are populated the sender can select or type in the list of recipients in the Contacts section (upper right corner in Figure 4). Thereafter, select the delivery types (or transport methods), i.e. email, SMS, web, for the message to be disseminated to the prescribed recipients via those channels.

Evaluation Method

With respect to the information flow cycle illustrated in Figure 1, the technology design is partitioned into a set of data collection, event detection, and alerting vertical segments as shown in Figure 5. The vertical segments are, further, horizontally partitioned into social, content, application, and technology layers (Figure 5).

The RTBP program was evaluated both subjectively and objectively. One aspect was determining the usability of the technology, affect on structural or process quality, investment and operational costs, problems associated with daily operational costs, and social consequences of introduction of technology. The qualitative and quantitative factors were measured through interviews, observations, and exercises. Many important findings of the RTBP was revealed after the technology was deployed and the processes were put in to actions.
Field testing was intended as a formative or summative approach, with formative studies aimed at producing insight as to the effectiveness of a component or sub-system (Figure 5) and identifying potential improvements to that component. Although the general attributes of evaluation is categorized as human resources, communication networks and computing resources\(^{(21)}\), the RTBP further categorized the attributes (i.e. horizontal layers in Figure 5) as social (human resources and policy implications), content, application (computing resources), and technology (communication networks).

Technology acceptability is a key defining attribute for the human element that leads to better understand to what extent the health care workers are willing and able to use the RTBP introduced ICTs. Technology Acceptance Model (TAM) introduces two theoretical constructs: perceived ease of use and perceived usefulness, to study factors influencing technology acceptance of ICTs by individuals within organizations\(^{(22)}\). In addition to TAM, a sector specific approach to acceptability has been developed by the US Center for Disease Control and focuses on measures involving various participation rates among health professionals\(^{(23)}\).

Other key aspects are qualitatively evaluating the organizational impact of introducing ICTs in Healthcare \(^{(24)}\). RTBP also absorbed formal subjective and objective quantitative and qualitative evaluation schemes defined for bioinformatics systems\(^{(25)}\).

As part of the cohort study, the project will execute simulations to evaluate the reliability and effectiveness of each of the verticals; in other instances will take the present system as a basis to assess the proposed system. The replication study would involve injecting data from the past and evaluating the systems performance with respect to known events in the past.

Discussion

**mHealthSurvey**

In the pilot areas, each medical officer examines over 75 patients from 8:00am to 1:00pm (total 4 – 5 hours) each day. Given the frequency at which the crowd needs to be cleared, health workers in Kurunegala District are unable digitize the data using the mHealthSurvey in real-time (i.e. send the Outpatient data before end of the same day). Sarvodaya Suwadana Centre Volunteers (abbreviated as Suwacevo for the purpose of this paper), either shadowing the physician to digitize the spoken out patient case information or extract from the hospital treatment chits. In some occasions when that day's work of digitizing the chits cannot be completed, the Suwacevo, complete them the next day. The delayed entries accounts for, only, 11 percent of the total records.

A certification exercise on the usability to install/configure the mobile application, submit data in a timely manner, and understand the standard operating procedures revealed these young field workers were far more comfortable with using the mHealthSurvey and understanding the processes than anticipated. In Sri Lanka 14 of field workers participated in the exercise and 12 of them scored over 70%.
Each day the databases records approximately 800 to a 1200 records from the 12 Hospitals in Kurunegala District. The signal to noise ratio of the data received through the mHealthSurvey for T-Cube analysis is defined as the number of records with misspelling or misleading as a proportion of all the case records in the database. The Suwacevo are not knowledgeable in medicine (i.e. knowing the terminologies and deciphering between disease and syndrome). Thus, the SNR is 45 percent. After six months in to the evaluation phase, regardless of the learning curve nosy data continues to grow in direct proportion to the count of patient visitations. This is mainly due to the absence of pre-populated disease-syndrome data in the mobile phone memory. However, over the past 6 months the database has grown from 100 to 250 diseases and over 500 syndrome vocabulary with variations of synonyms. The data received varies between British spelling and US spelling, e.g. “diarrhea” versus “diarrhoea”; abdominal pain versus pain in the abdomen; use of short codes such as RTI for Respiratory Tract Infection Given the generality of the T-Cube temporal screening and spatial scanning algorithms, statistically each of the synonyms are two separate statistics. The RTBP pilot has identified the need for a disease-syndrome standardized ontology with semantics and set vocabulary, which would cut across the three vertical components shown in Figure 5.

### T-Cube Web Interface

Statistical surveillance analyses rely on the robustness of data received from the health workers through the mHealthSurvey. The inconsistencies in data introduce noise which in turn can decrease event detection rates and increase rates of false positive alerts.

The present surveillance and notification system in Sri Lanka, besides the H544 notification and investigation process, also reports large number of cases by telephone to the Medical Officer of Health (MOH). The paper-based, manual consolidation method does not provide the real-time surveillance ability, and most importantly, it does not present the data in a navigable and searchable spatio-temporal format as does the TCWI. The computational efficiencies offered by the T-Cube data structure enable health surveillance officials and epidemiologists to quickly see the data visualized spatially on a map and temporally on time series plots. They can monitor evolution of disease in spatio-temporal fashion, either at a high-level of regional or national aggregation, or drilling down as needed to even single cases of diseases like malaria or dengue fever. Accessing the information collected throughout jurisdiction of their particular MOH, but also throughout the neighbouring jurisdictions or over an entire region, will provide better situational awareness and allow more effective responses.

An important requirement for TCWI in the RTBP project is to achieve high level of intuitiveness of its operation, even by users without a lot of specialized training in statistical disease surveillance. Not all health officials in Sri Lanka are familiar or could remember the statistical methods implemented in TCWI for detection analysis. Therefore, the full potential of TCWI may not be exploited initially.

Health officials in North Western Province are seeing the advantages of the real-time data analysis capabilities. Face-to-face interviews with these institution heads in Kurunegala district concluded that MOH, PHI, Regional Epidemiology Staff, and Planning/Programming Assistant would be the resource persons with TCWI training, a gradual and iterative process. Eventually the interface, technical documentation, and training regime, will converge to an end state of feasible level of practical capitalization on the potential of the implemented technology.
Data entry error by one of the Suwacevo members, instead of Worm Infestation, accidentally, sent 24 cases of Whooping cough. Analysts were able to detect this notifiable disease the very next day but later confirmed to be a false alarm since Whooping Cough has been eradicated in Sri Lanka for the last few decades. However, this exercise was a rewarding lesson to learn that T-Cube was able to pick out that abnormal incident in a very short time span. Other low priority incidence of Respiratory Tract Infection, and Acute Diarrhoea were also detected and also to be confirmed as true positives.

The studies of utility also involve designing the procedures for scheduling TCWI event detection algorithms to execute automatically in order to periodically generate a set of alerts which then can be communicated to the appropriate recipients. These alerts will include a list of current findings, their time spans, affected locations and probabilities or significance scores and they will be disseminated to targeted officials via SMS and Email. These officials upon receiving the alerts would use TCWI to further verify the validity of these potentially adverse events, and to make decisions regarding any necessary response activity required to mitigate the situation.

In order to validate the accuracy of the implemented statistical algorithms, RTBP team has extracted publicly available Weekly Epidemiological Returns published by Sri Lanka Epidemiology Unit, and used it to synthesize data of the shape and form identical to that being collected through mHealthSurvey. Analysing such data with TCWI, we have found reliable signals pertaining to leptospirosis and dengue fever events as early as two months ahead of time of their detection with the existing processes. In the case of dengue, in two occasions, the disease abnormal emerged during off season, once in August 2007 and once in April 2008.

Sahana Alerting Module
RTBP uses SMS, among other channels such as email and web postings, to alert the MOH, PHI, and MOIC, of any adverse events that are worthy of bringing to their attention. Such an all-media one-to-many text based alerting/situational-awareness technique is not used in Sri Lanka. Instead the present system uses voice telephony, facsimile, or the postal system, and in some instances the PHI makes a physical visit to MOH office to receive the notifications. Current practices, although cost ineffective and slow, are tolerable for low volume incidences; however, is inefficient for broadcasting to larger sets of health workers or health officials during a crisis. Moreover, the RTBP introduced alerting/situational awareness component is cost effective when compared with PHIs having to travel long distances to the MOH office for receiving instructions or consequences that are an affect of delayed communications.

Health officials see the benefits in the Sahana Alerting Module, which is structured to maintain recipient groups and reusable templates. The researchers through a series of administered face-to-face interviews have realized two templates: notifiable disease alerts/situational-awareness with operational guidelines and other communicable alerts/situational-awareness for informative purposes but does not require actions. This would increase the efficiencies and reduce costs in disseminating alerts/situational-awareness messages to targeted health workers and health officials. Relative to a voice call that is Rupees 6.00 – 10.00 and a text message is Rupees 0.50 – 1.00 depending on the length of the message.
Interviews with Suwadana Centre Community Health Workers in Sri Lanka indicate that they are unaware of local health risks and they usually get their information predominantly through the local Television channels, Radio stations or Newspapers, which they will come to know only if they tune in to those media. Word-of-mouth was another popular source for learning of health risk information. Whereas the health officials can incorporate community health workers such as the Suwadana Centre system by informing them through SMS alerts or situational awareness messages for public health education and preventive strategies.

This system, besides the day to day use of communicating localized alerts/situational awareness, can be adopted very easily during the time of a major wide spread epidemic; where conventional or existing communication channels are resource intensive and time consuming; i.e. Sahana Messaging can do one-to-many correspondences; where a single message can be disseminated to an entire province or district with the click of a button.

**Conclusion**

The health departments and health workers involved in the RTBP pilot see the benefits in the mHealthSurvey for real-time data collection, T-Cube Web Interface for near-real-time outbreak detection, and Sahana Alerting Module for real-time health risk information dissemination. Preliminary lessons to date indicate the need for more robust mobile application for data collection with complete ontology, semantics, and vocabulary in disease-syndrome for reduction of noise and increase of reliability in the datasets. More rigorous capacity building and frequent use is required for health officials to take advantage of the full potential of TCWI. The new paradigm of alerting and situational awareness opposed to the present practice of notification will be easily absorbed by the health system over time. Given that the system has been in preliminary use for over six months, it is anticipated that the usability issues will subside in time to come.

**Acknowledgements**

Authors wish to acknowledge the funding agencies: International Development Research Center (IDRC) of Canada for the grant (105130) in full support of the project titled – Evaluating a Real-Time Bio-surveillance Program. The authors are grateful for the enthusiastic participation and support of the Wayamba (North Western) Provincial Director of Health Services office and the Lanka Jathika Sarvodaya Society in Sri Lanka.

**References**


