Evaluating a Real-Time Biosurveillance Program: A Pilot Project

March 2008

LIRNE*asia*.NET

Learning Initiatives on Reforms for Network Economies

Submitted to IDRC

LIRNE*asia**
12 Balcombe Place,
Colombo 8 Sri Lanka
Tel: +94 11 452 7647 Fax: +94 11 452 7648
E-mail: samarajiva@lirne.net
Web site: www.lirneasia.net

LIRNEasia is a regional ICT policy and regulation capacity-building organization active across the Asia Pacific. Its mission is to "improve the lives of the people of the emerging Asia-Pacific by facilitating their use of ICTs and related infrastructures; by catalyzing the reform of laws, policies and regulations to enable those uses through the conduct of policy-relevant research, training and advocacy with emphasis on building in-situ expertise." Additional information is available at http://www.lirneasia.net.

> "Timely and complete statistics on disease incidence and prevalence help to facilitate rational and evidence-based health planning and allow resource allocation to be optimized. The geographic distribution of disease patterns is important management information which shows where specific measures or extra effort are needed. Notification of rapidly rising incidence of specific diseases, if received early can help to prevent their spread by alerting authorities to the need for increased preventive and treatment services. Emergencies can be better managed when their effects on disease pattern can be determined and in many cases the development of outbreaks can be prevented."[1]
>
> Dr. Khan Tun
> WHO representative to Sri Lanka

[1] The paragraph was extracted from the book titled "Surveillance Case Definition for Notifiable Diseases in Sri Lanka", Epidemiology Unit, Ministry of Health Sri Lanka 2005 -- http://www.epid.gov.lk/pdf/Final-Book.pdf

TABLE OF CONTENT

LIST OF TABLES

LIST OF FIGURES

1. ABSTRACT

"Can Biosurveillance Algorithms coupled with Pervasive Mobile Applications potentially be effective in the early detection of disease outbreaks?"

November 2006, Colombo: "Sri Lanka's health authorities were fighting to control a viral fever spreading in the country. A student of the Sri Lanka Law College died from the fever, and the Law College hostel was infected with the virus that has made over thirty students ill. Over one thousand viral fever patients had been reported from the Colombo, Mannar, Kalmunai, and Jaffna areas. Many students in these areas were unable to attend the national university entrance: Advance Level (A/L) exams as a result of being affected by the Mosquito born disease: Chickungunya viral fever." [2] This outbreak in Sri Lanka occurred immediately after India was hit by the Chickungunya viral illness.

Problem faced by the *Epidemiology-Unit*[3] Officials in the Region is receiving health information in a timely manner in order to prevent diseases reaching epidemic states as it has with the case of the Chickungunya viral fever, which is a communicable disease but does not require the Epidemiology Unit to be "notified". However, an increased number of any diseases must be notified. The current surveillance system does not provide the much needed "real-time" information flow and analysis to detect such an event of increased number of scattered cases. The real-time detection shortcomings can be easily overcome with reliable and robust Information Communication Technologies (ICTs) and Intelligent Software (SW).

Proposed Surveillance and Alerting Program will provide for the execution of decision analyses of the assessment and response problem faced by the State and Regional Epidemiology Unit personnel. Healthcare Worker[4] based Mobile Phone (MP) sensor system and deployment of advanced detection SW algorithms such as *spatio-temporal scanning*, *Bayesian modeling*, and *multivariate time series analysis* will provide the State/regional Epidemiology Units with the tools to combat the real-time detection and monitoring dilemma. Moreover, they program will couple governance and civil society through Wireless Local Loop (WLL) ICTs such as *mobile handheld* devices.

In the proposed Pilot Phase, Health-related information gathered through mobile handheld devices will be communicated through the WLL Networks to a central Database (DB) for analytics and monitoring[5]. The ICT system will use General Packet Radio Service (GPRS), Short-Message-Service (SMS) as transport technologies to enable access to SW applications to communicate information on surveillance and alerting. The technologies and SW applications will be assessed for suitability by measuring the performance in India and Sri Lanka.

---

[2] The FluTracker bulletin can be found here -- http://www.flutrackers.com/forum/showthread.php?t=12672

[3] The Epidemiology-Unit referred to in this proposal is the National Epidemiology Unit of Sri Lanka and the National Institute for Communicable Diseases in India, see Table 4, Appendix A

[4] The Healthcare-Worker referred to in the proposal is defined in Table 4, Appendix A.

[5] Mobile-forms (a.k.a openROSA consortium) is an open source consortium developing mobile applications (XForms) to gather health-related information; project plans to borrow and add-on to these freely available system -- http://code.dimagi.com/JavaRosa/

A Real-Time Biosurveillance Program (RTBP) will be developed around the ICT based surveillance and notification system. Pilot implementation in the two nations will be a research project aimed at evaluating the RTBP for a wider scale deployment in India and Sri Lanka with the possibility of extending the RTBP to the region.

2. KEYWORDS

Rapid-detection, infectious, disease, outbreak, bio-surveillance, public, e-health, disaster, policy, last-mile, mobile, wireless, information, communication, intelligent, detection, software, technology

3. OVERVIEW

"Problem that this program promises to solve is to strengthen existing disease surveillance and detection communication systems, reduce latencies in detecting and communicating disease information and set a standard interoperable protocol for disease information communication with National and International Health-related Organizations in the region"

This is a Multi-Partner[6] Pilot Project to be conducted over a period of 2 years where the research will be accomplished in 4 components – (I) Establish the Electronic Communication System, (II) Introduce Computer based detection system, (III) Implement the RTBP and (IV) Evaluate the Biosurveillance Program. Each component is to be assessed independently and cumulatively as the project matures.

Although it will not be part of this study to focus on regional interoperability and integration of e-Health systems, the project will keep these aspects in mind during when designing, implementing, and evaluating the proposed study.

The overall goal of the project is to improve the ability of India and Sri Lanka to collect and respond to clinical, spatial, and weather data to form an overall probabilistic assessment of threat, and combine this probabilistic output from the detection algorithms combined with the results of decision analyses to create an implemented response decision-making system for the Epidemiology Units in the National/State/Regional settings before the disease reaches a "tipping point".

Appendix B discusses the current disease surveillance and notification systems in both Sri Lanka and India.

[6] Partners – Ministry of Health State of Tamil Nadu (India), Ministry of Health, Epidemiology Unit – Government of Sri Lanka (www.epid.gov.lk) , Lanka Jatika Sarvodaya Shramadana Sangamaya (www.sarvodaya.org), National Center for Biological Sciences (www.ncbs.res.in), Indian Institute of Technology – Madras Rural Technology Business Incubator (www.iitm-rtbi.res.in) , Carnegie Mellon University Auton Laboratory (www.autonlab.org) , University of Alberta (www.extension.ualberta.ca/faculty/memb_gow.aspx) , Lanka Software Foundation (www.opensource.lk, and LIRNEasia (www.lirneasia.net)

4. RESEARCH PROBLEM

4.1. Need for Biosurveillance

There are already significant indications of disease burden occurring in the World as a result of climate change and population increase; there are many causative factors including infective agents like bacteria and viruses. What would happen if these micro organisms mutate and reappear in a different form in India or Sri Lanka? The disease can be carried into the country by two individuals who arrive from overseas but live in two different parts of the country. Is the present system able to distinctly identify the outliers with possibly similar symptoms reported by two different Health Workers in the different areas? If that is the case then it is important that the two cases with similar symptoms are centrally detected immediately before a mass contamination and spreading takes place.

Another significant risk factor that needs attention is the emergence of a novel and pandemic viruses; the frequent travel/migration of people to and from the areas which are currently affected by the avian influenza, which is known to be highly virulent and mutable virus. To add to the complexity the symptoms may be common enough to be considered unimportant and be neglected. Definitely, a human being does not have the capacity or diligence to search through all the hospital information strings to identify clusters of similar patterns in spatially distributed data sets.

"The latter part of 2007 and beginning of 2008 the Central Province of Sri Lanka, Anuradhapura District, was faced with a fever like disease. Although cases were reported, by over 100 people, the disease went unnoticed until it claimed 3 deaths. The abnormal disease surfaced during the rainy season when illnesses with fever are higher than normal." According to the report by "FluTracker"[7] each case had slightly alternate symptoms. The increase of incidents of flu during the rainy seasons is a normal experience in South Asian Region but incidents of Flu along with some cases with bleeding will not be a local irregularity but nationally scattered cases of flu along with bleeding and the rate of increase of national cases geospatially is a national anomaly.

4.2. Can Information Technology Help?

Statistical analysis of large datasets is time-consuming. A surveillance system must be ready to detect changes in complex, multivariate data very quickly (ideally, in the real time), while maintaining the ability to test a huge number of hypotheses regarding geospatial co-locations, temporal correlations of the individual cases and their demographic characteristics, in order to detect a possible outbreak. That requires advanced algorithms for detection of abnormal patterns which would efficiently handle large sets of multivariate data and effectively signal statistically significant departures from the expected. Advances in algorithms, data structures and artificial intelligence allow for practical applications of data-driven outbreak detection methods which can

[7] "Infectious disease with life threatening complications and at least 3 deaths reported", the report on the disease that threatened the Central Province of Sri Lanka can be found here -- http://www.flutrackers.com/forum/showthread.php?p=124572

handle the complexity of the task at hand by learning from examples in historical data and from real or simulated outbreaks recorded in it.

The socio-economists see the key problem is not software but accurate and timely entry of data by Healthcare-Workers. Therefore, the project will extend the user interfacing to the last-mile by using ICT networks that span the nation, opposed to waiting for broadband connectivity to be widely available; i.e. use WLL technologies that cover the last-mile. Belief is that the introduction of WLL applications opposed to traditional computer desktop applications with broadband connections will increase the early detection and warning of communicable diseases; the demand side ICT market study: *Teleuse on a Shoestring (two)*[8] reveals that coverage of the WLL market is far beyond that of the fixed line market in India and Sri Lanka. Adopting data acquisition software applications that work on MPs hopes to reduce the latencies in communicating "public health information".

In a hill-country village in Sri Lanka there may be an unusual number of people with hand, wrist, and arm injuries during the harvest season. Discoveries of this nature cannot be observed from traditional methods. Moreover, such cases are non communicable diseases and may not be reported at all. It requires some reliable reporting and intelligent detection systems to identify such anomalies within a huge database of health information. Detecting phenomenon of this nature by a Biomedical surveillance system, a discipline focused on finding solutions to prevent disasters and provide security to individuals, can collaborate within or with other national departments, such as the Agriculture department, to provide solutions and the necessary knowledge to the farmers to prevent such health hazards.

In both countries, the current paper based system does not feed the infectious disease information to central processing on a timely manner. It is evident from the *Weekly Epidemiology Report*[9] (WER) published by the Epidemiology Unit in Colombo is lagging by 3-4 weeks. It is not because the time taken to edit the information to fit the template, it is mainly because the IDCC in the Districts lag in processing the statistic and delivering timely information up the chain through the paper based system. In the handbook titled "Surveillance Case Definition for Notifiable Diseases in Sri Lanka"[10], section on "Notification System in Sri Lanka" explains the tedious processes associated with the prevention and notification processes.

[8] Teleuse on a Shoestring (2006 – 2007),
[9] The Sri Lanka Epidemiology Unit published weekly report can be found here -- http://www.epid.gov.lk/wer.htm
[10] The handbook can be found online here -- http://www.epid.gov.lk/pdf/Final-Book.pdf

5. GOALS AND OBJECTIVES

The proposed RTBP entails introduction of technology and processes to complement the existing disease surveillance and notification systems in the region

List of the objectives in point form

- Deploy mobile phone sensor system for gathering health-related information from the Community-Healthcare-Workers
- Train the Community-Healthcare-Workers on using the technology and adopting processes for submitting health-related information
- Evaluate the usability, adaptability, and effectiveness of the mobile phone based data acquisition process
- Deploy advanced detection software algorithms such as *spatio-temporal scanning*, *Bayesian modeling*, and *multivariate time series analysis* for statistical data mining for State/Regional Epidemiology Units to detect disease outbreaks
- Train the State/regional Epidemiology Units with the tools and processes to analyze the gathered health-related data to detect disease outbreaks in near real-time
- Evaluate the ability of the detection system to assist the National/State/regional epidemiology units with their task of discovering outbreaks a head of time, to efficiently carry out the analysis in a timely manner, reliably predicting disease outbreaks with minimal ambiguity, and interpreting the analyzed information with zero complexity.
- Deploy a disease outbreak notification software tool for State/Regional Epidemiology Units to use for notifying divisional and community Healthcare-Workers of a possible disease outbreak as well as monitor the situations with feedback reports on the response actions
- Train the State/Regional Epidemiology Units on the software tool and processes in notifying possible disease outbreak as well as instructions on protocols
- Evaluate the notification system for its usability, reliability, and effectiveness in managing communications with Healthcare workers during and emerging disease outbreak emergency situation
- Disseminate the outcome of the research for policy makers, practitioners, and researchers to study the lessons learned

5.1. Provide a Healthier life in the last-mile communities

The main goal is to provide the Healthcare-Workers in the last-mile (PHI in Sri Lanka and VHN in India) and Disease Control Centers (Regional Epidemiology Unit in Sri Lanka and NICD in India) a program based around ICT tools to submit health information in a timely manner in order to rapidly predict disease outbreaks.

The program will contribute to sustainable development through effective man-machine coexistence with an added benefit of freeing the intellectual minds of the health professionals through the introduction of ICTs. The State/Regional Epidemiology Units can give the professionals the time needed to tackle problems that cannot be solved by ICTs by getting the

ICTs to do the routine paper based work. Once freed-up, the Health officials can spend the excess time to focus on development of programs and enhancement of systems to give their communities a healthier life opposed to wasting time on meticulous processing and daily decision making.

Assume the scenario where there is a water born disease infecting the community water system. The current protocol in Sri Lanka requires that Public Health Inspector (PHI) give instructions to the community on preventive measures to avoid contamination of the infectious disease. Some communities are hard to reach and the households are distributed far apart, making the task of the PHI difficult to cover all the households and communities in a timely manner. However, if the PHI can provide proper instructions to a pre-identified responsible person in the community to carry out the propagation of instructions to the community households then the efficiency of the dissemination work is improved. There are Community-based Organization (CBO) Health programs being developed by Non Governmental Organizations (NGOs) in rural communities. These CBOs can be the focal point for the PHI to make the notifications. It is not necessary that the PHI physically approach the CBO to provide the instructions, the task can be achieved by either phone call, sending a text-message, or an email to the designated key person in that community.

Although the main outcome of the study is on productivity it will also closely look at the health outcomes related to human health such the ability to eradicate Malaria or mitigate Dengue in the two tropical countries the study is being conducted. There are other human health issues such as mental health diseases that are hidden in place like Sri Lanka and India due to lack of immunization at early childhood. The researchers anticipate reporting of all health-related information may surface such wide spread diseases that otherwise would go unnoticed or neglected. The program will not only focus on improving the productivity but also look into the ability of the system to assist with improving overall health in the nation.

5.2. Sequence of Project Components

Overall intention of the project is to develop the complete real-time detection and warning system, which comprises 4 Components:

(I) Establish the Electronic Communication System,
(II) Introduce Computer based detection and monitoring,
(III) Implement the RTBP
(IV) Evaluate the Biosurveillance Program, and

Component (I): Establish the Electronic Communication System

The biggest dilemma is the inadequacy of bandwidth and wire-line access to the internet that is affordable for the rural communities in India and Sri Lanka. However, GSM/CDMA networks are quite well established and cover over 70% of the respective nations. Therefore, WLL internet connectivity for accessing applications must be first proven to be robust. The project will ensure that the communities and Healthcare-Worker areas selected for the pilot study have reliable

connectivity to submit the data for analysis. Mobile application developing technology partner: *Indian Institute of Technology – Madras Rural Technology Business Incubator* (IITM-RTBI[11]) will provide the necessary data acquisition and data staging infrastructure. Handheld mobile SW applications that work both online and offline will be adopted to acquire the health-related information from the Last-Mile areas. All offline data will be transmitted to the central DB when connectivity is established; where as online data will be received by the central DB instantaneously. State/Regional Epidemiology Unit staff will be able to visualize the health information that is stored centrally. The initial silent-tests will evaluate the usability and adoptability of the ICT system in the RTBP by the Healthcare-Worker and the IDCC Staff.

"OpenROSA"[12] is a Free and Open Source Software (FOSS) consortium evolving around providing a framework for communicating health-related information using Java based (J2ME) mobile handheld applications. The software developed the OpenROSA consortium has only been lab tested but not field tested in a scale as proposed in this study. The project foresees that it will be necessary to enhance the readily available software before adopting in India and Sri Lanka. Prior to deploying the software with the Healthcare-Workers testing will be conducted in a smaller scale to validate the application through a series of silent-tests. Project will use the openROSA framework and set of available mobile applications to develop the necessary upstream (last-mile to datacenter) communications applications. The project will contribute to the openROSA initiative with the research findings and make recommendations to reform the the open source software.

Project will use an existing SMS based alerting application, which is another FOSS application made available by the *Sahana*[13] Disaster Management Messaging Module. *Lanka Software Foundation* (LSF) of Sri Lanka is the umbrella organization that provides the main infrastructure for the Sahana core software development team. The Sahana core development team will provide the basis to host and manage the source code of the RTBP, including the mobile applications. The software developed for the RTBP will also be made available as FOSS; namely under the *Sahana Biosurveillance Module*.

The Community-Healthcare-Workers (PHI in Sri Lanka and the VHN in Tamil Nadu) will be equipped with the handheld devices, software applications, and ICT network to supply health-related information. However, processes and protocols have to be put in place for the system to work properly; i.e. if the people (Healthcare-Worker) fail to engage then the upstream communication component will fail. For example, the PHI and VHN will use the MP to call each of the Health-related establishments in their designated area to gather the daily or weekly counts on the patient visits. Thereafter, the Community-Healthcare-Worker will submit the information to the central database via the MPs. One of the data acquisition applications: *Ping-Pong*, discussed in Appendix D, forces the Community-Healthcare-Workers to submit the daily or weekly information, which increases the reliability. Another application: *Enum-Cards*, also discussed in Appendix D, is a redundant application, in the event the J2ME application is

[11] The stakes and contributions of IITM-RTBI are addressed in section 8.2 – Organizations. Also see IITM-RTBI website. IITM is in the space of developing healthcare application that can be made available for rural populations such as the remote medical unit - http://www.tenet.res.in/Activities/Products/doc/medicalDiagnosticKit.php
[12] OpenROSA (a.k.a mobile-forms) consortium information can be found here – http://code.dimagi.com/JavaRosa/
[13] Information on "Sahana" FOSS "Messaging Module" can be found here – http://www.sahana.lk

accidentally removed or fails, then the user can use simple SMS and a short-code to send the daily health information.

Component (II): Introduce Computer based Detection and Monitoring

The analytical software will be implemented by the *Carnegie Mellon University Auton Lab*[14] (CMU-AutonLab), who will contribute to the design of the Health-related information acquisition database (Component I), if necessary, make recommendations to enhance the Data Acquisition applications to capture additional information necessary to fill the input requirements of the analytics algorithms. CMU-AutonLab will customize the existing software libraries already developed by them. These applications are being used predominantly in USA. The computer based analytical system will be tested, in these parts of the world, for reliability and effectiveness relative to the current paper based analytical systems; i.e. the ability to actually contribute to the detection and monitoring of diseases. CMU-AutonLab will also develop a *dashboard* (graphical-user-interface) for the State/Regional Epidemiology Units to visualize the analytics and run algorithms to fetch outliers to warn Health officials and/or the Public of probable outbreaks. Appendix D - Figure 5 shows the datacenter and housing of the analytics system components and Figures 6 and Figure 7 illustrate the software components and GUI of the computer based detection and monitoring system.

The intelligent software will use statistical data mining as a standard technique to auto-detect anomalies lying in the heath information produced by the Community-Healthcare-Worker. There are several types of ways of applying the analytics to geo-temporal, spatio-temporal, and demographic datasets. The outcome of the RTBP is to identify the necessary and sufficient set of algorithms and datasets for a comprehensive software based analytic system that requires minimal human intervention. The dashboard and database will be built in a way for all Healthcare-Worker and Researchers interested in running cross tabs to investigate correlations in health-related data. Thereby, other areas of Public Health, besides disease outbreak detection, can use the system to identify new areas for research and development for prevention and cure.

As in Component (I), the project will study the gaps through simulated trials and through usability testing done through the Healthcare-Workers. If necessary, new processes will be developed to fully operationalise the computer based detection and monitoring ICT system.

Component (III): Implementation of RTBP

The RTB is two fold: an ICT system to assist in the near real-time detection of disease outbreaks and a program created around the system with introduced processes to operationalize the system in a national public healthcare system. The actors comprise Healthcare Workers and Public who will use ICTs to maximize health security.

[14] Auton Lab has extensive experience in statistical data mining and providing solutions for disease surveillance. They have working solution in USA. They've contributed to the popular Real-Time Outbreak and Disease Surveillance (RODS) open source project.

The implementation process is governed by a participatory design process including the stakeholders: healthcare workers, public, experts and technology solution providers. At the very start all the partners (stakeholder groups) will meet to discuss and commit to their roles and responsibilities as well as assess the practicality of the project deliverable.

Two Research Assistants will be hired by the Indian and Sri Lankan epidemiology units. The Research Assistants will participate in all project activities including carrying out the necessary research activities to facilitate the forward progress of the design, development, implementation, and evaluation process of the research. In addition, the research assistants will document the outcomes of the research findings with the supervision of the senior researchers with the aim of disseminating the results.

Part of the research involves evaluating the RTBP in 16 community-healthcare-worker areas with another 16 acting as control groups. The combined 32 community-healthcare-workers will be associated with 32 villages with the presence of a Community-based Organization (CBO) in those villages. The national epidemiology unit staff and the CBO will collaborate on selecting 32 community-healthcare-worker/villages to participate in the pilot.

The first action of the ICT system design process involves the technology providers and healthcare-workers collaborating on developing a set of User Requirements specifications (URS). Upon completion of the URS the technology providers will demonstrate the prototypes to the healthcare-workers at a planning workshop to ensure that the URS have been captured and suits the healthcare worker requirements. The technology providers will then take the URS and the lessons learned from the demonstrated prototypes as a reference to develop a set of system requirements (or software requirement specifications – SRS). The SRS will be validated by the healthcare-workers and experts before actual development of the system can begin.

Once the technology providers finish developing the individual components they will be tested by the healthcare-workers and the other researchers. Upon acceptance the system components will be integrated. The technology providers will develop a set of User Manuals and the researchers will develop a set of Standard Operational Procedures (SOP). Through a workshop the user manuals and SOPs will be introduced to the healthcare-workers as part of a training exercise. The project manager and the research assistants will repeat the same training exercise onsite in each of the community-healthcare-worker areas and State/Regional epidemiology. The redundant onsite training session will be followed by a series of silent-tests involving the healthcare-workers to ensure they can operate the technology as well as adhere to the newly introduced processes.

A similar approach of centralize workshop based and onsite training will be given to the CBO villages participating in the research. They will also run through a round of silent-tests to be familiar with the exercises and protocols they need to carry out during the evaluation period.

Component (IV): Evaluation of Biosurveillance Program

The evaluation process is twofold: first is evaluating the project itself as to understand the feasibility of achieving the goals, second is evaluating the RTBP as to the performance and

outcomes of the program. The evaluation method is described in section 6.1, while the questions governing the evaluation are discussed in section 5.3.

Periodic evaluation through meetings, online, offline, and face-to-face, will be conducted to help the principal and agents better understand the project; how it is impacting the participants and target groups; and how it could strengthen the project to achieve its goals while it is underway. The team will meet monthly with its partners and external research collaborators via e-mail, messaging or audio/video conferencing over the Internet. The purpose of the meetings will be to provide an objective assessment from the project's point of view of the progress and achievements of the project during the preceding month, and, in particular, to identify any improvements and adjustments that should be considered for the subsequent periods. At these meetings, the schedule of agreed upon deliverables will be reviewed in order to ensure they are being followed by the project partners.

Although part of the project entails introduction of e-Health for health information gathering, detection of disease outbreaks and disease outbreak notification there is more to the project that involves:

- Understanding the social implications of introducing e-Health to an already functioning infectious disease surveillance and notification system such as the Healthcare-Workers rejecting the newly introduced technology based system
- The feasibility of replacing and sustaining the workable components nation wide
- Costs and benefits of an e-Health based RTBP versus the existing paper and legacy (Voice/Fax) ICT based program
- Existence of capacity and the capability of the national epidemiology unit to sustain the proposed RTBP
- Identify gaps in the current system to determine how e-Health can add value to improve the current system as well as complement other e-Health initiatives
- Does the proposed program complement the ongoing initiatives curative and preventive programs
- Look at the other health system outcomes such as better disbursement of public health funds or medicine

The project will not use a trial and error approach rather an evaluation of a selected sample where a set of Healthcare-Worker will be given the new tools and a control sample is not given the tools. The sample set will be selected based on a set of criteria and not a randomized sample. Evaluation process is twofold:

- Performance of the RTBP on the basis of reliability and effective before and after the RTBP
- Evidence-based outcomes of the RTBP on the basis impacts and benefits before and after the RTBP

The actual statistical methodology to assess the performance and quality assurance of the observed variation such as applying Kappa Statistical methods will be determined by the researchers during the research.

The evaluation will also be subject to *real-situations* as well as a series of *mock-drills*. The evaluation period will be at least 1 year. The real situation will depend on the normal daily activities. The project will examine the samples on a weekly basis to learn performance issues

and outcome of activities associated with the RTBP. The mock drills will be a series of randomized events that will be conducted over the evaluation period; where a hypothetical case of symptoms will be reported in various location with known disease spread patter to gauge whether the system can identify the disease before it spreads beyond a tipping point. The mock drills are further segregated as *silent-tests* and *live-exercises*. The silent-tests are simply testing each process but without executing the dependent processes. The live-exercises will entail evaluating the entire end-to-end set of processes of the program.

Evaluation will be targeted in four larger Healthcare government geographic areas (MOH Divisions in Sri Lanka and DDH Districts areas in India) where two geographic areas will have access to the RTBP and the other two Divisions will act as control sample of the research. All four geographic areas will be subjected to the real-situation and mock-drill evaluations.

The lessons learned from the Pilot study will be used to discuss the necessary reform required to policy in infectious disease surveillance and notification. Specifically look at the long term implications and sustainability of the RTBP in a wider scale implementation. RTBP will also be a catalyst to understanding the policy for interoperability of disease information between international domains.

The project is viewed as an opportunity to test methods for improving the effectiveness of reporting public health information but also as a demonstration project for an advanced workshop on policy implications for e-Health based reporting in the region. Issues to be considered in this phase include: local and regional governance, management of health-related ICT systems, gender issues, privacy protection, international cooperation and data sharing.

As part of a strategy to influence and engage in policy a final workshop will be arranged for all research partners and participants to come together to discuss the outcome of the research and the strengths and weaknesses of the proposed system. The lessons learned along with the policy implications will be disseminated to the public through a media event.

5.3. Specific Objectives

As in evaluating any program, we are keen in scrutinizing the stability of the RTBP in the given environments. Specific objectives of the project are –

1) Evaluating the effectiveness of the e-Health RTBP for detecting outbreaks
2) Evaluating the latencies of communicating disease information
3) Contribution of community organization and gender participation
4) Developing a Toolkit for assessing e-Health RTBPs

1) Evaluating the effectiveness of the e-Health RTBP for detecting outbreaks

This objective entails assessing the technology and the processes associated with the semi automated ICT based RTBP for outbreak detection. The two ICT components are the mobile handheld last-mile technology for communicating health information and centralized database with analytics algorithms for detecting outbreaks. Effectiveness will be measured by observing the parameters –

(i) *Geographical coverage of surveillance and notification with the use of last-mile ICTs* - An important aim for the study is to address the rural-urban distinction associated with the penetration of communications technology. Disparity in access to communications media in the rural-urban divide is a significant obstacle to achieving effective surveillance and notification information in the last mile of the RTBP. Access to resources and training is also split along this rural-urban divide; where effectiveness can be evaluated before the system is put in place and measure against after the system is in place.

2) Minimizing latencies of communicating disease surveillance information

Latency is a measure of efficiency. In process analysis, latencies are also measured with respect to the mean time to failure; i.e. what is the probability of the process failing. In this regard the processes involving the actors and the sequence of functions will be assessed to evaluate the ability of the proposed RTBP to perform faster and more accurately than the current disease surveillance and notification system. The lessons learned will be used to make recommendations for process-reengineering of both the RTBP and/or the current system. Basic questions related to this part of the assessment are -- does the surveillance data get communicated in the same or better organized manner? Do the Epidemiology Unit staff members receive valuable information? Does visualizing the central electronic data make Epidemiology Unit staffs job efficient?

Health-related information will be communicated by the Community-Healthcare-Worker in two ways:

(i) *Received case reports* – The current protocol requires that the Community-Healthcare-Worker investigate reported communicable disease and report the investigated case to the area Public Health Officer (MOH in Sri Lanka and DDH in India). If applicable, the case is reported to the National level through the State/Regional Epidemiology Unit. The Community-Healthcare-Worker can use the introduced MP with embedded SW to report the case immediately to the central National DB; instantaneously faster than the paper system propagates up the chain.

(ii) *Gathered case reports* – The Community-Healthcare-Worker (PHI in Sri Lanka and VHN in India) will use the MP to call (voice) each of the Healthcare facilities and Healthcare-Worker both government and private to obtain a daily report of the patient "counts" (cases with common symptoms and the number of cases). Thereafter, they will

consolidate the collected data and submit the information to the central database through the introduced MP and software application.

The proposed processes force the Community-Healthcare-Worker to submit all Health-related information to the National DB, which in turn will provide a good baseline for patient information. Confidentiality is preserved since the electronic information submitted is simply patient counts without actual names of people. Since it is the Community-Healthcare-Worker, a government appointed official, submitting the data, it can be assumed to be authentic and accountable. Moreover, if the set of data submitted by the Community-Healthcare-Worker proves to be efficient and effective, then implementation of the wider scale National system can be scaled for all Authorized Healthcare Providers to directly submit the information of cases, as in (ii), followed by the area Community-Healthcare-Worker investigating the case and authenticating the report online. It is further expected to provide a set of health-related information advocate further research and development, as in 3).

3) Contribution of Community organization and gender participation

The focus is to build a program based on the concept of "health care by the people for the people".

(i) *Community-based Healthcare* – Although private hospitals and private Healthcare Practitioners exist, Health services in a developing country such as India and Sri Lanka are, predominantly, a public service; where the government subsidizes one hundred percent of the government run healthcare services. Given the limitation of resources it is difficult for the state to provide the best globally available services to all the communities. There is also an urban and rural divide in accessibility to state of the art healthcare services. Therefore, the communities have taken initiatives to help themselves through CBOs that facilitate services such as *primary healthcare* services. This reduces the burden imposed on the state to commit resources to health services that can easily be facilitated by the Community themselves as well as reduce the costs of communities without state of the art facilities from having to travel long distances to cities to ascertain these services. Project will investigate whether the rural underprivileged communities can be provided an unbiased service by coupling governance and communities through ICTs.

(ii) *Advocate other e-Health services* - The project will evaluate whether the RTBP can be a catalyst to improving personalized health systems and services in the community such assisting the national initiative in managing disease outbreaks. Other services that have potential for improvement are remote patient monitoring, tele-consultation, tele-care, tele-medicine, and tele-radiology services.

(iii) *Influence national health statistic* - The national health system has the potential of acquiring a secondary set of primary healthcare information from the Healthcare CBOs. This gives the opportunity to improve the baseline data for detection and monitoring as well. The project will investigate the possibilities of integrating community level data to improve the national level database. This will also provide the means to learn the type of

information that the Health CBOs would like to see the National Health department supply downstream to the communities.

(iv) *Feed into other Community Healthcare services* - Family Health Bureau[15] (FHB) in Sri Lanka and the National Rural Health Mission[16] (NRHM) in India are predominantly geared to handle women's health-issues; especially *Midwives* who guide mothers' through pre and post natal care. Preliminary results of the *HazInfo*[17] project proved active participation of women in the community in Disaster Management projects. This is predominantly because the Men go out to work. One of the objectives is to evaluate the options of providing training to Midwifes and/or other Community advocated Healthcare Providers to address other home care issues besides child birth and maternal care. Although family healthcare services are not directly affiliated with the Government Epidemiology programs, the project will like to see how other Healthcare programs at the community level can benefit from e-Health initiatives.

4) Developing a Toolkit to assess e-Health RTBPs

The World Health Assembly adopted a new International Health Regulation (IHR) that came in to force on 2007 in 193 member countries of the World Health Organization (WHO). The goal of the IHR is to prevent the international spread of emerging infections. One of the core requirements of the IHR is for member countries to comply with the strengthening of core capacity to detect, report, and respond rapidly to public health events. There is not one right solution or strategy for what works for all countries equally. The unique features of each country require unique solutions. However, the overall outcome of the solution should adhere to the goals of IHR.

The purpose of the proposed toolkit is to provide local authorities with information that will allow them to evaluate the many options available to provide surveillance and notification programs. Based on *Operations Research* Analysis methods, this toolkit will be developed specifically with infectious disease detection, reporting, and response in mind; it may also be applicable to many other programs that could threaten the Health emergencies in a state. The objectives of this exercise will be to devise a set of methodologies and identify a set of expected (or tolerable) outcomes and results for any RTBP to evaluate the respective program. Methodology framework will introduced ways to setup test scenarios, indicators to measure, and identify key nodes in the program to monitor. The aim of the assessment results is to reveal the strengths and weaknesses of the system and processes of the program as well as specifically identify the points where change/improvement is needed. In the event the Toolkit is applied to more than one program that are different then the scoring scheme, based on a probabilistic indicators, will provide means to identify which of the RTBPs are best suited for the particular environment (i.e. rank each of the RTBPs on a scale of 0 to 1). The Toolkit will be piloted in this project, which involves two test cases: India and Sri Lanka. The project will fine tune the RTBP assessment Toolkit based on the findings from the two test cases.

[15] Family Health Bureau of Sri Lanka website can be found here -- http://www.familyhealth.gov.lk/home.php
[16] National Rural Health Mission website can be found here -- http://mohfw.nic.in/NRHM.htm
[17] Details of the LIRNEasia conducted HazInfo project can be found here -- http://www.lirneasia.net/projects/current-projects/evaluating-last-mile-hazard-information-dissemination-hazinfo/

6. PROJECT DESIGN AND IMPLEMENTATION PLAN

The RTBP is based on the concept of a *closed system*; system with feedback and error-correction. The input to the system is the Health CBO learning to take proper actions in reporting illnesses to the state medical facilities as well as learn of methods of proper follow-up procedures. The proposed program has envisaged the three elements to fulfill the input, output, and feedback components of the closed system.

This research project is also a comparative study of implementing a RTBP between the two South Asia countries: India and Sri Lanka. The research in India will be carried out in the state of Tamil Nadu, which has a similar environment and climate as Sri Lanka. While the project is still in its proposal stage, although the countries will remain the same, the actual areas: Districts and Divisions in each country where the trials are to be carried out will be determined by the partners, mainly the Government Health Ministry and Community-based Organization in each country.

6.1. Research Methodology

Evidence-based Assessment

The proposed RTBP is being conducted as a pilot rather than an implementation, it is necessary to deploy the RTBP in sufficient number of Government Healthcare-Worker Division to evaluate the performance and outcomes in different geographical, infrastructural and socioeconomic contexts. Proposed research design will allow us to compare the performance and outcome of the RTBP deployed in different Government Healthcare-Worker Divisional areas and in the absence or presence of the proposed RTBP. The control Divisional areas will shed light on the extent to which having a dedicated RTBP for surveillance contribute to effectively communication of infectious diseases and prevention of outbreak. This pilot will involve communities and CBOs embedded in those communities to help with the subjective assessments.

Setup of the Sample Space

The ‘last-mile’ in this Pilot will be confined to the Community-Healthcare-Worker Areas in the Divisional areas. Each Community-Healthcare-Worker Area is linked to a State/Regional Epidemiology Unit. A survey will be conducted among -Workers to ascertain their degree of organizational development and the infrastructure that is available before selecting the pilot areas. Based on the survey data, the sample of 4 Government Healthcare Divisions will be selected that reflect the diversity of the State or Nation.

Based on the population density in the area, the Division based convenience sample, shown in table 1, may occupy from 3 to 10 Community-Healthcare-Workers. The selection process will ensure that there is a uniform distribution of Community-Healthcare-Worker among each of the Divisions such that 4 Community-Healthcare-Workers are linked to each Division. The RTBP will be deployed in 2 Divisions where the remaining 2 Divisions will not be exposed to the

RTBP and shall be used as control groups. Altogether 16 Community-Healthcare-Worker (8 with RTBP and 8 Control) will participate in the pilot, see Table 1. The State/Regional level Epidemiology Units belonging to the group exposed to the RTBP will be provided with a computer to visualize the data and perform the analytics for their governing Divisions.

The 8 Community-Healthcare-Worker in the two Divisions exposure to the RTBP will receive a mobile handheld with a built in application as well as have access to normal modes to communicate health information. All Health Officials in Indian and Sri Lanka exchange official communications in English. Therefore, the upstream health-related information will be communicated in English. The Community-Healthcare-Workers in the Control Divisions will practice the existing method of communicating health information.

The software applications deployed in the handhelds will capture simple environmental attributes, which are features such as the season and the day of week that cause trends in the data, and response attributes, which are the remaining features such as syndrome, gender, and age. Project will provide training to the Community-Healthcare-Workers and State/Regional level Epidemiology Unit staff participating in the pilot; especially the Healthcare-Worker exposed to the RTBP.

Every Medical Practitioner in the selected Division attending on a patient suffering from a disease will immediately notify such a case to the Community-Healthcare-Worker of the respective area. This notification to the Community-Healthcare-Worker may even be made by the principal of a school (in case of a student) or even the Chief Occupant of a house, see Figure 4 in Appendix C for the state transitions.

In order to develop the baseline for statistical data mining, at eased two divisions must be employed to acquire sufficient data. Also, at least two divisions are required to cover a larger population and geographical area. Otherwise, there will be very little variation in the data.

Table 1: Matrix with 4 Divisions, 16 Community-Healthcare-Workers, and 32 Communities

	Exposed to RTBP				**Unexposed to RTBP**			
	Division 1		**Division 2**		**Division 3**		**Division 4**	
+	C01	C05	C09	C13	C17	C21	C25	C29
	H01	H03	H05	H07	H09	H11	H13	H15
-	C02	C06	C10	C14	C18	C22	C26	C30
+	C03	C07	C11	C15	C19	C23	C27	C31
	H02	H04	H06	H08	H10	H12	H14	H16
-	C04	C08	C12	C16	C20	C24	C28	C32

Note 1 In Table 1, H – denotes Community-Healthcare-Worker and C denotes Community. The cells with "magenta" background are the Organized Communities with the presence of a Community-based Healthcare facility (+) and the cells with "orange" background are the Communities that do not have a formal Community-based Healthcare facility (-). Basically each

Community-Healthcare-Worker ("yellow" cells) will cover an Organized Community and a Less Organized Community.

Each Community-Healthcare-Worker Area covers approximately 2 to 4 villages (communities). Therefore, the pilot will require that a total of 32 communities covering the 4 Divisions (i.e. 16 Community-Healthcare-Worker Areas) participate in the mock-drills. For the purpose of the mock-drills the pilot will partner with a NGO that has an embedded CBO in the selected communities of the respective Divisions. The selected communities will act as the "public" in this proposed research. Half (16) of the selected communities will be required to have a formal Non-governmental Community-based Healthcare Program (such as maternity care, child care, primary healthcare) where the other 32 of the communities will not have such a Community-based Health program, see Table 1. This disparity of communities with organized healthcare and absence of community healthcare will be use to compare the contribution of community organizational level to local and national healthcare; in this event mainly on prevention of disease outbreaks.

Evaluation process

The project will use members of the community to initiate mock illnesses and follow the normal procedure of consulting a health official. The pilot will assess the performance and outcome of the RTBP based on the simulated activities. In the case a particular Community-Healthcare-Worker Area does not have sufficient communities then neighboring participating community affiliated to the NGO will fill the void. The mock-drills involving Communities will be carried out only four times during the entire 1 year evaluation period; i.e. 1 mock-drill per quarter to reflect the season trends in south Asia; i.e. rainy season and dry season. The 2 mock-drills will cover the Southwest and Northeast monsoon seasons because most mosquito born diseases and flu like diseases surface during this time of the year.

The mock-drills will use a computer simulation to determine the spread of a predetermined diseases or variation of symptoms. The Communities will be asked to report the disease based on the geo-temporal frequency of the event prescribed by the computer simulation results. Although 32 Communities are recruited for the pilot not all of them will be required to report a case because the computer simulation will give a more realistic scenario where it will only show a few communities being infected with the disease. The reason to use real people is to understand the pragmatic social problems associated with the proposed RTBP. The mock-drills will also simulate both a fast and slow spreading infectious disease.

The RTBP will also collect daily reports from the entire state or country because the initial baseline information will require the national level data. For example the Sri Lanka Epidemiology Unit compiles a weekly report of identified diseases. The Research Assistants employed through the project will have to take on the responsibility of feeding this information to the database. The project will also use this information throughout the evaluation period to assess the RTBP.

A Community Team Leader will be appointed to each community; or one Community Team Leader to oversee a maximum of 4 Communities. They will be asked to randomly select 25 Households in their communities. Once a week the Community Team Leader will visit the selected households to inquire of any illnesses they had experienced in the family and whether they had sought Medical assistance. Instead of randomly sampling, the Community Team Leader may inform the entire community about the project and request the community households to report all the cases to him or her. The Community Team Leader will record this information. Thereafter, we will check the ICT system against this data to investigate whether the data had been communicated to the database. This assessment will be carried out over the one year evaluation period.

The criteria for evaluation will include the rapidity at which an action is taken after receipt of information, the organizational efficiency in coordinating collective action, the choice of appropriate response to the particular incident, the speed at which response plans are carried out, the proper functioning of the ICT and related technologies among others. In an abstract sense these will related to the assessment of the inputs, assessment of the activities/processes, assessment of the outputs, Impact assessment, and the benefits/changes to overall epidemiology system. The basis for the assessments will be guided by the 6 hypothesis below.

Hypotheses

1. Healthcare Workers in Divisions 1 & 2 exposed to the RTBP will respond more effectively to communicating disease to the respective Epidemiology Center than the Healthcare Workers in the Control Divisions 3 & 4 unexposed to the RTBP.

2. Epidemiology Units in Divisions 1 & 2 exposed to the RTBP will detect disease outbreaks accurately and contain the outbreak efficiently than Epidemiology Units in Control Divisions 3 & 4 unexposed to the RTBP.

3. Healthcare Workers and Epidemiology Units in Divisions1 & 2 exposed to the RTBP will show interest and recognize the benefits in adopting e-Health programs opposed to the Healthcare Workers and Epidemiology Units in the Control Divisions 3 & 4 unexposed to the RTBP.

4. Communities in Divisions 1 & 2 exposed to the RTBP will have confidence in the National Disease Surveillance and Notification programs more than the Communities in the Control Divisions 3 & 4 unexposed to the RTBP.

5. Healthcare Workers and Epidemiology Units in Divisions 1 & 2 exposed to the RTBP in addition to their RTBP function will leverage ICTs in other areas to enrich their daily activities more than the Healthcare Workers and Epidemiology Units in the Control Divisions 3 & 4 unexposed to the RTBP.

6. Communities that have non governmental Community-based Healthcare Organizations will perform better in monitoring, communicating, and containing disease outbreaks than communities that do not have a formal non governmental Community-based Healthcare Organization.

7. LIMITATIONS AND ANTICIPATED PROBLEMS

Although action will be taken if and when appropriate to mitigate the limitations and anticipated problems, it is best that they are addressed in the proposal in order to better plan the project.

- It is possible for the Government Healthcare-Worker will reject participating in the pilot project because they do not see any personal incentive to contribute. People are often not interested in transforming in to using new systems.

- Community-Healthcare-Worker does not have the capacity to engage in data entry and is uncooperative in providing timely accurate information, which may also be as a result of complexities in the applications. The transition from the old ways of doing things to the new ways may take a lot longer period as a result will not give the project enough time to fully evaluate the system.

- One or more of the partners breaching the contract will make the research fall short in evaluating an end-to-end set of processes. As a result may even affect the outcomes of the dependent processes of the RTBP.

- Either the Indian or Sri Lankan evaluation is not fully completed to make the evaluation comparison between two countries. Similarly if one of the countries doesn't fully cooperate then the Interoperability evaluations will also be crippled.

- The primary phase of the project will not consider the pharmaceutical and only use the non-pharmaceutical public health information. It is uncertain as to whether excluding the mentioned set of data will hinder the performance of the system, since some patients tend to seek self care before or without visiting a Healthcare Provider.

- There are no violations of the ethical codes of research since the analytics is based on quantitative information; therefore, does not reveal any personal information of patient information. Validity of the data generated by the research for analysis can be questioned in the absence of a person to associate with the data record.

- The existing surveillance and notification system in the country may not have the sufficient parameters to acquire the full set of national level data to train the algorithms and set the baseline. One such parameter is the season where South Asia has only the rainy and dry season opposed to the most other countries with the standard four seasons. The Healthcare-Workers may be reluctant to add new formal information and be susceptible to change.

- Project is not requesting financial support for dissemination activities for publications, attending symposiums/conferences until such time the project has accomplished Components (I), (II), and (III). Upon successful completion, if seen necessary, the project will request for financial support to engage in regional dissemination activities to share the knowledge of the research findings. There is a possibility that the donor may reject a supplementary proposal due to lack of funds or changing of priorities.

8. ACTIVITIES AND BUDGET

8.1. Project Plan

Project Launch – partner meeting to finalize research, recruit project staff and procure project staff equipment; *Design System* – Health Worker planning meeting to gather system requirements and develop a set of system requirement specifications; *Release Beta* – preliminary version of the ICT system for Health Workers to test; *Training* – workshops for Health Workers and Community members to provide the know how to operate system and engage in RTBP; *Go Live* – fully implement ICT system and setup RTBP with standard operational procedures to begin evaluations; *Analysis* – discuss the research findings and draw conclusions; *Disseminations* – web postings/forums, journals publications and conference proceedings.

Table 3 and 4 in Appendix E describes the project tasks and weekly schedule in detail with resource allocations; where Figure 1 below illustrates a summary of the tasks, timelines, and resource allocations.

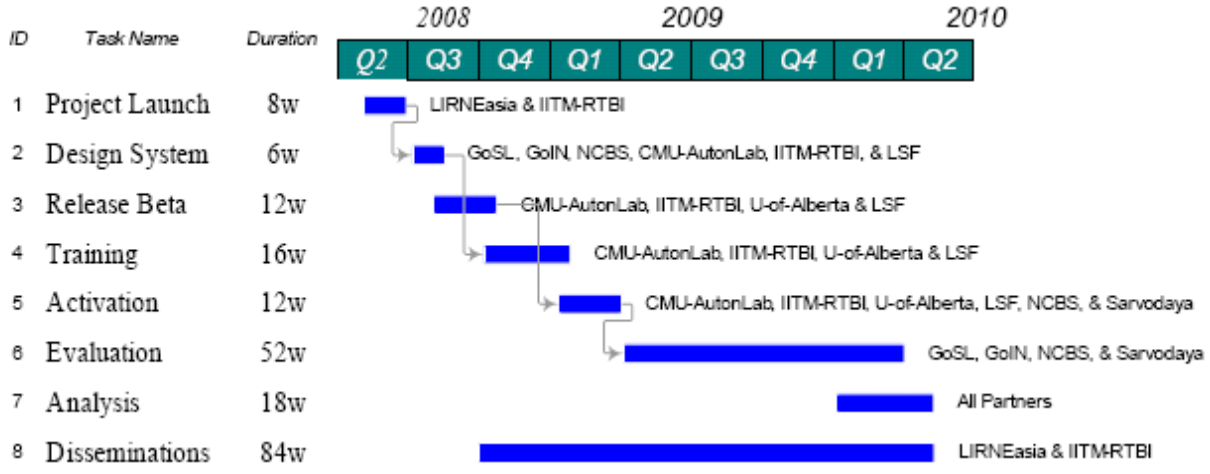

Figure 1: Gantt chart of Tasks, timelines, and resources over the 2 year project period

9. ORGANIZATIONAL CAPACITY

National Institute of Communicable Diseases, Ministry of Health and Family Welfare, Tamil Nadu, India (http://www.health.nic.in/NRHM/State%20Files/tamilnadu.htm)
The National Institute of Communicable Disease (NICD) was established on July 30, 1963, subsequent to the decision of the Government of India to expand and reorganize the activities of the Malaria Institute of India (MII) which remained in existence under different names since its inception in 1909. The objectives of the Institute broadly cover three activities viz., training, service and operational research in the field of communicable diseases and their prevention and control in the country. The broader objectives are – assist government with issues on prevention and control, initiate inquiries and undertake investigations in communicable diseases, assist state health authorities, undertake research on prevalence and spread of diseases, clinical investigation, and training.

Epidemiology Unit, Ministry of Health and Nutrition, Colombo, Sri Lanka (http://www.epid.gov.lk/) - The Epidemiological Unit was established in 1959 with assistance from World Health Organization (WHO) to strengthen surveillance of communicable diseases. Presently the Dr. Nihal Abeysinghe is the Chief Epidemiologist for the Nation and is assisted by Dr. Sumitha Ginige. In 1987, the Epidemiological Unit commenced computerization of surveillance data with the assistance of WHO. Major activities of the Epidemiology Unit are – disease surveillance, monitoring and evaluation of expanded programs on immunization, control of diarrhea, acute respiratory infections, dengue, Japanese encephalitis, and rubella, planning monitoring, and evaluation of new emerging and re-emerging diseases, provide feedback on selected communicable diseases and conduct research activities.

National Center for Biological Sciences, India (http://www.ncbs.res.in) – Research at NCBS (a part of the Tata Institute of Fundamental Research) uses experimental and computational approaches, a premier research institute with all the necessary facilities that a research scientist may need. Headed by Professor K. VijayRhavan (Director), the Institute's research interests of the faculty lay in the frontier areas of biology among which a strong involvement with Bioinformatics, both in development and in advanced usage. NCBS has experience in conducting research along side the government health institutions of India and has the capability to acquire the required support from the government to carry out the ground level activities. NCBS has the reach to the communities to advocate a participatory approach to research and development.

Carnegie Mellon University Auton Lab, USA (http://www.autonlab.org/) affiliated with The Robotics Institute at Carnegie Mellon University is a team of 20 highly skilled professionals (faculty, programmers and analysts, PhD students) led by Professors Artur Dubrawski and Jeff Schneider. The main topic of their research is applied statistical data mining that is scalable, accurate, reliable and fast. Recent deployments cover the areas of bio-surveillance, food safety, fleet health prognostics, astrophysics, homeland security, pharmaceutical, oil and machinery industries. Current and past sponsors include a number of the US government agencies (including – in the projects into public health surveillance - Centres for Disease Control and

Prevention, Department of Homeland Security, United States Department of Agriculture, Food and Drug Administration) and industrial partners including international and Fortune 100 companies. The Auton Lab has an extensive hands-on experience in designing and deploying data-driven methods of rapid detection of emerging patterns with applications to public health surveillance.

University of Alberta, Canada (http://www.extension.ualberta.ca/faculty/memb_gow.aspx) - Gordon Gow currently an Assistant Professor, Graduate Program in Communications and Technology (MACT), Faculty of Extension, University of Alberta, Edmonton, Alberta, Canada Currently, his research interests include development of WLL voice and data systems, especially with respect to regulatory concerns such as spectrum policy and management, telecom reform, technical standardization, public safety, and location based services. His research also looks at the development of electronic communications networks from a combined social and technical perspective, with the aim of expanding public understanding of and participation in policymaking. A primary concern is with the long term planning and management of critical infrastructure systems, especially those that support public alerting and emergency management activities. He participated in the development and testing of an all-hazards warning system in Sri Lanka, as well as sitting on the Common Alerting Protocol (CAP) Working Group for interoperability chaired by Industry Canada.

Lanka Software Foundation, Sri Lanka (http://www.sahana.lk) - Roshan Hewapathirana, MD, University of Colombo School of Computing, is a research fellow of the Lanka Software Foundation (LSF). He is a Medical Practitioner by profession with an M.Sc in Bioinformatics, a discipline under Information Technology. His contribution to the project will be in the capacity of an e-Health Expert. LSF is a Free and Open Source R&D Non-Profit organization with the objective of giving fellowships and internships to talented developers who would like to contribute to FOSS. Notable contributions from Lanka Software Foundation have been the Apache AXIS project, other Apache Web Services projects and the Sahana Disaster Management System, an open source research laboratory, attached to the University Of Colombo School Of Computing. LSF is experimenting with the concept of Community empowered disease monitoring for shared care, rapid diagnosis and outbreak detection in Pandemics: Extending Sahana framework as a social networking tool for disease Surveillance.

Sarvodaya, Sri Lanka (http://www.sarvodaya.org) **-** The Lanka Jatika Sarvodaya Shramadana Sangamaya (Sarvodaya) is Sri Lanka's largest and most broadly embedded people's organization, with a network covering: 15,000 communities, 345 divisional units, 34 Sarvodaya district offices; 10 specialist Development Education Institutes; built over 1000 Community Based Primary Healthcare Service Facilities, known as "Suwadana Centers", in Rural Sri Lanka. Dr. Vinya Ariyaratne, MD, Executive Director of Sarvodaya and Ravindra Kandage, Director of the Shanthi Sena Movement initiated the Suwadana Centers (mini-health centers) project, which is accordingly named as the Comprehensive Community Health Program (CCHP). The goal of the CCHP project is to address determinants of health rather than direct provision of health services.

Indian Institute of Technology – Madras, Rural Technology and Business Incubator, India (http://www.rtbi.in/index.html) - RTBI is a registered society established under the aegis of the Indian Institute of Technology, Madras and functions under the Department of Electrical Engineering, IIT Madras. RTBI's mission is to design, pilot and incubate business ventures with a specifically rural focus. Its vision is to facilitate rural-inclusive technology and business development in the rural space. Professor Ashok Jhunjhunwala is the Chair and Professor Timothy Gonsalves is a Director of RTBI, both faculty members. Currently, RTBI is engaged in the in the Health sectors and cares for the millions of rural people who are forced to suffer from preventable and curable illnesses because they do not have access to quality primary care; for women who die in childbirth and children who grow up in a web of malnutrition and disease. For patients of TB and malaria, victims of sudden accidents and injuries, and those who silently bear the stigma and burden of HIV and mental health problems.

LIRNEasia, Colombo, Sri Lanka (http://www.lirneasia.net) – Professor Rohan Samarajiva (Executive Director) leads team of Researchers at LIRNEasia, which is a regional information and communication technology (ICT) policy and regulation capacity-building organization active across the Asia Pacific. LIRNE*asia*'s program of actionable research seeks to identify the institutional constraints to effective use of ICTs to improve the lives of the people of the Asia Pacific, not simply in abstract terms but in country context, and to work collaboratively with multiple stakeholders to catalyze the changes conducive to greater participation by users and suppliers. LIRNE*asia*'s overall mission of capacity building seeks to contribute to building capacity for evidence-based intervention in the public-policy process by persons attuned to the specific national contexts within which policies are made and implemented. LIRNE*asia* partners with Research ICT Africa (RIA!), Regional Dialogue on the Information Society (DIRSI) and several European universities in LIRNE.NET and is a contributor to the World Dialogue on Regulation. It is a member of the Global Knowledge Partnership (GKP) and the Pacific Telecommunications Council.

10. APPENDIX A: TERMS, ABBREVIATIONS, AND ACRONYMS

Table 2: Abbreviations and Acronyms used in the proposal

CBO	Community-based Organization
DB	Database
DDH	Deputy Director of Health
DPH	Director of Public Health
Epidemiology Unit	Ministry of Health and Nutrition Epidemiology Unit of Sri Lanka
GPRS	General Packet Radio Service
GSM	Global System for Mobile communications
HW	Hardware
ICT	Information Communication Technology
IDCC	Infectious disease control centers
IDCC-N	Infectious Disease Control Center Nurse
MOH	Medical Officer of Health
MP	Mobile Phone
PHI	Public Health Inspector
RTBP	Real-Time Biosurveillance Program
SMS	Short Message Services
SMSC	SMS Controller
SW	Software
TCP/IP	Transmission Control Protocol/Internet Protocol
VHN	Village Health Nurse
WLL	Wireless Local Loop
NICD	National Institute of Communicable Diseases

Table 3: General terms used in the proposal to be consistent between in India and Sri Lanka

General Term used in proposal	Sri Lankan Term	Indian Term
Community-Healthcare-Worker	MO or PHI	MO or VHN
Health Worker	MO, PHI, or MOH, IDCC-N	MO, VHN, DDH, or DPH
Epidemiology Unit	National Epidemiology Unit	NICD
State/Regional Epidemiology Unit	Regional Epidemiology Unit	NICD Branch

11. APPENDIX B: CURRENT NATIONAL PROGRAMS

11.1. Sri Lanka Disease Surveillance and Notification System

The surveillance of *communicable disease*, in Sri Lanka, is based on the notification of certain diseases of priority through a Gazette notification. The Quarantine and Prevention of Diseases Ordinance of 1897 and its subsequent amendments provide the necessary legislation for the implementation of this system. The list of diseases to be notified includes the three diseases under the International Health Regulation and a second list which presently includes 25 diseases[18].

The National Epidemiology Unit is a department of the Ministry of Health and Nutrition of Sri Lanka. The unit is headed by a Chief Epidemiologist and supervises 9 Provincial Epidemiology Units and a series of Infectious Disease Control Centers in the districts belonging to the Provincial areas. The Government Health Ministry has further divided the Districts in to several Medical Officer of Health (MOH) Divisions such that each MOH covers approximately a population of 100,000. The MOH is responsible for the standard of the health in their area among, which disease surveillance notification is one of the tasks.

A Multi Disease Surveillance/Hospital Information System (MDS/HIS) is in place in a move to improve the disease Notification system in the Government Hospitals (GH) country-wide. The World Health Organization (WHO) in collaboration with the Government of Sri Lanka (GoSL) Ministry of Health Epidemiology Unit (Epidemiology Unit) has initiated a project on Multi Disease Surveillance / Hospital Information Systems in six selected hospitals in six Districts[19] of Sri Lanka. This computer based electronic information system was introduced in November 2005 at GH Ampara and GH Hambanthota. The system will soon be introduced to GH Karapitiya, GH Kalutara, BH Kalmunai and TH Batticoloa[20]. The MDH/HIS is purely a database to manage the bed tickets and inpatient/outpatient information but does not address the detection and monitoring of disease outbreaks as proposed in this paper. The proposed RTBP intends to extend the MDH/HIS functionality to assist the IDCC with rapid detection and monitoring of possible outbreaks as well as speed the timely reporting of infectious diseases.

A patient with an illness in a community may seek medical assistance from the Government Medical Office (MO), Private Medical Practitioner (PMP), and government hospital in the area. In most cases if it is a communicable disease, the MO or PMP may refer the patient to the nearest Government Hospital; else they would complete the required H544 form and submit it to the MOH. The Government Hospital would inform the MOH as well. Following the MOH informs the PHI to investigate the case and asks the PHI to produce a report to him/her. The MOH compiles a weekly report on the notifiable disease, which is submitted to the National Epidemiology Unit with a copy to the Regional Epidemiology Unit. If an IDCC Nurse is

[18] General Disease Surveillance, the Center for Prevention and Control of Diseases, Epidemiology Unit, Ministry of Health and Nutrition, Sri Lanka (http://www.epid.gov.lk/services.htm#surveillance)
[19] Sri Lanka's primary administrative divisions are 25 districts. The districts are subdivided into 324 "divisional secretary's divisions", The Department of Census and Statistics (DCS) of Sri Lanka, 2001.
[20] Programs currently running under WHO guidance, WHO, Sri Lanka

available in the hospital then they can communicate the disease to the National Epidemiology Unit as well. The report submitted on paper is then entered in to a computerized database at the National Epidemiology Unit to maintain statistics.

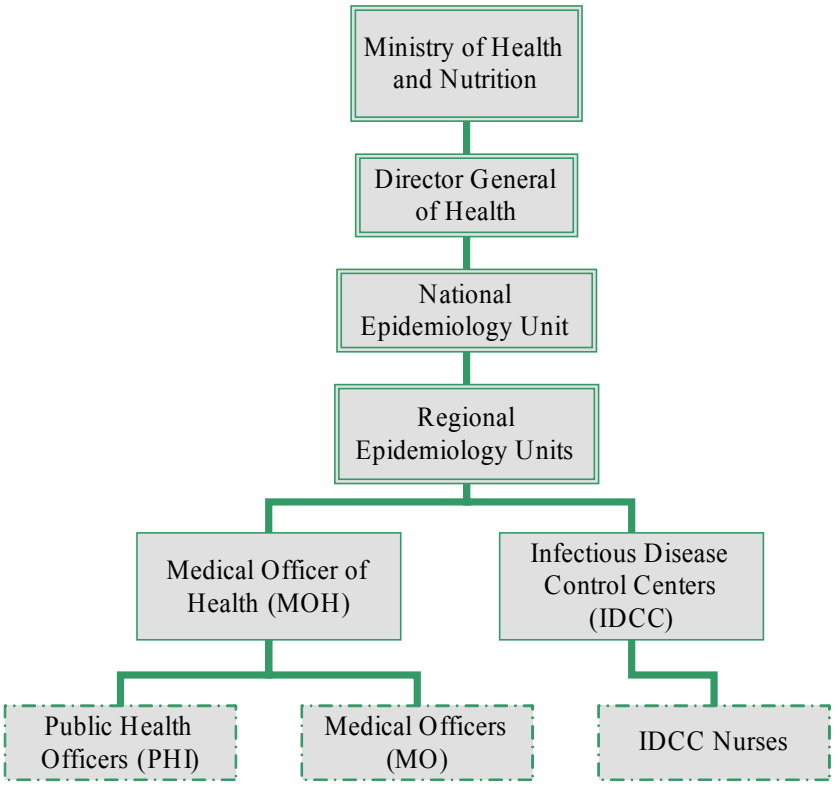

Figure 2: Sri Lanka government disease surveillance and notification organizational chart

11.2. Indian National Communicable Disease Control System

National Surveillance Program for Communicable Diseases (NSPCD) was initiated in 1998 as a pilot project under the hood of the National Institute for Communicable Diseases[21] (NICD), which is the body that supervises the districts and analyses the data for outbreaks in India. NICD was established on in 1963, to expand and reorganize the activities of the Malaria Institute of India (MII) which remained in existence under different names since its inception in 1909. The reorganized Institute was established to develop a national centre for teaching and research in various disciplines of epidemiology and control of communicable diseases. The Institute was envisaged to act as a centre par excellence for providing multi disciplinary and integrated expertise in the control of communicable disease. The Institute was also entrusted the task of developing reliable rapid economic epidemiological tools which could be effectively applied in the field for the control of communicable diseases. The experience from the pilot is subsequently being expanded to build the Integrated Disease Surveillance Program (ISDP) for India.

[21] A full description of the NIDC objectives are discussed here -- http://nicd.org/NICDObjectives.asp

National Surveillance Program for Communicable diseases has been launched to strengthen the disease surveillance system so that early warning signals are recognized and appropriate timely follow-up action is initiated. The main objective of the program is capacity building at district and state levels. "WHO is in the process of computerizing the surveillance system in the states of Tamil Nadu and Maharashtra. Computers have been provided to the districts and the relevant staff is being trained in computer applications vis-à-vis surveillance. This will result in faster transmission of information in both directions and prompt action in the management of outbreaks."[22]

A disease may be detected by a Government VHN or MO. The communicable disease is informed to the DDH designated to the area. Upon confirmation of the case, the DDH communicates the case to the DPH designated to the state of Tamil Nadu. The information is then entered in to a computerized database, which is shared with the NICD.

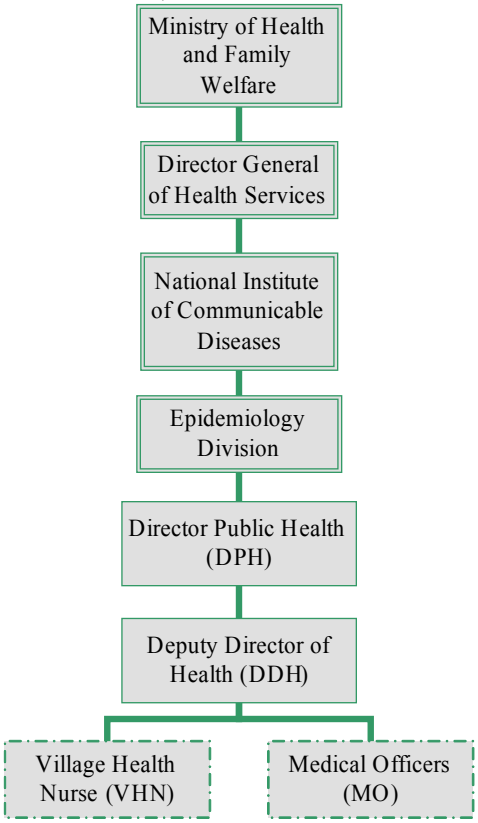

Figure 3: Indian Epidemiology and Control of Communicable Disease Organizational Structure

[22] WHO instigated initiative can be found here -- http://www.whoindia.org/EN/Section3/Section108.htm

12. APPENDIX C: COMMUNICATION STATES AND FLOW

The state transition diagram in Figure 4 illustrates the 4 entities: Community, Community Healthcare-Workers, Communications Providers, and National/State/Regional/District Epidemiology Unit. Decisions within the transitions are only taken by the Epidemiology Unit and by the Community Healthcare-Workers; where as all other transitions contribute to relaying the information and taking actions. The overall transition completes the cycle, meaning the information flow is initiated by the community and the information flows back to the community, other than when it is terminated at the National, State, Provincial, or District level when no events of interest are detected, or at the Community-Healthcare-Worker level.

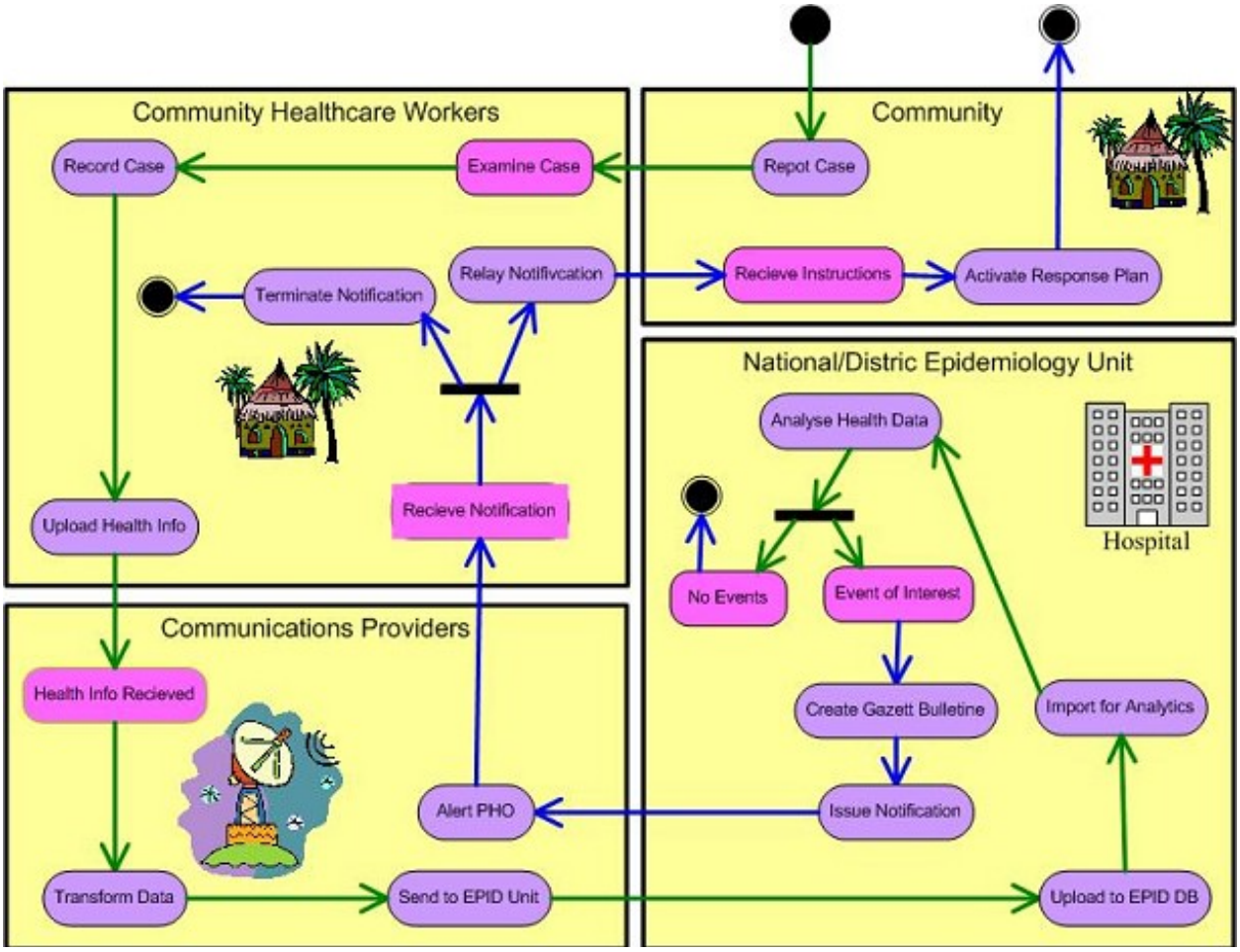

Figure 4: Sequence of disease surveillance and notification state transitions and the entities

(Purple) and (Pink) denote the states of the process of sensing information from the Community-Healthcare-Worker, uploading data to State/Regional Epidemiology Unit DB and extracting for statistical data mining, reporting any anomalies discovered by the State/Regional Epidemiology Units to Community-Healthcare-Worker at the Divisional levels, and finally Community-Healthcare-Worker deciding whether or not to relay message to their local communities. Green

arrows (→) are the upstream (data acquisition, analysis, and detection) and the Blue arrows (→) are the downstream (detection and notification) communication paths.

Description of the states

(i) *Report Case* – A member of the community visiting the local primary or secondary Healthcare-Provider at the Community Dispensary (Village Clinic) or local hospital in the nearest town to seek medical assistance. The patient information is recorded by the Medical Practitioner examining the case.

(ii) Examine Case – In the event the primary or secondary Healthcare-Provider reports the case as a communicable disease to Area Government Healthcare-Worker (MOH in Sri Lanka and DDH in India) the Community-Healthcare-Worker (PHI in Sir Lanka and VHN in India) follow the standard procedure, which is to examine the case to verify the report.

(iii) Record Case – The verified case will be recorded by the Community-Healthcare-Worker to be notified as per protocol using the standard form (e.g. H544 in Sri Lanka). The same information will be entered by the Community-Healthcare-Worker in to the introduced ICT application, either online or offline.

(iv) Upload Health Info: It may be the case that the Community-Healthcare-Worker is unable to connect to the network immediately because the location of the patient may be outside the WLL coverage area. Hence, will enter the data in to the system as a draft and submit as soon as the next connection is established or the application will automatically detect the strength of the connection to forward the data.

(v) Health Info Received: Information submitted such as SMS can get held up at the Provider's SMS Controllers due to heavy SMS traffic or after submitting the information the GPRS connection can get stalled as a result losing the information in thin air. Therefore, the Community-Healthcare-Worker must receive acknowledgement from the Central DB implying successful receipt of the data. In the event an acknowledgement is not received in the standard time frame the Community-Healthcare-Worker will try an alternate technology or alternate application to push the information to the central database.

(vi) Transform Data: The information that is sourced from as SMS, Email, Internet, and Voice must be transformed into a standard record structure before committing the information to the database. For example the SMS text will be transformed and stored as an XML file. The standard record format will depend on the relational database and the data visualization application used.

(vii) Send to Epidemiology Unit: The newly transformed and standardized data stream will be submitted to the National Epidemiology Unit's Central DB by the commercial ICT Provider.

(viii) Upload to Epidemiology Unit DB: Before importing the information in to the Epidemiology Unit DB for visualization, the records are staged to ensure referential integrity. The verified information is then synchronized with the existing dataset. The new records will update the baseline information of the algorithms.

(ix) Import for Analytics: Periodically datasets will be imported for analytics. This data will be a subset of the Epidemiology Unit DB based on geography and time. The information retrieved for analytics will be mostly counts of certain parameters of interest opposed to raw records of information.

(x) Analyze Health Data: State/Regional Epidemiology staff will run the algorithms and review the analytics. The analysis may be conducted on a daily basis. However, the frequency of analysis and review may increase during outbreak to monitor the spread of the disease and increase of cases.

(xi) No Events: If there is no cause for alarm; i.e. no abnormal events or zero detections, then the Epidemiology Unit will not notify of any outbreaks and terminate the process.

(xii) Event of Interest: If the analysis reveals anomalies in the data and the possibility of an outbreak; also termed as events of interest, by the Epidemiology Unit, then they will activate their response plans to initiate the notification process.

(xiii) Create Gazette Bulletin: State/Regional level Epidemiology Unit staff will apply the standard government procedure to communicate the outbreak to the Community-Healthcare-Workers in the State/Region. This may include consulting the necessary experts in Epidemiology, verifying the analysis, and initiating the notification preparation procedures to communicate the health risk information. Epidemiology Unit staff will prepare the necessary information to enter a proper notification into the system, which is a Gazette notification in Sri Lanka.

(xiv) Issue Notification: The message will need to be converted to a message that can be transmitted over the SMS or Email. Also the messages will need to be edited to suit the end device due to limitations in mobile handset displays. Epidemiology Unit staff will disseminate the message to the targeted Community-Healthcare-Workers based on the notification priority

(xv) Alert Community-Healthcare-Worker: At the time a priority channel for alerting emergency officials such as the Community-Healthcare-Workers is established by the Government and Commercial ICT Providers then the alert issued by the Epidemiology Unit will be treated as a priority notification by the ICT Provider; i.e. Avoid network congestion for timely delivery.

(xvi) Receive Notification: If the Community-Healthcare-Worker alert receiving mobile device is active then the Community-Healthcare-Worker can acknowledge by replying to sender. This is an indication of receiving the message. Until the Epidemiology Unit receives an acknowledgement, the Epidemiology Unit will consider the alert to be unsuccessful.

(xvii) Terminate Notification: if the nature of the event, severity, certainty, or urgency of the alert is not critical to the particular Community-Healthcare-Worker Area, then the Community-Healthcare-Worker will make the decision to terminate the alert and take no action but be on stand by for further updates.

(xviii) Relay Notification: In the case the event is of significance to the Community-Healthcare-Worker area, the Community-Healthcare-Worker may decide to relay the message to the local Health Officials and local community to take precautionary measures to prevent the community from getting infected. Community-Healthcare-Worker can further localize the message by translating and forwarding the message in the local language with personal comments. Community-Healthcare-Worker will be given access and measures to alert the local community and in this project it will be Health-related CBOs.

(xix) Receive Instructions: If the alert is not received by the CBO representative (i.e. local first responder) in the community, then the message is not effectively disseminated to the local community by the representative. Hence, the representative must inform the Community-Healthcare-Worker indicating receipt of message.

(xx) Activate Response Plan: Health-related CBO will consider the Community-Healthcare-Worker message and instructions to activate their local response plans to notify the community and make them aware of any emergency health risks.

13. APPENDIX D: COMMUNICATIONS ARCHITECTURE

The RTBP is an ICT based program. The sections below describe the architecture and behavior of the ICT system for data capture, analytics, and disseminations.

13.1. Network and Hardware Components

Community-Healthcare-Worker equipped with java enabled MPs will record the health information using the hand held MP. The information will travel through the GSM network. The Community-Healthcare-Worker has the option of using a menu driven J2ME application residing on the hand held to record the data. Depending on the available reliable transport, the application will select between GPRS or SMS import the data through a WAP enabled HTTP Server or an SMS Center. By default the application will be set to use GPRS; in the event the connection is doggy, the application will push the information through the SMS channel. The Community-Healthcare-Worker may want to provide additional instructions to the Epidemiology Unit and shall use Email or SMS text to send additional information. The Epidemiology Unit will configure and house the HTTP application with the WAP gateway and open the access through the Internet. The Community-Healthcare-Worker submitted data will travel through the Commercial ICT networks through the provider gateways directly to the Epidemiology Unit DB. The Epidemiology Unit Server will also be connected to a SMS Modem. Thereby, SMS text is directly received by the Epidemiology Unit. The Server will have an internal interface to strip the HTTP and SMS data to populate the central DB. The Epidemiology Unit Nurses entrusted upon disease surveillance can access the DB to retrieve surveillance data to run the analytics algorithms for outbreak detection.

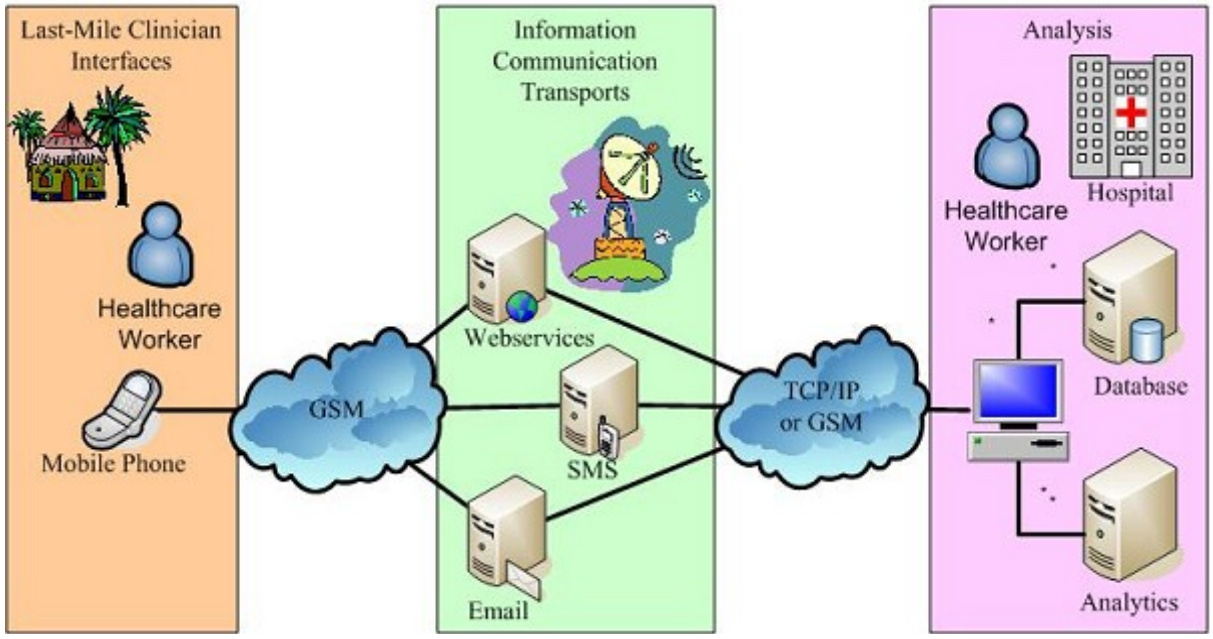

Figure 5: WLL Communications system architecture with mobile handhelds, TCP/IP-GSM Network, and Datacenter.

13.2.System Software Components

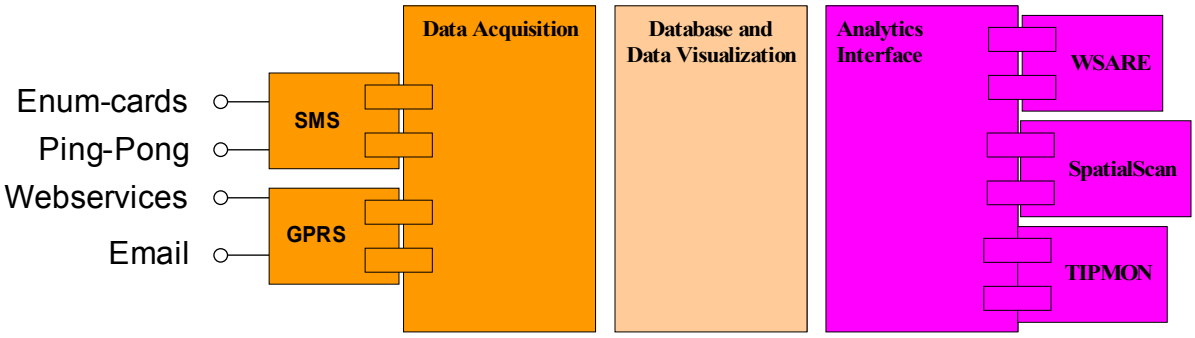

Figure 6: Set of software components necessary to communicate and analyze the health data.

Data Acquisition

The Community-Healthcare-Worker data is captured in four different ways though a collection of application components; namely a set of Mobile-Forms, Ping-Pong, Enum-Cards, and Email.

(i) *Mobile Forms* – is a set of menu driven electronic forms residing on the mobile hand held devices. The Community-Healthcare-Worker simply navigates through the series of forms to enter information and submit as individual records or bulk-send the entire set of records. The application will allow the user to select the communicable disease from a drop down menu,

and then prompt the Community-Healthcare-Worker to enter the syndrome, gender, age, and the number of cases. Once the record is committed then they follow the same steps to enter the next case information.

(ii) *Ping-Pong* – in the event the Community-Healthcare-Worker has not submitted any data before a certain time in the day, then the Community-Healthcare-Worker is prompted with a sequence of question to enter the same information as in the Mobile Forms application in (i). It may be the case that the Community-Healthcare-Worker accidentally deletes the onboard application or has forgotten to tell the system that there are no cases for the particular day. This SMS based application will alert and force the Community-Healthcare-Worker to submit data or indicate zero cases. Also in the event of an epidemic the Epidemiology Unit may want to collect additional information from all the Community-Healthcare-Workers and can use the Ping-Pong application to send a request to all the Community-Healthcare-Workers and prompt them to submit the information on-demand.

(iii) *Enum-Cards* – where a set of numeric codes for communicable diseases are listed in a reference card and the Community-Healthcare-Worker refer to the code to enter a short code SMS text to define the record for using the short code rather than having to type the long description of the communicable disease. It is a string of number and special characters such as 5*3*2 means "5" cases of "diarrhea" (3) in "Kalutara" District (2). Part of this project will be to construct this system. They same enumerated short codes can be used in the Mobile-Forms in (i) and Ping-Pong applications in (ii) in order to reduce the information payload.

(iv) *Email* – will be used in the event of the need to notify a set of unknown diseases or symptoms that can not be described through the any of the above applications. Also the Epidemiology Unit as means to push messages downstream to the Community-Healthcare-Worker with any alert bulletins or instructions. The RTBP is also interested in knowing whether Health Officials will use Email in their daily activates.

Data Visualization

The State/Regional Epidemiology Unit staff members will be trained to scrutinizing the imported disease information for unusual increase in reported cases or abnormal trends. They will use a "dashboard", as conceptually illustrated in Figure 7, to visualize the data from all possible angles; i.e. run cross tabs on geo-temporal records. The dashboard will display statistics on topographical maps as well as display trends. Nurses will also run queries on data sets to identify potential areas to focus on long term mitigation programs. The dashboard can also be used to monitor the progress of a preventive program or the effect of the emergency response plans applied to a certain epidemic.

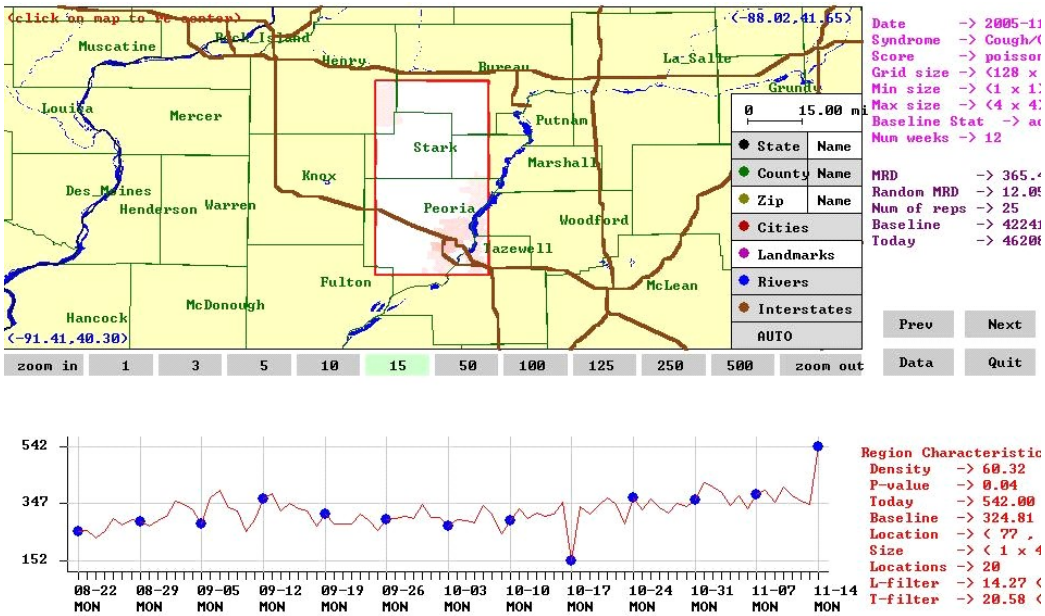

Figure 7: Dashboard visualizing analyzed data across a map with statistical trends

Analytics Software Interfaces

Traditional Biosurveillance algorithms detect disease outbreaks by looking for peaks in a univariate time series of health-care data. Current health-care surveillance data, however, are no longer simply univariate data streams. Instead, a wealth of spatial, temporal, demographic and symptomatic information is available. Analytics Interface will examine aggregate and identified data routinely collected by interfacing with the C-HCW input data storage system automatically and in real time for trends and anomalies suggestive of disease outbreaks. Using data that indicates spatial distributions of cases, software can assist Epidemiology Units to identify the location of an outbreak. A spike in hospital admissions from a neighborhood, or a message that something is strange, could merely indicate the presence of a flu bug -- or it could mean that a chemical or biological weapon has been released.

The research will begin with a few different algorithms to identify, which one or which combination of them works best in the context of the particular countries in South Asia. The use of particular algorithms will be determined mostly by the quantity and frequency of the available data. However, the project will not restrict the scope of only using the algorithms listed below, but it will also look at the options of adopting other algorithms or even developing a hybrid of the currently available collection.

(i) *WSARE* – This algorithm uses a multivariate approach to improve its timeliness of detection. WSARE (acronym for What's Strange about Recent Events) employs a rule-based technique that compares recent health-care data against data from a baseline distribution and finds subgroups of the recent data whose proportions have changed the most from the baseline data. In addition, healthcare data also pose

difficulties for surveillance algorithms because of inherent temporal trends such as seasonal effects and day of week variations. WSARE approaches this problem using a Bayesian network to produce a baseline distribution that accounts for these temporal trends. The algorithm itself incorporates a wide range of ideas originating in the field of Statistical Machine Learning, including association rules, Bayesian networks, hypothesis testing and permutation tests to produce a detection algorithm that is careful to evaluate the significance of the alarms that it raises. For WSARE to work, we need lots of patient records (typical data consists of a single line listing per patient visit to hospital. The longer the duration of collected data, the better it is.

(ii) *SpatialScan* - Here the focus is on the task of *spatial cluster detection*: finding spatial regions where some quantity is significantly higher than expected. For example, the goal may be to detect clusters of disease cases, which may be indicative of a naturally occurring epidemic (e.g. influenza), a bioterrorist attack (e.g. anthrax release), or an environmental hazard (e.g. radiation leak). The algorithm has two main goals: to identify the locations, shapes, and sizes of potential clusters, and to determine whether each potential cluster is more likely to be a "true" cluster or simply a chance occurrence. Primary motive of the application is *prospective disease surveillance*: detecting spatial clusters of disease cases resulting from a disease outbreak. In this application, the surveillance is typically performed on a daily basis, with the goal of finding emerging epidemics as quickly as possible. For this task, a number of cases of some given syndrome type (e.g. respiratory) in each spatial location (e.g. postal code) on each day. To be more precise, the system cannot typically measure the actual number of cases, and instead rely on related observable quantities such as the number of Patient visits. Thereafter, it detects those increases which may be indicative of emerging outbreaks, as close to the start of the outbreak as possible, while keeping the number of false positives low. In bio-surveillance, every hour of an earlier detection can translate into thousands of lives saved by allowing more timely mitigation activity, such as administration of antibiotics, and this has led to widespread interest in systems for the rapid and automatic detection of outbreaks. For spatial scan, even 3 months of past data may be enough to train the baseline models, provided that the quantity of data is reasonable.

(iii) *TIPMON* - The Tip Monitor estimates how likely it is for a newly reported patient case to be a close copy of some other case in the past data, if both have been generated by the same specific underlying cause, such as for instance a malicious contamination of raw food at a processing plant. This algorithm can work with small, sparse data but the models may need to be hand tuned to give reasonable alerts. TIPMON's job is to identify small subsets (even just pairs and triplets) of significantly correlated records in an incoming stream of multidimensional event-based data such as anonymous individual patient health events involving chief complaint strings, prescription orders, public safety hotlines or customer complaints. Very often, such data is sparse, noisy and it may contain very little and spotty evidence of potentially crucial coincidences. The last feature makes it very hard to detect important events with more traditional approaches used by biosurveillance analysts (such as spatial scan statistics, WSARE, or multivariate time series analysis), since they are designed to benefit from ample evidence. Suppose that among the chief complaint strings of two unrelated pa-

tients in the same city on the same date there was a mention of bloody stools in pediatric cases. The multiple mentions of "bloody stools" or "pediatric" might not be surprising, but the tying together of these two factors, given matching geographic locations and timings of reporting, is sufficiently rare that seeing only two such cases is of interest. Reliable automated detection of such signals in multivariate data requires new analytic approach such as that in TIPMON.

14. APPENDIX E: TASK DETAILS AND RESOURCE ASSIGNMENT

14.1. Tasks and description of activities

Table 4: Tasks and Notes of the Activities to build and evaluate the Data Acquisition and Visualization of the Surveillance System

Task	**Notes**
Project Launch	*Activities before launching the project*
• *Hold a Partner Meeting*	All the partners will meet in India (Chennai), to discuss the scope of the project, identify their tasks/roles, ensure contractual terms/payments adhere, and timelines are acceptable; give the partners the opportunity to adjust research if necessary
• *Recruit Research Assistants*	Hire a Research Assistant for each State/National Epidemiology Unit for the project duration; responsibilities are assisting with research, system implementation, and project coordination
• *Select Divisional Health Areas, Healthcare-Worker Personnel, and Communities*	Each participating country's Government Epidemiology Unit and Non Governmental Community-based Organizations will collaboratively select 4 Government Healthcare Divisions. Thereafter, identify the participating Healthcare-Workers and Communities to fill the requirement of the research matrix in Table 1.
• *Procure RA Equipment*	Procure the Project Staff equipment such as the Laptops and MPs
Design System	*Understand the local environment and collect system requirements*
• *Write Software Requirement specifications*	Hold series of online discussions with the National/State/Regional Epidemiology Units, Healthcare-Workers, and Technology Partners to collect the system requirements; i.e. system inputs and outputs. Based on the user requirements design the architecture, components, user interfaces, and database for a system that will work in both countries.
• *Demo Prototype*	The technology partners will develop a prototype of their respective ICT components to demonstrate to the project that the system adheres to the requirements and specifications collected. The prototype will be first demonstrated to the project staff and consultants, second to the

Task	Notes
	Healthcare-Worker during planning workshop in each country.
• *Organize Healthcare Worker Planning Workshop*	A planning workshop will be organized in each country to introduce the project to the Government Healthcare-Workers. The participating Epidemiology Units and Healthcare-Workers along with the Project Partners will come together to understand the scope of the project, identify their roles/responsibilities, and agree to the project schedule.
Release Beta	*The first release of the working system for internal testing*
• *Develop beta Release*	Technology partners will build the individual components such as the algorithms, graphical user interfaces, mobile applications, database, and interoperability components; the Research Assistants, e-Health Expert, ICT/Interoperability Expert and other selected resource persons will test the individual components to ensure they meet the standards and produce a bug report to identify the short comings; after initial bugs are fixed the individual components will be integrated as a single system; once system has been integrated the Research Assistants will implement the system with actual master data; the Research Assistants, Epidemiology Unit staff, and Experts will test the integrated system to identify any further bug and design changes
• *Procure System Equipment*	After beta release is complete, the MPs, internet connectivity, switch/access-point, and Server for the Community-Healthcare-Workers and State/Regional Epidemiology Units will be purchased
Training	*Healthcare Worker and Epidemiology Unit training*
• *Prepare Training Manuals and Operational guidelines*	Technology partners will develop a set of user manuals for the Epidemiology Units and Healthcare-Workers to refer to and learn to operate the ICTs; The e-Health and ICT expert will develop a set of "standard operational procedures" for the Healthcare-Workers and Epidemiology Units to follow when using the newly introduced RTBP
• *Organize Healthcare Worker Training Workshop*	Training event will be organized in each country to introduce the functional aspects of the system components to the Healthcare-Workers and State/Regional Epidemiology Unit users. RTBP design/development Partners will use the workshop, final demo, before the "final release"
Activation	*Finalize all system components of RTBP*
• *Deploy final release of ICT system*	Deploy system in each country to begin trials; this task will be done at the same time they conduct the user training; Research Assistants will finalize the master data implementation before releasing system to the Healthcare-Workers and Epidemiology Units; The Healthcare-Workers and State/Regional Epidemiology Unit staff will use the system over a

Task	Notes
	short period to test the usability and integrity of the system before beginning the trials. This will give the technology partners a final chance to fine tune any bugs
• *Conduct Interoperability Testing*	The Epidemiology units in each country along with the e-Health and Interoperability expert will run some drills to test the integrity of the system in terms of exchanging health-related information with other entities and between the two countries; giving the technology partners a final chance to fix last moment bugs.
• *Commission System for evaluation stage*	e-Health and ICT/Interoperability expert along with the assistance of the Researcher Assistants will produce a report ensuring the technology system and human participation is ready for the program to go live; i.e. to begin the "evaluation phase".
Evaluation	*Running of mock-drills over a 1 year period*
• *Plan simulation evaluation activities*	e-Health and ICT/InterOp expert in consultation with the Epidemiology Units, Researchers, Research Assistants and Communities will develop a simulation plan, which identifies the simulation intervals, mock epidemic, spread pattern, expected response plans, and parameters to observe
• *Design RTBP Evaluation Toolkit*	Develop a Toolkit based on the experience of designing and developing the RTBP and propose an evaluation scheme to be used in the mock-drills and evaluation process; developing data collection forms is part of the toolkit; will contain the methodology for processing the data; Toolkit go through a series of revisions before the final version is released before the project ends
• *Organize simulations planning workshop*	All participants: Healthcare-Workers, Epidemiology Units, & Community-based Organization members including the local partners in each country and selected experts will engage Divisional level training workshop; there will be 4 such events in each country; the workshop will discuss the steps and schedule of the simulation as well as the individual/group roles during the exercises
• *Hold a Public Lecture*	Hold a public lecture in each country on e-Health based surveillance and notification systems; an expert from the region will be invited to each country; the public lecture will be held to also announce the release of the pilot system and the beginning of the evaluation phase
• *Run simulated exercises*	The a series of silent-test and live-exercises will be held periodically to test the end-to-end functionality of the RTBP;
Analysis	*Assess the results and discuss the findings*
• *Discuss research findings at a workshop in*	All partners: Healthcare-Workers, Epidemiology Units, Experts, Technology Partners, Community members, and other Researchers will be invited to participate in a research findings discussion to interpret

Task	Notes
each country	the data and identify the key findings that can contribute to policy; a summary of these findings will be further discussed in the follow-up "Comparison Research Meeting"
• *Comparison Research Meeting*	Project partners and other invited experts will meet to discuss the research findings to conceptualize and interpret the results; the outcomes of the meeting will be documented in the final report
• *Final Report*	A written report containing all research findings will be produced for both the donor and the Researchers and Practitioners.
Research Disseminations	*Publications to be produced by Researchers on the topics below*
• *Epidemiology Policy*	Study each countries policy changes over history pertaining to epidemiology and e-Health, identify global policy that can enhance the current policy to embrace e-Health programs, and other regulatory aspects relevant to the topic of the RTBP
• *Epidemiology e-Health Programs*	Study the present day e-Health programs addressing epidemiology and systems that have been implemented; scrutinize the strengths and weaknesses of each program; identify state of the art systems that are available for implementation in India and Sri Lanka
• *Interoperability in Epidemiology through e-Health*	Study the interoperability issues at a national and regional international level; scrutinize the current global interoperability standards for communicating epidemiology information; develop a profile for the country based on an adopted standard or hybrid of standards to be used in the project
• *E-Health RTBP Implementation challenges*	Learn and document the challenges of implementing the proposed RTBP in the respective countries for future implementers or researchers to reference when developing similar programs.
• *Outcome of the RTBP in India and Sri Lanka*	Publish the lessons learned in the project based on the simulations (mock drills) addressing the behaviors of the ICT and Human systems.

14.2. Resources and Schedule of tasks

Table 5: Task Resource allocation Matrix

Task	*Gov-Health (IN/SL)*	*CMU-Auton-Lab*	*IITM-RTBI*	*NCBS*	*U-of-Alberta*	*LSF*	*Sarvodaya*	*LIRNE asia*	*Timeline*
Organize a Partner Meeting			■					■	Week 01 – 04 (Month 01)
Recruit Research Assistants	■		■	■		■		■	Week 05 – 08 (Month 01 - 02)
Select Divisional areas, Healthcare-Worker Personnel and Communities	■			■			■		Week 05 – 06 (Month 02)
Procure RA Equipment			■					■	Week 05 – 10 (Month 02)
Write Software Requirement Specifications		■	■			■			Week 08 – 14 (Month 02 - 03)
Demo Prototype		■	■						Week 12 – 18 (Month 03 - 04)
Organize Healthcare-Worker Planning Workshop	■		■	■				■	Week 12 – 18 (Month 03 – 04)
Develop Beta release		■	■			■			Week 18 – 28 (Month 04 – 06)

Task	*Gov-Health (IN/SL)*	*CMU-Auton-Lab*	*IITM-RTBI*	*NCBS*	*U-of-Alberta*	*LSF*	*Sarvodaya*	*LIRNEasia*	*Timeline*
Procure system equipment									Week 24 – 30 (Month 06 – 07)
Prepare Training Manuals and Operational Guidelines									Week 30 – 36 (Month 07 – 08)
Hold Healthcare-Worker Training Workshop									Week 36 – 40 (Month 08 – 09)
Deploy final release of ICT system									Week 36 – 42 (Month 08 – 10)
Conduct Interoperability Testing									Wee 42 – 44 (Month 10)
Plan Simulation and evaluation activities									Week 40 – 44 (Month 09 – 10)
Commission ICT system for evaluation stage									Week 44 – 46 (Month 10)
Design RTBP Evaluation Toolkit									Week 40 – 46 (Month 10)

Task	*Gov-Health (IN/SL)*	*CMU-Auton-Lab*	*IITM-RTBI*	*NCBS*	*U-of-Alberta*	*LSF*	*Sarvodaya*	*LIRNEasia*	*Timeline*
Organize Simulations planning Workshop	■			■			■	■	Week 46 – 52 (Month 10 – 12)
Hold Public Lecture			■					■	Week 48 – 52 (Month 12)
Run simulation exercises	■			■			■		Week 53 – 82 (Month 13 – 20)
Organize to discuss research findings of each country			■					■	Week 80 – 86 (Month 20 – 21)
Organize Comparison Research Meeting								■	Week 86 – 92 (Month 21 – 22)
Final Report								■	Week 92 – 104 (Month 22 – 24)
Research Disseminations	■	■	■	■	■	■	■	■	Throughout
Project management								■	Throughout