Nivārana: System design and implementation focused on rapid response to epidemics

Authors

W.M.A.S.B. Wickramasinghe, University of Colombo School of computing, Sri Lanka, asangawick@gmail.com

Dr. S.M.K.D. Arunatileka, University of Colombo School of Computing, Sri Lanka, sda@ucsc.cmb.ac.lk

Dr. R.G. Premaratne, Epidemiology Unit of Ministry of Health, Sri Lanka, risintha@yahoo.com

Abstract: Communicable disease surveillance was identified as one of the most important activity in responding to disease outbreaks. Early recognition of impending outbreaks followed by rapid response is the key for controlling communicable diseases in a community. Nivārana is a communicable disease surveillance and analysis system focused on swift data collection, efficient data processing and rapid information dissemination in order to control and prevent the spread communicable diseases. While other electronic disease surveillance initiatives and systems focus on data collection and analysis, we believe that swift information dissemination also plays a significant role in responding to disease outbreaks. The system is designed based on the object oriented methodology and implemented based on java technologies and using free and open source software libraries and tools. The system is designed with minimum implementation and recurrent costs which is critical for developing countries like Sri Lanka. As the name suggests Nivārana is conceptualized with the saying "Prevention is better than cure" and will no doubt is a lifesaver for many innocent people.

1 Introduction

Communicable disease surveillance has evolved as an important factor of public health systems (Rolfhamre et al., 2004) and is carried out by every country with the aim of detecting the outbreaks as early as possible (Hashimoto et al., 2000). Detecting communicable diseases early as possible increases the ability to control the spread of diseases (Rolfhamre et al., 2004; Hashimoto et al., 2000).

Epidemiology Unit (EU) is the premier institution in the Ministry of Health care and Nutrition which carries out surveillance, prevention and control of communicable diseases in Sri Lanka. There are several factors that hinder the capability of EU in disease surveillance activities. Therefore, Information and Communication Technology (ICT) could be used as the means of breaking these barriers in disease surveillance. Countless opportunities could be gained, if the surveillance system at present is integrated in to an ICT solution.

Nivārana is jointly initiated by the University of Colombo School of Computing (UCSC) and the EU to address the prevailing issues and to enhance the disease surveillance capacity of the

EU. Nivārana is conceptualized having the saying "prevention is better than cure" in mind and it is expected to facilitate swift public health action to minimize the spreading communicable diseases in Sri Lanka thereby, preventing ill health and suffering in the community and to conserve much needed resources otherwise spent on curative healthcare.

Hence to suggest the most suitable ICT solution for this issue, requirement gathering was carried out along with a literature review. Interviews with epidemiologists at the EU and a self administrated questionnaire which was distributed among public health staff working at divisional levels were used in the requirement gathering.

The system was designed using the object oriented methodology and was based on a set of design considerations. According to the system design, it is capable of been extended to the level of an enterprise application if needed. System includes seven modules and out of these data collection module, data analysis and outbreak detection module and information dissemination module were seen as the most important in disease prevention and control activities.

This paper gives an insight to literature on existing disease surveillance systems, system requirements and elaborates on system design and implementation aspects.

2 Related Work

Nine communicable disease surveillance initiatives were investigated and critically reviewed. The review was carried out based on their functionality and applicability of the features in the Sri Lankan context.

Scalable Bio-Surveillance Architecture is a system proposed to the Department of Defence (DoD) and the Civilian Public Health Authorities in disease surveillance in United States (Chang et al., 2001). DoD initiated this project to identify bio-terrorist attacks and meanwhile identifying suspected disease outbreaks. SMINET-2 is an Internet based communicable disease surveillance system developed for Sweden (Rolfhamre et al., 2006). Main objectives of SMINET-2 are simplifying reporting, improving completeness and increase timeliness (Rolfhamre et al., 2006). The Integrated Public Health Information System (iPHIS) is a component in Canadian Integrated Public Health Surveillance programme (Public Health Agency of Canada, 2007). Both SMINET-2 and iPHIS carries the same objectives. Communicable Disease Reporting and Surveillance System (CDRSS) is a system developed to assist disease surveillance in New Jersey (Hamby et al., 2004). It was developed in adhering to the CDC (Centre for Disease Control) standards and guidelines. Patient level tracing of communicable diseases could be seen as the most important aspect in these systems.

When data collection aspect is considered, most of them collect data from multiple data sources and on daily basis. Disease surveillance initiatives like Real-time Outbreak and Disease Surveillance (RODS) (Espino et al., 2004), National Electronic Disease Surveillance System (NEDSS) initiative (CDC, 2001; CDC, 2006) gather information from hospitals, laboratories and other health care institutes. RODS (Espino et al., 2004) and Scalable Bio-

Surveillance Architecture (Chang et al., 2001) uses other data sources such as absenteeism and pharmacy sales data to improve the automatic epidemic detection. Scalable Bio-Surveillance Architecture reaches to the extent of using web quires to improve the automatic detection of epidemics.

Outbreak detection is one of the highlighted aspects. Systems like Scalable Bio-Surveillance Architecture (Chang et al., 2001) propose highly sophisticated Bayesian network which uses 14 nodes to represent an individual, for outbreak detection while initiatives like RODS (RODS Laboratory, n.d.) investigate on statistical approaches. Number of other useful outbreak detection algorithms and approaches like Serfling method (Serfling, 1963), Recursive-least-square (RLS)(Moore et al., 2002), Autoregressive Moving Average (ARIMA)(Sebastiani & Mandl, 2003), Wavelet-based Anomaly Detector (WAD) (Zhang et al., 2003), Hidden Markov Model (HMM) (Sebastiani & Mandl, 2003) ,CuSUM (Sebastiani & Mandl, 2003) and use of detection filters(Reis et al., 2003) were also investigated.

As identified, existing disease surveillance systems do not focus on information dissemination aspect, which we think the key to initiate actions that leads to controlling communicable diseases.

3 System requirements

Identifying system requirements is critical to deliver a usable and an effective system. Interview was selected as the main requirement gathering methodology. Literature review and self administrated questionnaire were also employed for collecting information.

The identified requirements could be broadly categorized in to 5 groups, namely a) Data collection b) Reporting c) Analysis d) Information dissemination and e) Administrative requirements.

3.1 Data collection

When data collection is concerned, the system should be able to collect Weekly Return of Communicable Diseases forms electronically and provide appropriate interfaces for verification and validation. Initially this will be the sole source of data that will be fed in to the system.

3.2 Reporting

Identifying suitable formats for reporting is a very important in decision making. Both Horvitz & Barry (1995) and Montazemi & Wang (1989) have highlighted this fact in their work. By analysing the nature of communicable diseases three basic metadata could be identified mainly aimed at reporting; namely a) Personal b) Geographical and c) Temporal. Since the weekly return only consists of geographical and temporal parameters the reporting functionality is based on these two aspects.

It was identified that for data representation tables, graphical charts and information plotted on maps needed to be utilised. Furthermore, it was identified that maps are the best information representation format since the particular disease spread is of spatiotemporal in nature. The report formats need to be constantly reviewed with the epidemiologist to ensure the usefulness and ease of decision making.

3.3 Analysis

The information analysis is more focussed on reducing the human intervention in identifying suspicious epidemic situations. The system should automatically detect the suspicious situations based on the count and disease density for a given geographical area. Suspicious events need to be notified immediately to the responsible officials.

3.4 Information dissemination

Information dissemination is paramount important in preventing and controlling epidemics. There are two types of information a) information which requires prompt action and less detailed and b) information which are more detailed mainly important for decision making. The system should support dissemination of these types of information to the appropriate public health officers.

3.5 Administrative requirements

One of the main important aspects in the administrative module is to support geographical divisions of the Ministry of Health which changes at different time intervals. The Medical Officer of Health (MOH) divisions are based on the population in a given area and they are subjected to change with the change of population. User activity log and weekly return report are also required for administrative purposes.

The identified requirements were translated carefully to the system design. For the system evolution, the system should be designed in a way that supports effortless management and seam less integration of new features.

4 System Design

Good information system design will result in increased maintainability which is important for the sustainability and the usability. Prior to designing, set of design considerations were identified in order to align the design to the operational and implementation goals. The design considerations were 1) low cost, 2) sustainability, 3) ease of use, 4) configurability and 5) extendibility.

System was designed by following object oriented methodologies because it could model the real world behaviors and properties accurately. The Nivārana architecture, as depicted in the Figure 4., consists of 7 modules, namely 1) data collection module, 2) analysis module, 3) reporting module 4) information dissemination module, 5) user authentication module 6) activity logging module, and 7) system management module. Among these modules data collection module, analysis module, reporting module and information dissemination module implements the business logic. The core classes which are used in the system are illustrated in the Figure 4.. All the above mentioned modules depend on these classes to perform their respective functionalities.

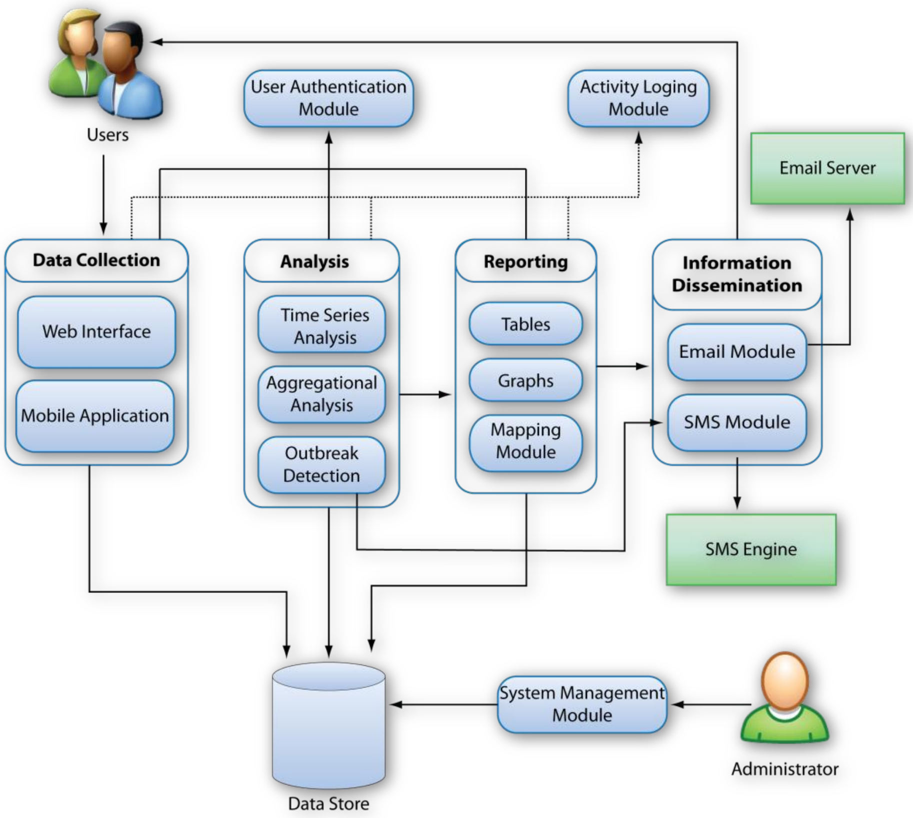

Figure 4.: Architecture of Nivārana consisting of four core modules and three administrative modules

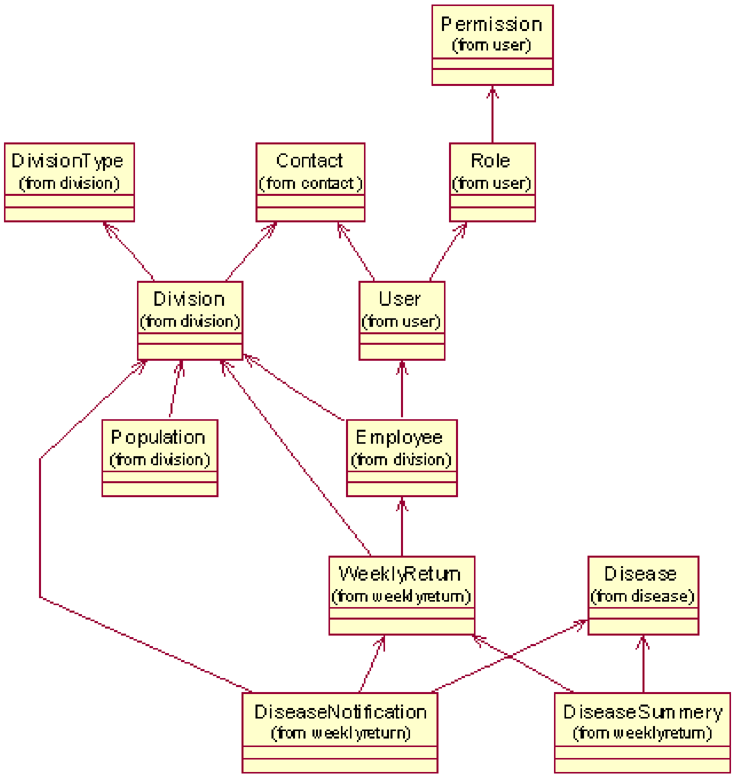

Figure 4.: Class diagram representing core classes

4.1 Data collection module

Data collection module could be seen as the heart of the system. This module manages weekly return of communicable disease data collection and is the sole source of data for the initial stage of implementation of the system. This module enables validation and verification of the submitted weekly returns with the view of increasing the accuracy.

4.2 Analysis module

Analysis module facilitates the analyzing of data collected from the data collection module and it closely associates the reporting module. In this module statistical analysis is more focused which also incorporates trend analysis. The outbreak detection sub-module tries to detect the suspicious incidents and alerts the relevant public health officials via the information dissemination module. In the outbreak detection module, users are able to add monitors for particular disease occurring at a given geographical area. The suspicious situations are detected based on these monitors. The monitors indicate whether to use disease counts or disease densities for detection of the outbreaks.

4.3 Reporting module

Mainly three kinds of report structures are produced by the reporting module, namely 1) tabular reports 2) reports that include graphs and 3) maps. To display large amount of information tabular reports are employed. Graphs display the information in the graphical manner which is important for rapid decision making. As stated in the *section 3* maps are used to depict the spatiotemporal relationship of a particular disease more effectively. Maps are capable of displaying the information spatially in a dynamic manner which could be very effective in decision making.

4.4 Information dissemination module

Information dissemination module supports sending of emails and SMS messages to health authorities in a timely manner.

5 Implementation

Nivārana has a three tier architecture. It is implemented with full featured web based client and a desktop client for efficient data collection. The web based system enables the users to access the system irrespective of the location and reduces the training efforts. In this section, programming languages and technologies are elaborated, followed by a discussion of the implementation of important features.

5.1 Programming Language and technology selection

It is important to select the most appropriate language, technologies, tools and libraries for information system development. Cost and the availability of libraries were the dominant factors in the selection process. Benchmark which included five criterions was carried out to select the most appropriate programming language. According to the bench mark as illustrated in *table 1* Java resulted as the well suited programming language.

Table : Benchmark of programming languages

Technology	Cost	Stability	Availability of rich libraries(Free)	Development effort	Deployment effort
.net	High	High	Low	Low	Medium
php	Low	Medium	Medium	Medium	Low
Java	Low	High	High	Low	High

The backend of Nivārana is implemented using Enterprise JavaBean 3.0 (EJB 3.0). For the front end, JavaServer Faces technology was used with RichFaces framework to implement the web interface and Eclipse Rich Client Platform (RCP) was used to implement desktop client. The desktop client will be deployed using the Java web start technology. EJB3.0 has built-in database abstraction which is referred as Java Persistent Architecture (JPA) which allows seamless integration with verity of database management systems. JBoss is used to host the system because it is a light weight, highly scalable and stable Java application server.

Alerting using SMS is an important functionality in Nivārana. For SMS sending Kannel SMS engine was selected. It is a free and open-source software which is capable of communicating with both the Short Message Service Centres (SMSC) and with GSM/GPRS modems. Use of Kannel makes this functionality more scalable.

These selected implementation technologies reduce the system development effort and more importantly it made the system cost effective while preserving the required system performance.

5.2 Implementation of features

Implementation hurdles are there in every software development project. These challenges have to be carefully handled. Implementation of outbreak detection module and mapping module were seen as the most important challenges in this project.

Outbreak detection and alerting module has a high impact and could be seen as a key to preventing the spread of the disease. Upon submitting the weekly communicable return of communicable disease report outbreak detection module automatically analyzes the newly submitted weekly return asynchronously against the available monitors for the given geographical area. When suspicious situation is detected it is notified to relevant responsible health care officials via SMS and full detailed report is sent via email. Figure 5. represents the disease monitor administration screen and automatically generated SMS message.

Challenging task in implementing mapping module is the selection of most appropriate approach. Maps are represented in a compressed Scalable Vector Graphic (SVG) files because SVG provides both the functional and non-functional requirements. Apache Batik SVG toolkit was used to manipulate the SVG at the server side and the map is sent as a raster image making the client much more independent. The Figure 5. represents the information represented on the map for the first week of January 2010 for Dengue.

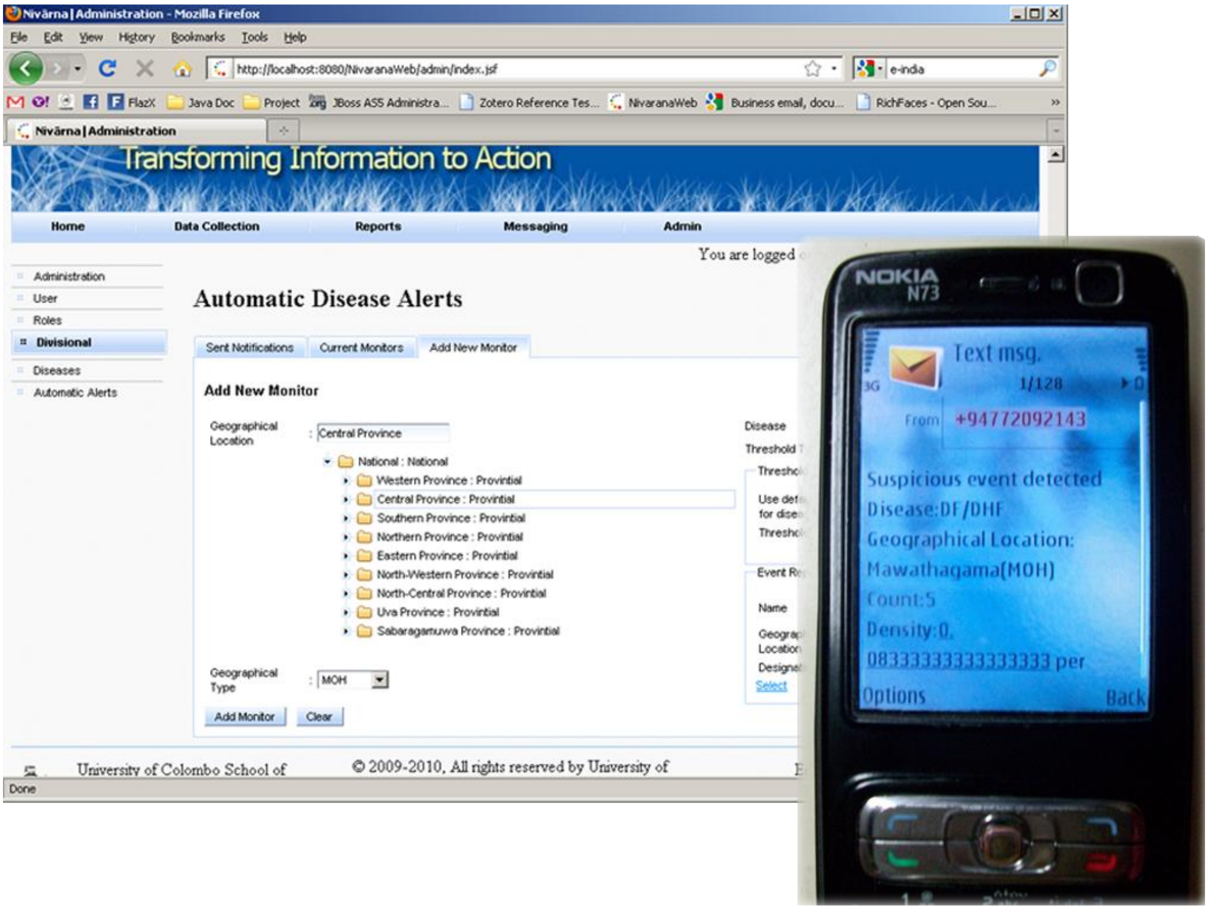

Figure 5.:Disease monitor administration interface and automatically generated SMS message

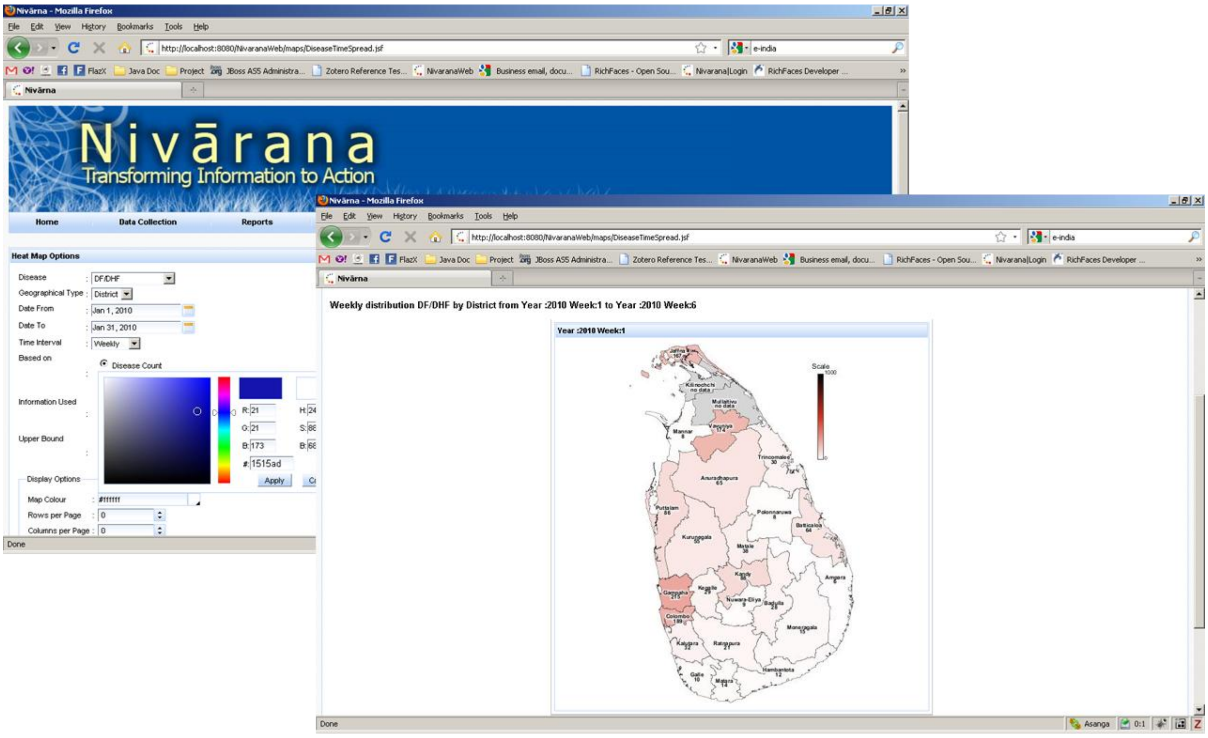

Figure 5.: The report criteria interface and the information represented on the map

6 Discussion and conclusion

We believe that rapid dissemination of information is the key to take swift preventive measures, although existing communicable disease surveillance systems do not focus on this aspect.

Figure 6. illustrates how the opportunity to control communicable diseases is reduced with delayed response and Figure 6. illustrates how this opportunity to control could be increased with rapid response. In Nivārana, we try to exploit this opportunity to control communicable disease by transforming information for quick actions.

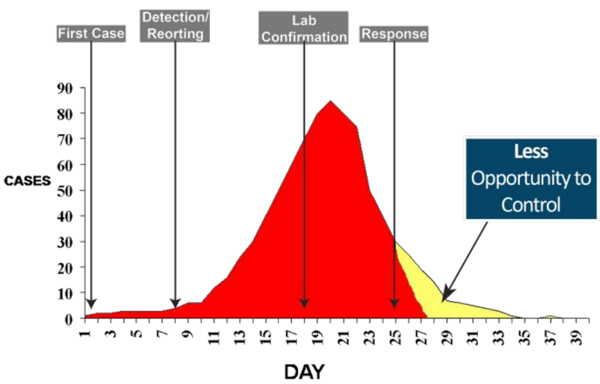

Figure 6.: Outcome of delayed response to outbreaks

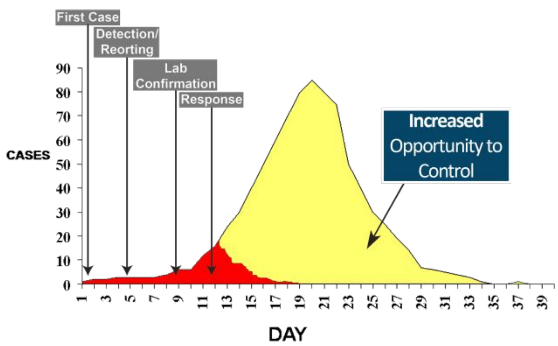

Figure 6.: Outcome of rapid response to epidemics

Keeping the main focus on swift information dissemination Nivārana system design and implementation also supports efficient data collection, efficient information processing and quick decision making.

Nivārana tries to empower regional level public health officers with more information and analysis capabilities for better decision making regarding their region, which is currently lacking in the manual surveillance system. With the implementation of the system we believe that the ground level officers will be able to increase their contribution in outbreak prevention.

System is designed and implemented focused on cost effectiveness. Free and open-source libraries and tools have greatly contributed to achieve this. Extendibility is one of the main focuses of the design. Hence, the system could easily be extended to the level of an enterprise application. When implementing, technologies are selected to minimize the switching cost by the means of adhering to open standards and using the tiered architecture.

Thus, capabilities of the EU and peripheral level to respond rapidly to communicable diseases are leveraged which result in effective and efficient outbreak control and prevention actions. Hence, Nivārana will no doubt is a lifesaver for many people.

References

CDC, 2001. Description of the NEDSS Base System.
Available at: http://www.cdc.gov/NEDSS/BaseSystem/NEDSSBaseSysDescription.pdf
[Accessed December 11, 2009].

CDC, 2006. National Electronic Disease Surveillance System (NEDSS).
Available at: http://www.cdc.gov/NEDSS/About/overview.html
[Accessed December 11, 2009].

Chang, J., Cheng, O., Ganapathiraju, M. & Luo, A., 2001. Scalable Bio-Surveillance System Architecture.
Available at: http://www.cs.cmu.edu/~luluo/Courses/17755BSS.pdf
[Accessed December 7, 2009].

DataDyne.org, 2009. EpiSurveyor | DataDyne.org.
Available at: http://www.datadyne.org/episurveyor
[Accessed December 7, 2009].

Espino, J.U., Wagner, M.M., Tsui, F., Su, H., Olszewski, R.T., Liu, Z. et al., 2004. The RODS Open Source Project: Removing a Barrier to Syndromic Surveillance.
Available at: http://openrods.sourceforge.net/
[Accessed December 9, 2009].

Hamby, T., McHugh, L., Octania-Pole, S., Robertson, C.A., Tsai, S., Bresnitz, E.A. et al., 2004. Communicable disease surveillance in New Jersey. *New Jersey medicine : the journal of the Medical Society of New Jersey*, 101(9 Suppl).
Available at: http://www.nj-ptc.org/njmed2/pdf/3_CommunicableDisease.pdf
[Accessed December 7, 2009].

Hashimoto, S., Murakami, Y., Taniguchi, K. & Nagai, M., 2000. Detection of epidemics in their early stage through infectious disease surveillance. *International Journal of Epidemiology*, 29(5), 905-910.

Horvitz, E. & Barry, M., 1995. Display of information for time-critical decision making. In Montreal.
Available at: http://eprints.kfupm.edu.sa/35364/
[Accessed March 13, 2010].

Montazemi, A.R. & Wang, S., 1989. The effects of modes of information presentation on decision-making: A review and meta-analysis. *Journal of Management Information Systems*, 5(3), 101-127.

Moore, A., Cooper, G., Tsui, R. & Wagner, M., 2002. *Summary of Biosurveillance-relevant statistical and data mining technologies*,
Available at: http://www.autonlab.org/autonweb/14638/version/3/part/5/data/moore-biosurv.pdf?branch=main&language=en
[Accessed December 13, 2009].

Public Health Agency of Canada, 2007. Canadian Integrated Public Health Surveillance (CIPHS)- Public Health Agency of Canada.
Available at: http://www.phac-aspc.gc.ca/php-psp/ciphs-eng.php
[Accessed December 7, 2009].

Reis, B.Y., Pagano, M. & Mandl, K.D., 2003. Using temporal context to improve biosurveillance.
Available at: http://www.pnas.org/content/100/4/1961.full
[Accessed December 13, 2009].

RODS Laboratory, Real-time Outbreak and Disease Surveillance Laboratory.
Available at: https://www.rods.pitt.edu/site/
[Accessed December 9, 2009].

Rolfhamre, P., Grabowska, K. & Ekdahl, K., 2004. Implementing a public web based GIS service for feedback of surveillance data on communicable diseases in Sweden. *BMC Infectious Diseases*, 4(1), 17.

Rolfhamre, P., Janson, A., Arneborn, M. & Ekdahl, K., 2006. SmiNet-2: Description of an internet-based surveillance system for communicable diseases in Sweden. *Eurosurveillance*, 11(5).
Available at: http://www.eurosurveillance.org/ViewArticle.aspx?ArticleId=626
[Accessed December 7, 2009].

Sebastiani, P. & Mandl, K.D., 2003. Biosurveillance and Outbreak Detection.
Available at: http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.93.7893
[Accessed December 13, 2009].

Serfling, R.E., 1963. Methods for current statistical analysis of excess pneumonia-influenza deaths. *Public Health Reports*, 78(6), 494–506.

Zhang, J., Tsui, F., Wagner, M.M. & Hogan, W.R., 2003. Detection of outbreaks from time series data using wavelet transform. In Omni Press CS, pp. 748–752.