An Overview of the Medical Data Surveillance System

Mr. William M. Pugh
2717 Curie Place,
San Diego, CA 92186-5122
USA

wpugh1@san.rr.com

Ms. Jamie K. Pugh
Space and Naval Warfare Systems Center
49490 Lassing Road
San Diego, CA 92152 – 6167
USA

jpugh@san.rr.com

ABSTRACT

In theater, it is essential that the services have a common medical system; therefore, a program was initiated to design, develop, and implement an automated system to process medical data on deployed personnel. Once it was demonstrated that medical information could be efficiently processed in theater, a research effort was undertaken to automatically analyze the data in order to detect deviations from historic illness patterns. Consequently, a system called the Medical Data Surveillance System (MDSS) was designed and developed as a Web enabled system for data analysis and reporting. The analysis capability of MDSS allows the user to compare the relative number of cases in one set of categories to those in another set across two different time intervals for a selected military unit. Or, the relative number of cases for two different military units can be compared during the same time interval. A unique set of signal detection algorithms called Dynamic Change-Point Detection (DCD) allows the user to select and analyze the entire population or those cases associated with a particular medical treatment facility (MTF), those cases associated with a particular Military Unit, or those cases associated with a group of Military Units that used a specified set of MTFs. Results are returned in a tabular format and as colored bar graphs. The analyses are conducted on the cases that have been reported, and on illness rates per thousand per day. Increasing and decreasing trends are identified in MDSS tables and graphs. The shifts and trends as well as bursts and outliers are labeled in the tables. Finally, MDSS generates a daily alert matrix for each MTF by automatically running the DCD algorithms for the illness and injury categories produced by MDSS.

1.0 INTRODUCTION

In the U.S. Navy, theater medical surveillance has its roots in a series of studies of illness among shipboard personnel that began in the late 1960s and continued through the 1970s. These studies were an effort to identify the factors related to the onset of illness among the crews deployed aboard U. S. Navy ships

Paper presented at the RTO HFM Symposium on "NATO Medical Surveillance and Response, Research and Technology Opportunities and Options", held in Budapest, Hungary, 19-21 April 2004, and published in RTO-MP-HFM-108.

[1-4]. Initially, it was necessary to collect the medical data by manually transcribing clinical information to data sheets, which were used to enter the data into an electronic database. As the shipboard studies continued, the manual data collection evolved into a semi-automated system where each patient encounter was documented on a form that could be optically scanned. Participating ships forwarded these sheets to the Naval Health Rsearch Center (NHRC) where they were scanned, entered into the database, and used to compile the Monthly Medical Services and Outpatient Morbidity Report for participating ships [5]. As automation of medical information in the field became more feasible, research efforts by NHRC led to the design and development of the Medical Data Tag for the capture and storage of medical information in theater [6]. The Field Medical Surveillance System was designed and developed by NHRC to assist with the diagnosis and tracking of illness incurred by personnel in field environments [7].

In theater, as elsewhere, it is essential that the different branches of the armed services have a common medical system. Therefore, the Office of the Secretary of Defense for Health Affairs (OSD [HA]) initiated the Theater Medical Information Program (TMIP) to design, develop, and implement an automated system to process medical data in theater. This effort began with each service determining its functional requirements. In the case of the Navy, SSC-SD was recognized as the Navy activity with the mission and necessary expertise for system engineering and development. Hence, SSC-SD was tasked to determine and document the naval theater medical information requirements. From the requirements that had been provided to OSD (HA), SSC-SD developed the Theater Medical Core Services (TMCS) as a prototype technology designed to meet many of the identified requirements. TMCS was able to capture selected clinical data, medical supply data, and staffing data; communicate that information up the chain of command; and use it to compile a variety of status reports.

Once it was demonstrated that medical information could be efficiently processed in theater, however, new questions appeared. How could one review all the reports that were submitted and be sure every anomaly was investigated? And, how can deviations from historic rates be detected as the data are being received? These questions were the subject of an independent research effort by J Pugh at SSC-SD. The culmination of this effort was a set of signal detection algorithms known as the Dynamic Change-point Detection (DCD) algorithms. The DCD algorithms provide the ability to evaluate new data upon receipt and:

- Rapidly set illness and injury baselines
- Set threshold values for various types of deviations from the baseline
- Flag changes from the baseline for a disease or disorder
- Determine when a change from the baseline began and ended

By systematically processing incoming data using the DCD algorithms, large quantities of data could be analyzed and deviations from the baseline could be flagged for investigation.

2.0 MDSS CAPABILITIES

To transition the array of surveillance tools to the Fleet, a research effort was initiated to create a system for medical surveillance of Navy and Marine Corps deployed forces. The primary objective of the system, called the Medical Data Surveillance System (MDSS) was to rapidly detect medical threats through the analysis of routinely collected patient data. Working as a team, personnel from NHRC contributed epidemiological expertise and technology developed through their field medical technologies program, and

SSC-SD personnel contributed statistical expertise and technologies developed through their intelligence, surveillance, and reconnaissance program. The team recognized that the key to conducting an effective medical surveillance program is to be able to rapidly aggregate and process data from a wide range of sources and then quickly disseminate the results. Thus, MDSS is a Web-enabled system. It is designed to accept data from the U.S. Navy's Shipboard Non-tactical ADP (SNAP) Automated Medical System (SAMS) and from the U.S. Military's Composite Health Care System's Ambulatory Record System. As data are acquired by MDSS, they are aggregated in several ways. These data are used to form the Disease and Non-Battle Injury (DNBI) reporting categories, categories for ill-defined conditions, Chemical-Biological-Radiological (CBR) groupings, key symptom categories, and the major categories of the International Classification of Diseases, Ninth Revision (ICD-9) currently used to document diagnoses in U.S. military medical records. The DNBI categories support the standard monitoring of deployed forces, the CBR groupings are designed to highlight effects of a CBR exposure, and the ill-defined condition and key symptom categories allow separate symptoms such as fever or diarrhea, to be trended and analyzed from data derived entirely from diagnostic codes.

After the data are received and stored, they can be used to automatically generate the DNBI report that summarizes DNBI incidence counts, incidence rates, lost work days, days of light duty and numbers of outpatients and inpatients. Similar reports are generated for ill-defined conditions, the CBR categories, key symptoms, major ICD-9 categories and reportable conditions. In addition, information needed for the Medical Event Report is produced for any patient with a condition included on the Tri-Service Medical Event List. A list of the specific individuals who contribute to the counts on these reports can be readily determined by clicking on the category of interest. Moreover, detailed information on any one of these individuals can be obtained by clicking on the person's name. Users can also create Ad Hoc reports by downloading the patient data into an Excel spreadsheet. Then using the Excel Wizard, the user can create tables and graphical displays of the information.

The analysis capability of MDSS includes DCD and 3 non-parametric statistics. The non-parametric statistics were imported from the Enhanced Consequence Management Planning and Support System developed with support from the Defense Advanced Research Projects Agency. The non-parametric functions allow the user to compare the relative number of cases in one set of categories to those in another set, across two different time intervals for a selected military unit. Or, the relative number of cases for two different military units can be compared during the same time interval. The DCD capability allows the user to select for analysis the entire population; or, the subset of cases associated with a particular medical treatment facility (MTF), those cases associated with a particular Military Unit, or those cases associated with a group of Military Units that used a specified set of MTFs. Then the user specifies the time period of interest and the illness category of interest. Results are returned in a tabular format and as colored bar graphs. The analyses are conducted on the number of cases that have been reported, and on illness rates per thousand per day if Population At Risk information has been provided to the system. Increasing and decreasing trends are identified in MDSS tables and graphs. The shifts and trends as well as burst and outliers are labeled in the tables. On colored graphical displays green is used for days within normal limits, yellow is used for days that show some increase from the background rate, and red is used to flag those days that are two standard deviations above the baseline, or when a trend or shift is statistically significant. Black is used for days that are three or more standard deviations above the baseline. Finally, the beginning point and ending point of trends are indicated when the color changes to or from green.

Because there is a difference in the operation of MTFs on weekends and holidays, there is a dramatic difference in the number of patients seen. MDSS adjusts for this phenomenon by developing and using separate baseline and threshold values for those days. Finally, MDSS generates a daily alert matrix for each MTF by automatically running the DCD algorithms for the illness and injury categories produced by MDSS, and it displays the results for each of the past 10 days. As a result, at the beginning of each day the user can view the status of each illness and injury category and see if there is any single day that differs from the others, and see if there are any temporal patterns in the recent past.

3.0 THE DCD METHOD

In this section the development of the DCD is discussed to help provide an understanding of this analytic technique and its value for medical surveillance. The approach starts with viewing medical surveillance as being analogous to a manufacturing problem where a production line is monitored for defective parts. From this perspective, an increase in the rate of defects would indicate the manufacturing process has been degraded. In an analogous manner, an increase in patients would indicate that the health protection process has been degraded. Further, in the development of DCD a distinction was made between two types of changes. One is a sharp change, a stark departure from the historic levels that often, just as rapidly, returns to the norm. This is known as a burst or outlier. The second type of change is generally a smaller deviation from historic levels but is sustained over an extended period of time. This type of change is known as a shift or trend. The first change may reflect a break in the process, while the second may reflect the need for an adjustment in the process or recalibration.

The recognized method for detecting bursts and outliers is the Shewhart test, and when detecting shifts and trends the CUSUM technique is used [8]. For the medical application, however, the CUSUM needed some extensions. First, once a change is detected, it is important to know when the change began and when it ended. Knowing the day that an increase in patient visits started can be useful in determining the cause of the change. Also, it is necessary to know if the number of cases decreased following the implementation of an intervention, when assessing the efficacy of the intervention.

In addition to establishing start and stop points, the CUSUM needed to be extended to small samples. Specifically, the need was to extend the CUSUM statistic so that that stable baseline and threshold values could be established in a few days rather than requiring 1 or 2 months of data before these parameters could be established. This is particularly important when deploying to a new area where no information is available on disease rates. This capability is also valuable when conducting Ad Hoc analyses because various subgroups can be readily investigated with respect to their unique historic values. The desired extensions to the CUSUM technique were derived by starting with the Neiman-Pearson fundamental theorem of statistics [9], thereby, generating the uniformly most powerful test. The resulting test was combined with the Shewhart test to form the DCD algorithms.

The advantage that DCD had over other methods was determined by conducting a Monte Carlo study where a series of random data sets were generated. Various types of signals (slopes, single points, square waves, saw tooth.) were prepared and embedded in random data (i.e., noise). Then, DCD and each of the alternative methods, including regression tests, moving average techniques and F-ratios, were used to find the signal within the noise. In each case the threshold was set so that random deviations exceeded the threshold 5% of the time. The power of each method was assessed by the frequency that the embedded signal was detected by the test and how long the signal was present before it was detected. In the case of DCD, separate thresholds were set for the Shewhart and CUSUM statistics. The results showed, if the signal was present during the entire period of observation, DCD was as good or better than any other method. However, if the signal began and/or ended one or more times during the interval, the DCD algorithms were superior to all other methods for all types of signals.

4.0 EXERCISES AND IMPLEMENTATIONS

Although various surveillance components have been demonstrated and tested over a period of years, MDSS as an integrated system was first exercised in Cobra Gold 2001. In that exercise it demonstrated immediate utility. Data were captured using SAMS and were supplied to MDSS every 15 minutes. In less than one week a baseline value and threshold were established from the incoming data. During the exercise unusual occurrences of dermatological problems was flagged by the system. Using the patient information report, the personnel located at the Joint Task Force Headquarters were able to determine that many of the cases were occurring within a particular unit. A team was sent to investigate and they found that the problem was caused by mite bites. It was found that because many of the personnel from the affected unit had not sprayed their mosquito netting, mites were getting through. Once the netting was treated properly the problem subsided. Because of the rapid detection, the ability to identify the population affected, and the intervention taken, the elevation in dermatological problems lasted just 2 days. As a result of this experience the patient drill-down feature was added to MDSS so that the individuals contributing to a particular total can be found simply by clicking on the display.

MDSS was also exercised in Kernel Blitz during 2001. On the one hand, users found that the system was useful for processing the quantities of data that were received. By using MDSS, environmental health personnel could ensure all deviations from baseline values would be flagged for investigation. On the other hand, a need was identified for an alerting mechanism that would prompt a person to run an analysis on the data being received. Consequently, the alert matrix was implemented so that a standard set of DCD analyses would be performed each day and displayed, thereby, providing an automatic alert.

Early in 2001 the US CINCPAC Surgeon recognized the need for a medical surveillance system in the Korean Theater of Operations (KTO). Therefore, a Theater Health Assistance Team was sent to South Korea to assess the potential of new and innovative technologies for enhancing health surveillance. As a result, a plan for developing a system that built upon the extant infrastructure was devised. Basically, the concept was to feed MDSS with the Standard Ambulatory Record Data generated by the CHCS Ambulatory Record System. An initial capability was developed and tested at the Naval Medical Center, San Diego. Then, in September 2001 a team implemented the system at the 121st evacuation hospital in Korea. Soon, it was found that the difference in hospital operations during weekends and holidays created an artifact that interfered with data analysis and interpretation. This problem was resolved by separating full staff days from reduced staff days and analyzing the data sets separately so that each one is evaluated with respect to its own baseline and threshold.

Currently, MDSS is being used routinely in KTO for medical surveillance and, in September 2003, MDSS was formally transitioned to the Theater Medical Information Program. The technology described in this paper is the subject of one or more pending patent applications and is available for licensing from the U.S. Navy. Licensing inquiries may be direct to Ms. Jamie Pugh.

REFERENCES

[1] Environments. II. Demographic, Social Background, and Occupational Factors. *Military Medicine* 135:453-458.

[2] Erickson, J.M., Dean, L.M. and Gunderson, E.K.E. 1976. Morbidity as a Factor in the Operational Effectiveness of Combat Ships. Naval Health Research Center, San Diego, Calif. Report No.76-8.

[3] Pugh, W.M., Gunderson, E.K., Erickson, J. M., Rahe, R.H. and Rubin, R.T. 1972. Variations of Illnes Incidence in the Navy Population. *Military Medicine* 137:224-227.

[4] Harvey, R.A. and Hermansen, L.A.1978. Determinants of Gastrointestinal Disorders Aboard Navy Ships. *Military Medicine* 143:635-638.

[5] Hermansen, L.A., Jones, A.P. and Butler M.C. 1980. Development of an Outpatient Medical Treatment Reporting System for Shipboard Use. *U.S Navy Medicine* 71:16-21.

[6] Tropeano, A.M. and Pugh, W.M. 2001. Review of the Naval Health Research Center's Development of Medical Information Systems for Far-Forward Echelons of Care, 1983 to 1997. *Military Medicine* 166:656-663.

[7] White, M.W., Pugh, W.M., Hanson, K., Angus, J.E. and Show, I.T. 1996. A Field Medical Surveillance System for Deployed Forces: A Conceptual Model. Naval Health Research Center, San Diego, Calif. Report No. 96-12.

[8] Ryan, T.P. 1989. *Statistical Methods for Quality Improvement*. John Wiley & Sons, New York, New York.

[9] Brunk, H.D. 1975. *An Introduction to Mathematical Statistics*, Xerox, College Publishing, Lexington, Mass.