A Syndromic Surveillance Model for Influenza-Like-Illnesses and Intentional Release of Biological Agents Based on Sequential Bayesian Control Technique

K. D. Zamba
College of Public Health
Department of Biostatistics
The University of Iowa
200 Hawkins Dr C22C GH
Iowa City, IA 52242
gideon-zamba@uiowa.edu

Panagiotis Tsiamyrtzis
Department of Statistics
Athens University
of Economics and Business
76 Patission Str,10434
Athens, Greece
pt@aueb.gr

Douglas M. Hawkins
School of Statistics
University of Minnesota
313 Ford Hall
224 Church Street
Minneapolis, MN 55455
doug@stat.umn.edu

SUMMARY. Protecting the US population against bioterrorism has been an important task facing US officials, policy makers, health care providers and the Center for Disease Control (CDC) following September 2001. The period after September 11 has raised the level of awareness of incorporating medical based intelligence functions such as Influenza-Like-Illnesses (ILI) observed during visits to emergency rooms (ER). Developing a control technique for ILI however is a complex process which involves the unpredictability of the time of emergence of influenza, the severity of the outbreak and the effectiveness of influenza epidemic interventions. Furthermore the need of detecting the epidemic in an on-line fashion makes any influenza-based control even more challenging. This complexity and uncertainty around influenza have kept many scientists away from tackling

preparedness based on ILI. In this work, we present a Bayesian model for the course of ILI. This model uses a recursive and sequential update by finding the posterior distribution at each stage and setting it as a prior distribution of the next stage to chart the discrepancy between the observed and the predicted percentage of ILI. The prior was coupled with a threshold model to account for the seasonality in the distribution of ILI and the severity of the epidemic.

KEY WORDS: Bayesian Dynamically Updated Mixture; Control Chart; Error management; Historical Data Set, Phase I,II; Statistical Process Control.

1 Introduction and Background

Among many other potential threats, bioterrorism, is the new frustration that has turned to be a new wave of research interest and research opportunity. Surveillance usually relies on intelligence functions. Intelligence functions are sources of information usually given by individuals, animals, measurement instruments and many other sources such as disease syndromes observed on people and reported to sentinels. Disease surveillance is critical for detecting and responding to natural outbreaks as well as biological terrorism, and for addressing serious public health concerns. A surveillance system can help identify the source and cause of exposure and reduce consequences by directing health officials to a rapid intervention. There are currently many surveillance programs in the US such as BioWatch, Guardian, and the Real-time Outbreak Disease Surveillance (RODS). The aspect of surveillance that looks for signs and symptoms characterizing an abnormality in a given population takes the name of syndromic surveillance. Syndromic surveillance applies to surveillance using health-related data that precede diagnosis and signal a sufficient probability in a case of an outbreak to warrant further public health response. Syndromic surveillance relies on intelligence functions such as ER observations, over-the- counter(OTC) sales, veterinary data, agricultural data, medical and public health information to provide valuable measure on outbreaks and intentional releases; Greenko et al (2003); Cochrane et al (2003); Barthell et al (2004). The

ideal syndromic surveillance system must shorten the delay between outbreak and intervention to allow timelier intervention to remove the threat, immunize the population and minimize casualties. Syndromic surveillance is an emerging field with very few analytical and statistical tools. Buehler et al (2003) provides an overview of the use of Syndromic surveillance compared to clinicians' reports to yield a diagnosis in the event of bioterrorist attack; Green and Kaufman (2002); provides overviews and examples of syndromic surveillance systems; Pavlin (2003) describes the steps of disease outbreak investigation and syndromic surveillance. Mandl et al (2003) stresses the importance of surveillance system quality and the integration of syndromic surveillance with public health response. A thorough discussion and literature on this emerging field, the technologies and the decision support systems can be found in Bioterrorism Preparedness and Response (2002) or at `http://www.ahrq.gov/clinic/epcsums/bioitsum.htm`. Currently, the Federal Government relies on the CDC which receives information from the National Electronic Telecommunication Surveillance System (NETSS) for surveillance purposes. The overall capabilities of these systems to detect biological attack however are low and the period of time between the first reported case and identification of the problem sometimes exceeded two weeks; see for example Armstrong et al (2004). The efficiency of most surveillance systems however is contingent upon the biological agent the systems are designed to detect; Kosal (2003). Also, much controversy had evolved around the use of environmental senors in surveillance; knowing that not all biological agents have the same size or the same dissemination potential and not all can be detected by a sensor. Many authors have warned about the danger of having absolute reliance on sensors and have suggested that surveillance techniques should consider medical data as intelligence function. These authors reason around the fact that sensors are more efficient in battlefield environments and in war scenarios, where they are highly sensitive to the chemical cloud resulting from chemical attack, when only meteorological conditions and the geometry of the field allow it. Their argument was supported by the fact that an Environmental Protection Agency (EPA) sensor was placed just blocks away from the World Trade Center towers, but following the collapse of the towers on 9/11, it did not register the incident-only when the wind direction changed on September 12 did the sensor became

aware of the incident; Armstrong et al (2004); National Institute of Justice (NIJ) Guide 101-00, December 2001, 23-25. This paper is a technical work; as such, describing the effect and interaction of biological agents will be beyond its scope. For a good review on biological agents, their size, their reproductive machinery, the weapons they create after release, see Campbell and Reece (2002); Meltzer et al (2001); Armstrong et al (2004). Our work does not dismiss the use of sensors in surveillance system. It is a data driven technique that uses as intelligence function %ILI data.

2 Influenza-Like-Illnesses: Use of Medical and ER data

What makes medical data appealing is their reliance on symptom recognition. Medical diagnosis give more information about outbreaks and can be an appealing intelligence function. Medical surveillance reporting ranks high and its use seems more particularly prudent since not only it helps reject a null of no intentional release in case of attack but also it has the potential for answering general public health concerns such as emerging infectious diseases. The CDC has standardized a list of the most likely bioterrorism agents. The top six with highest dissemination potential make up the category A and are: anthrax, botulism, plague, smallpox, tularemia, and viral hemorrhagic fever. In the October 19th, 2001 Morbidity and Mortality Weekly Report (MMWR), the CDC had provided illness patterns and diagnostics clues that might indicate an unusual infectious disease outbreak associated with intentional release of biological agents–most of which have early symptoms similar to influenza-fever, dyspnea cough and chest discomfort for anthrax; fever cough hemoptysis and chest pain. For detailed information, see MMWR vol.50/ No 41 and relevant publications such as Arnon et al (2001); Inglesby et al (2000); Henderson et al (1999); Dennis et al (2001); Inglesby et al (1999). In order to target data providing positive identification of pathogens for a known agent the focus has been on data related to influenza reports from ER since the first signs of bioterrorism activity will be sensed at emergency rooms, from X-ray cases, fever cases, and cough related illnesses that cluster in one geographic location. Based on real-time gathering and analysis of data on patients seeking care daily with certain syndromes one can find early signs

of intentional release. A number of such medical based surveillance are being developed in the District of Columbia Department of Health, in conjunction with the Maryland and Virginia state health departments. The simulated results obtained from the study of these data were not quite encouraging Stoto et al (2004).

2.1 ILI data and their own Challenges: Need for Sophisticated Tools

ILI data present their own challenges thereby demanding more sophisticated analysis tools. The study of influenza related data is complicated by the fact that the course of the flu differs from year-to-year and that the three types of influenza (A,B,C) antigenic properties of their H and N molecules can mutate to create a new variant every year. These changes in the subtypes and the combinations they create explain why the flu vaccine must be reformulated each year and also provide another reason why the flu vaccination program cannot eradicate influenza. Any attempt of influenza eradication can be likened to shooting a target that is constantly moving. The ILI historical data set is far from providing a consistent estimate of the true %ILI because surges seen in ILI end with its specific year; the beginning of each flu season will come with its own challenges and will generate %ILI specific to that year–brief, a new data process to be studied. Even in case the same subtype of influenza hits a given geographic area over consecutive years, it does neither guarantee having the same maximum %ILI nor the same distribution parameters as the previous years due to the fact that the susceptible and the exposed populations are dynamic. It is to be noted too that like every other infectious disease, influenza follows either the classical SIR or the more general MSEIR epidemiological transfer model with immune class M, susceptible class S, exposed class E, infective class I and recovered class R. However, the parameters of these models are far from being fixed; they will be random due to the above mentioned reasons. Figure 1 is generated from data provided by the sentinel program in the US. One can obviously see the irregularity, unpredictability, and randomness of the %ILI through these figures.

Figure 1 about here

To be optimistic, we will focus on their common feature such as their individual starting low %ILI, their peak corresponding to a peak infective time, and their decrease nearly symmetrical to their rise. Monitoring these readings can be likened to a year-to-year control problem with unknown, unstable or partially known parameters. Another complication comes with a potential jump even under the hypothesis of no intentional release. The historical data set gives us only some information about the shape of the %ILI and its irregular and non-stationary trend. Our proposed work uses advanced quality control surveillance techniques based on Bayesian reasoning and is adapted to these data sequentially. It is a technique capable of giving enormous improvements in sensitivity by carrying information from one time period to the next to more clearly show changes in prevalence. We set up our model to act in detect-to-warn fashion. Our technique, although controlling the probability of type I error will also be sensible to the probability of type II error and consequently to an out-of-control condition.

2.2 Previous Work

There are many works on infectious diseases and their mathematical modeling using the classic SIR (Susceptible Infectives Recovered) epidemic and endemic models that consider variations in population characteristic over time as intelligence function. Works such as Hethcote (1989; 2000) approach infectious diseases modeling by solving differential equations on a population dynamics. A complete literature review on these works can be found in Hethcote (2000). Quality control methodologies applied to infectious disease have received less attention in the papers. It is to be noted that the CDC has routinely applied the cumulative sum (cusum) technique to laboratory-based data for outbreak detection; see for example Hutwagner et al (1997). A thorough review on the theory of cusums can be found in Hawkins and Olwell (1998). Another quality control work in connection with infectious disease is the compound smoothing by Stern and Lightfoot (1999). Attention has been drawn to timely assessment of Influenza deaths through the use of ARIMA models in Simonsen et al (1997). Hutwagner et al (2003, 2005) used aberration detection techniques on simulated data to study intentional biological activities. One advantage of quality

control techniques is that they can be applied both off and on-line.

3 Our Solution: The Bayesian Approach

In traditional quality control, one gathers a substantial data set for data cleaning process called Phase I study. Phase I data are used for a phase II analysis consisting of on-line charting. The problem we are facing cannot be solved by a traditional method due to the challenges we have mentioned earlier. Besides, our setting forces us to do an inspection of each datum observed and test it to see whether it is the result of an intentional release. Thus, an appropriate control technique for individual measurements with unknown or partially known baseline information is needed. There are many new statistical tools to handle individual observations in low volume productions and start up processes. Among them are works in change point methodologies by Hawkins et al (2003), Hawkins and Zamba (2005,a,b); that rely on i.i.d normal assumption of the readings. On Bayesian arena, few works have been conducted in process control area. Some works are Woodward and Naylor(1993); Tsiamyrtzis and Hawkins (2005). In case of random baseline like the one we witness with the %ILI data, to our knowledge, no quality control work has been done so far.

3.1 Modeling

Any control theory can be likened to hypothesis testing problem. The one fitting our context is defined as:

$$\begin{cases} H_0: & \text{Normal course of flu} \\ H_1: & \text{Some unusual activity.} \end{cases} \tag{1}$$

At time t, the observed %ILI (y_t) is available. As we receive these data sequentially, our goal is to detect the beginning of an epidemic or unusual activities as fast as possible. The true %ILI in the

population denoted by θ_t will be modeled as follow:

$$\theta_t = \theta_{t-1} + \epsilon_t, \qquad \text{where} \qquad \epsilon_t \sim \begin{cases} N(0, \ \sigma^2) & \text{with probability} \quad p \\ N(\delta, \ \sigma^2) & \text{with probability } 1-p \end{cases}$$

σ^2 represents the time-to-time variation and δ a jump in the %ILI occurring with probability $1-p$. Thus, we model the %ILI as a random walk on which we superimpose a possible jump indicating the beginning of an epidemic. We have considered positive jumps only for reasons we will explain later. Given that we consider only positive jumps in the model above we can use it only to determine whether an epidemic have started and when it has started. The model will not be able to detect the full course of the ILI i.e. how the epidemic will evolve during the flu season. Had we introduced both positive and negative jumps, we should be able to detect the sinusoidal shape of ILI; with the tradeoff to add more complexity to our model. In simple words this means that the proposed model can be used as a surveillance tool able to sound an alarm as soon as an unusual activity starts. It will also be able to follow the course of the flu to its peak infective time. Of course, we are not able to observe the true parameter θ_t directly; instead, y_t is available and assume

$$y_t|\theta_t \sim N(\theta_t, \tau^2);$$

where τ^2 represents the variability related to inaccuracy of recording ILI patients. ILI has been defined as fever greater than 100^o F, and cough and/or sore throat. From a medical perspective, some report of ILI might be subjective; not only variability comes from doctors but also from health care providers. Also to be noted that many other diseases host around ILI. Thus, it sounds reasonable that τ^2 quantifies these variabilities.

At time $t = 0$, (i.e. before we see any data points) we define an initial prior distribution:

$$\theta_0 \sim N(\zeta, \sigma_0^2).$$

Following Tsiamyrtzis and Hawkins(2005) the posterior distribution of $\theta_t|(y_1, \ldots, y_t)$ at every $t = 1, 2, \ldots$ is given by

$$p(\theta_t|y_1, \ldots, y_t) = \sum_{i=0}^{2^t-1} \alpha_i^{(t)} N\left(\theta_i^{(t)}, \hat{\sigma}_t^2\right)$$

with the variance, weights, and means obeying the following recursive equations:

$$\hat{\sigma}_t^2 = (1 - K_t)\tau^2 = K_t \left(\sigma^2 + \hat{\sigma}_{t-1}^2\right)$$

and for $j = 0, 1, \ldots, 2^{t-1} - 1$

$$\begin{array}{rclcrcl} \alpha_{2j}^{(t)} & = & p\, \alpha_j^{(t-1)} m_j(y_t)/NC & & \theta_{2j}^{(t)} & = & K_t \theta_j^{(t-1)} + (1 - K_t) y_t \\ \alpha_{2j+1}^{(t)} & = & (1-p)\, \alpha_j^{(t-1)} m_j^*(y_t)/NC & & \theta_{2j+1}^{(t)} & = & K_t \left(\theta_j^{(t-1)} + \delta\right) + (1 - K_t) y_t \end{array}$$

where:

$$K_t = \frac{\tau^2}{\tau^2 + \sigma^2 + \hat{\sigma}_{t-1}^2}, \quad m_j(y_t) = \frac{exp\left\{-\frac{\left(y_t - \theta_j^{(t-1)}\right)^2}{2\left(\tau^2 + \sigma^2 + \hat{\sigma}_{t-1}^2\right)}\right\}}{\sqrt{2\pi\left(\tau^2 + \sigma^2 + \hat{\sigma}_{t-1}^2\right)}}, \quad m_j^*(y_t) = \frac{exp\left\{-\frac{\left(y_t - \theta_j^{(t-1)} - \delta\right)^2}{2\left(\tau^2 + \sigma^2 + \hat{\sigma}_{t-1}^2\right)}\right\}}{\sqrt{2\pi\left(\tau^2 + \sigma^2 + \hat{\sigma}_{t-1}^2\right)}},$$

$$NC = \sum_{j=0}^{2^{t-1}-1} \left[p\, \alpha_j^{(t-1)} m_j(y_t) + (1-p)\, \alpha_j^{(t-1)} m_j^*(y_t)\right]$$

Given that the model is a mixture of two Normal distributions, the posterior after t data points become available is obtained recursively as a mixture of 2^t Normal components. As t grows the number of components in the mixture increases exponentially. Using 2^t Normal components when we have t data points is probably an overkill. Most of these Normal components have tiny differences on the means and almost all will have negligible weights. This redundancy can be reduced by approximating the exact distribution with another distribution that has fewer components. This can be done for example by an algorithm proposed by West (1993), where the idea is "to collapse or cluster mixture components by simply replacing nearest neighboring components with a single average component".

3.2 Inference

As y_t becomes available ($t = 1, 2, \ldots$) our focus will be on the following sequence of hypotheses:

$$\left\{ \begin{array}{ll} H_0: & \text{There is regular occurrence of ILI at week } t \\ H_1: & \text{Something unusual started at week } t \end{array} \right\}.$$

This can be easily translated to

$$\left\{ \begin{array}{ll} H_0: & \theta_t \leq K \\ H_1: & \theta_t > K \end{array} \right\};$$

where K is some pre-specified constant indicating a major public health signal of an epidemic (usually this is set around 2%). We use a Bayesian sequentially updated approach to perform these hypotheses testing. At each step we find the posterior distribution upon which we base our decision to accept H_0 or H_1. Regarding the tests, the setting will dictate what we should be more concerned with; whether type I errors are costlier than type II. In case we are more concerned with type II error than with type I we assess accordingly c_I and c_{II} to be their costs respectively and we base our decision on the following Bayes test:

$$\left\{ \begin{array}{l} \text{accept } H_0: \theta_t \leq K, \quad \text{if} \quad P(\theta_t \leq K | y_1, \ldots, y_t) > \frac{C_{II}}{C_I + C_{II}} \\ \\ \text{reject } H_0: \theta_t \leq K, \quad \text{if} \quad P(\theta_t \leq K | y_1, \ldots, y_t) < \frac{C_{II}}{C_I + C_{II}} \end{array} \right\}.$$

Implementing a decision rule based on the Bayes factor (Jeffreys, 1948) is also plausible. If we denote $P(H_i|y)$, $i = 0, 1$ the posterior probabilities of the hypothesis H_i, $i = 0, 1$ when the data y were observed and $P(H_i)$ the prior probabilities of H_i, $i = 0, 1$ then the Bayes factor, B, is defined as the ratio of the posterior odds of H_0 to the prior odds of H_0:

$$B = \frac{P(H_0|y)/P(H_1|y)}{P(H_0)/P(H_1)}$$

Jeffreys (1948) provides a table of cutoff values for B to be used when deciding about the rejection of H_0. In our study at every time t we have available both the posterior $\theta_t|(y_1, \ldots, y_t)$ and the prior $\pi(\theta_t)$ which we can use to calculate the Bayes factor B and decide whether the mean has shifted above the upper threshold value K.

3.3 Handling Non-Fixed p : Use of Threshold Model

For our proposed model (1), the probability of not having a jump p is constant over time. Under this model assumption, updating the posterior of θ at time t to the prior of θ at time $t+1$ gives:

$$\pi(\theta_{t+1}|y_1,\ldots,y_t) \sim \sum_{i=0}^{2^{t-1}-1} \left[p\ \alpha_i^{(t)} N\left(\theta_i^{(t)}, \sigma^2+\hat{\sigma}_t^2\right) + (1-p)\ \alpha_i^{(t)} N\left(\theta_i^{(t)}+\delta, \sigma^2+\hat{\sigma}_t^2\right)\right]$$

The reality is, even under the assumption of no intentional release, the course of the flu presents some challenges due to natural outbreak. A close look at the pattern of the flu reveals that the first sign of natural outbreak yielding a jump in the %ILI is followed by successive jumps until the epidemic reaches its peak infective value. This suggests the idea of updating the probability of no jump as a function of t since p will not remain constant all the way. The update can be done based on either theoretical ground or on expert opinion. In the Bayesian framework, one way to handle the unknown p is by putting a hyperprior on this nuisance parameter. The conjugate choice of Beta distribution seems particularly prudent due to the fact that it covers an extremely large number of possible scenarios. One can always play around its location and shape parameters to cover specific prior beliefs about p. In more practical words, any expert opinion or prior belief about p will find appropriate parameters (α,β) in the beta family distribution that will match its distribution as function of time. So we have the following:

Lemma 1 *If $\pi(p) \sim Beta(\alpha,\beta)$ will be used as a hyperprior for p, then the updated prior distribution of θ at time $t+1$ is given by :*

$$\pi(\theta_{t+1}|y_1,\ldots,y_t) = \sum_{i=0}^{2^t-1}\left[\frac{\alpha}{\alpha+\beta}\right]\alpha_i^{(t)} N\left(\theta_i^{(t)},\ \ \sigma^2+\hat{\sigma}_t^2\right) + \sum_{i=0}^{2^t-1}\left[\frac{\beta}{\alpha+\beta}\right]\alpha_i^{(t)} N\left(\theta_i^{(t)}+\delta,\ \ \sigma^2+\hat{\sigma}_t^2\right)$$

Proof of this Lemma can be found in Appendix A. As we observe the updated prior is an exact replica of what we have obtained earlier had we started out by a constant p (replace p with $\alpha/(\alpha+\beta)$ and $1-p$ with $\beta/(\alpha+\beta)$). This carries a twofold message; the first of which is an equivalence of two reasonings: putting a beta conjugate on p or starting out with a constant probability of no jump does not change the distributional properties of $\pi(\theta_t \mid \ldots Y_{t-1})$. The second message is that the iteration formula of the posterior and its relationship with decision making properties will

be globally preserved when switching from one value of p to another. In the light of the above remarks, we use a threshold model based on successive difference of posterior coverage probabilities as a mean of updating the probability of no jump. The reasoning behind this idea finds its root in a fact. Before a natural outbreak, the posterior coverage probability of the alternative hypothesis will be relatively low. The coverage probability will have a sudden jump at the first time of natural outbreak and will increase afterward. By putting a threshold on successive differences in posterior coverage, we also allow the model to accept natural outbreaks as part of its null with the tradeoff to have a low sensitivity if an intentional release occurs at the same time as natural outbreak. The choice of threshold ($h = 0.15$) on successive differences is based not only on expert opinion but also upon observing prior data on how the epidemic curve slopes up around its inflection point. Defining p the following way seems fitting:

$$p = p_t = \begin{cases} p_1 & \text{if } P_{\theta,t-1} - P_{\theta,t-2} \quad \leq 0.15 \\ p_2 & \text{if } P_{\theta,t-1} - P_{\theta,t-2} \quad > 0.15 \end{cases}$$

$P_{\theta,k}$ is the posterior coverage probability of the alternative hypothesis at time k; p_1 and p_2 are selected according to prior data on %ILI and are such that $p_1 > p_2$.

3.4 Predictive Distribution and Error Management

Another interesting aspect of the proposed methodology is that at any given week in $\{0, \ldots, t_p - 1\}$, where t_p is the time the epidemic reaches its peak, one can obtain the predictive distribution of the future observable(s). The predictive distribution takes into account uncertainties about both the unseen observation and the parameter. In case we are interested for one step ahead predictions one can easily derive

$$\begin{aligned} P(y_t \mid y_1, \ldots, y_{t-1}) \quad \sim \quad & \textstyle\sum_{i=0}^{2^{t-1}-1} p \;\; \alpha_i^{(t-1)} N\left(\theta_i^{(t-1)}, \;\; \tau^2 + \sigma^2 + \hat{\sigma}_{t-1}^2\right) \\ & + \textstyle\sum_{i=0}^{2^{t}-1} (1-p) \;\; \alpha_i^{(t-1)} N\left(\theta_i^{(t-1)} + \delta, \;\; \tau^2 + \sigma^2 + \hat{\sigma}_{t-1}^2\right). \end{aligned}$$

The predictive distribution can be used to provide point/interval estimates for the next observable.

Regarding surveillance issues, a very challenging problem will be to detect unusual activity in epidemic time. The error management approach is proposed as a solution to this problem. The technique consists of charting the discrepancy between the observed and the one-step-ahead predicted value for aberration detection; assuming our model has good sensitivity and specificity.

4 Applying to real epidemic data

Figure 1 shows 10 years of weekly %ILI through the sentinel program in the US. We use the series from 1994 until 2002 to elicit the nuisance parameters of our model and apply the model to the last 3 years to judge how well it could approximate sequential real-time or near real-time ILI data. Based on the data from 1994-2002, we elicited: $\zeta = 1, \sigma_0^2 = \sigma^2 = 0.05$ and $p_1 = 0.9$ and $p_2 = 0.3$. Concerning the size of the jump, apart from studying the successive differences of the "epidemic" weeks of the previous years we were interested in picking a value that will be "big" enough to matter and at the same time "small" enough to be immediately seen. Based on this philosophy we chose $\delta = 5\sigma$. For the variability $\tau^2 = 0.3$ external info was used, while for K we picked 2% baseline level.

Applying the proposed model (keeping all components in the mixture) we obtained the exact posterior distribution of the % ILI at each week of the three years under study. A time series plot of the data along with the posterior mean (Bayes estimate under squared error loss) can be seen in Figure 2.

Figure 2 about here

The posterior coverage probabilities of the alternative hypothesis ($H_1 : \theta_t > 2$) along with the Bayes factors that can be used to judge when the epidemic started are provided in Table 1.

Table 1 about here

For the 2002-03 year in week 13 we have a significant increase of the posterior coverage probability which is further supported from the Bayes factor where for first time we drop below 1 (evidence

against H_0 according to Jeffreys, 1948). Regarding the 2003-04 period in week 7 we obtain sufficiently large posterior coverage of H_1 and again this is in agreement with the Bayes factor results. Finally for the 2004-05 year week 12 seems to ring the bell for the beginning of the epidemic with 0.3004 as posterior coverage probability of H_1 and a Bayes factor of 0.6761.

Apart from inference though our model can be used for prediction purposes as well. So at each stage of the process the one step ahead forecasting distribution is available in closed form. In our setup we are particularly interested in assessing at each week t, $P(y_{t+1} > K|y_1, \ldots, y_t)$ i.e. the predictive probability that the next observable will exceed the prespecified threshold value. A time series plot of these probabilities for the three successive years under study can be seen in Figure 3 where our predictions are in good agreement of what we established with our decision scheme.

Figure 3 about here

Sensitivity analysis regarding the effect of misspecifications of the nuisance parameters values will be examined extensively in the simulation study.

5 Simulation Study and Sensitivity Analysis

The aftermath of an intentional release may depend on the biological agent, its incubation time, its dissemination potential, the method of dispersion, the population dynamics/contact, the quantity of agents released, the quality of intervention techniques and many other parameters. Note that we are neither estimating transmission rate nor building an epidemiological model for a biological agent in our context. We are building a fast detection technique if the population is under siege. We do not consider the latent period of time between release and infection; after all, we will not know the exact time of release and will only take action after seing the effects resulting from infection. Thus, our focuss is mainly on the time between infection and intervention; during that time, patients are infectious and show flu-like symptoms; they will be infectious for only a period of time and either die or recover after appropriate intensive care. We also operate under the assumption that infection is reflected on the %ILI. In that time, infectious individuals will continue to transmit the

disease and a cumulative number of deaths and ILI cases will occur before the outbreak comes to the attention of health care officials. We also assume that once an outbreak has been identified, further transmission of the disease will stop or slow down since health care officials would have taken eradicative actions. Thus, we use the following approach to simulate the temporal evolution of the epidemics and to assess the sensitivity analysis.

5.1 Simulation Setting

One thousands (1,000) epidemics have been simulated to study the model's performance for various combinations of nuisance parameters. In the following we define, n_1 (true model value) as the first time that the true (known from simulation) %ILI exceeds $K = 2\%$; and n_2 (decision scheme value based on posterior) as the time our model signals that the %ILI exceeded the decision threshold K. The study of $n_2 - n_1$ will offer useful insight on the performance of our approach. Another parameter we will define at this point is n^*, the starting time of the epidemics. Even though n^* might seem random, in order to have sound comparison in our simulation we have assumed it to be fix by choosing $(n^* = 5, 8)$. So, we conduct a pair of simulations with the same combination of nuisance parameters and differing only by the value of n^*. Another important issue to clarify at this point is the choice of the peak infective time as the Historical Data Set (HDS) did not give any clear picture about where the peak will be from year to year. The randomness observed in the HDS should cause our series to be of varying length. We decide to make the series length fixed for the sake of the simulation to have a realistic ground for comparison. From the HSD available to us, it is reasonable to expect the peak infectivity around time point 15. We therefore decided to fix the series length at 15.

To generate the dataset corresponding to the true parameter θ_t, we start picking out n^* and followed these steps:

- For $t = 1, \ldots, n^*$

$$\theta_t | \theta_{t-1} \sim N(\theta_{t-1}, \sigma^2)$$

with initial prior

$$\pi(\theta_0) \sim N(\zeta, \sigma^2).$$

- For $t = n^* + 1,\ldots,15$

$$\theta_t|\theta_{t-1} \sim N(\theta_{t-1} + \delta, \sigma^2)$$

Once $(\theta_1, \theta_2, \ldots, \theta_{15})$ are generated we added on each a random $N(0, \tau^2)$ error corresponding to the measurement error component to produce the data $(y_1, y_2, \ldots, y_{15})$. These data will be used to run our model. We will call True Settings (T.S.) the hyperparameter values used to generate the θ's and the y's. We consider the following combinations for T.S: $\sigma = (0.1, 0.2)$, $\tau = (0.4, 0.6)$ and $\delta = (5\sigma, 7\sigma)$ yielding a $2 \times 2 \times 2 = 8$ different parameter combinations. The probability of not having a jump p is 1 for time 1 till n^* and 0 from time n^*+1 until 15 in the true setting. Our model uses values of $p = (0.7, 0.9)$ as starting values for the sensitivity analysis. Each of these p will be updated to 0.3 the first time successive differences of posterior coverage probability of the alternative is bigger than 0.15. These create a $2 \times 2 \times 2 \times 2 = 16$ different combinations that we call Decision Settings (D.S.) (8 combinations for each p). In each of the (T.S., p) combinations, we misspecify one of the three T.S. parameters to see how robust our model will be to parameter misspecification. To illustrate, if $(\sigma_1, \tau_1, \delta_1), (\sigma_1, \tau_1, \delta_2), \ldots, (\sigma_2, \tau_2, \delta_2)$ are the different combinations of T.S. and $(\sigma_1, \tau_1, \delta_1, p = 0.7)$ is the true setting for $p = 0.7$, then the remaining seven i.e. $(\sigma_1, \tau_1, \delta_2, p = 0.7), \ldots, (\sigma_2, \tau_2, \delta_2, p = 0.7)$ are misspecifications. For each of the 8 T.S. parameters combinations 1000 epidemics of length 15 were generated. Each of them was run against 7 misspecifications and one correct choice of σ, τ and δ. In sum, for each value of n^* we have a total of $8 \times 8 \times 2 = 128$ different scenarios.

We also define:

- Correct Alarm (C.A.) cases where $n_2 = n_1$, i.e. we signal at the right time
- Missed Alarm (M.A.) cases where $n_2 > n_1$, i.e. we signal with a delay

- False Alarm (F.A.) cases where $n_2 < n_1$, i.e. we signal earlier than the event has actually occurred.

In Table 2 we present the simulation results for the $2 \times 8 = 16$ cases where T.S. values coincide with the D.S values (i.e. there are no model misspcifications). The settings where the parameters are misspecified are not tabled but are rather commented on as these settings will add to the length of the paper.

Table 2 about here

To have further insight in both the missed and false alarm cases the bar graphs of $n_2 - n_1$ are provided in Figure 4 for all the cases of Table 2.

Figure 4 about here

5.2 Summarizing the Simulation Results

- There is only a tiny difference between the $p = 0.7$ and $p = 0.9$ blocks (that is why the $p = 0.7$ was the only case presented in Table 2). After all, they have been both updated to the same value of 0.3 after successive posterior coverage probabilities have exceeded the threshold.

- The model seems to have a slightly better performance for early activities; the case $n^* = 5$ has an overall faster detection than $n^* = 8$. These differences can be observed in the C.A. and F.A. departments. It seems like for larger values of σ and/or τ, the F.A. percentage increases in the 8^{th} week compared to that of the 5^{th}, thereby dropping the C.A. percentage. This can be explained by the fact that if the beginning of the epidemic is moving to the right, then it is more likely to switch from the original value of p to 0.3; thus making our model to react faster to even tiny jumps and causing a slight increase in the F.A. department.

- Under the correct model specifications where the D.S. parameter values coincide with the T.S. values the performance of our model depends mostly on the ratio of the σ/τ and to a much lower extend on the value of the δ. More precisely as σ decreases and/or τ increases then

our model's performance gets worst. This is something which is expected since for "high" values of τ related to "small" values of σ the true underline process will be masked by the measurement error variability affecting the model's performance in detecting jumps. This is in direct connection with the effect of δ where the larger the jump the easier will be to be detected, even for very noisy epidemics.

- The simulation results for the various model misspecifications showed that the model is more sensitive to wrong estimation of σ/τ than to δ. The overall information across these simulations is that when σ/τ is underestimated, the percentage of C.A. decreases causing the F.A. and M.A. to increase. The model's performance is in general robust to δ misspecifications except the cases where σ/τ is very small, where underestimation of δ leads to a decrease of the C.A. and a corresponding increase of the percentage of F.A..

- We have also explored the possibility for the measurement errors to be less than the model error. This scenario in fact yields a setting where the %ILI recording is done more accurately. This setting is expected to produce the least noisy epidemics. Our simulation results were impressive in this case. The F.A rate dropped dramatically yielding to an increase in the C.A. and M.A. departments. The higher M.A. rate observed in these settings might look misleading at first blush, but nearly all M.A and F.A. occurred within one observation of n_1.

6 Concluding Remarks

The preceding proposal, by considering only positive jumps, might seem like a partial solution to the surveillance problem. It is not the case. In fact, any unusual activity outside the flu season will be highly suspicious; the same suspicion holds for any surge in %ILI when the outbreak's infective proportion should be decreasing toward zero as a function of time. The portion of the curve that is more likely to be favorable to a flu-like biological activity without being seen by a naked eye is the increasing portion. Another reason why this portion of the epidemic curve is more important to us is that according to the hallmark of a typical epidemic outbreak (flu included), when the

epidemic reaches its peak, it will decrease to zero roughly symmetrically to its rise. By then, and having accumulated enough data, one can thus fit a sinusoidal function to the data and carry a one step ahead prediction along with an error management that consists of charting the discrepancy between the observed and the predicted response. One of these models is the well known

$$Y_t = \beta_o + \beta_1 t + \beta_2 t^2 + \beta_3 \sin\left(\frac{2\pi t}{52.167}\right) + \beta_4 \cos\left(\frac{2\pi t}{52.167}\right) + \epsilon_t$$

where 52.167 is the average number of week per year and ϵ_t is the error term. Even though this model proposed by Simonsen (1997) works for death due to influenza, it will work as well for ILI activity reports through some minor adjustments.

Our method is a data-driven technique; as such, it relies heavily on the availability of %ILI data on regular basis. The reality on the field however is different due to the delay observed in data gathering and its recording. The District of Columbia Department of Health, in conjunction with the Maryland and Virginia state health departments are aware of the need of daily logs and send emergency room data to the health department on daily basis. Our hope is these data will be available for daily analysis in the near future so that we can use new techniques such as the above to act in an early detect-to-warn fashion. We hope this technique will be applied not only to ILI surveillance and intentional release of biological agents but also will have a potential to respond to the serious danger of emerging infectious diseases.

ACKNOWLEDGMENTS

We are grateful to the Iowa Department of Public Health (IDPH), to the Michigan Department of Public Health (MDPH) and to the Center for Disease Control and Prevention(CDC) for their unconditional support during data collection and by giving us clear explanations on the course of influenza in the US.

REFERENCES

Armstrong R., Coomber P., and Prior S. (2004), "Looking for Trouble: A Policymaker's Guide to Biosensing ," National Defense University, Center for Technology and National Security Policy, June 2004.

Arnon S.S, Schechter R, Inglesby T.V, et al. (2001), "Tulinum toxin as a biological weapon: medical and public health management," Journal of the American Medical Association (2001); 285:1059–70.

Barthell E.N., Aronsky D., Cochrane D.G., Cable G., Stair T. (2004), "The Frontlines of Medicine Project progress report: standardized communication of emergency department triage data for syndromic surveillance," Annals of Emergency Medicine Sept; 44 (3): 247-252.

Bioterrorism Preparedness and Response: Use of Information Technologies and Decision Support Systems. Summary, Evidence Report/Technology Assessment: Number 59, July 2002. Agency for Healthcare Research and Quality, Rockville, MD.
(http://www. ahrq.gov/clinic/epcsums/bioitsum.htm)

Buehler J.W., Berkelman R.L., Hartley D.M., Peters C.J. (2003) "Syndromic surveillance and bioterrorism-related epidemics," Emerging Infectious Diseases. 2003,9 pp.1197-1204.

Campbell, N.A. and Reece J.B(2002), Biology, 6th edition, San Francisco, CA Benjamin Cummings

Cochrane D.H., Allegra J.R., Rothman J. (2003), "Comparison of physician's choice of charting template to ICD9 codes for biosurveillance using an emergency department electronic medical records database," Academic Emergency Medicine 2003; 10(5): 525.

Dennis D.T, Inglesby T.V, Henderson D.A, et al. (2001), "Tularemia as a biological weapon: medical and public health management," Journal of the American Medical Association;

285:2763–73.

Green M.S., Kaufman Z. (2002), "Surveillance for early detection and monitoring of infection disease outbreaks associated with bioterrorism," Israel Medical Association Journal Jul; 4(7): 503-6.

Greenko J., Mostashari F., Fine A., Layton M. (2003), "Clinical Evaluation of the Emergency Medical Services (EMS) Ambulance Dispatch-Based Syndromic Surveillance System", New York City. Journal of Urban Health. June; 80 Suppl 1:I50-I56.

Hawkins, D.M., Qiu, P., Kang, C.W. (2003), "The Change point Model for Statistical Process Control" Journal of Quality Technology 35, 355-365.

Hawkins, D.M., and Olwell, D.H., (1998), Cumulative Sum Charts and Charting for Quality Improvement, Springer Verlag, New York.

Hawkins, D.M., Zamba, K.D. (2005), "A Change point Model for Statistical Process Control with Shift in mean or Variance", Technometrics 47(2), 164-173.

Hawkins, D.M., Zamba, K.D. (2005), "Change point Model for a Shift in Variance", Journal of Quality Technology 37(1), 21-31.

Henderson D.A., Inglesby T.V., Bartlett J.G., et al.(1999), "Smallpox as a biological weapon: medical and public health management," Journal of the American Medical Association 1999;281:2127–37.

Hethcote, H.W. (2000), " The Mathematics of Infectious Diseases ," SIAM Review, Vol 42, number 4, 599-653

Hethcote, H.W. (1989) " Periodicity in Epidemiological Models", Applied mathematical Ecology, Springer-Verlag, Berlin, pp 193-211

Hutwagner, L., et al. (1997) " Using Laboratory-based Surveillance data for Prevention: An algorithm to detect Salmonella Outbreaks", Emerging Infectious disease, 1997, 3, 395-400

Hutwagner, L., et al. (2003). "The Bioterrorism Preparedness and Response Early Aberration Reporting System (EARS)". Journal of Urban Health 80, Supplement 1, 89-96.

Hutwagner, L., et al. (2005). "A Simulation Model for Assessing Aberration Detection Methods Used in Public Health Surveillance for Systems with Limited Baselines". Statistics in Medicine 24, 543-550.

Inglesby T.V., Dennis D.T., Henderson D.A., et al.(2000), "Plague as a biological weapon: medical and public health management," Journal of the American Medical Association 2000; 283:2281–90.

Inglesby T.V., Henderson D.A., Bartlett J.G., et al.(1999), "Anthrax as a biological weapon: medical and public health management" Journal of the American Medical Association 1999;281:1735–963.

Jeffreys, H., (1948), Theory of Probability, Second Edition, University Press, Oxford.

Kosal M. E.(2003), "The Basics of Chemical and Biological Weapons Detectors," Center for Nonproliferation Studies, Monterey Institute of International Studies, CA, November 24, 2003 http://cns.miis.edu/pubs/week/031124htm

Mandl K.D., Overhage J.M., Wagner M.M., Lober W.B., Sebastiani P., Mostashari F., Pavlin J.A., Gesteland P.H., Treadwell T., Koski E., Hutwagner L., Buckeridge D.L, Aller R.D., Grannis S.(2003), "Implementing Syndromic Surveillance: A Practical Guide Informed by the Early Experience". Journal of the American Medical Informatics Association, 2003, Nov 21.

Meltzer M.I., Damon I., Leduc J.W., Millar J.D. (2001), "Modelling Potential Responses to Smallpox as a Bioterrorist Weapon" Emerging Infectious Diseases 7(6), 959–969.

Morbidity and Mortality Weekly Report (MMWR), vol 50, number 41.

National Institute of Justice, "An Introduction to Biological Agent Detection Equipment for Emergency First Responders," NIJ Guide 101-00, December 2001, 23-25.

Pavlin J.A. (2003), "Investigation of Disease Outbreaks Detected by "Syndromic" Surveillance Systems", Journal of Urban Health 2003; 80: i107-i114.

Pollak, M. and Siegmund, D. (1991), "Sequential Detection of a Change in Normal Mean when the Initial Value is Unknown" Ann. Statist, 19, 394-416.

Sartwell, P.E. (1949),"The distribution of Incubation Periods of Infectious Disease" American Journal of Epidemiology, Vol 141, 5, 386–394.

Simonsen, L. (1997) " A Method for Timely Assessment of Influenza Associated Mortality in the United States," Epidemiology, Vol 8, number 4.

Stern L. and Lightfoot D. (1999) " Automated outbreak detection: A Quantitative Retrospective Analysis" Epidemiology of Infectious disease, 122, 103-110.

Stoto M.A., Schonlau M., Mariano L.T., (2004), "Syndromic Surveillance: Is it Worth the Effort?" Chance; 17(1):19-24.

Tsiamyrtzis, P. and Hawkins, D., M., (2005), "A Bayesian Scheme to Detect Changes in the Mean of a Short Run Process", Technometrics, 47, 446-456.

West, M., (1993),"Approximating posterior distributions by Mixtures",Journal of Royal Statistical Society, Series B, Vol. 55, No. 2, 409-422.

Woodward P.W. and Naylor J.C. (1993) " An Application of Bayesian Methods in SPC, " The Statistician, 42, 461-469

APPENDIX A

Proof of Lemma 1

$$\left\{\begin{array}{rcl} \theta_{t+1}|\theta_t,p, & \sim & pN(\theta_t,\ \sigma^2)+(1-p)N(\theta_t+\delta,\ \sigma^2) \\ \pi(\theta_t) & \sim & \sum_{i=0}^{2^t-1}\alpha_i^{(t)}N\left(\theta_i^{(t)},\ \hat{\sigma}_t^2\right) \\ \pi(p) & \sim & Beta(\alpha,\beta) \end{array}\right\}\Rightarrow$$

$$\begin{aligned} \Rightarrow \pi(\theta_{t+1}) &= \int\int \pi(\theta_{t+1},\theta_t,p)dp\ d\theta_t = \\ &= \int\int \pi(\theta_{t+1}|\theta_t,p,)\pi(\theta_t)\pi(p)dp\ d\theta_t = \\ &= \int \pi(\theta_t)\left[\int \pi(\theta_{t+1}|\theta_t,p)\pi(p)dp\right]d\theta_t = (*) \end{aligned}$$

but,

$$\begin{aligned} I_1 &= \int \pi(\theta_{t+1}|\theta_t,p)\pi(p)dp = \\ &= \int\left[\frac{p}{\sqrt{2\pi\sigma^2}}exp\left\{-\frac{(\theta_{t+1}-\theta_t)^2}{2\sigma^2}\right\}+\frac{1-p}{\sqrt{2\pi\sigma^2}}exp\left\{-\frac{(\theta_{t+1}-\theta_t-\delta)^2}{2\sigma^2}\right\}\right] \\ &\quad\times\left[\frac{1}{Be(\alpha,\beta)}p^{\alpha-1}(1-p)^{\beta-1}\right]dp = \\ &= \frac{Be(\alpha+1,\beta)}{Be(\alpha,\beta)}\frac{1}{\sqrt{2\pi\sigma^2}}exp\left\{-\frac{(\theta_{t+1}-\theta_t)^2}{2\sigma^2}\right\}+\frac{Be(\alpha,\beta+1)}{Be(\alpha,\beta)}\frac{1}{\sqrt{2\pi\sigma^2}}exp\left\{-\frac{(\theta_{t+1}-\theta_t-\delta)^2}{2\sigma^2}\right\} \end{aligned}$$

and

$$\frac{Be(\alpha+1,\beta)}{Be(\alpha,\beta)}=\frac{\alpha}{\alpha+\beta} \qquad\qquad \frac{Be(\alpha,\beta+1)}{Be(\alpha,\beta)}=\frac{\beta}{\alpha+\beta}=1-\frac{\alpha}{\alpha+\beta}$$

Thus (after some algebra) we get:

$$\pi(\theta_{t+1}) = \sum_{i=0}^{2^t-1}\left[\frac{\alpha}{\alpha+\beta}\right]\alpha_i^{(t)}N\left(\theta_i^{(t)},\ \sigma^2+\hat{\sigma}_t^2\right)+\sum_{i=0}^{2^t-1}\left[\frac{\beta}{\alpha+\beta}\right]\alpha_i^{(t)}N\left(\theta_i^{(t)}+\delta,\ \sigma^2+\hat{\sigma}_t^2\right)$$

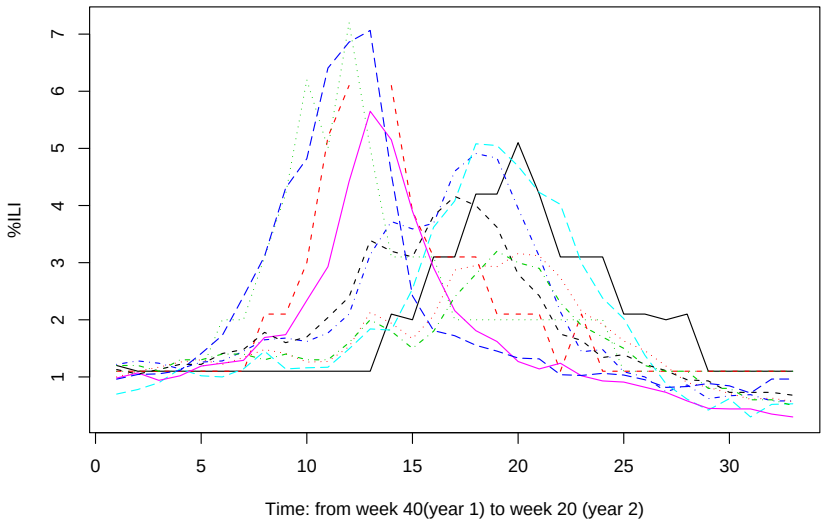

Figure 1: Percentage of ILI in the US between years 1994 and 2004

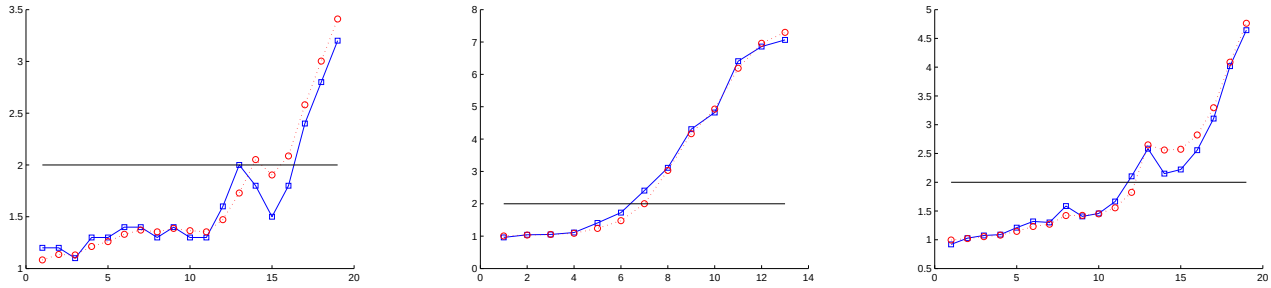

Figure 2: Time series plot of the data (square points - solid line) along with the posterior mean (circle points - dashed line) for the three successive periods 2002-03, 2003-04 and 2004-05 respectively

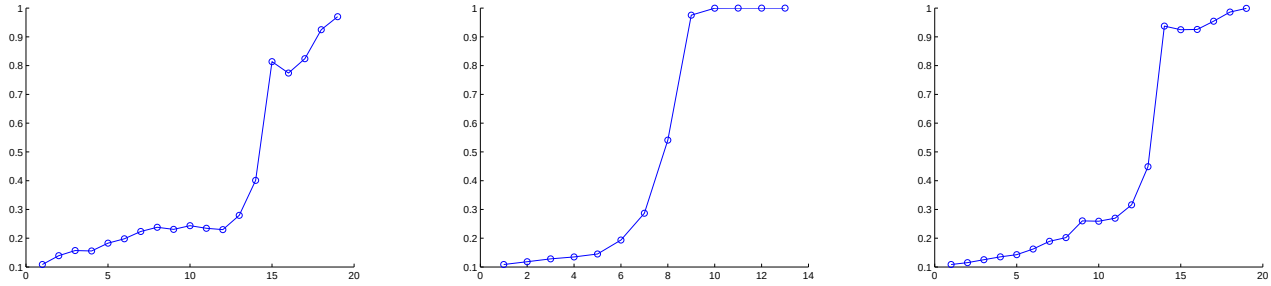

Figure 3: Time series plot of the predictive probability $P(y_{t+1} > 2|y_1, \ldots, y_t)$ for the three successive periods 2002-03, 2003-04 and 2004-05 respectively

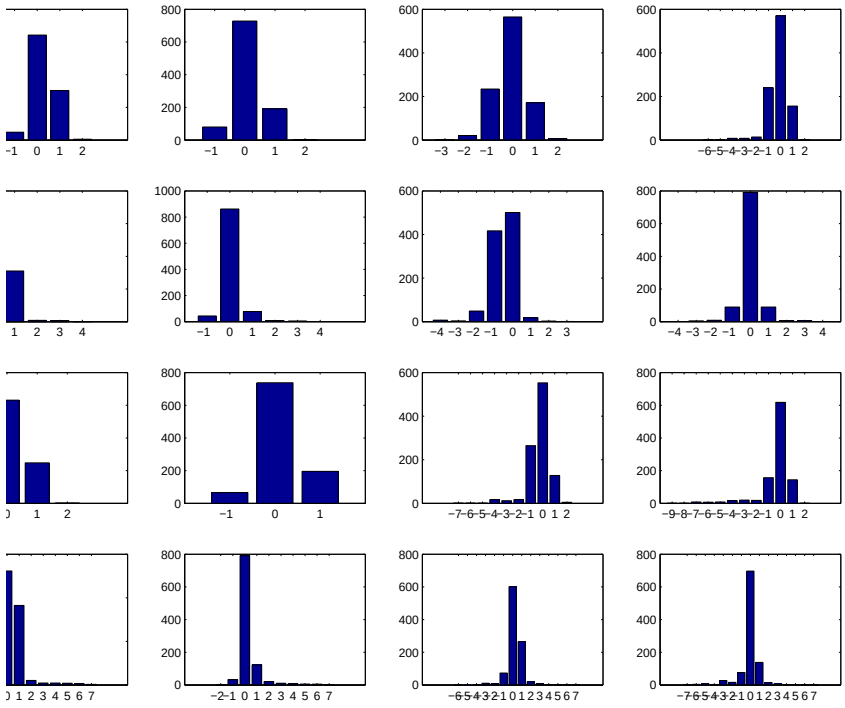

Figure 4: The barcharts of the $n_2 - n_1$ differences for the 16 cases (by rows) reported in Table 2, where the zeros correspond to the CA, and the negative (positive) cases refer to the FA (MA)

Table 1: Posterior Probabilities of the alternative hypothesis H_1: $\theta_t > K$ at the end of week t along with Bayes Factors (BF) for the three years 2002-2005 under study

	2002-03		**2003-04**		**2004-05**	
Stage	$P(H_1\|y)$	BF	$P(H_1\|y)$	BF	$P(H_1\|y)$	BF
1	0.0136	5.0498	0.0056	12.5067	0.0047	14.6714
2	0.0168	5.4877	0.0086	9.1801	0.0076	9.6198
3	0.0129	8.0955	0.0091	8.9663	0.0097	8.1997
4	0.0244	4.0737	0.0117	7.3390	0.0109	7.9445
5	0.0285	4.3024	0.0339	2.6942	0.0172	5.2731
6	0.0409	3.2513	0.1047	1.1964	0.0276	3.8344
7	0.0463	3.3837	0.4788	0.2847	0.0299	4.3108
8	0.0374	4.5894	0.9886	0.0133	0.0701	1.8883
9	0.0462	3.4414	1.0000	0.0001	0.0586	3.4183
10	0.0379	4.6201	1.0000	0.0000	0.0638	2.9772
11	0.0354	4.6126	1.0000	0.0000	0.1085	1.7655
12	0.0768	1.9790	1.0000	0.0000	0.3004	0.6761
13	0.2284	0.7856	1.0000	0.0000	0.9165	0.0598
14	0.5310	0.4414			0.9000	3.5854
15	0.4008	8.6755			0.9031	2.6980
16	0.5591	3.4304			0.9619	0.9976
17	0.8848	0.8130			0.9974	0.1434
18	0.9839	0.3815			1.0000	0.0009
19	0.9991	0.1068			1.0000	0.0000

Table 2: Simulation results when true model parameters are $\sigma = (0.1; 0.2)$ $\tau = (0.4; 0.6)$ $\delta = (5\sigma; 7\sigma)$ and there are no model misspecifications

Case	n^*	p	σ	τ	δ	C. A.	M. A.	F.A.
1	5	0.7	0.1	0.4	5σ	64.2%	30.8%	5.0%
2	5	0.7	0.1	0.6	5σ	72.8%	19.2%	8.0%
3	5	0.7	0.2	0.4	5σ	56.5%	17.8%	25.7%
4	5	0.7	0.2	0.6	5σ	57.0%	15.8%	27.2%
5	5	0.7	0.1	0.4	7σ	64.5%	32.6%	2.9%
6	5	0.7	0.1	0.6	7σ	86.2%	9.4%	4.4%
7	5	0.7	0.2	0.4	7σ	50.1%	2.2%	47.7%
8	5	0.7	0.2	0.6	7σ	79.1%	10.4%	10.5%
9	8	0.7	0.1	0.4	5σ	63.1%	25.0%	11.9%
10	8	0.7	0.1	0.6	5σ	73.8%	19.6%	6.6%
11	8	0.7	0.2	0.4	5σ	55.3%	13.3%	31.4%
12	8	0.7	0.2	0.6	5σ	61.8%	14.5%	23.7%
13	8	0.7	0.1	0.4	7σ	52.3%	42.1%	5.6%
14	8	0.7	0.1	0.6	7σ	79.1%	17.5%	3.4%
15	8	0.7	0.2	0.4	7σ	60.2%	30.3%	9.5%
16	8	0.7	0.2	0.6	7σ	69.7%	16.9%	13.4%