



NAVAL POSTGRADUATE SCHOOL

MONTEREY, CALIFORNIA

THESIS

**ASSESSING THE EFFECTIVENESS OF THE EARLY
ABERRATION REPORTING SYSTEM (EARS) FOR
EARLY EVENT DETECTION OF THE H1N1 ("SWINE
FLU") VIRUS**

by

Katie S. Hagen

September 2010

Thesis Advisor:

Ronald D. Fricker, Jr.

Second Reader:

Rachel T. Johnson

Approved for public release; distribution is unlimited

THIS PAGE INTENTIONALLY LEFT BLANK

REPORT DOCUMENTATION PAGE			<i>Form Approved OMB No. 0704-0188</i>	
Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instruction, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188) Washington DC 20503.				
1. AGENCY USE ONLY (Leave blank)		2. REPORT DATE September 2010	3. REPORT TYPE AND DATES COVERED Master's Thesis	
4. TITLE AND SUBTITLE Assessing the Effectiveness of the Early Aberration Reporting System (EARS) for Early Event Detection of the H1N1 ("Swine Flu") Virus			5. FUNDING NUMBERS	
6. AUTHOR(S) Katie S. Hagen				
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Naval Postgraduate School Monterey, CA 93943-5000			8. PERFORMING ORGANIZATION REPORT NUMBER	
9. SPONSORING /MONITORING AGENCY NAME(S) AND ADDRESS(ES) N/A			10. SPONSORING/MONITORING AGENCY REPORT NUMBER	
11. SUPPLEMENTARY NOTES The views expressed in this thesis are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government. IRB Protocol number ____ N.A. ____.				
12a. DISTRIBUTION / AVAILABILITY STATEMENT Approved for public release; distribution is unlimited			12b. DISTRIBUTION CODE A	
13. ABSTRACT (maximum 200 words) The Monterey County Health Department (MCHD) in California uses the Early Aberration Reporting System (EARS) to monitor emergency room and clinic data for biosurveillance, particularly as an alert system for various types of disease outbreaks. The flexibility of the system has proven to be a very useful feature of EARS; however, little research has been conducted to assess its performance. In this thesis, a quantitative analysis based on modifications to EARS' internal logic and algorithms is assessed. Logic is used as a counting tool for potential cases of outbreak, and the Early Event Detection (EED) algorithms are used to determine whether or not an outbreak is about to occur. The EED methods are compared by assessing their ability to detect the presence of a known H1N1 outbreak in Monterey County. This research found the cumulative sum (CUSUM) detection method to be the most reliable in signaling the H1N1 outbreak, across all combinations of logic explored.				
14. SUBJECT TERMS Biosurveillance, Syndromic Surveillance, Early Event Detection (EED), Early Aberration Reporting System (EARS), C1, C2, C3, Cumulative Sum (CUSUM), Adaptive Regression			15. NUMBER OF PAGES 89	
			16. PRICE CODE	
17. SECURITY CLASSIFICATION OF REPORT Unclassified	18. SECURITY CLASSIFICATION OF THIS PAGE Unclassified	19. SECURITY CLASSIFICATION OF ABSTRACT Unclassified	20. LIMITATION OF ABSTRACT UU	

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. Z39-18

THIS PAGE INTENTIONALLY LEFT BLANK

Approved for public release; distribution is unlimited

**ASSESSING THE EFFECTIVENESS OF THE EARLY ABERRATION
REPORTING SYSTEM (EARS) FOR EARLY EVENT DETECTION OF THE
H1N1 (“SWINE FLU”) VIRUS**

Katie S. Hagen
Lieutenant, United States Navy
B.S., Tulane University, 2004

Submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE IN OPERATIONS RESEARCH

from the

**NAVAL POSTGRADUATE SCHOOL
September 2010**

Author: Katie S. Hagen

Approved by: Ronald D. Fricker, Jr., PhD
Thesis Advisor

Rachel T. Johnson, PhD
Second Reader

Robert F. Dell, PhD
Chairman, Department of Operations Research

THIS PAGE INTENTIONALLY LEFT BLANK

ABSTRACT

The Monterey County Health Department (MCHD) in California uses the Early Aberration Reporting System (EARS) to monitor emergency room and clinic data for biosurveillance, particularly as an alert system for various types of disease outbreaks. The flexibility of the system has proven to be a very useful feature of EARS; however, little research has been conducted to assess its performance. In this thesis, a quantitative analysis based on modifications to EARS' internal logic and algorithms is assessed. Logic is used as a counting tool for potential cases of outbreak, and the Early Event Detection (EED) algorithms are used to determine whether or not an outbreak is about to occur. The EED methods are compared by assessing their ability to detect the presence of a known H1N1 outbreak in Monterey County. This research found the cumulative sum (CUSUM) detection method to be the most reliable in signaling the H1N1 outbreak, across all combinations of logic explored.

THIS PAGE INTENTIONALLY LEFT BLANK

TABLE OF CONTENTS

I.	INTRODUCTION.....	1
A.	EARLY EVENT DETECTION (EED).....	1
B.	BIOSURVEILLANCE	1
1.	Syndromic Surveillance.....	3
2.	Early Aberration Reporting System (EARS) Syndromic Surveillance System	8
3.	EARS and Monterey County Biosurveillance.....	9
C.	2009 H1N1 “SWINE FLU” VIRUS.....	12
D.	ORGANIZATION OF THIS THESIS.....	16
II.	SYNDROMES: DEFINING AND CALCULATING DAILY COUNTS	17
A.	BACKGROUND	17
B.	CHIEF COMPLAINT DATA COLLECTION AND REPORTING.....	17
C.	CHANGES IN LOGIC	19
1.	Symptom Aliases	19
2.	Syndrome Definitions	20
3.	Text Matching Algorithms.....	22
4.	Summary of Changes in Logic.....	24
III.	EARLY EVENT DETECTION METHODS	29
A.	EXISTING EED METHODS: EARS	29
B.	AN ALTERNATIVE EED METHOD BASED ON THE CUSUM	31
1.	Basic CUSUM.....	31
2.	CUSUM Applied to Biosurveillance Data.....	31
3.	Application of the CUSUM to MCHD Data	33
IV.	RESULTS	37
A.	ILI COUNT COMPARISONS	37
B.	ORIGINAL CDC LOGIC (BASE CASE)	39
C.	EXPANDED MCHD LOGIC (VARIANT 1A)	41
D.	RESTRICTED MCHD LOGIC (VARIANT 2A)	43
E.	A CLOSER LOOK AT CUSUM 1.....	45
V.	CONCLUSIONS	49
A.	EXERCISE CAUTION WHEN CHANGING LOGIC.....	49
B.	ALGORITHMS NEED GOOD DATA.....	49
C.	RESTRICTED LOGIC IS PREFERRED.....	50
D.	FUTURE RESEARCH OPPORTUNITIES.....	51
APPENDIX A.	R CODE FOR “ENHANCED NPS LOGIC”	53
APPENDIX B.	MATLAB CODE FOR CALCULATING CUSUM	55
APPENDIX C.	STANDARDIZED RESIDUAL PLOTS AND QQ PLOTS OF HISTORICAL DATA.....	57
APPENDIX D.	R CODE TO ESTIMATE CUSUM THRESHOLD	61

LIST OF REFERENCES	63
INITIAL DISTRIBUTION LIST	67

LIST OF FIGURES

Figure 1.	Results of changes to EARS logic, symptom aliases, and syndrome definitions, as applied to MCHD data from August 1, 2008–July 31, 2009.. xvi
Figure 2.	Civilian Funding of Biodefense by Fiscal Year, FY2001–FY2011 in \$millions (From Franco & Sell, 2010).....3
Figure 3.	Illustration of the various subsets of biosurveillance, to include epidemiologic surveillance and syndromic surveillance. 4
Figure 4.	Illustration of how biosurveillance is intended to improve Early Event Detection (EED) and Situational Awareness (SA) (From Fricker, 2010) 5
Figure 5.	Four main components of a biosurveillance system: data collection, data management, analysis, and reporting. (From Fricker & Hanni, 2010).....8
Figure 6.	MCHD implementation of the EARS biosurveillance system (After Fricker & Hanni, 2010).....10
Figure 7.	Illustration of the reporting component of a biosurveillance system, the Monterey County DOSE Report for Friday, May 28, 2010 (From MCHD, 2010b) 11
Figure 8.	Confirmed and probable H1N1 case rate by age group from April 15–July 24, 2009 (From Hanni, 2009) 12
Figure 9.	Estimated pandemic H1N1 flu hospitalization rate in the United States by age group from April 15 to July 24, 2009 (From Hanni, 2009).....13
Figure 10.	CDC Testing Recommendations for Pandemic (H1N1) 2009 Influenza Virus, last updated September 08, 2009 (From MCHD, 2009).....15
Figure 11.	Number of confirmed H1N1 cases from May 2009 to April 2010 in Monterey County (From Hanni, 2010) 16
Figure 12.	Percentage of visits for ILI reported by the U.S. Outpatient ILI Network, National Summary 2008–09 and previous two seasons (From CDC, 2009)...18
Figure 13.	Three versions of the ILI syndrome definition: original EARS (CDC), expanded MCHD, and restricted MCHD.....21
Figure 14.	Emergency Room Influenza-Like Illness Visits for Monterey County and California, 2008–2009 Season (From Hanni, 2009).....22
Figure 15.	Examples of how “Enhanced NPS Logic” works for text-matching keywords with symptoms 24
Figure 16.	Qualitative results of changes to EARS logic, symptom aliases, and syndrome definitions, as applied to MCHD data from August 1, 2008–July 31, 2009.....25
Figure 17.	Smoothed ILI counts from MCHD during August 1, 2008–July 31, 2009 (excluding weekends) using five different combinations of text matching logic, symptom aliases, and syndrome definitions. 27
Figure 18.	Smoothed ILI count comparisons of Base Case, Variant 1A, and Variant 2A logic from September 30, 2008 to October 29, 2009.....38
Figure 19.	Comparison of actual Monterey County H1N1 cases with the results of five EED algorithms under the original CDC logic (Base Case), from September 30, 2008 to October 29, 2009 40

Figure 20.	Comparison of actual Monterey County H1N1 cases with the results of five EED algorithms under the “expanded” MCHD logic (Variant 1a), from September 30, 2008 to October 29, 2009.....	42
Figure 21.	Comparison of actual Monterey County H1N1 cases with the results of five EED algorithms under the “restricted” MCHD logic (Variant 2a), from September 30, 2008 to October 29, 2009.....	44
Figure 22.	CUSUM 1 comparisons across the various sets of logic during the first H1N1 flu wave, April 16–June 10, 2009	47
Figure 23.	CUSUM 1 comparisons across the various sets of logic during the active summer months, June 12–September 1, 2009	48

LIST OF TABLES

Table 1.	Comparison of the categories of data used in biosurveillance, epidemiologic surveillance, and syndromic surveillance (From Fricker, 2010)	4
Table 2.	Typical syndromes used in syndromic surveillance systems.....	6
Table 3.	Examples of actual chief complaints taken from Monterey County’s clinic data. (From Hanni, 2009a)	7
Table 4.	Subset of the CDC “symptom_code” file, as used by EARS to map keywords to symptoms (e.g., sore throat, cold, cough)	19
Table 5.	CUSUM parameters used in Matlab code for monitoring the ILI syndrome ..	35
Table 6.	Number of alarms corresponding to five EED algorithms under “Base Case” logic from March 2 to October 29, 2009.....	41
Table 7.	Number of alarms corresponding to five EED algorithms under “expanded” MCHD logic from March 2 to October 29, 2009	43
Table 8.	Number of alarms corresponding to five EED algorithms under “restricted” MCHD logic from March 2 to October 29, 2009.....	45

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF ACRONYMS AND ABBREVIATIONS

ATFS	Average Time to False Signal
CBRN	Chemical, Biological, Radiological, and Nuclear
CDC	Centers for Disease Control and Prevention
CUSUM	Cumulative Sum
DOSE	Daily Observational and Situational Evaluation
EARS	Early Aberration Reporting System
EED	Early Event Detection
ER	Emergency Room
HSPD-21	Homeland Security Presidential Directive 21
ICD-9	International Classification of Diseases 9 th Edition
IFFGD	International Foundation for Functional Gastrointestinal Disorders
ILI	Influenza-Like Illness
MCHD	Monterey County Health Department
PPE	Personal Protective Equipment
Q-Q	Quantile-Quantile
SA	Situational Awareness
SPC	Statistical Process Control
WHO	World Health Organization

THIS PAGE INTENTIONALLY LEFT BLANK

EXECUTIVE SUMMARY

The Monterey County Health Department (MCHD) in California uses the Early Aberration Reporting System (EARS) to monitor emergency room and clinic data for biosurveillance, particularly as an alert system for various types of disease outbreaks, both natural and man-made (e.g., bio-terrorism). The concept behind Early Event Detection (EED) is to quickly detect abnormalities from the normal trends so that public health authorities can take the appropriate action to deal with them. In particular, the intention is to expedite the detection of disease outbreaks by using data based on pre-diagnosis “syndromes with the hope that at least for some outbreaks there will be a sufficiently strong signal in the data that the outbreak can be detected using a statistical algorithm in advance of the first case diagnosis by a medical professional.

Monterey County’s EARS system uses data from six public clinics and four hospitals located throughout the county. The data received from MCHD include the date of patient visit, age, sex, and home ZIP code of the patient, chief complaint, and diagnosis code for the clinics only. Daily counts for various syndromic categories are calculated based on the presence or absence of key words in the chief complaints. For example, existence of either the word “flu” or the phrase “fever and cough” in an individual's chief complaint would result in that individual being included in the Influenza-Like Illness (ILI) syndrome count for that day.

The flexibility of the system has proven to be a very useful feature of EARS; however, little research has been conducted to assess the performance of EARS. Specifically, are there any changes that can be made to EARS’ logic and/or settings that would maximize the system’s ability to detect disease outbreaks? Also, how do these changes affect EARS’ ability to detect a particular, known outbreak, such as the novel 2009 H1N1 virus?

To answer these questions, a quantitative comparison was conducted by implementing modifications to EARS' logic and assessing the affect on daily counts, which is one of the key measures used by EARS to monitor for outbreaks. Logic modifications were compared by evaluating counts of the ILI syndrome over a one year period.

As shown in Figure 1, out of 153,696 total patient records from August 1, 2008 to July 31, 2009, the logic encoded in the unmodified EARS system (the "Base Case") flagged 9,093 records for the ILI syndrome.

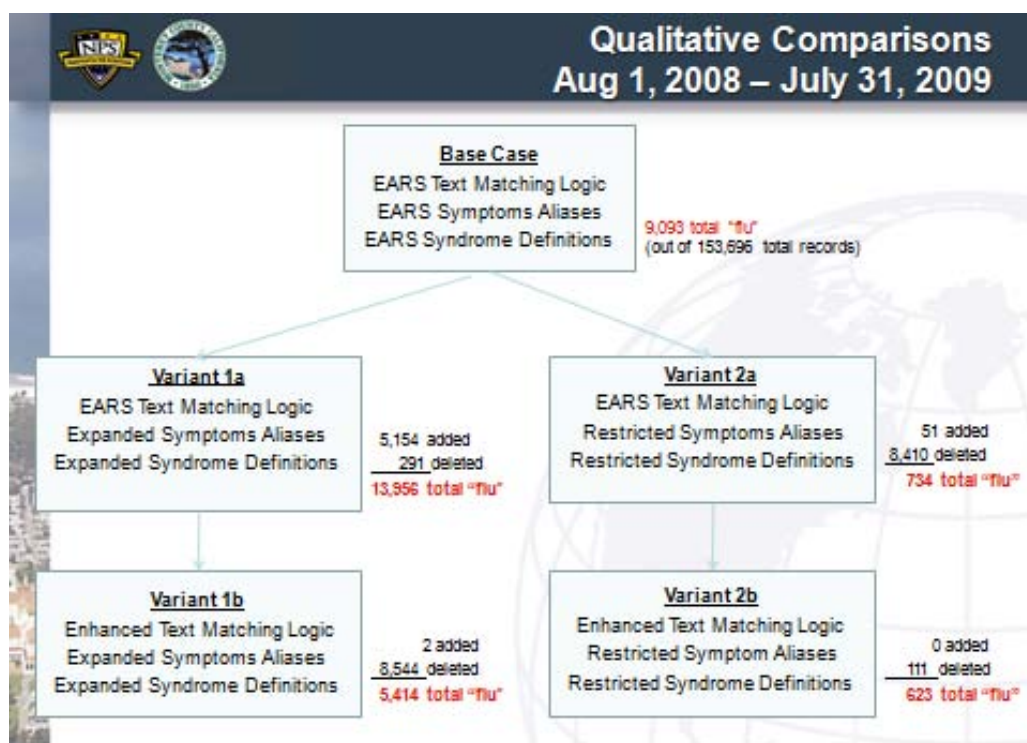


Figure 1. Results of changes to EARS logic, symptom aliases, and syndrome definitions, as applied to MCHD data from August 1, 2008–July 31, 2009

The second tier in Figure 1 illustrates the number of ILI syndrome counts when EARS' symptom aliases and syndrome definitions are modified. Specifically, the "Variant 1a" box on the left is based on an expanded ILI syndrome definition as well as a more robust symptom alias list used by MCHD. Variant 1a modifications to the EARS logic resulted in a 53% increase in the number of records flagged for the ILI syndrome.

In comparison, the box on the right labeled “Variant 2a” used a restrictive symptom alias list and restrictive syndrome definitions subsequently employed and resulted in a reduction of the number of records flagged for ILI by 92% of the original “Base Case.” This illustrates the dramatic impact that changes in EARS logic can have on the daily syndrome counts.

The bottom tier of Figure 1 illustrates the effects of changing the text-matching logic, which resulted in similarly large swings in the number of coded ILI syndromes. For example, the only difference between “Variant 1a” and “Variant 1b” is the change in text matching logic, which results in a 62% decrease (13,956 down to 5,414) in the number of records flagged for ILI.

Using the various ILI counts that result from the logic variants, EARS performance was assessed by determining the system’s ability to detect a known outbreak. To evaluate this, the ILI counts produced by the Base Case, Variant 1a, and Variant 2a logic were then used as inputs into the EARS’ system. In addition to the modified logic, alternative EED methods based on the cumulative sum (CUSUM) were tested. Lastly, all methods were compared by assessing their ability to detect the presence of a known H1N1 outbreak in Monterey County.

The CUSUM EED method proved the most reliable at signaling alarms prior to and throughout the time when Monterey County was experiencing H1N1 cases. Currently, EARS does not utilize the CUSUM algorithms. When testing the current EED methods, Variant 2a logic was shown to have the best performance in terms of signals triggered prior to an outbreak. Surprisingly, under original and Variant 1a sets of logic, EARS methods were of little to no value in signaling an outbreak.

THIS PAGE INTENTIONALLY LEFT BLANK

ACKNOWLEDGMENTS

First and foremost, I would like to thank Professor Ron Fricker for inspiring me to pursue research in the field of biosurveillance and for introducing me to the public health community. I have learned a great deal from you, and it was truly rewarding to apply my newly mastered Operations Research skills towards a real life problem. Your mentorship, professionalism, and enthusiasm will not soon be forgotten.

I would also like to thank two very special individuals at the Monterey County Health Department, Dr. Krista Hanni and Ms. Susie Barnes. Thank you for all of the time, energy, and resources you have devoted towards helping me in my thesis research. Without you, I could not have done it.

Lastly, I would like to thank my parents, Dr. Dean Hagen and Sharon Howes. Mom and Dad, you have always inspired me to pursue opportunities and to not be afraid of challenges. You have my deepest heartfelt thanks for always being there for me.

THIS PAGE INTENTIONALLY LEFT BLANK

I. INTRODUCTION

A. EARLY EVENT DETECTION (EED)

Health emergencies can either be naturally occurring (e.g., influenza), accidental (e.g., fire-related illnesses), or intentional (e.g., bioterrorism). Given the possible life-threatening nature of these situations, decision makers require timely diagnosis and reporting to reduce the negative impact to public health. The concept behind Early Event Detection (EED) is to quickly detect abnormalities from the normal trends so that public health authorities can take the appropriate action to deal with them. More formally, the Centers for Disease Control and Prevention (CDC) defines EED as “supporting the early detection of health events including determining and monitoring the size, location and spread of health events, and providing situational awareness to assist in the investigation and management of health events” (CDC, 2006). This research evaluates EED methods found within a specific biosurveillance system known as the Early Aberration Reporting System (EARS) on actual H1N1 flu data from Monterey County, California.

B. BIOSURVEILLANCE

Shmueli and Burkom (2010) define *biosurveillance* as “the practice of monitoring data to detect, investigate, and respond to disease out-breaks.” Homeland Security Presidential Directive 21 (HSPD-21, 2007) further defines *biosurveillance* as:

...the process of active data-gathering with appropriate analysis and interpretation of biosphere data that might relate to disease activity and threats to human or animal health—whether infectious, toxic, metabolic, or otherwise and regardless of intentional or natural origin—in order to achieve early warning of health threats, early detection of health events, and overall situational awareness of disease activity.

Before the late 1990s, traditional biosurveillance generally took a retrospective approach for determining the cause of disease outbreaks. Such outbreaks were generally identified only after one or more patients had been diagnosed by a medical professional and then subsequently reported to the appropriate public health authorities. After diagnostic medical and public health data had been collected and analyzed on a disease,

sometimes it would take weeks or months to report these findings. The problem with such delayed reporting is that it is more difficult for medical and public health decision makers to take mitigating measures, such as establishing quarantines for infected individuals and/or regions.

Modern biosurveillance systems are intended to drastically shorten the time it takes to analyze and report data of interest, with the goal of facilitating the proactive detection and management of outbreaks. By using less specific, aggregated, syndromic data, modern biosurveillance systems can now search for earlier outbreak signals, often in advance of an actual case being diagnosed, which may lead to more successful public health interventions (Shmueli & Burkom, 2010).

While biosurveillance systems are most often used to detect and monitor natural diseases, they can also be used to detect bioterrorism events. Bioterrorism, as defined by Evans (2010), “refers to the intentional release of organisms that can cause sickness or death.” Shmueli and Burkom (2010) caution that complications can arise from the intended dual use of biosurveillance systems for detecting natural outbreaks and bioterror-related illnesses. Specifically, it is difficult to define “normal behavior,” from which to derive appropriate baseline information, for both purposes. For example, if a bioterrorism pathogen such as tularemia were released during peak flu season, a dual-use biosurveillance system may not be able to detect the bioterrorism attack. While the issue of dual-use biosurveillance systems is beyond the scope of this research, it is likely to play a continuing and significant role in the detection and monitoring of disease outbreaks (Fricker, Hegler & Dunfee, 2010).

Figure 2 illustrates that since 2001, the U.S. government has spent substantial resources on preparing the nation against a bioterrorist attack, including a proposed increase in funding of \$271.3 million in the President’s FY2011 budget (Franco, 2010). In 2004, President Bush’s Project BioShield sought to address the challenges of potential chemical, biological, radiological, and nuclear (CBRN) terrorism attacks. To see more information on Project BioShield, refer to the Congressional Research Service Report for Congress (Gottron, 2009).

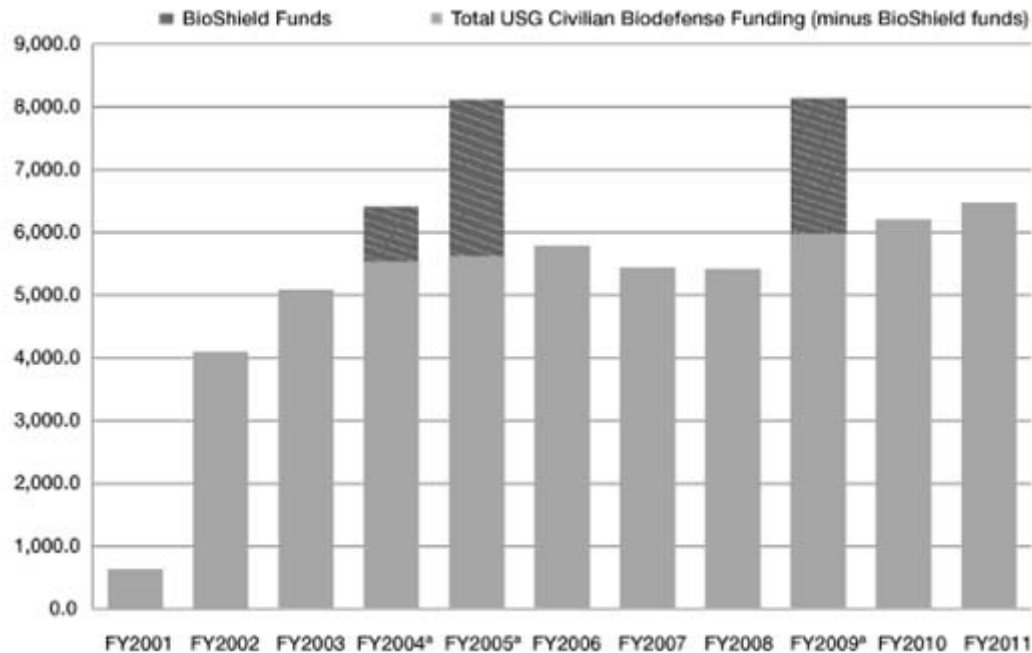


Figure 2. Civilian Funding of Biodefense by Fiscal Year, FY2001–FY2011 in \$millions
(From Franco & Sell, 2010)

1. Syndromic Surveillance

This research assesses the performance of three syndromic surveillance EED methods that are implemented by EARS. Figure 3 illustrates how this type of surveillance fits into the broader category of biosurveillance. To begin, *epidemiologic surveillance* addresses biosurveillance as it applies to human beings. Even more specialized, *syndromic surveillance* is defined as the “the ongoing, systematic collection, analysis, interpretation, and application of real-time (or near-real-time) indicators of diseases and outbreaks that allow for their detection before public health authorities would otherwise note them” (Sosin, 2003). Table 1 provides a more detailed comparison of the various types of data used in biosurveillance, epidemiologic surveillance, and syndromic surveillance. Notice that syndromic surveillance uses the least medically specific data, which is often derived from people who explain their symptoms, otherwise known as chief complaints, to hospitals or clinics (Fricker et al., 2010).

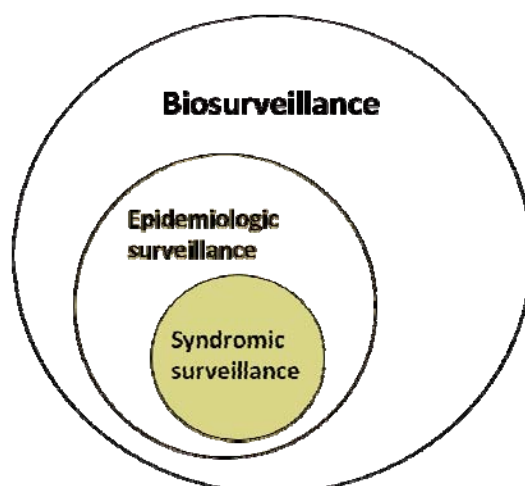


Figure 3. Illustration of the various subsets of biosurveillance, to include epidemiologic surveillance and syndromic surveillance.

Category of Data	Biosurveillance	Epidemiological Surveillance	Syndromic Surveillance
Pre-diagnosis			
- ER chief complaint	X	X	X
- OTC medicine sales	X	X	X
- EMS call rates	X	X	X
- Absenteeism records	X	X	X
- Lab results	X	X	X
- Other	X	X	X
Medical diagnoses	X	X	
Lab results	X	X	
Water and air monitoring	X		
Zoonotic	X		
Agricultural	X		

Table 1. Comparison of the categories of data used in biosurveillance, epidemiologic surveillance, and syndromic surveillance (From Fricker, 2010)

Figure 4 illustrates the EED improvements that medical and public health communities hope to achieve with a biosurveillance system. In particular, the intention is to expedite the detection of disease outbreaks by using data based on pre-diagnosis “syndromes,” with the hope that at least for some outbreaks there will be a sufficiently strong signal in the data that the outbreak can be detected using a statistical algorithm in advance of the first case diagnosis by a medical professional. As defined by the International Foundation for Functional Gastrointestinal Disorders (IFFGD), a *syndrome*

is “a set of symptoms or conditions that occur together and suggest the presence of a certain disease or an increased chance of developing the disease” (IFFGD, 2010).

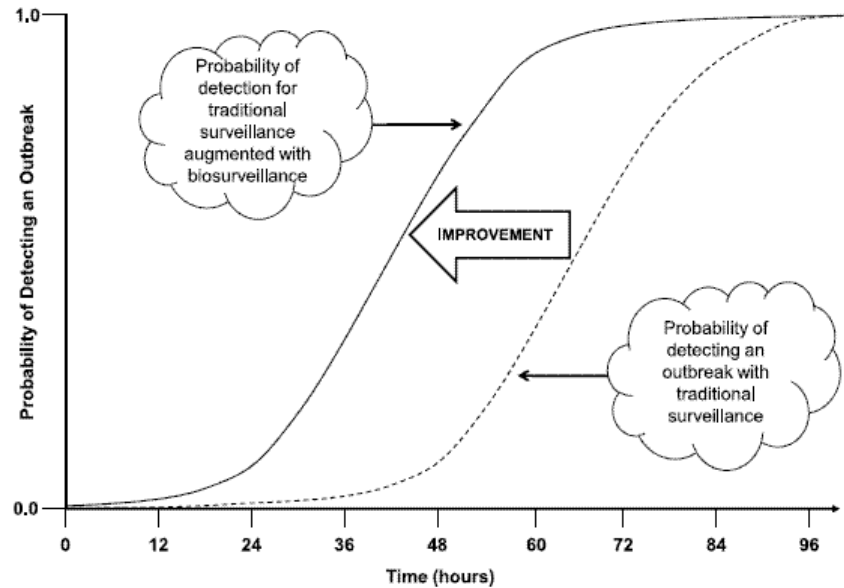


Figure 4. Illustration of how biosurveillance is intended to improve Early Event Detection (EED) and Situational Awareness (SA) (From Fricker, 2010)

Some examples of Monterey County syndromes are listed in Table 2, while Table 3 offers examples of actual chief complaints taken from Monterey County clinic data. As described in Fricker, et al. (2010):

Syndromes are frequently derived from emergency room chief complaint data. A chief complaint is a brief summary of the reason or reasons that an individual presents at a medical facility. Written by medical personnel, chief complaints are couched in jargon, acronyms, and abbreviations for use by other medical professionals. To distill the chief complaints down into syndrome indicators, the text is searched and parsed for key words, often of necessity including all the ways a particular key word can be misspelled, abridged, and otherwise abbreviated.

Fever

Gastrointestinal (GI)

Hemorrhagic Lesion	Upper-respiratory Lower-respiratory
Neurological	Influenza-like Illness (ILI)

Table 2. Typical syndromes used in syndromic surveillance systems.

A biosurveillance system has four main components: data collection, data management, analysis, and reporting (Fricker et al., 2010). Mandl, Overhage, Wagner, Lober, Sebastiani, Mostashari, and Pavlin (2004) provide the following discussion and guidance about how to implement these components:

- **Data Collection:** Electronically stored data sources are necessary because they allow for robust syndromic grouping and are typically readily available. It is usually the case that the data received has already been collected for other purposes. Unfortunately, implementing a new process is deemed as cost prohibitive and administratively taxing. The use of pre-existing database systems does have the benefit of ensuring the availability of baseline data, which is important for algorithm development. Public health officials must then determine which disease and associated syndromes should be tracked.
- **Data Management:** The next step is to acquire and manipulate the data, which can either be done manually or automatically. Manual acquisition may require personnel resources from various clinics, hospitals, etc to transfer data.
- **Analysis:** The next step is to logically group the data in some way that provides useful information. Free-text chief complaints can be grouped into syndromes using statistical algorithms to analyze the data for possible outbreaks over space and time. Rolka (2006) notes that analysis should be of “sufficient sensitivity to provide signals within an actionable time frame while simultaneously limiting false positive signals to a tolerable level.”
- **Reporting:** The final step in the biosurveillance process is to report the findings to appropriate medical and public health communities. Having sufficient and timely information is essential for designated authorities to take necessary action, such as conducting a public health investigation.

FU ANEMIA	chdp
book per angie	FEVER,PHLEGM
fever x3 days cough	4WK FU OB
FEVER,WHEEZING	FEVER
VOMITING,FEVER,POSS EAR INFECT	PAP
CHDP	COUGH,DXd W/ ASTHMA
WI C/O HA//MM	CHDP
PAP per Kennedy	DEPO
BOOK PER MD ROB	4WK FU OB ..OVBK
COLD	ABD PAIN CONJESTION
WALK IN BURN TO R-HAND	new born with mom
FU RESULTS	RASH
FU OB	FU WT CHECK
SHLDR PAIN FOR 1 WK	PAP
WCC 2 MONTH..ok	WCC 2 MONTHS
4WK FU OB..OVBK	VMX/IC
mom and baby	NP/MUDGE/JR
POST PARTUM/BOOK	FU VST URGENT FEET SWOLLEN
NP MED REFILLS	PRE-OP VST
VOMITING AND COUGH	walk-in hospital fu
F/U ASTHMA, FEVER AND COUGH	APPT 830 OK PER DR WELL CHILD EXAM

Table 3. Examples of actual chief complaints taken from Monterey County's clinic data.
(From Hanni, 2009a)

Figure 5 illustrates the four main biosurveillance system components described above. First, raw health-related data is collected from various sources. Next, the incoming data is processed into databases by data management experts and software. Statistical algorithms will then analyze the data for possible outbreaks over space and time. Lastly, the information must be communicated to medical and public health communities in order to support EED and SA efforts.

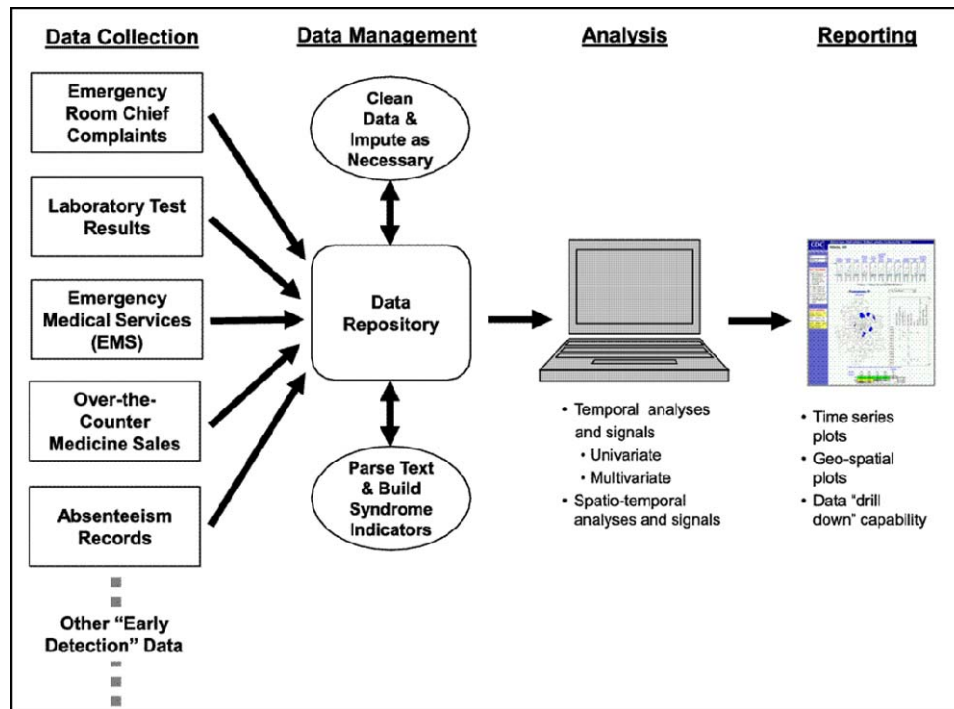


Figure 5. Four main components of a biosurveillance system: data collection, data management, analysis, and reporting. (From Fricker & Hanni, 2010)

2. Early Aberration Reporting System (EARS) Syndromic Surveillance System

Although a number of syndromic surveillance systems are available, EARS uses aberration detection models to identify deviations in current data when compared with a moving baseline of recent data (Lawson & Kleinman, 2005). EARS was originally developed by the CDC as a method for monitoring large-scale bioterrorism attacks in locations with little to no baseline data (e.g., less than 7 days) (CDC, 2010b). For example, EARS was used for syndromic surveillance at the Democratic National Convention in 2000, the Super Bowl and World Series in 2001 (Hutwagner, 2003), and Hurricane Katrina in 2005 (Toprani, 2006). Following the terrorist events of 11 September 2001, EARS has also been used as a routine health surveillance system by various city, county, and state public health officials.

EARS is primarily focused on providing public health care officials with a means for early event detection (EED). It is important to remember that EED does not

necessarily mean that an outbreak is occurring. Rather, EED provides a signal that an outbreak *may* be occurring and potential justification for expending resources to further investigate. As a by-product of this investigative process, enhanced SA about the specified syndrome will more than likely be achieved. It is important to note that biosurveillance systems are not the sole means of detecting an outbreak. It may very well be the case that clinicians or sentinel physicians will be faster at detecting an outbreak than a biosurveillance system. Essentially, detection depends on the specific circumstances and in some cases, luck (Fricker et al., 2010). Sosin (2003) conveys the idea that biosurveillance systems can act as a safety net, should the traditional detection methods, such as clinical diagnosis, fail.

3. EARS and Monterey County Biosurveillance

In 2004, Monterey County staff received training in some of the available biosurveillance systems. Ultimately, the county decided to use EARS because of the system's flexibility and allowance to keep the data local. In particular, Monterey County Health Department (MCHD) liked the fact that it could develop its own syndromes for unique, local circumstances such as agriculture pesticide spraying and fire-related illness tracking (Fricker & Hanni, 2010).

Figure 6 (a tailored version of Figure 5) illustrates how MCHD implements the EARS biosurveillance system. Notice that the final reporting step corresponds to the Daily Observational and Situational Evaluation (DOSE) report which is updated daily and posted on the Internet¹. Figure 7 is an example of what the Monterey County DOSE report looks like. The varying levels of “alert” correspond to a color-scheme of green, yellow, orange, and red. A green block indicates that there were no alert flags from the previous day (e.g., no health concerns) while a red block indicates multiple alert flags (e.g., highest level of concern). Alert colors other than green (e.g., action items) are usually discussed at the beginning of the report. While EARS is primarily responsible for

¹ The most current DOSE report can be found by visiting the MCHD Web site at: http://www.co.monterey.ca.us/health/healthalerts/pdf/MC_DOSE.pdf.

analyzing the data collection from clinics, hospitals, and ambulance reports, the DOSE report also encompasses other data categories as specified in Figure 5.

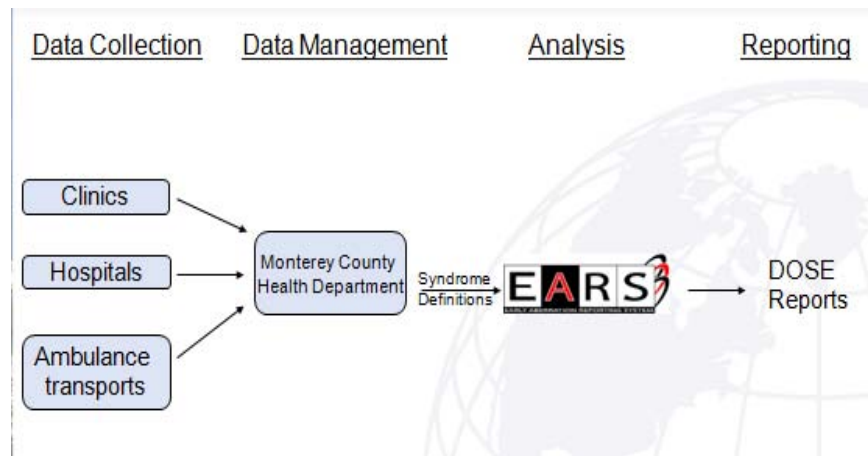


Figure 6. MCHD implementation of the EARS biosurveillance system (After Fricker & Hanni, 2010)

Prior to the 2009 H1N1 pandemic, MCHD had gained valuable experience in using syndromic surveillance to track Influenza-Like Illness (ILI) and to improve response plans. Local hospitals and clinics also benefitted from having access to these daily reports, and thus their compliance with MCHD's data requirements improved. Once the H1N1 virus began to affect the Monterey County population, these pre-established relationships helped with mutual response needs, such as planning for and responding to personal protective equipment (PPE) requests (Fricker & Hanni, 2010).

 Monterey County Daily Observational and Situational Evaluation (DOSE) Report Friday, May 28, 2010 2:50 PM					
This is a report of surveillance activities from the Monterey County Health Department's Public Health Bureau.					
Alert Levels: See below specifications for each indicator. *Data from 5/26/2010 #VALUE!					
Low		Elevated		High	
Low		Elevated		High	
Action Items: The number of patients exhibiting Gastrointestinal symptoms was at the ELEVATED alert level on 5/26. All indicators are listed at the LOW alert level.					
ED Census and Clinic Services Census*:				05/26/10	05/27/10
Green	No alert flags on the previous day	Total Hospital Census	305	315	
Yellow	One alert flag on the previous day	Clinic Services Visits	505	609	
Orange	Two alert flags on the previous day	*Note that the hospital ED aggregate and syndromic surveillance only includes data from facilities providing information for the reported days.			
Red	Three alert flags on the previous day or three consecutive days of single or double alert flags				
Syndromic Surveillance: County hospital aggregate (# of patients)				05/26/10	05/27/10
Green	No alert flags on the previous day	Gastro Intestinal	55	48	
Yellow	One alert flag on the previous day	Influenza-Like Illness	10	8	
Orange	Two alert flags on the previous day	Lower Respiratory	4	11	
Red	Three alert flags on the previous day or three consecutive days of single or double alert flags	Upper Respiratory	20	22	
		Neurological	28	23	
Deaths reported through EDRS: Pulled by date of death, 2:10 PM 05/28/10				Number Filed:	
Green	≤ 1 Standard deviations (over 6 mth avg)		05/26/10	3	
Yellow	> 1 and < 2 Standard deviations (over 6 mth avg)		05/27/10	1	
Orange	> 2 and < 3 Standard deviations (over 6 mth avg)				
Red	> 3 Standard deviations (over 6 mth avg)				
Poison Control Calls* (by date call received)				Number of Calls:	
Green	≤ 1 Standard deviations (over 6 mth avg)		05/26/10	6	
Yellow	> 1 and < 2 Standard deviations (over 6 mth avg)		05/27/10	4	
Orange	> 2 and < 3 Standard deviations (over 6 mth avg)				
Red	> 3 Standard deviations (over 6 mth avg)				
Air Quality: (particulates shown as 1-hr average, µg/m³)				05/27/10	05/28/10
Green	Good = 0 - 40	Monterey	1 @ 2 PM	9 @ 12 PM	
Yellow	Moderate = 41 - 80	Salinas	8 @ 2 PM	10 @ 12 PM	
Orange	Unhealthy for Sensitive Groups = 81 - 175	King City	1 @ 2 PM	1 @ 12 PM	
Red	Unhealthy to Hazardous = Over 175	[Source: http://www.mbusapcd.org/inaps]			
Weather and 3-day Forecast:				Alert Level:	
Source: http://www.arh.noaa.gov/mesowest & http://www.kabw.com					
Highs		Temp. @ 2:00 PM		Inland Coastal	
Green	≤ 84 ° F	Monterey	80 ° F	05/29/10	82 ° F 76 ° F
Yellow	85 - 89 ° F	Salinas	84 ° F	05/30/10	84 ° F 72 ° F
Orange	90 - 94 ° F	Bradley	87 ° F	05/31/10	82 ° F 70 ° F
Red	≥ 95 ° F				
Severe Weather Alerts: There are no weather alerts as of 2:50 PM 05/28/10					
Number of Ambulance Transports (AMR)*:				Number of Transports:	
Green	≤ 1 Standard deviations (over 6 mth avg)		05/26/10	46	
Yellow	> 1 and < 2 Standard deviations (over 6 mth avg)		05/27/10	43	
Orange	> 2 and < 3 Standard deviations (over 6 mth avg)				
Red	> 3 Standard deviations (over 6 mth avg)				
Other:					
- CD Unit is following 3 outbreaks. - Student absenteeism surveillance: On average, 1.2% of the Monterey County sentinel school student population was absent each day with influenza-like illness during the school week of May 3, 2010. - There are currently 0 wildfires reported for Monterey County. - Rain advisory in effect for all Monterey County beaches. The public is advised not to have ocean water contact for three days after the last rain event.					
DOSE completed by Monterey County Health Department, Epidemiology and Evaluation					
Questions or comments may be directed to (831)755-4588.					
DOSE based on DSA7 report from San Mateo County Public Health Department.					
Report can be accessed through the following link: http://www.co.monterey.ca.us/health/HealthAlerts/pdf/IMC_DOSE.pdf					

Figure 7. Illustration of the reporting component of a biosurveillance system, the Monterey County DOSE Report for Friday, May 28, 2010 (From MCHD, 2010b)

C. 2009 H1N1 “SWINE FLU” VIRUS

Cases of swine flu had been reported in literature prior to 2009; however, two children in southern California in March of that year were the first U.S. introduction to the current pandemic. Within a week of the CDC determining that the strains were genetically similar, there were reports of widespread severe flu activity in Mexico and even more cases in the United States and possibly Canada. By the second week after local health departments were alerted, and widespread H1N1 activity was reported across North America and into Europe. This pandemic spread considerably faster than expected, taking only six weeks to go from a local outbreak to a pandemic (personal communication with Hanni, 2009).

When the novel H1N1 flu outbreak was first detected in mid-April 2009, the CDC began working with states to collect, compile, and analyze information. From April 15 to July 24, 2009, states reported a total of 43,771 confirmed and probable cases of the H1N1 infection. Of these cases reported, only 12 percent were either hospitalized or died (CDC, 2010c). Illustrated in Figure 8, the number of cases reported during this timeframe per 100,000 people was highest among the 5 to 24 year age group (26.7 per 100,000) and lowest in people 65 years and older (1.3 per 100,000) (Hanni, 2009).

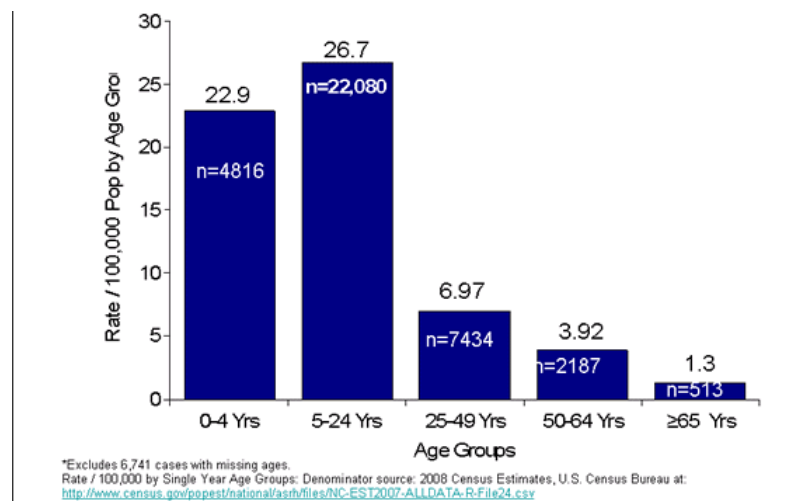


Figure 8. Confirmed and probable H1N1 case rate by age group from April 15–July 24, 2009 (From Hanni, 2009)

Figure 9 illustrates the estimated pandemic H1N1 flu hospitalization rate in the United States by age group from April 15 to July 24, 2009. These estimates are based on the 4,738 hospitalizations that were reported to the CDC during this time period. The reported hospitalization rate per 100,000 people was highest among children 0 to 4 years of age (4.5 children per 100,000) and lowest among people in the 25 to 49 years of age group (1.1 per 100,000) (personal communication with Hanni, 2009).

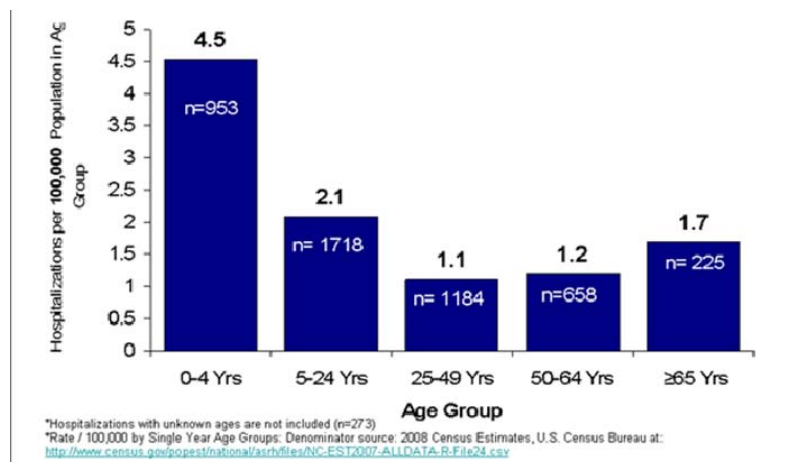


Figure 9. Estimated pandemic H1N1 flu hospitalization rate in the United States by age group from April 15 to July 24, 2009 (From Hanni, 2009)

On July 24, 2009, confirmed and probable case counts were discontinued after the CDC deemed the virus “widespread” across the United States (CDC, 2010c). In order to approximate the number of novel H1N1 flu cases in the US, a CDC model was developed that took the number of cases reported by states and adjusted the figure to account for underestimation. For instance, not all people with the virus sought medical care, and of those who did, some may not have been specifically tested for H1N1. The CDC model estimated that more than one million people became ill with novel H1N1 flu between April and June 2009 in the United States (CDC, 2010c).

Monterey County relates the symptoms of the pandemic H1N1 2009 influenza to symptoms of regular, seasonal flu. According to the MCHD website as of July 2010, people usually exhibit one or more of the following symptoms:

- | | |
|-----------------------------|----------------------------|
| --Fever greater than 100° F | --Sore throat |
| --Cough | --Headache |
| --Body aches | --Fatigue and/or dizziness |
| --Chills | --Vomiting and diarrhea |

Federal agencies working on pandemic influenza planning guidance understood the effects that H1N1 cases would have on the delivery of patient care. This understanding, however, had to be balanced with the availability of scarce resources (Hanfling & Hick, 2009). Perhaps Dr. Hanni (2009) said it best:

Our surveillance tools are many, but we also have potential for a strain on our resources, both staffing and supplies for the coming flu season statewide and locally. We have identified new risk groups that will need to be monitored and it is apparent that we will be needing to monitor for several strains of influenza and other respiratory viruses, especially given the fact that our vaccine for H1N1 will be available later in the season. To that end, there have been some changes in reporting requirements that have also resulted in some changes to what data we are collecting locally.

Monterey County often reviewed the CDC's guidance for determining if a patient should be tested for the H1N1 virus. These guidelines, as outlined in Figure 11, were first developed by the CDC in mid-June 2009 and last updated in mid-September 2009. During this particular period of increased concern over H1N1, MCHD looked at each alarm and engaged in daily discussions with the Infection Control Practitioners at their four hospitals (personal communication with Hanni, 2010). In other words, in order to achieve sensitivity for detecting H1N1 in Monterey County, MCHD was willing to tolerate a high false positive rate.

As for diagnosis, the MCHD lab was able to process samples within a few days to a week which indicated whether a person had the generic Influenza A virus (i.e., not a known H1 or H3 virus). If a person tested positive, that was good enough to proceed as if the case was positive for H1N1 while the sample was sent to the state lab, which meant physicians and their staffs would wear the appropriate PPE with that person in a room. Due to the large influx of samples, however, state lab testing was sometimes delayed by as much as several weeks (personal communication with Hanni, 2010).

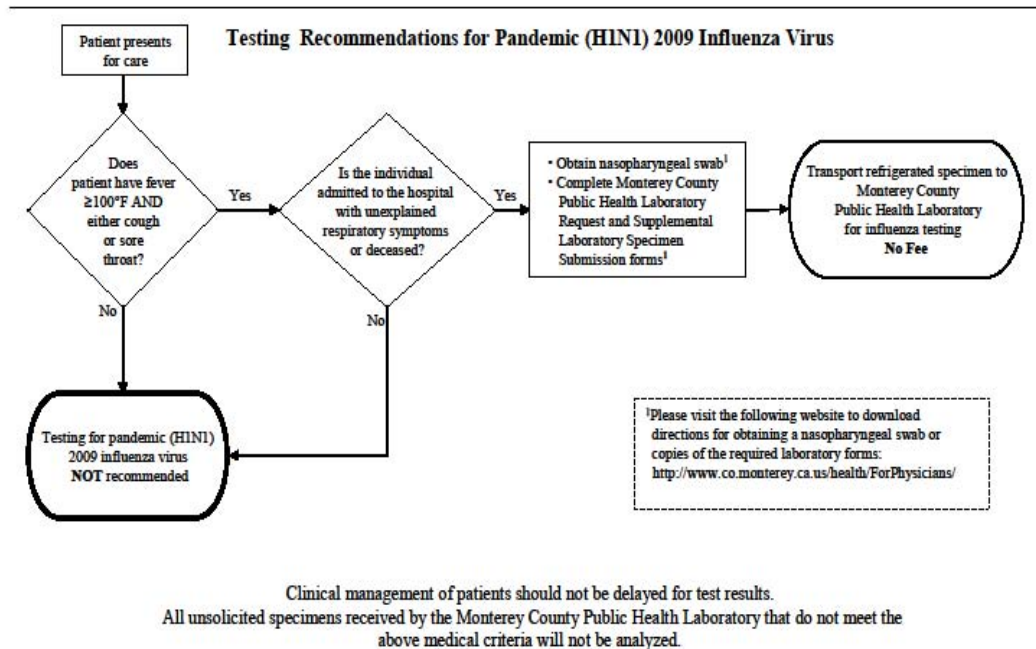


Figure 10. CDC Testing Recommendations for Pandemic (H1N1) 2009 Influenza Virus, last updated September 08, 2009 (From MCHD, 2009)

Figure 11 shows the cumulative total of confirmed H1N1 cases in Monterey County from May 10, 2009 to April 5, 2010. With 47 cases, September 2009 saw the highest number of confirmed H1N1 counts. August and October 2009 were the next highest months with counts in the mid-thirties.

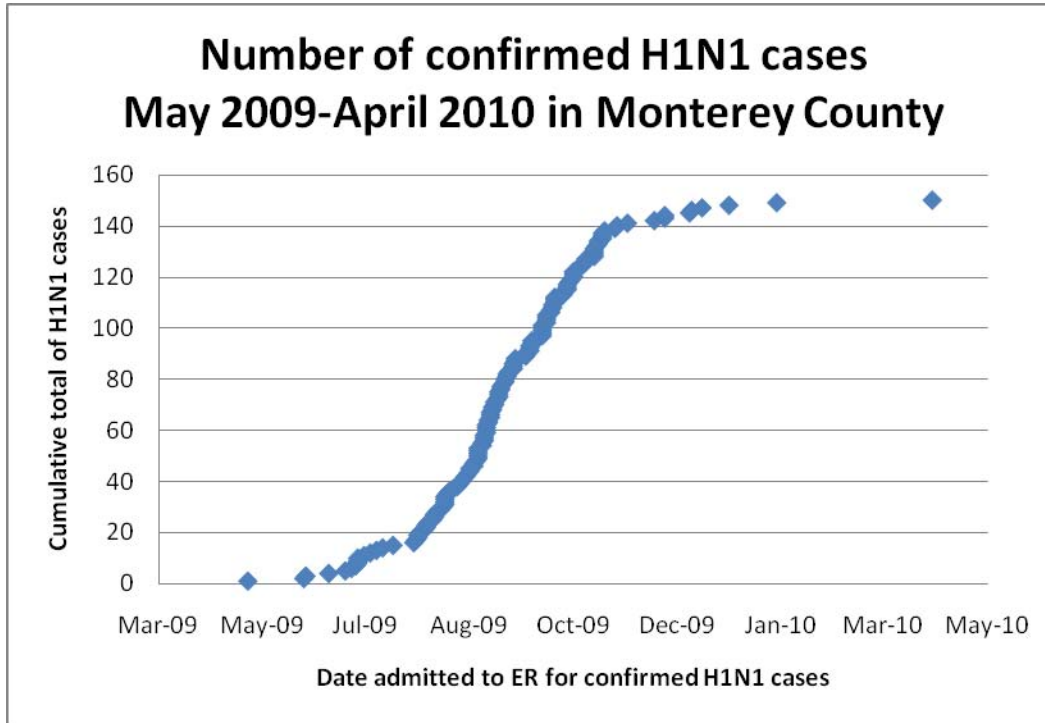


Figure 11. Number of confirmed H1N1 cases from May 2009 to April 2010 in Monterey County (From Hanni, 2010)

D. ORGANIZATION OF THIS THESIS

There are two main components of EARS that are evaluated in this research, the logic and the EED algorithms. As employed before the first H1N1 outbreak, EARS was deficient in detecting signals in the data. The following chapters are organized as follows. Chapter II describes how modifications to EARS' logic and/or settings can affect the system's EED performance. Chapter III describes the algorithms used to evaluate the relative performance of the various EED methods studied against confirmed cases of the H1N1 virus, as well as a description of how various input and threshold values were chosen. Chapter IV summarizes the results of the evaluation and makes recommendations for future EED improvement.

II. SYNDROMES: DEFINING AND CALCULATING DAILY COUNTS

A. BACKGROUND

EARS was originally intended to serve as a drop-in surveillance system; however, it is increasingly being used as a routine health surveillance system by local public health departments. Unfortunately, little research has been conducted to verify the EED performance of EARS; specifically, whether there are any changes that can be made to EARS' logic and/or settings that would maximize the system's EED performance. To gain more insight, this chapter describes a quantitative comparison of how modifications to EARS' logic would affect daily counts (e.g., as outlined in the DOSE report) of the ILI syndrome. The following three areas of possible logic modifications were explored: syndrome definitions, symptom aliases, and text matching algorithms.

B. CHIEF COMPLAINT DATA COLLECTION AND REPORTING

Monterey County's EARS system uses data from six public clinics and four hospitals located throughout the county. The data received from MCHD include the date of patient visit; age, sex, and home ZIP code of the patient; chief complaint and diagnosis code for the clinics only. All clinic visits that occurred during the previous work day are electronically transmitted daily to MCHD. When clinic offices are closed during the weekends and select holidays, data transmission does not occur.

Daily counts for various syndromic categories are calculated based on the presence or absence of key words in the chief complaints. For example, existence of either the word "flu" or the phrase "fever and cough" in an individual's chief complaint would result in that individual being included in the ILI syndrome count for that day. Refer back to Tables 2 and 3 for examples and follow-on discussion regarding syndromes and chief complaints.

Dr. Hanni (2009) notes that MCHD tracks influenza in the population by using reports of ILI from providers and clinics, both locally, statewide, and nationally. Here, it is important to identify several crucial dates and events from such reports. As depicted in

Figure 12, the y-axis shows the percentage of visits (by week) that are due to ILI, beginning in October 2008 and ending in October 2009. The dashed horizontal line is the national baseline, above which classifies as an epidemic situation.² Comparing 2009 data (in red) to previous years, Monterey County experienced a relatively mild influenza season. However, in late April and early May 2009, there was an unexpected increase in the percentage of outpatients that were being seen for ILI, which continued throughout the summer and again peaked in late October and early November 2009. Also of note is that on April 17, 2009, the CDC determined that the current pandemic H1N1 flu virus was active in the United States (personal communication with Hanni, 2009). On June 11, 2009, the World Health Organization (WHO) declared that H1N1 was a global pandemic (WHO, 2010).

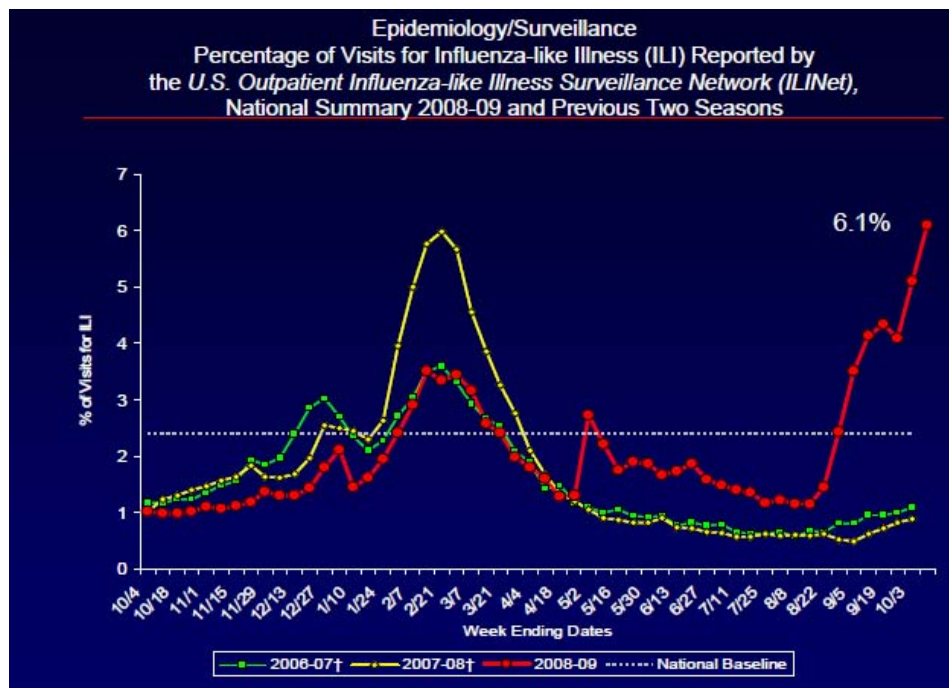


Figure 12. Percentage of visits for ILI reported by the U.S. Outpatient ILI Network, National Summary 2008–09 and previous two seasons (From CDC, 2009)

² For a detailed discussion on the ILI national baseline, refer to http://scienceblogs.com/effectmeasure/2009/09/trying_to_make_sense_of_flu_se.php.

C. CHANGES IN LOGIC

1. Symptom Aliases

In order for EARS to produce meaningful results, raw data such as text-based chief complaints must be turned into numerically-coded variables, such as syndrome indicators or demographic variables. As an example, the following pseudo-code creates an indicator variable “ili_ind” for the ILI syndrome by searching for keyword substrings within chief complaint text:

```
loop from i = 1 to number of data records
set ILI_ind(i) = 0
loop from j=1 to number of ILI keywords
  if ILI_keyword(j) is a substring in chief_complaint(i)
    then ILI_ind(i) = 1
```

The above example identifies whether a set of keywords or symptoms aliases is contained in a free-form text block. If so, these symptoms will become matched with a particular syndrome, in this case ILI. Computer logic requires a list of keywords to be searched for, including abbreviations, acronyms, and common misspellings; however, caution should be exercised when creating this list. If an acronym is too generic, for example, specificity may be jeopardized.

To illustrate, the original CDC definition for the ILI syndrome is defined by the following symptoms: “sore throat” or “cold” or “cough.” Table 4 shows a subset of the actual EARS’ “symptoms_code” file, which maps keywords to their respective symptoms.

<u>Keywords</u>	<u>Symptom</u>	<u>Keywords</u>	<u>Symptom</u>	<u>Keywords</u>	<u>Symptom</u>
SROETHROAT	SORETHROAT	COL	COLD	C0UGH	COUGH
SSORE THROAT	SORETHROAT	NOSE	COLD	C9UGH	COUGH
ST	SORETHROAT	URI	COLD	CCOUGH	COUGH
TBROAT	SORETHROAT	EAR PAIN	COLD	CIUGH	COUGH
THROAT	SORETHROAT	DISCH	COLD	CKUGH	COUGH
TH4OAT	SORETHROAT	OM	COLD	CLUGH	COUGH

Table 4. Subset of the CDC “symptom_code” file, as used by EARS to map keywords to symptoms (e.g., sore throat, cold, cough)

With the EARS' SAS code that MCHD is currently using, if a keyword (as found in the "symptoms_code" file) is found anywhere within the free-text chief complaint field, it will match to the respective syndrome. This simplified matching logic can be problematic, as in the case where "ASTHMA" is mapped to the SORETHROAT symptom because of the keyword "ST." Other examples of inappropriate matching include:

- "URINARY INFECTION" mapping to COLD because of keyword "URI"
- "COLPOSCOPY" mapping to COLD because of keyword "COL"
- "MOM NEEDS FOLLOW-UP" mapping to COLD because of keyword "OM"

The above examples demonstrate that without a more sophisticated text matching algorithm, the above pseudo code is subject to spurious false positives. A better approach might involve setting ili_ind=1 when, for example, "ST" occurs as a separate word in the chief complaint text (as in "ST AND FEVER") but not as part of a longer word (as in the "ASTHMA" example above). The algorithm must also be flexible enough to map root words found within the text string (e.g., "COUGH" within the text "COUGHING").

In addition to concerns associated with the text coding algorithm, the "symptoms" file itself must be heavily scrutinized. For example, the following inappropriate mappings were discovered when using the original CDC symptoms file:

- "NOSE BLEEDS" maps to COLD because of keyword "NOSE"
- "VAGINAL DISCHARGE" maps to COLD because of keyword "DISCH"
- "4 YEAR OLD WCC" maps to COLD because of keyword "OLD"

Thus, in an effort to improve the overall specificity of the EARS program, aliases "NOSE," "DISCH," and "OLD" were removed from the "symptoms" file.

2. Syndrome Definitions

Figure 13 illustrates three possible ILI syndrome definitions. According to the original CDC EARS definition, a record is flagged for ILI when a patient complains of any one or more of the following symptoms: "sore throat" or "cold" or "cough." As

shown in Figure 13, MCHD created an expanded definition of the ILI syndrome by allowing for many more symptom possibilities. In doing so, the goal was to increase the chances of correctly classifying someone with the flu, though this strategy comes at the cost of also increasing the number of false positives (i.e., counting those without the flu in the ILI syndrome).³

ILI Syndrome Definitions	
• EARS (CDC) – Sore throat or – Cold or – Cough • Restricted (MCHD) – Fever & cough – Fever & sore throat – Fever & cough & sore throat – Flu and – Not shot	• Expanded (MCHD) – Cold or – Cough or – Fever or – Chills or – Muscle pain or – Headache or – Flu and – Not shot

Figure 13. Three versions of the ILI syndrome definition: original EARS (CDC), expanded MCHD, and restricted MCHD

Monterey County now currently uses a more restricted definition of the ILI syndrome which has substantially lowered the number of records flagged.⁴ Instead of simply using one symptom to flag ILI, they now require more “evidence,” in the sense that to flag for ILI someone has to have two or more symptoms, such fever *and* cough, for example. By using the restricted definition, the current goal is to limit the chances of incorrectly counting an individual in the ILI syndrome who does not actually have the

³ Medical and public health professionals would say that the expanded definition is intended to increase sensitivity at the cost of decreasing specificity.

⁴ “Not shot” under the Restricted (MCHD) ILI syndrome definition means that the word “shot” is not included in the chief complaint field. This ensures that a chief complaint containing the text “flu shot” will not be included as an ILI syndrome indicator because of the existence of the word “flu” in the text.

flu, though this strategy comes at the cost of a greater chance of missing some true positives (i.e., failing to count those with the flu in the ILI syndrome).⁵

To illustrate the implications of different syndrome definitions, Figure 14 shows the estimated aggregate ER ILI activity for Monterey County and California since fall 2009. The upper line is the result of using the expanded ILI syndrome definition as compared to the original CDC definition (sore throat, cold, or cough) used by California Sentinel Providers. At a close look, the two plots seem to have similar trends and signals. This correlation may imply that it doesn't really matter which definition you use; however, it could be that the larger cycle in the upper line could mask a real signal.

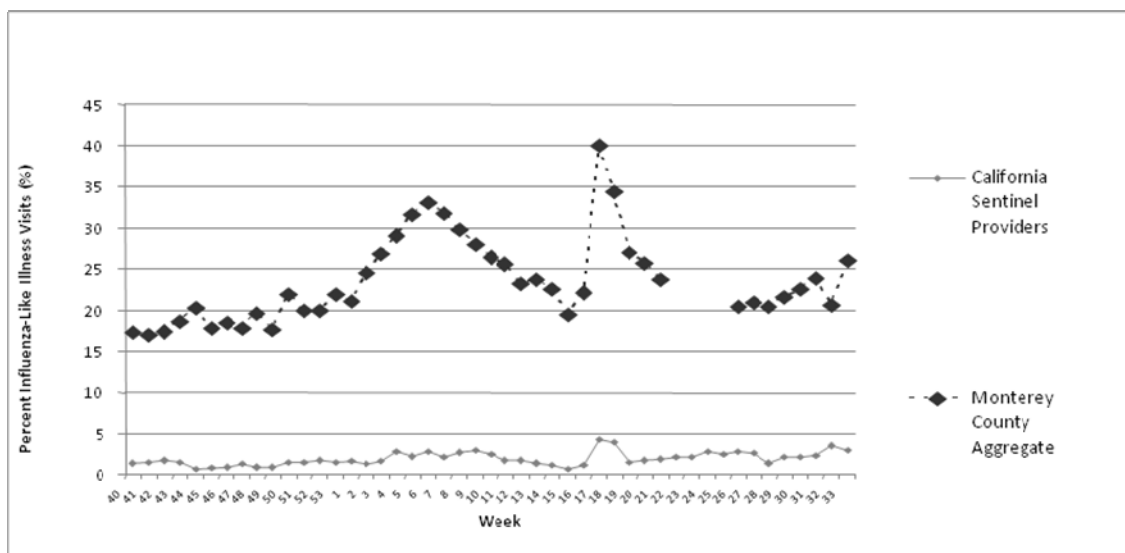


Figure 14. Emergency Room Influenza-Like Illness Visits for Monterey County and California, 2008–2009 Season (From Hanni, 2009)

3. Text Matching Algorithms

Changes to the EARS' text matching logic can also have a large and significant impact on the number of individuals coded with a given syndrome. In the context of the ILI syndrome, the original EARS (CDC) text-matching logic basically says the following: if an ILI keyword is found *anywhere* within the chief complaint field (even in

⁵ For the restricted definitions, medical and public health professionals would say that it increases specificity at the cost of decreasing sensitivity.

the middle of a word), then it will be flagged as an ILI indicator. By using such simple logic, words like “COLPOSCOPY” will be coded as a “COLD” symptom because of the keyword “COL.” In a similar example, if the letters “MI” map to the “CARDIAC” symptom, one would not want the word “VOMIT” to be associated with cardiac symptoms just because the letters “MI” were contained within that word.

In an effort to mitigate inappropriate symptom coding, a more sophisticated approach would be to revise the text matching logic. Under the proposal known as “enhanced NPS logic,” the text-matching algorithm only allows symptom matches if ILI keywords are at the beginning or end of a word (or matches the word exactly). The idea is that for any keyword that is less than or equal to *three* letters long, there cannot be any letters before or after it. Otherwise, it will not count as a symptom indicator. For example, by using enhanced NPS logic, the keyword “ST” in chief complaint “TEST” would not be flagged as an ILI indicator, nor would the keyword “COL” in chief complaint “RE-COLPO” be flagged as an ILI indicator. The reason in both cases is because there are letters on either side of the keyword. See Figure 15 for a graphical depiction of the examples above.⁶

⁶ “Red” or dark colored blocks indicate that text-matching logic will prevent keywords from mapping to symptoms if additional letters or symbols are present. “Powder blue” or light colored blocks indicate that additional letters or symbols are allowed by text-matching logic in mapping keywords to symptoms.

Enhanced (NPS) Logic

- For short words (≤ 3 characters)
 - No variations on alias words allowed
 - Example: ST
 - NP FOR HIV TEST PER VERONICA/CHART MADE/LM
 - Example: COL
 - RTNRE-COLPO/LM
- For longer words (≥ 4 characters)
 - Variations on one side of the alias word is allowed
 - Example: COUGH
 - PT COUGHING FOR 4 DAYS
 - Example: OUGH
 - R/S PREV APPT CALL NOT GOING THROUGH

Figure 15. Examples of how “Enhanced NPS Logic” works for text-matching keywords with symptoms

For words greater than or equal to four characters using enhanced NPS logic, variations on the symptom keywords are allowed but only on one side of the word or the other. For example, the word “COUGH” can have many variations, such as COUGHS, COUGHED, or COUGHING. In which case, it is appropriate to flag such variations of the word “COUGH” as being an ILI indicator. Unfortunately, using the enhanced logic alone does not necessarily guarantee that all inappropriate mappings will be eliminated. For example, keyword “OUGH” would allow the following chief complaint to be inappropriately flagged for ILI: “PREV APPT CALL NOT GOING THROUGH”. This example also highlights the importance of carefully reviewing the “symptoms” file. See Appendix A for the R code on implementing “enhanced NPS logic.”

4. Summary of Changes in Logic

After exploring the three areas of logic, Figure 16 illustrates the quantitative results of the analysis performed. The top box in Figure 16 refers to the original CDC

text matching logic, symptom aliases, and syndrome definitions as the “Base Case.”

When running these algorithms together, out of 153,696 total records, 9,093 records (almost 6%) were flagged as being ILI syndromes.

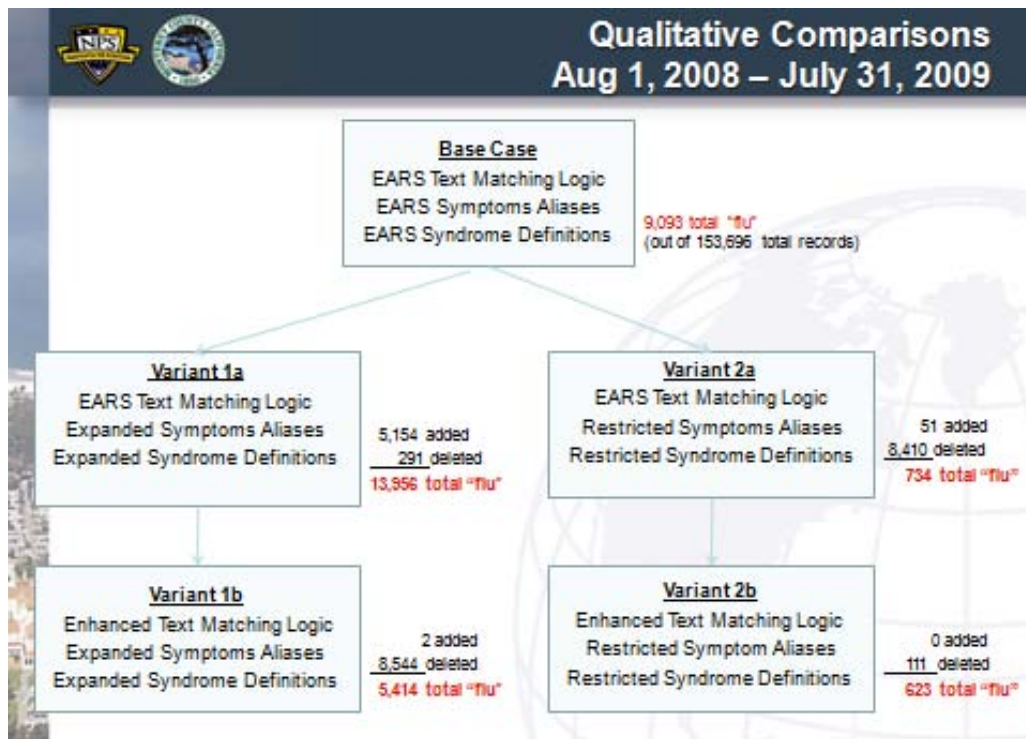


Figure 16. Qualitative results of changes to EARS logic, symptom aliases, and syndrome definitions, as applied to MCHD data from August 1, 2008–July 31, 2009

In the second tier of Figure 16, the EARS text matching logic remained the same, while the symptom aliases and syndrome definitions were altered. The box on the left indicated as “Variant 1a” used the expanded ILI syndrome definition (refer to Figure 13) as well as a more robust symptom alias list, which has been previously demonstrated to yield an increase in spurious matches. The expanded aliases and syndrome definitions resulted in a 53% increase in the number of records flagged for the ILI syndrome. In comparison, the box on the right labeled “Variant 2a” used the restrictive symptom alias list and restrictive syndrome definitions, which resulted in a reduction of the number of records flagged for ILI by nearly 92% of the original “Base Case.”

The bottom tier of Figure 16 illustrates the significance of changes to the text-matching logic, which results in similarly large swings in the number of coded ILI

syndromes. For example, the only difference between “Variant 1a” and “Variant 1b” is the change in text matching logic, which results in a 62% decrease (13,956 down to 5,414) in the number of records flagged for ILI.

Figure 17 is a visual representation of smoothed ILI counts over time using the five sets of logic described above. The Base Case (solid black line) and Variants 1a and 1b (dashed lines directly above and below the solid line) basically look to be the same curve shifted up or down, which suggests those variants are simply adding in or subtracting out some sort of average level of noise. On the other hand, there are some smaller differences (in terms of the “spikes”), which may turn out to be significant differences.

Variants 2a and 2b (the two lowest curves), look different from the other three, not just because of the significantly lower average daily counts, but also because it looks like some of the trends differ. For example, the time series from 1a, 1b, and the Base Case all show a clear spike between times 0–50, followed by another spike between times 50–100. In contrast, Variants 1b and 2b only seem to show a spike from times 50–100.

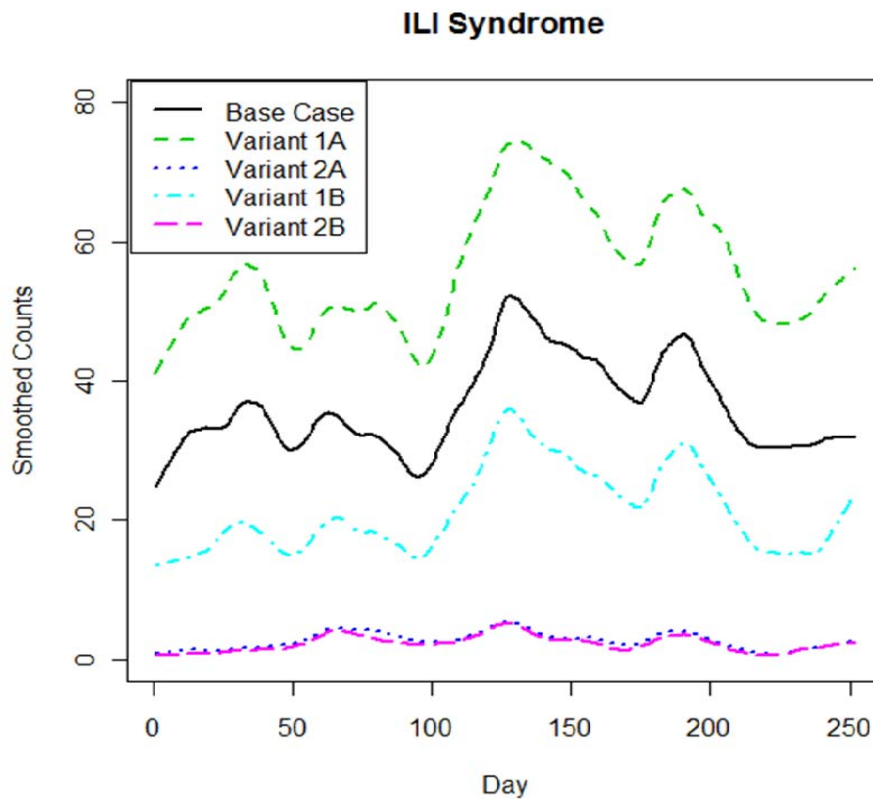


Figure 17. Smoothed ILI counts from MCHD during August 1, 2008–July 31, 2009 (excluding weekends) using five different combinations of text matching logic, symptom aliases, and syndrome definitions.

In Chapter III, the ILI counts produced by the Base Case, Variant 1a, and Variant 2a logic will be used as input into EARS and other EED algorithms. Unfortunately, it was simply not possible to choose the best set of daily counts using a “gold standard.”⁷ In a study performed by Espino and Wagner (2010), the classification performance of International Classification of Diseases 9th Edition (ICD-9) based ⁸ detectors was measured against the human classification of cases and found that:

⁷ As defined by The Free Dictionary (<http://www.thefreedictionary.com/gold+standard>), gold standards are “the supreme example of something against which others are judged or measured.”

⁸ As defined by The Free Dictionary (<http://medical-dictionary.thefreedictionary.com/ICD-9-CM>), ICD-9 codes are “A standardized classification of disease, injuries, and causes of death, by etiology and anatomic localization and codified into a 6-digit number, which allows clinicians, statisticians, politicians, health planners and others to speak a common language, both U.S. and internationally.”

ICD-9-coded diagnoses offer no advantages in positive predictive value and specificity. However, because the diagnosis codes are significantly delayed (on average by 7.5 hours), it is clearly the case that detection systems should focus on the chief complaint data, when it is available.

Since clinics adhere to ICD-9 codes for insurance billing purposes, it would be plausible to assume that these codes could be used as the “gold standard.” Unfortunately, there were many obvious issues that made ICD-9 codes unsuitable for use. Specifically, the code V20 appeared quite often as a catch-all for many seemingly unrelated illnesses (e.g., flu, women's wellness exams, HPV shots, etc).⁹ For these reasons and despite having ICD-9 codes in the MCHD dataset, it was determined not to use them as a “gold standard” and instead use documented H1N1 cases. Finally, Chapter IV will compare the results of various algorithms to see if one or more can do better at detecting a known H1N1 outbreak in Monterey County.

⁹ V codes are considered to be ‘Supplementary Classification of Factors Influencing Health Status and Contact With Health Services,’ where the V20 code is listed as ‘health supervision of infant or child’ under the subcategory of ‘Persons Encountering Health Services In Circumstances Related To Reproduction And Development.’

III. EARLY EVENT DETECTION METHODS

This chapter describes the EED methods that are tested in conjunction with the various sets of logic presented in Chapter II. These EED algorithms include the methods currently used in the EARS system as well as a CUSUM-based method first described and evaluated in Fricker et al. (2008). The goal of these methods is to monitor the syndromic data as it comes into MCHD over time and signal when the data deviates significantly from trends in the recent past.

A. EXISTING EED METHODS: EARS

As described in Fricker et al. (2008), the current EARS' detection methods are called "C1," "C2," and "C3." The C is likely an abbreviation for the cumulative sum (CUSUM) methodology from which the EARS documentation claims these methods were derived. However, as Section B makes clear, this description is incorrect because none of these methods are actually based on or derived from the CUSUM.

The C1 method calculates a standardized syndrome daily count for day t using the sample average and sample standard deviation estimated from the previous 7 days of daily counts,

$$C_1(t) = \frac{Y(t) - \bar{Y}_1(t)}{S_1(t)} \quad (1)$$

where

- $Y(t)$ is the observed syndrome count for day t
- $\bar{Y}_1(t)$ is the sample mean based on the previous 7 days of data,

$$\bar{Y}_1(t) = \frac{1}{7} \sum_{j=t-7}^{t-1} Y(j), \text{ and}$$

- $S_1(t)$ is the sample standard deviation based on the previous 7 days of data,

$$S_1(t) = \frac{1}{6} \sum_{j=t-7}^{t-1} [Y(j) - \bar{Y}_1(j)]^2$$

As implemented in EARS, the C1 method signals an alarm at time t when the $C_1(t)$ statistic exceeds a fixed threshold h , which occurs when the observed count exceeds three sample standard deviations from the sample mean: $C_1(t) > 3$.

The C2 method is very similar to C1 but differs in that it uses a 2-day lag before calculating a standardized value using the previous 7 days worth of data. Specifically,

$$C_2(t) = \frac{Y(t) - \bar{Y}_3(t)}{S_3(t)} \quad (2)$$

where

- $Y(t)$ is the observed count for period t
- $\bar{Y}_3(t)$ is the moving sample mean, $\bar{Y}_3(t) = \frac{1}{7} \sum_{j=t-3}^{t-9} Y(j)$, and
- $S_3(t)$ is the moving sample standard deviation, $S_3(t) = \frac{1}{6} \sum_{j=t-3}^{t-9} [Y(j) - \bar{Y}_3(j)]^2$

The C2 method also signals at time t when the $C_2(t)$ statistic exceeds a fixed threshold h , which occurs when the observed count exceeds three sample standard deviations from the sample mean: $C_2(t) > 3$.

Lastly, the C3 method combines current and historical data from day t and the previous two days, and signals an alarm when $C_3(t) > 2$. It calculates the statistic at time t as follows:

$$C_3(t) = \sum_{j=t}^{t-2} \max[0, C_2(j) - 1] \quad (3)$$

Of particular note and concern, EARS' C1, C2, and C3 algorithms factor in “zeros” for days when clinics are not open for business (e.g., weekends and holidays). Given that the sample mean and sample standard deviation are based on the previous 7-9 days worth of data, exceeding the alarm thresholds will prove difficult, as shown in Chapter IV.

B. AN ALTERNATIVE EED METHOD BASED ON THE CUSUM

1. Basic CUSUM

The CUSUM is a statistical process control (SPC) methodology often used in managing the quality of manufactured items. See Montgomery (2001) for an introduction and Hawkins and Olwell (1998) for a comprehensive treatment of the CUSUM. The most commonly used CUSUM is of the form,

$$C_{t+1} = \max[0, C_t + Y_{t+1} - \mu_0 - k], \quad (4)$$

where X_{t+1} is the count at time $t+1$, $X_{t+1} \sim N(\mu_0, \sigma^2)$ is the desired state of the process, and k is referred to as the “reference interval.” For the CUSUM designed to detect the shift in the mean of a normal distribution from μ_0 to μ_1 , the reference interval is defined as

$$k = \frac{\delta}{2} \sigma = \frac{\sigma |\mu_1 - \mu_0|}{2}, \quad (5)$$

where μ_1 is the mean shift that is desired to be detected quickly. Note that this is a one-sided CUSUM designed to detect increases in the mean. In industrial quality control applications, two CUSUMs are often employed—one to look for increases in the mean and another to monitor for decreases. In syndromic surveillance, however, only employing a single CUSUM (such as Equation 4) to look for increases is appropriate since detecting decreases in disease incidence is generally not of interest.

2. CUSUM Applied to Biosurveillance Data

In biosurveillance, the data is unlikely to be stationary since syndromic surveillance daily counts often have uncontrollable systematic effects and trends such as an annual influenza seasonal cycle and day-of-the-week effects. Yet, the CUSUM of Equation 4 and its use in quality control is based on the assumption of stationary data. It is therefore inappropriate to apply the CUSUM directly to biosurveillance data.

Instead, per Montgomery (2001), an appropriate approach is to model the systematic trends of the data and then apply CUSUM to the prediction errors from the model. Fricker et al. (2008) found the “adaptive regression with sliding baseline” of Burkom (2007) to be an effective modeling technique for removing systematic trends in syndromic surveillance data and this work will use the same approach.

As described by Fricker et al. (2008), the basic idea behind adaptive regression is as follows. Let X_t be the observation on day t , say the number individuals presenting at a clinic or emergency room with a particular syndrome. For $t \geq n$, where n (the “baseline period”) is some fixed number of time periods, model the most recent n daily syndrome counts, $X_t, X_{t-1}, \dots, X_{t-n+1}$, as a linear function of time relative to day t :

$$Y_i = \beta_0 + \beta_1 \times (i - t + n) + \beta_2 I_{\text{Mon}} + \beta_3 I_{\text{Tues}} + \beta_4 I_{\text{Wed}} + \beta_5 I_{\text{Thurs}} + \varepsilon \quad (6)$$

where, for $i = t, \dots, t-n+1$, β_0 is the intercept term, β_1 is the slope, the I s are indicator functions ($I = 1$ on the relevant day of the week and $I = 0$ otherwise) and ε is the error term to account for random variability. Following the approach of Fricker et al. (2008) and in spite of the non-normal time series data, the model is fit using ordinary least squares regression and is re-fit each day by using the most recent n observations as the sliding baseline.

Once fit, the model is used to estimate the predicted count for the current day ($t+1$),

$$\hat{Y}_{t+1} = \hat{\beta}_0(t) + \hat{\beta}_1(t) \times (n+1) + \hat{\beta}_j(t), \quad (7)$$

where $\hat{\beta}_0(t), \hat{\beta}_1(t)$ and $\hat{\beta}_j(t)$ are the estimated model coefficients from the regression fit at time t , and where $\hat{\beta}_j(t)$ is the relevant estimated day-of-the-week coefficient. Given the daily count at time $t+1$, X_{t+1} , the predicted count is then used to calculate the prediction error at time $t+1$,

$$\Delta_{t+1} = Y_{t+1} - \hat{Y}_{t+1}, \quad (8)$$

which is then standardized using the estimated standard deviation of the prediction errors from the last m time periods, $Z_{t+1} = \Delta_{t+1} / \hat{\sigma}_t$, for $\hat{\sigma}_t = \sqrt{\sum_{i=t-m+1}^t (\Delta_i - \bar{\Delta}_i)^2}$ so that $Z_{t+1} \sim N(0,1)$. For biosurveillance, the CUSUM of Equation 4 thus becomes

$$C_{t+1} = \max[0, C_t + Z_{t+1} - k]. \quad (9)$$

3. Application of the CUSUM to MCHD Data

Before applying Equation 9 directly to the MCHD dataset, the assumption of normality of the standardized residuals was assessed using a “historical” set of data.¹⁰ Using this subset of data, a baseline period of $n=35$ days and an additional $m=10$ days, standardized residuals were calculated from an adaptive regression. See Appendix B for the MATLAB code used to fit the adaptive regressions and calculate the standardized residuals. Normal quantile-quantile (Q-Q) plots demonstrated that the standardized residuals were reasonably normally distributed. See Appendix C for Q-Q plots and standardized residual plots of the historical data set.

Although Fricker et al. (2008) and Burkom (2007) used an 8-week sliding baseline ($n=56$ days, based on a 7-day week), preliminary research on the MCHD data found a 7-week sliding baseline ($n=35$ days, based on a 5-day week) to be preferred across all algorithm variants. This preference stemmed from evaluations of the Q-Q plots of the residuals, where for smaller and larger n the residuals for some of the algorithm variants began to show departures from normality. Fricker et al. (2008) cautions that depending on the particular outbreak of interest, there is a trade-off between the amount of historical data used and the predictive accuracy of the model.

In circumstances where it is important to detect a mean increase quickly, such as when MCHD was on high alert for H1N1, one might reasonably want to detect a one standard deviation increase in the mean. Given that $Z_{t+1} \sim N(0,1)$, $\mu_0=0$ and thus the reference interval from Equation 5 for the standardized residuals becomes

¹⁰ From August 8, 2008-January 12, 2009 equals approximately 1/3 of the entire MCDH data set (or 119 days worth).

$$k = \frac{\mu_1}{2}. \quad (9)$$

For the purposes of this thesis, the k for detecting this magnitude of shift in the mean is called an “aggressive” reference interval where, for $n = 35$, $k = 0.56$.

Besides choosing the reference interval, implementing the CUSUM also requires choosing a threshold h . The choice of threshold is based on the Average Time to False Signal (ATFS) metric, where assuming the CUSUM is re-set after signals, the ATFS is the average time between false signals.¹¹ Thus, the threshold is set to achieve the smallest ATFS that can be reasonably accommodated, given the finite resources available to further investigate the resulting rate of false positive signals. In the case where MCHD was on high alert for H1N1 symptoms, perhaps an ATFS of once every few days would be acceptable. Here, an “aggressive” ATFS was set to be 5 days.¹²

In order to determine threshold h based on a known reference interval k and known ATFS, Hawkins and Olwell (1998) recommend using Siegmund’s approximation

$$ATFS \approx \frac{e^{-2k(h+1.166)} + 2k(h+1.166) - 1}{2k^2} \quad (11)$$

This approximation is not accurate for small ATFS values and so the threshold was instead estimated via simulation. Appendix D contains the R code for the simulation. For a given k , one iteratively runs this simulation for various values of h starting with a small number of iterations, variable “ x .” As one gets closer to the desired ATFS with h , increase “ x ” until the standard error becomes small enough. Under the “aggressive” CUSUM parameters for looking at H1N1 cases, using $h = 0.296$ with $k = 0.56$ achieves an ATFS = 5 (s.e. = 0.0045).

¹¹ In SPC terminology, the ATFS when the biosurveillance algorithm is reset is functionally equivalent to the Average Run Length (ARL).

¹² Five days is one full week (Monday-Friday) since clinics are not open during the weekends.

For comparison purposes, this thesis also looked at more “routine” parameters for monitoring the ILI syndrome with ATFS=20 and $k = 1.06$ (an approximate 2σ shift in the mean). Using the simulation found in Appendix B and under these “routine” parameters, $h = 0.62$. See Table 5 for a summary of the two CUSUM algorithm parameters, and note their labels: CUSUM 1 (aggressive) and CUSUM 2 (routine).

		Parameters		
Type	Label	ATFS	k	h
"Aggressive"	CUSUM 1	5	0.56	0.296
"Routine"	CUSUM 2	20	1.06	0.62

Table 5. CUSUM parameters used in Matlab code for monitoring the ILI syndrome

THIS PAGE INTENTIONALLY LEFT BLANK

IV. RESULTS

This chapter summarizes the detection performance of EARS (C1, C2, and C3) and CUSUM (aggressive and routine) algorithms against actual H1N1 cases observed in Monterey County from September 30, 2008 to October 29, 2009. The subsections are organized by the ILI counts produced by the various sets of logic (e.g., Base Case, Variant 1a, and Variant 2a) as described in Chapter II.

A. ILI COUNT COMPARISONS

The time periods (e.g., seasonal flu, 1st H1N1 wave, 2nd H1N1 wave) labeled in Figure 18 refer to the *national* ILI outpatient trends in Figure 12 from September 30, 2008 to October 29, 2009, with an overlay of the smoothed MCHD ILI count comparisons of the Base Case, Variant 1a, and Variant 2a logic. Specifically, the national 2009 seasonal influenza outbreak began a steady upward trend beginning in late-January 2009 and peaked in late-February into early-March 2009. The vertical line above the date 3/2/09 in Figure 18 represents the end of the “historical” period,¹³ with the first possible EED signal occurring on March 2, 2009.

¹³ September 30, 2008, to February 27, 2009.

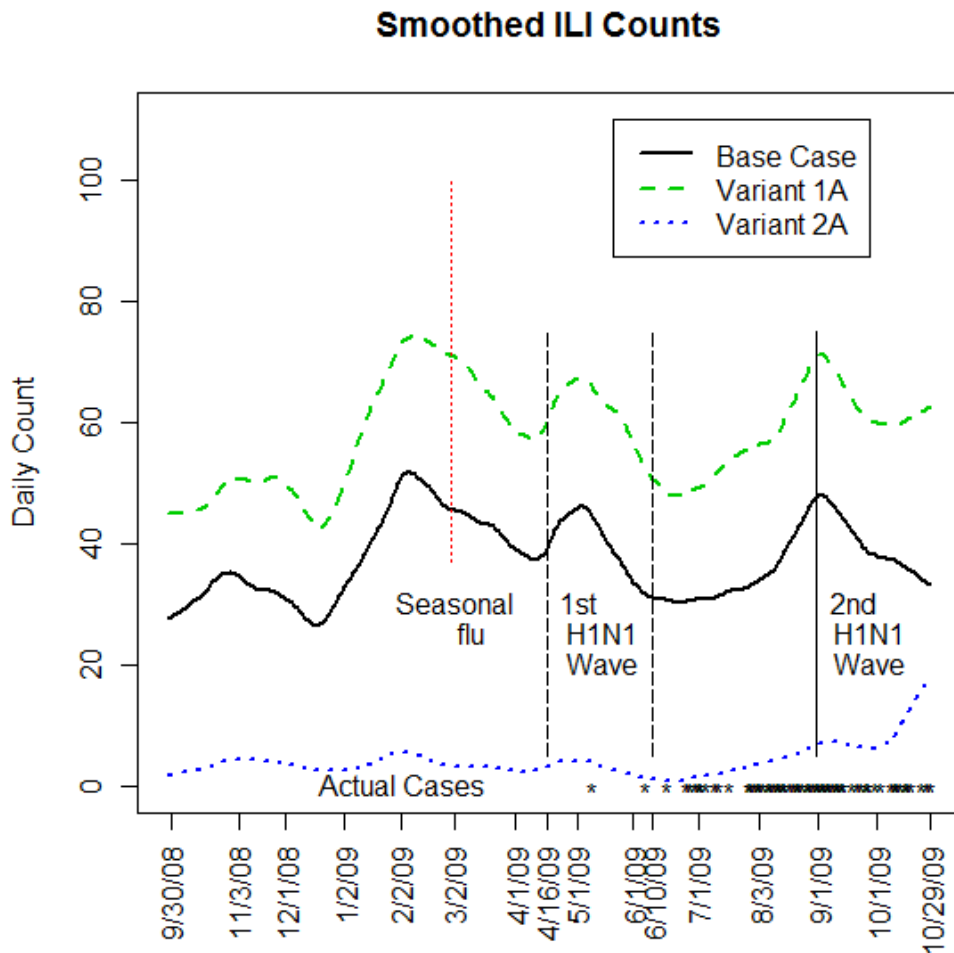


Figure 18. Smoothed ILI count comparisons of Base Case, Variant 1A, and Variant 2A logic from September 30, 2008 to October 29, 2009

The next period of interest is the first H1N1 flu wave (April 16 – June 10, 2009), which is depicted by the middle two vertically dashed lines in Figure 18. After June 10, an upward trend continued throughout the summer and worsened at the time when children across the nation were going back to school (late August–early September). As for the second H1N1 flu wave, the solid, right-most vertical line in Figure 18 represents the beginning (September 1, 2009) of that upward ILI count trend at the national level.¹⁴

¹⁴ For a detailed explanation of 2009 national ILI trends and timeline, refer to discussion in Chapter II Section B, “Chief Complaint Data and Collection Reporting.”

The overall trends for the three sets of logic in Figure 18 are fairly similar, with the exception of the 2nd H1N1 wave period. Notice how the Base Case trend line is downward sloping for that period, while Variants 1a and 2a show a convincing increase towards the end of the dataset (in alignment with the increasing national average). Given the small scale of Figure 18 for Variant 2a, the dramatic increase is even more intriguing when compared to the Base Case, which could possibly imply that the larger counts were masking the true signal (e.g., more ILI cases).

B. ORIGINAL CDC LOGIC (BASE CASE)

Figure 19 compares actual Monterey County H1N1 cases with the results of five EED algorithms under the original CDC “Base Case” logic. Notice that Figure 19 has some additional features from that of Figure 18. Of note, the small “circles” on the graph represent the aggregated daily flu counts for Monterey County on specific days. For example, the “circle” at the first peak above 11/3/08 indicates there were 50 aggregated flu counts for that day. Also, note that the five EED algorithms and corresponding “|” marks indicate that the algorithm thresholds had been exceeded for that particular day (e.g., signaled an alarm).

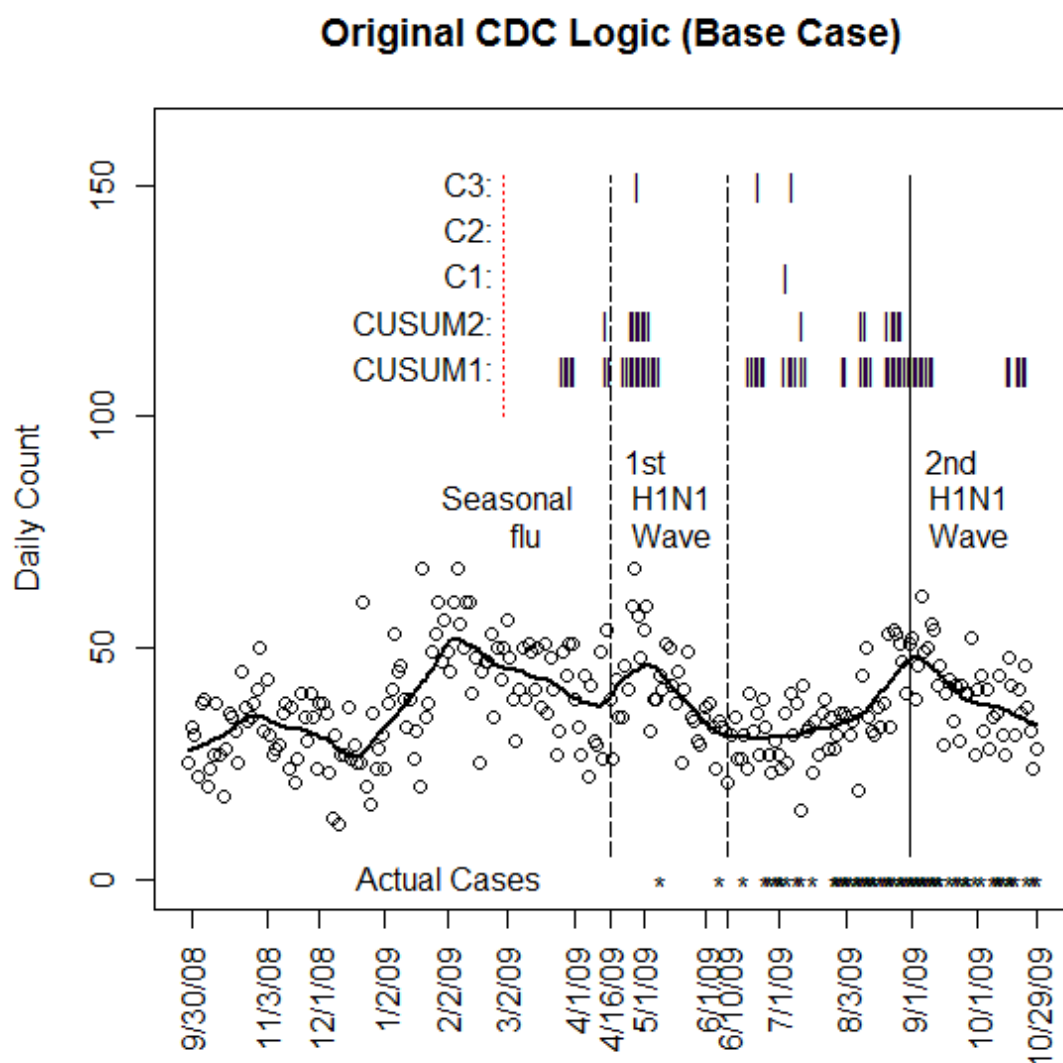


Figure 19. Comparison of actual Monterey County H1N1 cases with the results of five EED algorithms under the original CDC logic (Base Case), from September 30, 2008 to October 29, 2009

Surprisingly, the C1, C2, and C3 methods were of little to no value in signaling an outbreak, unlike CUSUM 1 and CUSUM 2 (aggressive and routine parameters, respectively) which performed much better at signaling alarms prior to the first H1N1 case on May 10.¹⁵ Ultimately, CUSUM 1 proved to be the best EED algorithm because it signaled consistently for 11 straight days up until the first actual case. On the other

¹⁵ Since May 10, 2009 falls on a Sunday, the “case” appears on the graph as Monday, May 11.

hand, CUSUM 1 also signaled between the two H1N1 waves (e.g., summer months) when the smoothed daily counts appear nearly flat, which makes those signals look like false positives. Yet, despite the seemingly flat trend line, the CUSUM 1 alarms do correspond to the high number of actual H1N1 cases in Monterey County during that same timeframe. Table 6 breaks down the number of alarms corresponding to each EED algorithm by time period, under “Base Case” logic.

Algorithm	Time Period			
	Seasonal flu (3/02-4/15)	1st H1N1 wave (4/16-6/10)	Summer (6/11-8/31)	2nd H1N1 wave (9/1-10/29)
C3	0	1	2	0
C2	0	0	0	0
C1	0	0	1	0
CUSUM2	1	7	8	0
CUSUM1	6	13	25	15

Table 6. Number of alarms corresponding to five EED algorithms under “Base Case” logic from March 2 to October 29, 2009

C. EXPANDED MCHD LOGIC (VARIANT 1A)

Figure 20 compares actual Monterey County H1N1 cases with the results of five EED algorithms under the “expanded” MCHD logic (Variant 1a). Notice that during the seasonal flu period, the “aggressive” parameters (ATFS=5, $k=0.56$ and $h=0.296$) of CUSUM 1 continued to signal even during the steady decline of ILI counts later in the season, unlike EARS’ methods which gave no signals. That is, given the ATFS is set at 5 days, CUSUM 1 is going to signal often, whether or not an outbreak actually exists. On the other hand, CUSUM 1 is also the most reliable at signaling when there are increases in the smoothed count line. Given MCHD’s desire for a high sensitivity EED system, frequently occurring false alarm rates were acceptable to decision makers. Table 7 breaks down the number of alarms corresponding to each EED algorithm by time period, under “expanded” MCHD logic.

Expanded MCHD Logic (Variant 1A)

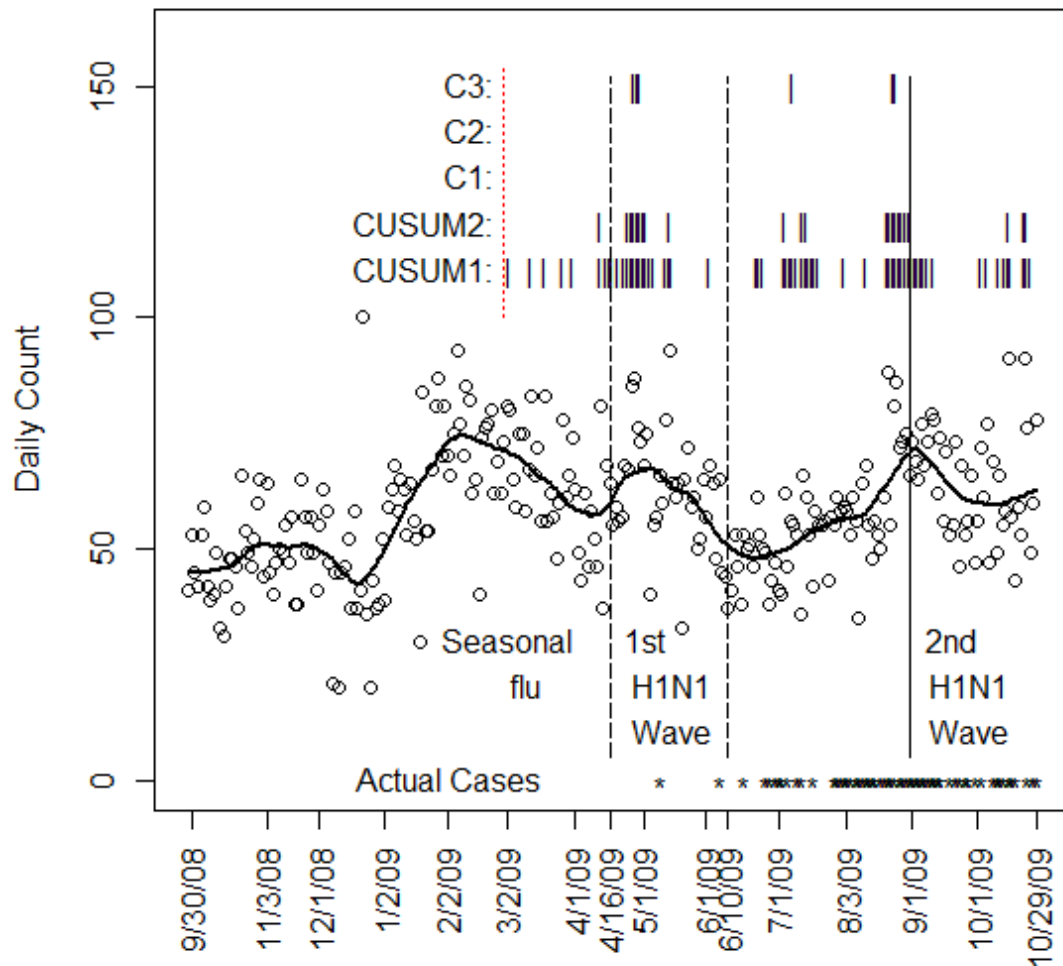


Figure 20. Comparison of actual Monterey County H1N1 cases with the results of five EED algorithms under the “expanded” MCHD logic (Variant 1a), from September 30, 2008 to October 29, 2009

Algorithm	Time Period			
	Seasonal flu (3/02-4/15)	1st H1N1 wave (4/16-6/10)	Summer (6/11-8/31)	2nd H1N1 wave (9/1-10/29)
C3	0	3	3	0
C2	0	0	0	0
C1	0	0	0	0
CUSUM2	1	8	10	3
CUSUM1	7	17	23	17

Table 7. Number of alarms corresponding to five EED algorithms under “expanded” MCHD logic from March 2 to October 29, 2009

While C3 and CUSUM 2 did in fact signal during the first H1N1 flu wave, it was CUSUM 1 that signaled consistently for 10 straight days prior to the first H1N1 case on May 10, 2009. EED methods C1 and C2 failed to signal any alerts for the entire outbreak period, rendering them completely ineffective to decision makers. Lastly, notice the Monterey County ILI trends using the expanded logic do not appear to match the national trends during the second H1N1 wave (e.g., as the national ILI count went up, MCHD counts went down). Perhaps the key takeaway here is that the detection methods (EARS and CUSUM) are working off the syndrome data, so they should signal when that data shows an increase in ILI counts. Then, separately, the syndrome data should show “peaks” around actual cases. Ideally, the syndrome data would show the underlying H1N1 outbreak as evidenced by known cases and the detection methods would signal given that count increase.

D. RESTRICTED MCHD LOGIC (VARIANT 2A)

Figure 21 compares actual Monterey County H1N1 cases with the results of five EED algorithms under the “restricted” MCHD logic (Variant 2a).

Restricted MCHD Logic (Variant 2A)

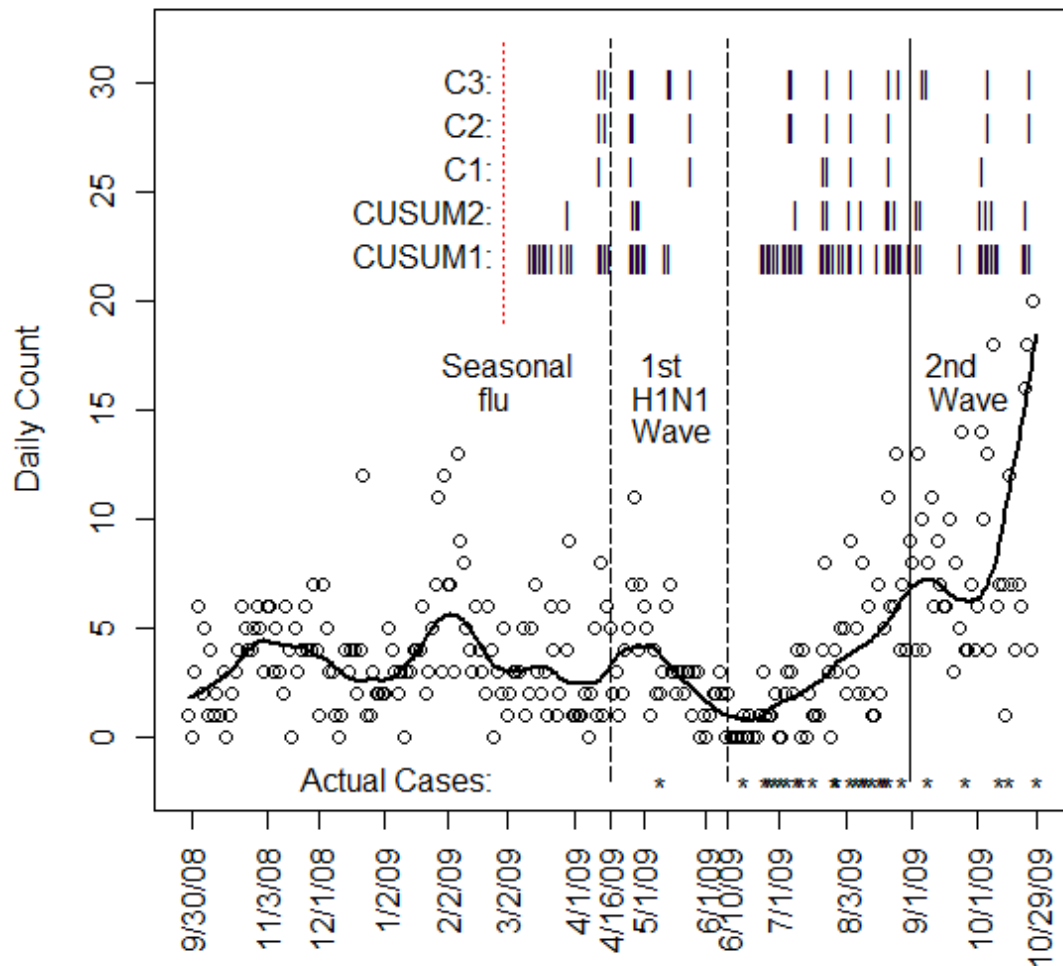


Figure 21. Comparison of actual Monterey County H1N1 cases with the results of five EED algorithms under the “restricted” MCHD logic (Variant 2a), from September 30, 2008, to October 29, 2009

Notice that during the latter part of the seasonal flu period, CUSUM 1 remained highly sensitive to “bumps” in the data yet it also signaled six alarms (from April 28 to May 5) prior to the first Monterey County H1N1 case on May 10th. Alternatively, these signals may have all been false positives. Nonetheless, all EED algorithms signaled at least once within 10 days prior to the first H1N1 case, giving credibility to the remaining algorithms. In fact, CUSUM 2 alarms appear to mimic those of EARS (C1, C2, and C3).

On May 26, all EARS methods signaled prior to the second confirmed H1N1 case on June 8, whereas the closest CUMSUM 1 signal came on May 18 (e.g., far in advance of the actual outbreak). Table 8 breaks down the number of alarms corresponding to each EED algorithm by time period, under “restricted” MCHD logic.

Algorithm	Time Period			
	Seasonal flu (3/02-4/15)	1st H1N1 wave (4/16-6/10)	Summer (6/11-8/31)	2nd H1N1 wave (9/1-10/29)
C3	2	5	6	4
C2	2	3	5	2
C1	1	2	4	1
CUSUM2	1	3	8	6
CUSUM1	13	9	30	14

Table 8. Number of alarms corresponding to five EED algorithms under “restricted” MCHD logic from March 2 to October 29, 2009

Overall, the five EED algorithms performed the best (e.g., signaled with increasing ILI counts) under the “restricted” logic of Variant 2a, as compared to the “expanded” logic of Variant 1a and the “original” logic of the Base Case. While CUSUM 1 is by far the most sensitive, the other methods appear to have signaled at the leading edge of most CUSUM 1 alarm clusters. This raises the interesting issue of tradeoffs, in terms of a continuous signal versus an alarm “reset.” That is to say, do the continuing signals provide additional information about the existence of an outbreak? While the goal of this research is to highlight the implications in choices of logic, this is a question best answered by public health officials.

E. A CLOSER LOOK AT CUSUM 1

With CUSUM 1 as the preferred EED method across all sets of logic, Figure 22 compares CUSUM 1 signals during the first national H1N1 flu wave, April 16–June 10, 2009. Notice that prior to the beginning of the flu season, CUSUM 1 signaled across all sets of logic. Prior to the first H1N1 case, as indicated by the asterisk on May 11, CUSUM 1 signaled consistently for 11 straight days using the Base Case logic (as

indicated by the black tick marks). Under Variant 1a and Variant 2a logic, CUSUM 1 signaled for seven and six consecutive days, respectively, but quit just three days prior to the first actual case.

Now turn to Figure 23 to see how CUSUM 1 performed during the active summer months, starting June 12 and ending September 1, 2009, just prior to the second *national* H1N1 flu wave. Here, the smoothed ILI counts for all three sets of logic appear to be on an upward trend, corresponding to the numerous H1N1 counts during this summer period. It is hard, however, to visually determine which logic best “matches” the H1N1 cases in Monterey County. In other words, given the CUSUM 1 EED methodology, there does not appear to be a clear “winner” for which set of logic should be used. Since identifying the leading edge of an outbreak is usually of most importance to public health officials, look to the period June 25 – July 14 (indicated by the dashed horizontal lines). During this time, there were ten confirmed H1N1 cases reported in Monterey County, of which, the CUSUM 1 algorithm signaled every day except for one using the Variant 2a logic (as indicated by the lowest level tick marks).¹⁶ However, there is evidence that the Base Case data results in more sensitivity, in the sense that for this one comparison, it signaled four days prior to the restricted case (and thus closer to the first actual case).

¹⁶ The asterisk on July 7, 2009 is representative of three H1N1 cases.

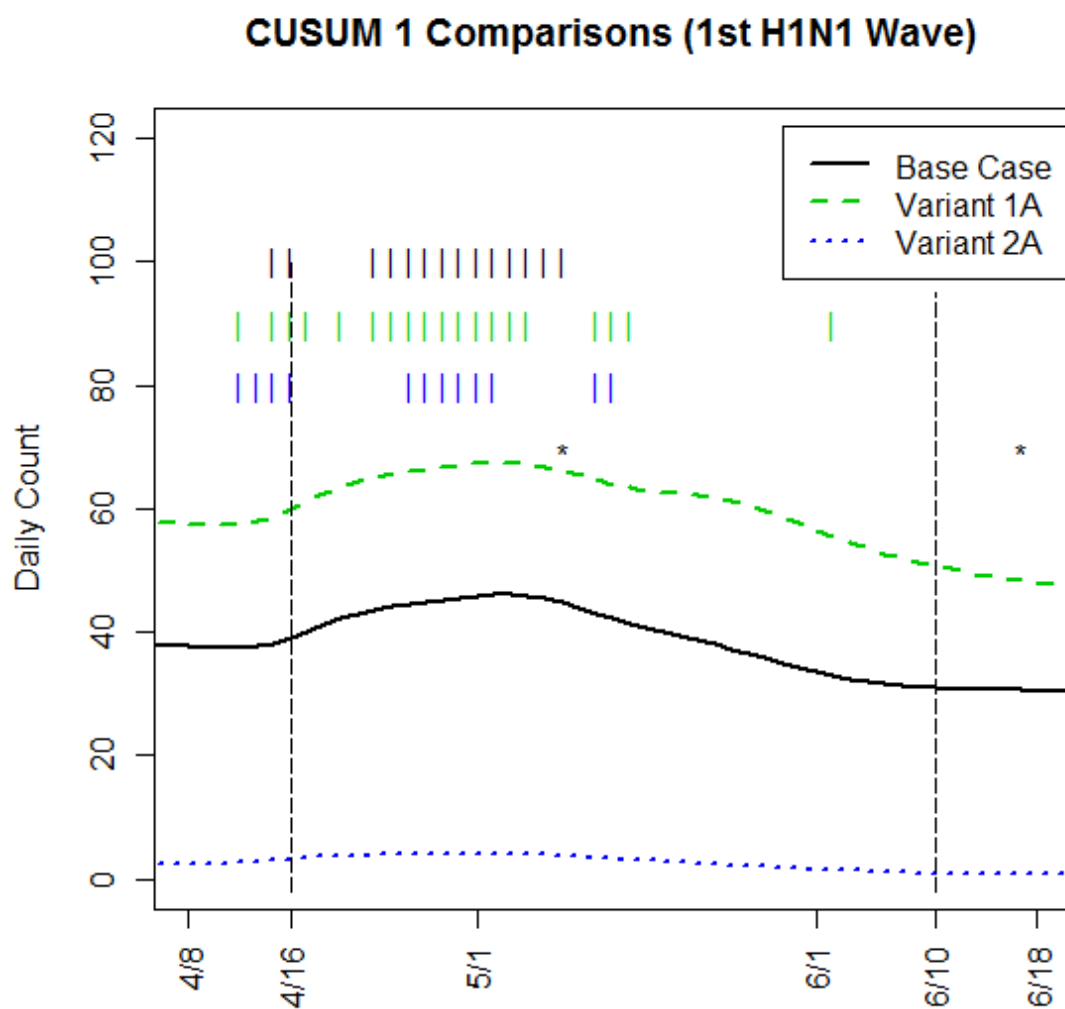


Figure 22. CUSUM 1 comparisons across the various sets of logic during the first H1N1 flu wave, April 16–June 10, 2009

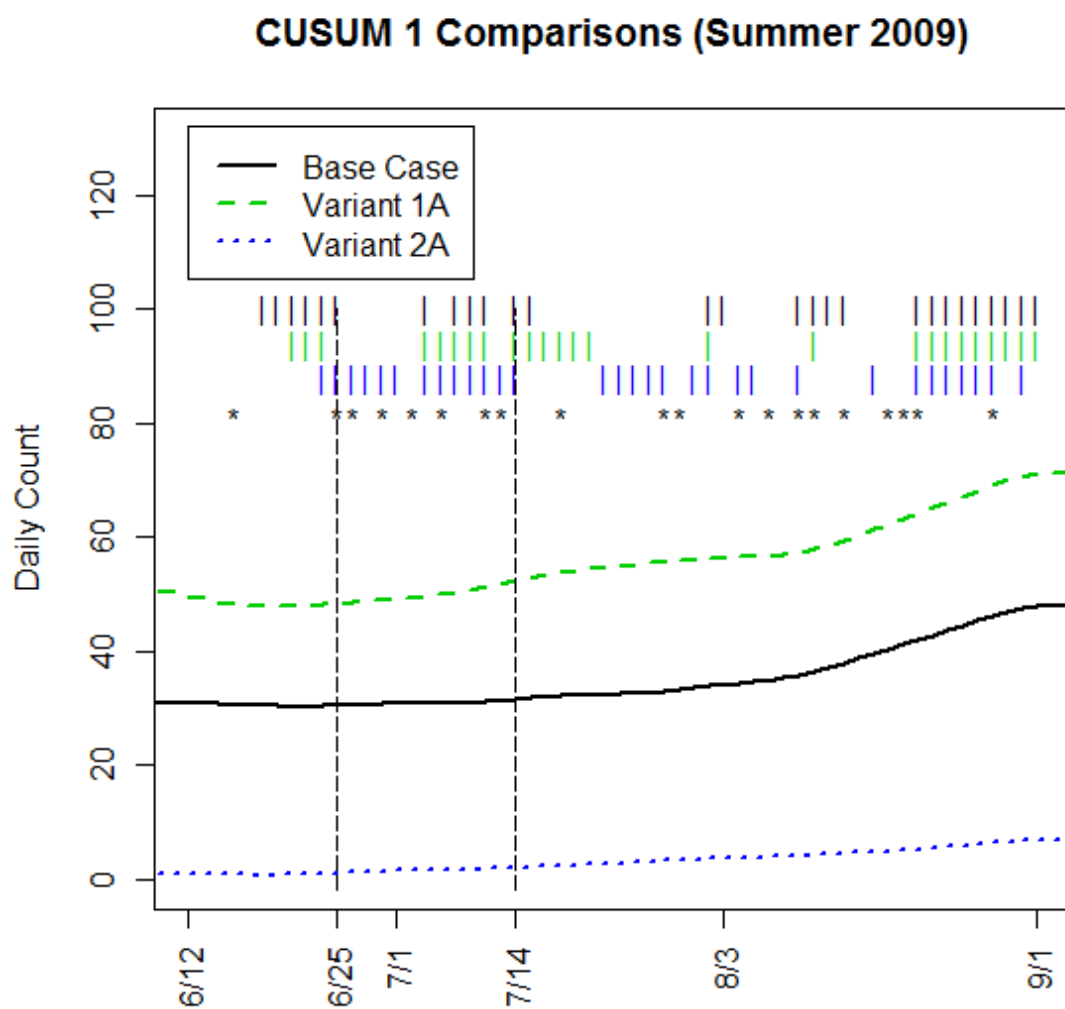


Figure 23. CUSUM 1 comparisons across the various sets of logic during the active summer months, June 12–September 1, 2009

V. CONCLUSIONS

This chapter summarizes the conclusions of this research and makes recommendations for future EARS improvement and follow-on study.

A. EXERCISE CAUTION WHEN CHANGING LOGIC

EARS' users, such as MCHD, have the option to modify three areas of logic in order to alter the performance of the CDC's original EARS system. While the system's flexibility is considered a tremendous benefit to many local health departments, as evidenced in Chapter II, small changes in logic can have large, poorly understood effects on the resulting syndrome counts and, hence, the performance of the EARS system. The results from Chapter II can largely be summarized by Figure 16, which illustrates smoothed ILI counts using five different combinations of text matching logic, symptom aliases, and syndrome definitions.

Under the "Base Case," out of 153,696 total records, 9,093 records (almost 6%) were flagged as being ILI syndromes. The expanded aliases and syndrome definitions (Variant 1a) resulted in a 53% increase in the number of records flagged for the ILI syndrome, while the restricted aliases and syndrome definitions (Variant 2a) resulted in a 92% decrease from the original "Base Case." In order to figure out which combination of methods worked best, an attempt was made to compare results with ICD-9 codes. For the reasons specified in Chapter II and despite having ICD-9 codes in the MCHD dataset, it was determined not to use them as the "gold standard" and instead use documented H1N1 cases, as provided by MCHD.

B. ALGORITHMS NEED GOOD DATA

A detection algorithm is only as good as the data. Greater emphasis, therefore, should be focused on improvements in data collection, management, and syndrome definitions. In other words, biosurveillance systems need quality data and a precise way to measure that quality (e.g., standards for sensitivity and specificity). Currently, there does not appear to be a "gold standard" for measuring the accuracy of EARS until *after*

an outbreak has already occurred, such as the case with this research. It is recommended that the public health community take the lead in demanding better quality symptom data from the various sources available.

Perhaps another area for improvement lies in the secondary use of chief complaint data for the purposes of using text-matching algorithms. Stated differently, chief complaints are serving a purpose for which they were not originally intended. Ideally, one should ask what information would be most useful and then do the best at gathering it rather than just rely on what data is collected by others. If chief complaint data is going to continue to be used (and there isn't much of an alternative), then the text matching logic and alias lists must be improved. Further, they likely need to be tailored to local conditions, conventions, and practices.

In the course of this research, it was also discovered that for Monterey County, EARS' C1, C2, and C3 algorithms factor in "zeros" for days when clinics are not open for business (e.g., weekends and holidays). Given that the sample standard deviation is based on the previous 7–9 days worth of data, it is no wonder that a 3 sigma threshold fails to signal as often as it should for C1 and C2. Figures 19 and 20 illustrate this point. It is strongly recommended that these "zero" data points be discarded before implementing EARS or that the EARS programming logic allow the local user the flexibility to redefine the workweek from seven days to whatever local conditions dictate.

C. RESTRICTED LOGIC IS PREFERRED

Surprisingly, under "original" and "expanded" sets of logic, EARS methods were of little to no value in signaling an outbreak. Ultimately, it was CUSUM 1 that proved the most reliable at signaling alarms prior to and throughout the time when Monterey County was experiencing H1N1 cases. Given that EARS does not utilize the CUSUM algorithms, however, it is clear from the results of Figure 21 that the "restricted" logic of Variant 2a is to be preferred for use by MCHD, at least in comparison to the two other options evaluated. Of note, EARS signaled at the leading edge of most CUSUM 1 alarm clusters under these conditions. Given that CUSUM 1 is by far the most sensitive across all variations in logic, it brings about the issue of tradeoffs, in terms of a continuous

signal versus an alarm “reset.” It is reasonable to question if these continuous signals provide value-adding information about the existence of an outbreak. While the goal of this research is to highlight the implications in choices of logic, this is a question best answered by public health officials.

D. FUTURE RESEARCH OPPORTUNITIES

It would be interesting to observe how EARS would perform given the non-zero data point entries, as discussed in Section B above. One might also want to assess the performance of EARS for thresholds other than those currently fixed in the program where, for alternate thresholds, EARS may be able to signal “appropriately.” Alternatively, one could observe how EARS methods would perform once adjusted for seasonal trends, such as found within ILI data.

Finally, more research should be devoted towards exploring the issue of false positives. As an example, in Chapter IV it was determined that the “aggressive” parameters of CUSUM 1 ($h = 0.296$, $ATFS = 5$, $k = 0.56$) performed reasonably well across all three sets of logic: CDC’s original (Base Case) logic, MCHD’s expanded (Variant 1a) logic, and MCHD’s restricted (Variant 2a) logic. Even after looking closely at the CUSUM 1 signals during the first H1N1 flu wave and during the peak summer months; however, there did not appear to be a clear “winner” for which set of logic should be used. Then again, this research only focused on one defined syndrome, ILI. Future research is certainly recommended to measure the robustness of these results across a variety of other syndromes.

THIS PAGE INTENTIONALLY LEFT BLANK

APPENDIX A. R CODE FOR “ENHANCED NPS LOGIC”

Build a function (“finder”) to look for matches of “str” (string) in “vec” (vector). Note that this is more sophisticated than the SAS coding that MCHD is currently using, which only looks for a simple match anywhere in “vec.” Finder, on the other hand, only allows matches if “str” is at the beginning or end of a word (or matches vec completely) for “str” longer than three characters. It requires even more restrictive matching for “str” of length three or fewer characters.

```
finder <- function (str, vec)
{
  noletters <- "[^A-Z]"
  vec <- paste (" ", vec, " ", sep="")
  lefty <- paste (noletters, str, sep="")
  righty <- paste (str, noletters, sep="")
  shorty <- paste (noletters, str, noletters, sep="")
  if(nchar(str)>3)
    {regexpr (lefty, vec, ignore.case=TRUE) > 0 | regexpr (righty,
vec, ignore.case=TRUE) > 0 | toupper(str) == toupper(vec)}
    else{regexpr (shorty, vec, ignore.case=TRUE) > 0 | toupper(str)
== toupper(vec)}
}
```

THIS PAGE INTENTIONALLY LEFT BLANK

APPENDIX B. MATLAB CODE FOR CALCULATING CUSUM

```
%Initialize variables
k= 1.06; %Placeholder until k can be established
h= .62; %Placeholder until h can be established
baselinePeriod = 35;
startupPeriod = 45; %Placeholder for #days of residuals to use
cusum = 0;
alarmCount = 1;

actualData = dlmread('unbiased.CDC1counts.csv');
matX1 = [ ones(baselinePeriod,1) (baselinePeriod:(-1):1)'];

dayOfAlarm = (1:1:size(alarmCount)); %Initialize vector to track days
in which cusum>=h
today = baselinePeriod;
tomorrow = today+1;
residuals = (1:1:length(actualData));
stdResidualVector = (1:1:length(actualData));

%Calculate residuals

for i=1:1:startupPeriod
    recentData= actualData((today:(-1):(today-baselinePeriod+1)),1);

    matX2= actualData((today:(-1):(today-baselinePeriod+1)),2:5);
    matX3 = [matX1,matX2];
    b= regress(recentData, matX3);
    predCount= ([1 (baselinePeriod+1),actualData((today+1),2:5)])*b;

    residuals(tomorrow)= actualData((tomorrow),1)-predCount;

    today= today+1;
    tomorrow= today+1;

end

while (today<length(actualData)-1)
    if (cusum>=h)
        dayOfAlarm(alarmCount) = today;
        alarmCount = alarmCount+1;

    end
    recentData= actualData((today:(-1):(today-baselinePeriod+1)),1);

    matX2= actualData((today:(-1):(today-baselinePeriod+1)),2:5);
    matX3 = [matX1,matX2];
    b= regress(recentData, matX3);
    predCount= ([1 (baselinePeriod+1),actualData((today+1),2:5)])*b;
```

```
residuals(tomorrow)= actualData((tomorrow),1)-predCount;  
stdDev = std(residuals(tomorrow:(-1):tomorrow-startupPeriod));  
  
stdResidual = residuals(tomorrow)/stdDev;  
stdResidualVector(today) = stdResidual;  
  
cusum=max(0,(stdResidual-k+cusum));  
%cusum  
  
today= today+1;  
tomorrow=today+1;  
  
end
```

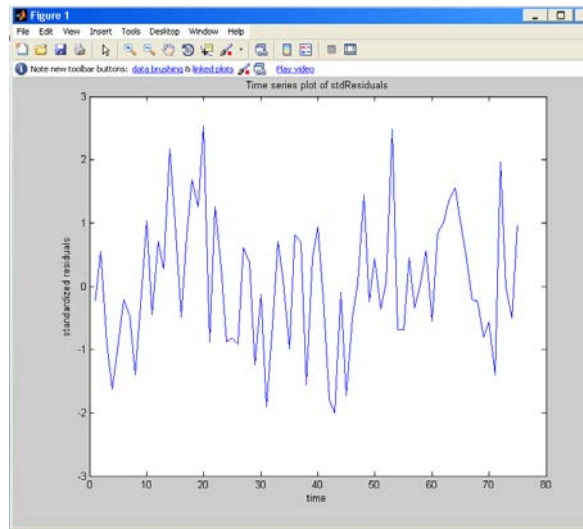
APPENDIX C. STANDARDIZED RESIDUAL PLOTS AND QQ PLOTS OF HISTORICAL DATA

To plot residuals in Matlab:

```
>X=stdResidualVector(45:119) %with baseline 35 days and “startup” 10 days; 119 = 1/3  
of 1.5yrs  
>plot(X)  
>xlabel('time')  
>ylabel('standardized residuals')  
>title('time series plot of stdResiduals')
```

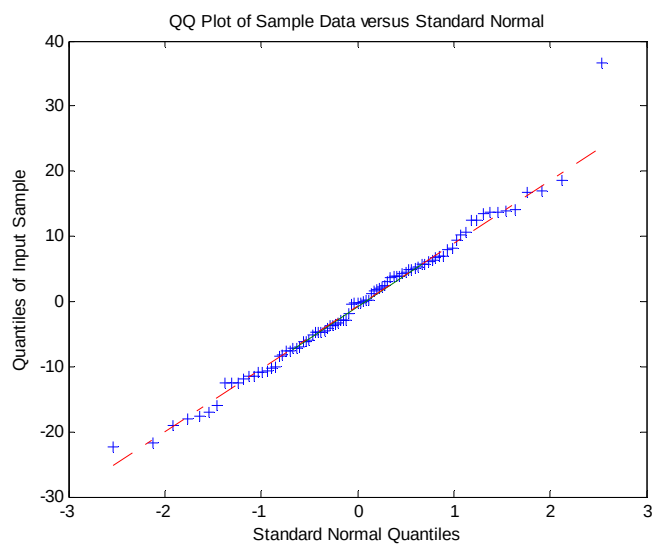
CDC1 “Base Case”

```
>mean(X)  
= .0429
```

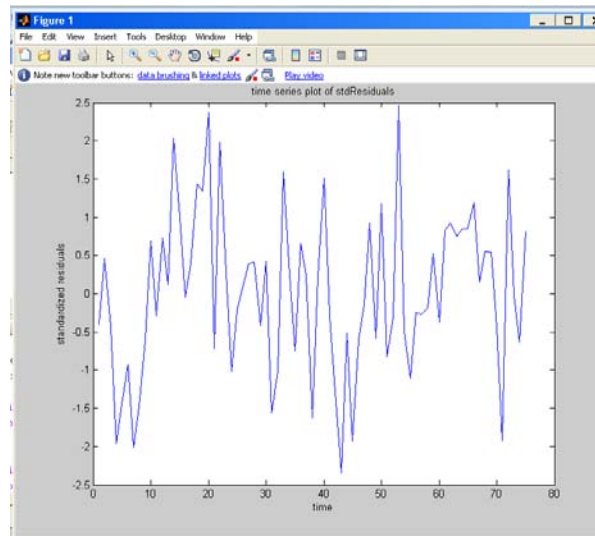


CDC1 “Base Case”

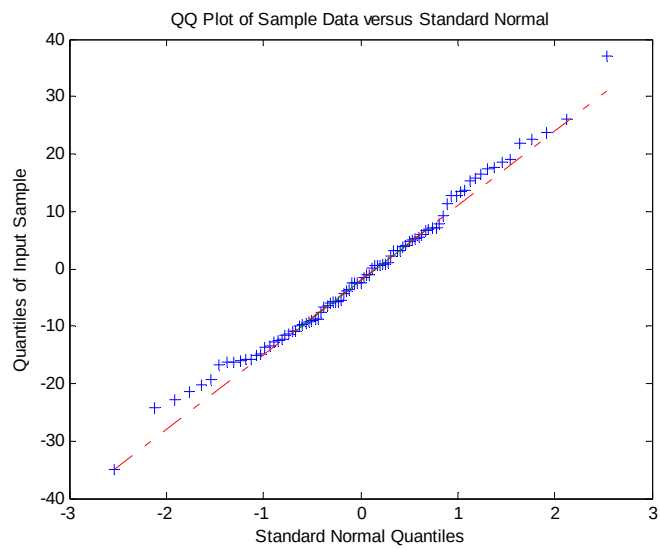
```
> std(X)  
= 1.1036
```



MCHD1 “Variant 1A”
 $> \text{mean}(X)$
 $= 0.0211$



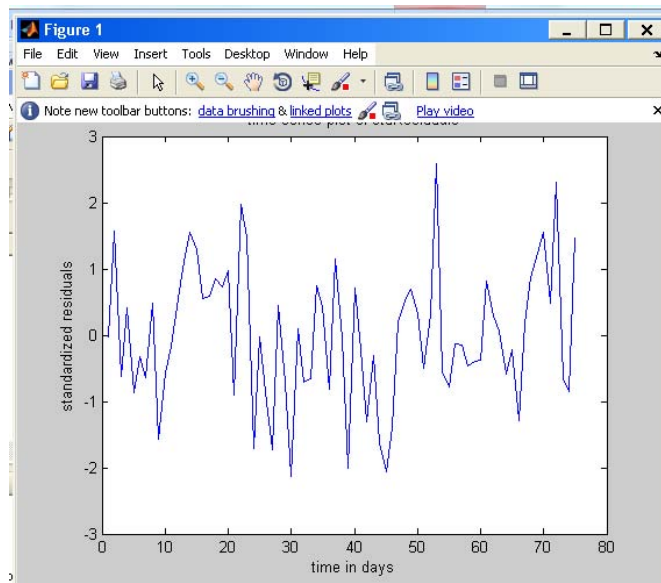
MCHD1 “Variant 1A”
 $> \text{std}(X)$
 $= 1.0073$



MCHD2 “Variant 2A”

>mean(X)

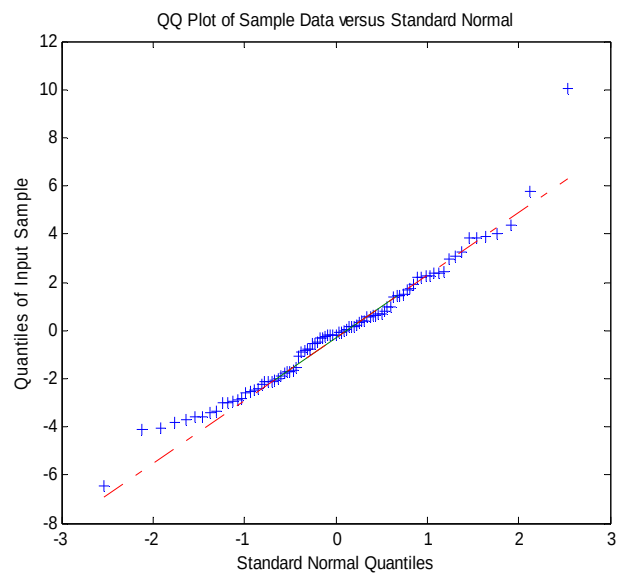
= 0.0068



MCHD2 “Variant 2A”

> std(X)

= 1.0987



THIS PAGE INTENTIONALLY LEFT BLANK

APPENDIX D. R CODE TO ESTIMATE CUSUM THRESHOLD

The output is the estimated ATFS and its standard error.

```
IC.ARL.EST.func <- function(x,h,k){  
  runs <- rep(0,x)  
  CUSUM <- 0  
  for(i in 1:x){  
    while(CUSUM<h){  
      CUSUM <- max(0, (CUSUM + rnorm(1)-k))  
      runs[i] <- runs[i] + 1  
    }  
    CUSUM <- 0  
  }  
  print(c(mean(runs),sd(runs)/sqrt(x)))  
}
```

To get an ATFS=5 (s.e.=0.0045) with k=0.56, set h=0.296. Here's the output:

```
> IC.ARL.EST.func(1000000,0.296,0.56)  
[1] 5.004290000 0.004450843
```

To get an ATFS=20 (s.e.=0.00) with k=1.06, set h=0.62. Here's the output:

```
> IC.ARL.EST.func(1000000,0.622,1.06)  
[1] 20.01101900 0.01944600
```

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF REFERENCES

- Burkom, H., Murphy, S., & Shmueli, G. (2007). Automated time series forecasting for biosurveillance, *Statistics in Medicine*, 26(22), 4202–4218.
- CDC. (2006). Public health information network early event detection and situational awareness functional requirements. Retrieved on May 2, 2010, from www.cdc.gov/phinf/library/documents/pdf/EED_RSv1.0.pdf
- CDC. (2009). Influenza epidemiology and surveillance. Retrieved on August 29, 2010, from <http://www.cdc.gov/vaccines/recs/acip/downloads/mtg-slides-oct09/12-2-flu-vac.pdf>
- CDC. (2010a). Biosurveillance coordination. Retrieved on July 17, 2010, from www.cdc.gov/osels/ph_surveillance/bc.html
- CDC. (2010b). Early aberration reporting system. Retrieved on June 5, 2010, from www.bt.cdc.gov/surveillance/ears
- CDC. (2010c). 2009 H1N1 early outbreak and disease characteristics. Retrieved on July 7, 2010, from www.cdc.gov/h1n1flu/surveillanceqa.htm
- Evans, S., Kleinman, & K., Pagano. (2010). Statistics in defense and national security: bioterrorism and biosurveillance. Retrieved on July 15, 2010, from <http://www.amstat.org/sections/sdns/amstat4.pdf>
- Espino, J., & Wagner, M. (2010). Accuracy of ICD-9-coded chief complaints and diagnoses for the detection of acute respiratory illness. Retrieved on July 14, 2010, from <http://www.ncbi.nlm.nih.gov/pubmed/11833477>
- Franco, C., & Sell, T. (2010). Federal agency biodefense funding, FY2010-FY2011. *Biosecurity and Bioterrorism*. Volume 8, Number 2. Baltimore: Mary Ann Liebert, Inc.
- Fricker, R. D., Jr., Hegler, B. L., & Dunfee, D. A. (2008). Comparing biosurveillance detection methods: EARS' versus a CUSUM-based methodology, *Statistics in Medicine*, 27, 3407–3429.
- Fricker, R. D., Jr. (2010). *Introduction to statistical methods for biosurveillance (Draft)*. Retrieved on May 1, 2010, from <http://faculty.nps.edu/rdfricke/OA4910.htm#book>

- Fricker, R., & Hanni, K. (2010). Biosurveillance: detecting, tracking, and mitigating the effects of natural disease and bioterrorism. Online presentation, Military Operations Research Society, Monterey, CA, February 10, 2010.
- Gottron, F. (2009). Project BioShield: purposes and authorities. *Congressional Research Service*. Retrieved on July 6, 2009, from <http://www.fas.org/sgp/crs/terror/RS21507.pdf>
- Hanfling, D., & Hick, J. (2009). Hospitals and the novel H1N1 outbreak: the mouse that roared. *Disaster Medicine and Public Health Preparedness*. Vol. 3, Supplement 2.
- Hanni, K. (2009). Pandemic H1N1 (2009) global to local surveillance. Keynote speaker, Monterey County health education consortium H1N1 panel, Monterey, CA, September 9, 2009.
- Hawkins, D. M., & Olwell, D. H. (eds). (1998). Cumulative sum charts and charting for quality improvement. New York: Springer.
- Homeland Security. (2007). Homeland Security Presidential Directive 21: public health and medical preparedness. Retrieved on July 26, 2010, from http://www.dhs.gov/xabout/laws/gc_1219263961449.shtm
- Hutwagner, L., Thompson, W., Seeman, G., & Treadwell, T. (2003). The bioterrorism preparedness and response Early Aberration Reporting System (EARS). *Journal of Urban Health: Bulletin of the New York Academy of Medicine*, 80(2(Supplement 1)), 89i–96i.
- IFFGD, 2010. Glossary. Retrieved on August 9, 2010, from <http://www.iffgd.org/site/learning-center/glossary>
- Lawson, B., Fitzhugh, E., Hall, S., Hutwagner, L., & Seeman, G. (2004). From implementation to automation - a step-by-step approach to developing syndromic surveillance systems from a public health perspective. *MMWR. Morbidity and Mortality Weekly Report: Syndromic Surveillance*, 53, 248.
- Lawson, B., & Kleinman, K. (eds). (2005). *Spatial & Syndromic Surveillance for Public Health*. Chichester, UK: John Wiley & Sons.
- Mandl, K. D., Overhage, J. M., Wagner, M. W., Lober, W. B., Sebastiani, P., Mostashari, F., Pavlin, J. A., et al. (2004). Implementing syndromic surveillance: a practical guide informed by the early experience. *The Journal of the American Medical Informatics Association*, 11, 141–150.

- MCHD. (2010a). H1N1 (swine) flu information. Retrieved on July 6, 2010, from <http://www.co.monterey.ca.us/health/healthalerts/pdf/BasicH1N1points.pdf>
- MCHD. (2010b). Monterey County Daily Observational and Situational Evaluation (DOSE) Report. Retrieved on January 27, 2010, from http://www.co.monterey.ca.us/health/healthalerts/pdf/MC_DOSE.pdf
- Montgomery, D. C. (2001). *Introduction to Statistical Quality Control* (4th ed.). New York: Wiley.
- Rolka, H. A. (2006). Data analysis research issues and emerging public health biosurveillance directions.
- Shmueli, G., & Burkom, H. (2010). Statistical challenges facing early outbreak detection in biosurveillance. Retrieved on July 15, 2010, from www.rhsmith.umd.edu/faculty/gshmueli/web/images/statchallengesbiosurveillancerevised-iii.pdf
- Sosin, D.M. (2003). Syndromic surveillance: the case for skillful investment view. *Biosecurity and Bioterrorism: Biodefense Strategy, Practice, and Science*, 1, 247–253.
- Toprani, A., Ratard, R., Straif-Bourgeois, S., Sokol, T., Averhoff, F., Brady, J., Staten, D., Sullivan, M., Brooks, J. T., Rowe, A.K., Johnson, K., Vranken, P., & Sergienko, E. (2006). Surveillance in hurricane evacuation centers - Louisiana. *Morbidity and Mortality Weekly Report*, 55, 32–35.
- WHO. (2010). Global alert and response, pandemic (H1N1) 2009, frequently asked questions: what is phase 6? Retrieved on August 29, 2010, from http://www.who.int/csr/disease/swineflu/frequently_asked_questions/levels_pandemic_alert/en/index.html
- Wilson, A., Wilson, G., & Olwell, D.H. (eds), *Statistical methods in counterterrorism: Game theory, modeling, syndromic surveillance, and biometric authentication*. New York: Springer, 101–107.

THIS PAGE INTENTIONALLY LEFT BLANK

INITIAL DISTRIBUTION LIST

1. Defense Technical Information Center
Ft. Belvoir, Virginia
2. Dudley Knox Library
Naval Postgraduate School
Monterey, California
3. Professor Ronald D. Fricker, Jr.
Operations Research Department
Naval Postgraduate School
Monterey, California
4. Assistant Professor Rachel T. Johnson
Operations Research Department
Naval Postgraduate School
Monterey, California
5. Krista Hanni, PhD
Monterey County Health Department
Salinas, California
6. Susie Barnes
Monterey County Health Department
Salinas, California