DISSERTATION

ONE HEALTH IN THE U.S. MILITARY:

A REVIEW OF EXISTING SYSTEMS

AND RECOMMENDATIONS FOR THE FUTURE

Submitted by:

Rebecca I. Evans

Department of Clinical Sciences

In partial fulfillment of the requirements

For the Degree of Doctor of Philosophy

Colorado State University

Fort Collins, Colorado

Summer 2014

Doctoral Committee:

Advisor: Mo Salman

Michael Lappin
Stephen Hayne
Jennifer Peel

Copyright by Rebecca I. Evans 2014

All Rights Reserved

ABSTRACT

ONE HEALTH IN THE U.S. MILITARY:

A REVIEW OF EXISTING SYSTEMS

AND RECOMMENDATIONS FOR THE FUTURE

Background:

The merging of the former U.S. Army Veterinary Command (VETCOM) with the former U.S. Army Center for Health Promotion and Preventive Medicine (USACHPPM) into the U.S. Army Public Health Command (USAPHC) in 2011 created an opportunity for the military to fully embrace the One Health concept. That same year, the USAPHC began work on a Zoonotic Disease Report (ZDR) aimed at supporting critical zoonotic disease risk assessments by combining zoonotic disease data from human, entomological, laboratory, and animal data sources. The purpose of this dissertation is to facilitate the creation of a military Zoonotic Disease Surveillance program that combines disease data from both military human and animal sources.

Methods:

Five of the most commonly used human military medical data systems were systematically reviewed using a standardized template based on Centers for Disease Control and Preventive Medicine (CDC) guidelines. The systems were then compared to each other in order to recommend the one(s) best suited for use in the USAPHC ZDR. The first stage of the comparison focused on each system's ability to meet the specific goals and objectives of the ZDR, whereas the second stage applied capture-recapture methodology to data system queries in order to evaluate each system's data quality (completeness).

A pilot study was conducted using Lyme borreliosis to investigate the utility of military pet dogs as sentinel surveillance for zoonotic disease in military populations. Canine data came from 3996 surveys collected from 15 military veterinary facilities from 1 November 2012 through 31 October 2013. Surveys simultaneously collected *Borrelia burgdorferi* (Bb) seroprevalence and canine risk factor data for each participating pet dog. Human data were obtained by querying the Defense Medical Surveillance System for the same 15 military locations and the

same time period. The correlation of military pet dog Bb seroprevalence and military human Lyme disease (borreliosis) data was estimated using the Spearman Rank Correlation. The difference between military pet dog data and civilian pet dog data was examined through the use of the chi-squared test for proportions. Multivariable logistic regression was then used to investigate the potential for identified risk factors to impact the observed association.

Results:

The comparison of human military medical data systems found the Military Health System Management Analysis and Reporting Tool (M2) data system most completely met the specific goals and objects of the ZDR. In addition, completeness calculation showed the M2 data source to be the most complete source of human data; 55% of total captured cases coming from the M2 system alone.

The pilot study found a strong positive correlation between military human borreliosis data and military pet dog Bb seroprevalence data by location ($r_s = 0.821$). The study showed reassuring similarities in pet dog seroprevalence by location for the majority of sites, but also showed meaningful differences between two locations, potentially indicating military pet dogs as more appropriate indicators of Lyme disease risk for military populations than civilian pet dog data. Unfortunately, whether canine Bb seroprevalence is influenced by the distribution of identified risk factors could not be determined due to limited study power.

Conclusions:

Based on this study M2 was recommended as the primary source of military human medical data for use in the Public Health Command Zoonotic Disease Report. In addition, it was recommended that Service member pet dog data be incorporated as a sensitive and convenient measure of zoonotic disease risk in human military populations. The validity of the data, however, should be evaluated further with either larger sample sizes and/or a zoonotic disease with higher prevalence.

TABLE OF CONTENTS

ABSTRACT ii
Chapter 1: Introduction 1
References 4
Chapter 2: Literature Review 5
Public Health Surveillance 5
Definitions 5
Types of Surveillance Strategies 6
Evaluating Surveillance Systems 15
Surveillance System Attributes 15
Applications of System Attributes in Surveillance Programs 16
Comparing Surveillance System Data 22
Potential Differences in Data 22
Approaches to Comparing Data 23
The Use of Animals as Sentinels in Zoonotic Disease Surveillance 27
One Health 27
Animals as Sentinels 28
Research Demonstrating Animals as Sentinels 29
References 38
Chapter 3: Review of Data Systems used in Public Health Surveillance 44
Introduction 44
Public Health Surveillance 44
Public Health Surveillance in the Military 45
Describing Data Systems Used in Public Health Surveillance 45
Background 46
A Systematic Approach to Describing Data Systems Used in Public Health Surveillance 46
Section I: US and Global Data Systems Used in Public Health Surveillance 49
National and Global Public Health Surveillance 49
Methods 50
GHO: WHO's Data and Statistics Portal 50
EMPRES: FAO's Emergency Prevention System 54
WAHIS: OIE's Web Interface for Animal Health Data 58
NAHRS: USDA APHIS's Animal Health Reporting System 63
SCS: USDA APHIS's Surveillance Collaboration System 67

Section II: Military Data Systems Used in Public Health Surveillance 74
Public Health Surveillance in the Military 74
Methods 75
MDR: Military Health System Data Repository 75
Defense Medical Surveillance System (DMSS) 80
HL7 86
ESSENCE 91
DRSi 95
References 101
Appendices 104
Chapter 4: Comparison of the Human Medical Data Systems Most Commonly Used for Public Health Surveillance in the U.S. Army 106
Introduction 106
Background 107
Evaluating Military Human Medical Data Systems 108
A Brief Description of the Human Medical Data Systems Most Commonly Used by the U.S. Army 108
Section I: A Comparison of Military Medical Data System Attributes 110
Zoonotic Disease Report Criteria 111
Zoonotic Disease Report Query Requirements 111
System Performance 111
Methods 112
Results 115
Discussion 118
Conclusions 120
Section II: An Evaluation of Data System Quality 120
Capture-Recapture Methodology 120
Log-linear Modeling 124
Methods 130
Results 133
Discussion 137
Conclusions 141
Overall conclusion 142
Recommendations 143
References 144
Appendices 146
Chapter 5: A Model Using Service Member Pet Dogs as Sentinels in Zoonotic Disease Surveillance 157

Introduction 157
American Veterinary Medical Association One Health Imperative 157
One Health in the Military 157
Background 159
Animals as Sentinels 159
Military Pets as Sentinels 159
Lyme disease and Military Relevance 160
Dogs as Sentinels for human Lyme disease 161
Structure of this Chapter 162
Section I: Determining the Association between Military Pet Dog and Human Borreliosis Data 164
Studies Using Canine Bb Seroprevalence data in Sentinel Surveillance for Human Borreliosis in Civilian Populations 164
The Utility of Military Pet Dog Seroprevalence data in Sentinel Surveillance for Human Borreliosis in Military Populations 165
Methods 166
Results 171
Discussion 176
Conclusions 178
Section II: A Comparison of Borrelia burgdorferi Seroprevalence in Military and CivilianPet Dogs 179
Civilian Pet Dog Borrelia burgdorferi Seroprevalence 179
Military Pet Dog Borrelia burgdorferi Seroprevalence 179
Methods 180
Results 181
Discussion 182
Conclusions 184
Section III: Evaluating the Military Pet Dog Borellia burgdorferi Seroprevalence and Location Association: Investigation of Confounders 184
Risk Factors for Borrelia burgdorferi 184
Methods 186
Results 190
Discussion 198
Conclusions 201
Overall Conclusions 201
Recommendations 203
References 204
Appendices 207

Chapter 6: Recommendations and Future Directions 249
Research Recommendations 249
Military Human Disease Data System Review 249
Military Pet Dogs as Sentinels in Lyme Disease Surveillance 250
Future Directions 251
Incorporating ROVR in the ZDR 252
A Zoonotic Disease Surveillance System Combining Human and Animal Disease Data 253
References 255

Chapter 1: Introduction

The One Health concept recognizes that human, animal, and ecosystem health are inextricably linked. It seeks to promote and improve the health and well-being of all species by enhancing cooperation and collaboration between physicians, veterinarians, other scientific health and environmental professionals (1). In 2009, the U.S. Army Surgeon General directed the creation of a new medical command that truly embraces this One Health concept (2). Effective July 2011, the U.S. Army Public Health Command (USAPHC) was established. The new command was created by merging the former U.S. Army Veterinary Command (VETCOM) and the former U.S. Army Center for Health Promotion and Preventive Medicine (USACHPPM) (2). The merger formally joined the technical skills of veterinarians, physicians, entomologists, laboratory specialists, epidemiologists, and many others. The mission statement for this new command is: "Promote health and prevent disease, injury, and disability of Soldiers and military retirees, their Families, and Department of the Army civilian employees; and assure effective execution of full spectrum veterinary service for Army and Department of Defense Veterinary missions" (http://phc.amedd.army.mil/).

Zoonoses account for 63% of all infectious reportable medical events within the Department of Defense (3). For this reason, knowledge of zoonotic pathogen presence creates opportunities for improved preventive medicine strategies. In 2011, the Epidemiology and Disease Surveillance Portfolio, Disease Epidemiology Program within the USAPHC initiated the development of a Zoonotic Disease Report (ZDR). The aim of the report is to provide U.S. Army public health personnel with critical health information regarding the presence and spread of zoonotic pathogens by combining zoonotic disease information from extant military human disease data systems, tick and mosquito borne disease as well as rabies specimen testing data from the Laboratory Services Portfolio, with animal data from intergovernmental organizations public access databases. Ideally the animal data will include zoonotic diseases diagnosed at U.S. Army veterinary facilities once the web-based electronic medical record known as the Remote Online Veterinary Record (ROVR) is fully deployed.

The purpose of this dissertation is to advocate for the formalized use of animal disease data from military veterinary facilities in sentinel surveillance for zoonotic diseases in military personnel. The specific aim is to demonstrate the

utility of a Zoonotic Disease Surveillance program that combines zoonotic disease data from both human and animal sources into a comprehensive ZDR. One step towards meeting this aim is to conduct a systematic comparison of the five human medical data systems most commonly used for public health surveillance in the U.S. Army in order to recommend the one(s) most suited for use in the USAPHC ZDR. Another step is to demonstrate the use of military pet dogs as sentinel surveillance for zoonotic diseases in military populations. These two steps are hereafter referred to Project 1 and Project 2, respectively. The overall approach is outlined in Figure 1.

Specifically, the chapters of this dissertation are as follows:

Chapter 2 is an extensive review and critical assessment of previously published studies and technical reports that pertain to the topics discussed in this dissertation. Broadly, such topics fall into the three main categories: a description of public health surveillance strategies, the evaluation of public health surveillance systems, and the use of animals as sentinels in zoonotic disease surveillance.

Chapter 3 is a systematic review of extant data systems used in public health surveillance. The chapter discusses the need for a standardized approach in describing public health data systems. Next it briefly reviews selected national and global public health surveillance systems, all of which have a zoonotic disease capability. The focus of the chapter, however, is an extensive description of the five most commonly used human military medical data systems.

Chapter 4 is an in-depth comparison of the five most commonly used human military medical data systems. The chapter discusses the goals of the ZDR and the specific medical data system criteria needed to meet these goals. Because the comparison is conducted in two separate stages, the chapter is split into two distinct sections. Section I focuses on a systematic comparison of the military medical data system attributes, whereas Section II emphasizes the evaluation of data system quality. The chapter concludes with an overall recommendation for the best data systems for use in the ZDR.

Chapter 5 provides the details of a pilot study using Lyme borreliosis to investigate the utility of military pet dogs as sentinel surveillance for zoonotic disease in military populations. Specifically, canine *B. burgdorferi* seroprevalence

data collected from participating military veterinary facilities from 1 November 2012 through 31 October 2013 are compared to human Lyme disease data from extant military medical data systems for the same time period. These data are used to determine the correlation between *B. burgdorferi* seroprevalence in military pet dogs and Lyme disease cases in military personnel. The application of the study as a model for the ZDR is then discussed.

Chapter 6 summarizes the findings of this dissertation, recommending the human medical data system that is best suited for use in a zoonotic disease report and advocating the formalized incorporation of animal disease data from military veterinary facilities. The chapter discusses considerations for implementing these changes and concludes with recommendations for the future of zoonotic disease surveillance in the military.

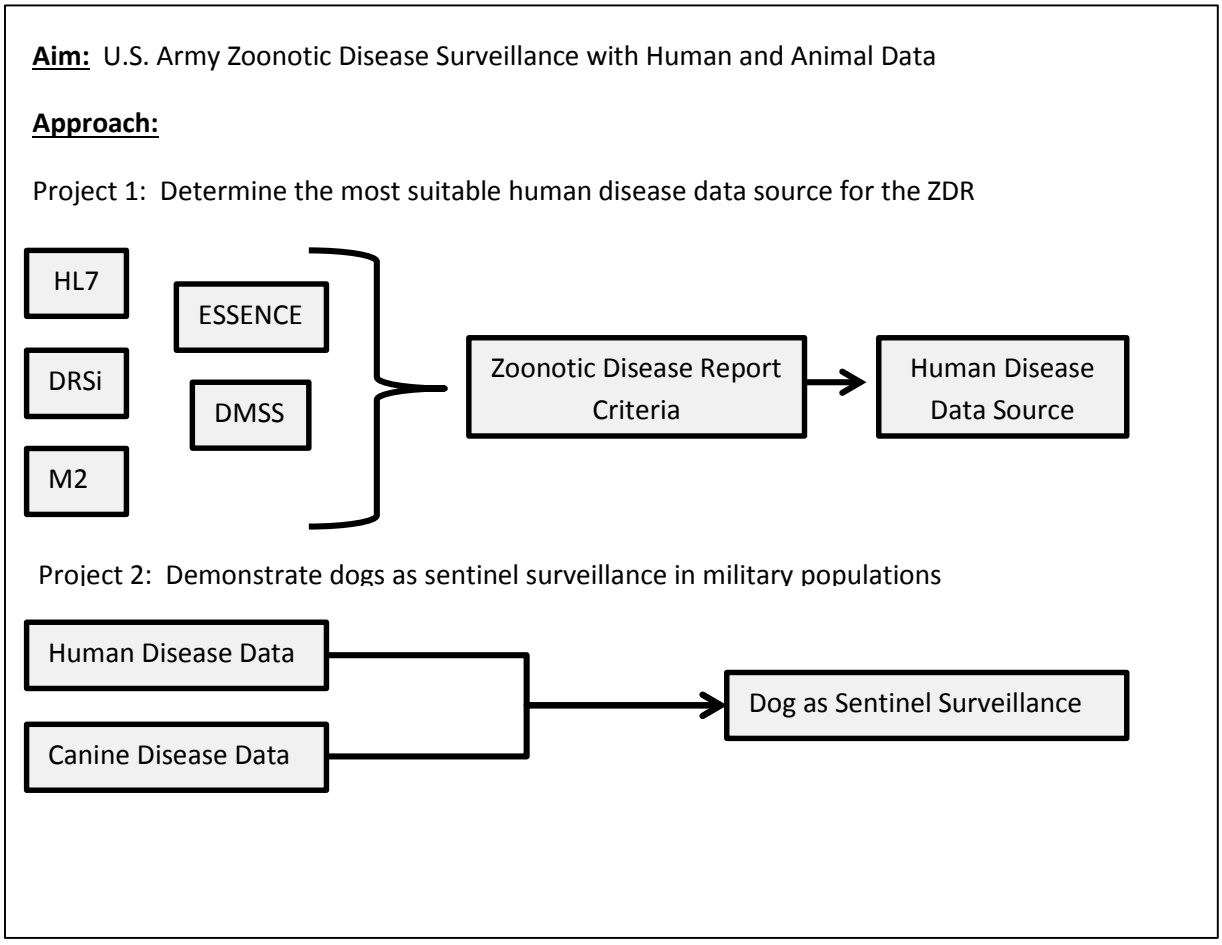

Figure 1.1: Overall Approach

References

1. One Health Initiative website. Available at www.onehealthinitiative.com. Accessed April 10, 2014.

2. Permanent Order 275-002, 2 October 2009, Headquarters U.S. Army Medical Command Fort Sam Houston Texas, TX 782345-6000.

3. Armed Forces Reportable Medical Events Guidelines and Case Definitions, Armed Forces Health Surveillance Center (AFHSC), March 2012.

Chapter 2: Literature Review

Public Health Surveillance

Definitions

In public health, the term **monitoring** refers to the ongoing efforts directed at assessing the health status of a given population (1). The sampling of individuals may be continuous or repeated. The health status monitored may be a specific infectious disease or a general condition/symptom/outcome. The population can be defined on many different levels, both temporally (birth cohort) and spatially (national, regional, family group/herd) as well as categorically (occupation, race, species, breed). The term **survey** refers to the systematic collection of information in support of an investigation (1). The sampling is aimed at answering a particular question or hypothesis, therefore the frequency and duration of sampling, along with the specific health status and population are clearly stated. In comparison to the previous two terms, the term **surveillance** refers to something more active, a system that not only collects health status data, but also involves a specified action when a set threshold is exceeded (1)

The World Health Organization defines **public health surveillance** as the ongoing, systematic collection, analysis, interpretation, and dissemination of population based data regarding a health-related event for use in public health action to reduce morbidity and mortality and to improve health (2). Such data can be used to measure the burden of disease, identify populations at risk, and monitor trends in disease. It can also be used to initiate immediate action, prioritize resource allocation, justify policy changes, and implement new public health programs (2, 3, 4). Animal **health surveillance** is defined by the United States Department of Agriculture as the ongoing systematic collection, collation, analysis, and interpretation of data and dissemination of information to those who need to know so that action can be taken (5). The purposes of such surveillance is rapid detection of introduced diseases and emerging issues, monitoring and providing actionable information for endemic diseases, and measuring regional prevalence of trade-significant diseases (5).

As demonstrated by these two definitions, surveillance has many different utilities. In addition to the actions mentioned above, surveillance can be used in disease control and disease eradication programs (6). Disease control programs (DCPs) combine monitoring and surveillance activities with disease control and intervention strategies over a prolonged period of time with the aim of reducing the frequency of a specific disease. Disease eradication programs (DEPs) are a special case of DCP where the objective is to eliminate the specific disease (6). The success of each of these programs relies on strict definition of three specific **surveillance components**: the disease **monitoring system** being used, the disease **threshold** that will trigger action, and the **action** that will be taken (6).

Types of Surveillance Strategies

Surveillance may be conducted in many different ways depending on the specific goals, approaches, and methodologies used; herein referred to cumulatively as **surveillance strategies** (7). This section will briefly review several (but not all) of the currently employed surveillance strategies; providing examples of each as available. Although the descriptions of each strategy may appear clear-cut, the application of the strategies is not; many of examples discussed below employ more than one surveillance strategy resulting in significant overlap.

Public Health Surveillance

Public health surveillance generally refers to a broad scale surveillance strategy that relies on a variety of different sources to draw conclusions on disease status for a population. There are two main approaches: 1) health officials submit case reports for specified diseases/conditions to designated agencies, or 2) researchers collate data from multiple sources in order to make generalizations about a disease status/disease burden.

The agency-driven process usually requires health officials to report cases of "notifiable" disease (as defined by receiving agencies) to local health departments; health department officials verify the disease reports, monitor incidence, identify possible outbreaks, and forward their findings to the next level (receiving agency: state, region, federal agency, country representative). At this level the information may be collated and independently analyzed, or forwarded further to organizations that monitor health at a global level. Examples of agencies that employ this

method of surveillance include the World Organization for Animal Health (OIE), the World Health Organization (WHO), the Centers for Disease Control and Prevention (CDC), the United States Department of Agriculture (USDA), and the Department of Defense (DoD) (8, 9, 10, 11). An example of a publication that utilizes this sort of data is the WHO's annual The World Health Statistics (12). The publication compiles statistics for key health indicators monitored by the WHO into an analytical report aimed at addressing cross-cutting topics such as women and health (in the 2013 issue) and a summary of the progress made towards achieving the health-related Millennium Development Goals (MDGs) and associated targets. The statistics in the publication are based on data reported by its 194 Member States. In order to ensure the best use of this country reported data and to maximize the comparability of the statistics across countries and over time, adjustments are made to deal with missing values and to correct for known biases (12). The Defense Reporting System-internet (DRSi) is one system used by the DoD for the purpose of public health surveillance. The system is designed to capture reportable medical events data as defined by the Armed Forces Reportable Medical Events Guidelines and Case Definition (13, 14, 15). The U.S. Army Public Health Command uses the data to produce daily and monthly reportable event reports for Army populations world-wide. The value of the agency-approach is being able to provide a snap-shot of disease status/burden at a population based level. The disadvantage is the lack of population details which limits the generalizability of the findings thereby inhibiting the ability to make public health policy based on the data.

The researcher-driven approach generally involves meta-analysis of existing literature or data. In 2004, Crump et al conducted a population-based study of the global burden of typhoid fever by searching multilingual scientific literature for studies confirming diagnoses through blood cultures between 1966 and 2001 (16). A total of 22 eligible studies were identified, representing 21 regions and 13 countries. Of these 21 regions, only 6 (29%) contained countries with national typhoid fever surveillance systems that routinely included confirmation by blood culture (16). For regions where there was no data, country level extrapolation was done based on geographical proximity and socioeconomic conditions followed by extrapolations based on age-incidence curves (16). In 2010, Majowicz et al. employed a similar approach to estimate the global burden of nontyphoidal Salmonella gastroenteritis (17). Synthesizing existing data from the literature, special studies, and laboratory-based surveillance from 21 regions, they extrapolated disease burden estimates to countries and regions where data were lacking (17). These studies, although useful in providing rough estimates of disease burden globally, highlight many limitations to

this approach of public health surveillance. The meta-analysis researcher has no control over the accuracy or validity of the data provided by each data source. This heavy reliance on a small number of observations may lead to significant over or under-estimates of disease burden. In addition, the researchers are limited to the details provided by the data sources, therefore in depth analysis of risk factors or other contributing variables could not be conducted.

Syndromic Surveillance

The term syndromic surveillance describes a surveillance strategy aimed at providing early warning of a disease occurrence before epidemic or outbreak status is attained (7). The strategy uses the health-related data that precede the actual diagnosis to determine if a sufficient probability of a case or outbreak exists to warrant further public health response. It is the 'real-time' monitoring of non-specific, pre-diagnostic indicators for disease outbreaks. Examples of operational syndromic surveillance systems include the Real Time Outbreak and Disease Surveillance System (RODS) used by several states to gather data on the symptoms of emergency room patients, the National Retail Data Motor (NRDM) created by the University of Pittsburgh laboratory to examine sales of over-the-counter health-care products, the federal BioSense program designed to aggregate data relevant to bioterrorism and other public health threats within the United States, and the Electronic Surveillance System (ESSENCE) operated by the DoD to allow epidemiologists to track syndromes from military hospital feeds in real-time (7).

Although the notion of syndromic surveillance as an early warning system makes sense in concept, the application of this strategy still remains largely unproven (7). In 2006, a French study showed that influenza data collected from real time syndromic surveillance at emergency departments correlated significantly with data collected from a network of sentinel general practitioners (18). Although the results showed that the data complement each other, it could not prove that syndromic surveillance was any better suited for use in early detection of the disease (18). Analysis of the United Kingdom's national syndromic surveillance system, intended to serve in early detection of infectious disease and bioterrorist attacks has indicated many sudden rises in the 11 monitored syndromes, but has never predicted an actual biological or chemical attack (19). Instead, the data accurately tracked rises of community morbidity of already identified risks such as influenza-like illnesses during flu season and heat-related disease

during heat waves (19). In 2014, local syndromic outbreak signals from telephone triage centers and over-the-counter antidiarrheal sales were compared to nine known waterborne and foodborne outbreaks that occurred in Sweden between 2001 and 2007 (20). Although syndromic outbreak signals were identified for the four largest outbreaks, the findings are difficult to generalize because the signals only represented large outbreaks with more than 1000 cases (20).

The inability of syndromic surveillance to detect rare, small-scale events such as localized biological attacks or the initial cases of a newly imported or emerging disease has been discussed in global forums (7). The studies discussed here, conclude that the syndromic systems were helpful because they were able either to accurately reproduce data generated by existing specific systems or to document excess mortality following an already identified risk. However, none demonstrated a real added capacity to detect events that preceded regular surveillance activities or events that would otherwise have been missed. When using syndromic surveillance, it must be in an atmosphere that doesn't penalize people for getting it wrong, an atmosphere where false positives are tolerated in favor of increased sensitivity such as broad situational awareness. Unfortunately, it is impossible to increase the singular attribute of sensitivity, specificity, or timeliness of syndromic detection without decreasing the other two (7).

Real-time and Batched Reporting

The terms 'real time' and 'batched' for disease surveillance refers data collection strategies, but can be used to mean different things by different authors. In general 'batched' data refers to data that has in some way been assembled and or processed; batched collections of health indicator data, batched processing of indicator data, batched reporting of data that triggers some set rule or alert, batched reporting of data at predetermined time intervals (7). Inherent in these descriptions is a time delay between data collection and batched reporting. 'Real-time', on the other hand, implies the data collected is simultaneously available for use in surveillance activities. Many hospitals have automated information systems based on the Health Level Seven (HL7) format. These systems provide a comprehensive framework for the exchange, integration, sharing and retrieval of electronic health information such as clinical observations, laboratory or pharmacy orders, patient admissions and transfer data, and billing information. The monitoring of HL7 data streams is often the closest to real time that hospitals can get. Many

surveillance systems claim 'real time' reporting, but few are actually suited to do so; most data collection systems require some form of processing. In truth, surveillance reports delivered in real time, in the form of automated alerts, tend to have considerable problems as processing is needed to reduce entry and transmission errors (7). The Real-time Outbreak and Disease Surveillance (RODS) system is a computer-based public health surveillance system for early detection of disease outbreaks that uses the HL7 message protocol. Although the name claims the system to function in real time, the extensive data processing steps involved in transforming the data into utilizable formats nullifies this statement (21). The RODS system was used during the 2002 Winter Olympics and is advocated as a resource for implementing, evaluating, and applying new methods of public health surveillance (21). Public health officials in Sydney, Australia used the 2003 Rugby World Cup (RWC) to test the viability of a near real-time syndromic surveillance system using automated data routinely collected from Emergency Departments (EDs) in the Sydney metropolitan area (22). The system did not identify any major public health threats during the RWC, which was consistent with evidence from other sources. Interestingly, however, the system also did not detect two known outbreaks that were already in progress before the tournament. This was likely due to limited baseline data early in the monitoring process which prevented the system from automatically identifying these ongoing outbreaks (22). Because many influenza infections are subclinical, Wu et al investigated the use of 'real-time' sero-surveillance rather than syndromic surveillance as a method of estimating outbreak infection severity (23). Using sera samples from the first wave of the 2009 Hong Kong pandemic, infection attack rates (IAR) and infection-hospitalization rates (IHR) were estimated. The study found IAR and IHR estimates to be one to two weeks behind the pandemic peak; assuming data had been available weekly in real-time (23). Computer simulations for future pandemics indicated 300 serum specimens per week would be required to yield reliable estimates of IAR and IHP; assuming the true IAR was about 6% (23). These findings show both the utility and the limitations of sero-surveillance in detecting emerging influenza pandemics. While the data may be indicative of the severity of the outbreak, the reliability of the system is dependent on the existence of a well-established sero-surveillance program. Another issue this study highlights is the misuse of the term 'real-time' when discussing surveillance.

Animal Disease Surveillance

There are many reasons for the surveillance of disease in animal populations. Such surveillance can serve to safe guard livestock industries and subsequent national and international trade. The USDA conducts such surveillance for the purpose of protecting the U.S. livestock industry (5). Surveillance in support of livestock production in developing countries is the priority for the Food and Agriculture Organization of the United Nations (FAO) (24). For the OIE the focus is on animal diseases that impact international trade (8). Animal disease surveillance can also play a vital role in public health surveillance efforts, especially as it applies to emerging and infectious zoonotic diseases. In 2001, a comprehensive literature review determined that 61% of infectious organisms known to be pathogenic to humans are zoonotic and 75% of the emerging pathogens are zoonotic (25). Zoonotic diseases include those that can be transmitted directly from animal to human (dog bites transmitting Rabies, some intestinal parasites), require a shared vector (*Ixodes* ticks and Lyme disease, *Culex* mosquitoes and West Nile Virus), result from animals serving as reservoirs for disease transmission (fecal contamination of food and water with *Salmonella*), or come from shared environmental exposures (*Anthrax* spores in soil). Unfortunately, there are no organizations or agencies dedicated to the surveillance of infectious zoonotic disease (26) and systematic review of surveillance systems used for emerging zoonoses revealed that only 19% incorporate information from both human and animals (27). In particular, there is a significant lack of infrastructure that supports the surveillance of such diseases in companion animals (26), yet market estimates from the American Veterinary Medical Association showed that for 2012 there were 70 million pet dogs kept by 37% of households and 74 million pet cats kept by 30% of households (28). Instead, the value of animal disease surveillance in public health is demonstrated in a series of independent studies, some of which will be discussed in more detail in the section of this chapter titled “Zoonotic Disease Surveillance and Use of Animals as Sentinels”.

Surveillance Networks

Traditional public health surveillance, based on reports and medical practitioners, relies on information to travel up or down a predetermined public health hierarchy: farm to State department, hospital to public health officials, national to international, and vice versa. Unfortunately, these seemingly orderly scenarios are often patchy and

erratic due to differences in priorities and available resources. Electronic medical records (EMR) have significantly improved surveillance capabilities, but a lack of uniformity in existing EMR systems and the absence of their existence in many developing countries make reliance on these systems impractical. Surveillance networks provide an alternative method of surveillance that incorporates data from nontraditional sources (e.g. Internet sites, news outlets, federal registries, universities, non-medical observer accounts) to fill in the gaps of missing medical data from traditional sources. These networks serve to both collect and disseminate data relevant to public health surveillance at a global scale (7). Examples of such networks include ProMED-mail (PMM), the Global Public Health Information Network (GPHIN), HealthMap, and the WHO's Global Outbreak Alert and Response Network (GOARN). Geared towards integrating information on infectious diseases from a variety of sources and disseminating such intelligence widely and rapidly to the user community, the utility of these systems should be discussed in terms of impact of alerts generated by them rather that literature published based on them.

The Program for Monitoring Emerging Diseases, ProMED, was founded in 1993 at a conference co-sponsored by the Federation of American Scientists (FAS) and the WHO. The ProMED-mail was created in 1994 for the purpose of enhancing communication between ProMED working group members the initial list-serve contained 40 subscribers from seven countries (29). The first e-mail outbreak report was sent later that year, reporting a laboratory infection of Brazilian hemorrhagic fever (Sabia virus) in the United States (29). In November 1994, PMM was opened to the public (29). ProMED-mail currently reaches over 60,000 subscribers in at least 185 countries and since 1999 operates as an official program of the International Society for Infectious Diseases, a nonprofit professional organization with 20,000 members worldwide (30). ProMED-mail is open to all sources and is free of political constraints. Sources of information include media reports, official reports, online summaries, local observers, and others (29, 30, 31). Reports are often contributed by ProMED-mail subscribers. A team of expert human, plant, and animal disease moderators screen, review, and investigate reports before posting to the network (29, 30). Reports are distributed by email to direct subscribers and posted immediately on the ProMED-mail web site (30). These open communications and discussions amongst subscribers aided in the detection of the Ebola virus outbreak in Zaire (1995), West Nile virus in the United States (1999), SARS in China (2002), and H5N1 avian influenza in Indonesia (2003) (30, 31).

The Global Public Health Information Network (GPHIN) is an electronic public health early warning system developed by Canada's Public Health Agency, and is part of the World Health Organization's (WHO) Global Outbreak and Alert Response Network (GOARN) (32). This system monitors internet media, such as news wires and websites, in nine languages in order to help detect and report potential disease outbreaks or other health threats around the world (32, 33). The advantage of this approach is that it does not rely on the gathering and dissemination of information to country level public health officials. Reliance on public health officials is possible in most developed countries however, in countries where the public health infrastructures are rudimentary, deteriorating or non-existent; reporting of many public health threats is considerably less than adequate. Furthermore, the reluctance of some countries or authorities to report potential threats due to the negative impact on trade and tourism, or to gain a tactical advantage, has also resulted in limited exchange of information between authorities. The information feeds are filtered for relevancy by an automated process which is then complemented by human analysis (32). Analysts at the GPHIN sort and score the more than 2,000 preliminary reports and news articles received on a daily basis (32, 33). Scoring merely applies to a decision tree used to determine whether or not to post the report; GPHIN does not systematically validate the data (32). Subscribers to GPHIN receive alerts by e-mail or when logging on to the system. The GPHIN audience includes ministries of health and departments of agriculture from several nations as well as FAO, OIE, NATO (North Atlantic Treaty Organization). The GPHIN prototype demonstrated its effectiveness as an early-warning system during the 2002 SARS outbreak, where it gathered information about an unusual outbreak occurring in Guangdong Province, Mainland China, as early as November 27, 2002 (32).

In 2000 the WHO established the Global Outbreak Alert and Response Network (GOARN) to connect existing surveillance networks internationally for the purpose of improving the coordination of outbreak responses and providing an operational framework for the delivery of support to these countries (34). The network focuses technical and operational resources from scientific institutions in WHO Member States, medical and surveillance initiatives, regional technical networks, networks of laboratories, United Nations organizations (e.g. UNICEF, UNHCR), the Red Cross (International Committee of the Red Cross, International Federation of Red Cross and Red Crescent Societies and national societies) and international humanitarian nongovernmental organizations (e.g. Médecins sans Frontières, International Rescue Committee, Merlin and Epicentre) (34). Participation is open to technical institutions, networks and organizations that have the capacity to contribute to international outbreak alert

and response. Since 2000, WHO and GOARN have responded to over 50 events worldwide with over 400 experts providing field support to some 40 countries (34). The GOARN was instrumental in initiating the global response to the 2002 SARS outbreak (34).

HealthMap is a freely available, web-based surveillance network that provides a global view of infectious disease outbreaks as reported by the WHO, PMM, Google News, and Eurosurveillance (35). Created in 2006, it attempts to bridge the gaps in current open-source electronic surveillance networks by aggregating and integrating the information to produce graphic, continuously updated models of global disease outbreaks over space and time (35). Alerts are displayed on a global map that can be viewed at a wide range of resolutions and are linked to source sites that provide details of the outbreak and information on the particular disease (7). Through an automated process, updating 24/7/365, the system monitors, organizes, integrates, filters, visualizes and disseminates online information about emerging diseases in nine languages, facilitating early detection of global public health threats (35). The system does not have a formal verification process; therefore any reporting biases that may exist in the HealthMap data sources remain (7).

The strength of these surveillance networks is to consolidate public health information at a global level for the purposes of international awareness. By employing both traditional and non-traditional data sources, these networks are able to fill in official surveillance gaps in order to flesh out missing data. They also are able to circumvent political filters that may prohibit truthful and honest reporting from certain governments/countries. Such networks can serve as sensitive indicators for potential outbreaks and disease events. In addition, they provide a forum for health officials to discuss and share concerns of emerging and current health threats. There are several limitations to such networks, however. For example, the lack of formal validation processes may lead to inaccurate conclusions and the implementation inappropriate (and potentially costly) public health actions. In addition these networks may have significant economic consequences for countries that may be labeled (either accurately or inaccurately) as harboring a feared infectious disease.

Evaluating Surveillance Systems

The purpose of evaluating public health systems is to ensure that the problems of public health importance are being monitored efficiently and effectively (3). Specifically, the focus should be on how well the system operates to meet the stated purpose and objectives of the given surveillance program. In general, the evaluation should involve an assessment of system attributes such as simplicity, flexibility, acceptability, data quality, sensitivity, specificity, positive predictive value, representativeness, timeliness, and stability (3). This section will briefly discuss several (but not all) of these attributes as defined by established CDC guidelines (3), providing examples of their use and demonstrating their impact on public health strategies.

Surveillance System Attributes

Simplicity refers to the structure of the system and its ease of operation; specific categories include methods of data collection (number and type of reporting sources), level of integration with other systems, methods for data analysis and data dissemination, time required to maintain and update the system, and human resource/staffing requirements. Overly complicated systems may be prone to system failures or may have extremely high human resource requirements. **Flexibility** is important as it represents the ability of a system to accommodate changing information needs or operation conditions with little additional time, personnel, or funding; the ability to adapt and evolve to meet new requirements. These two factors (simplicity and flexibility) contribute to the overall **acceptability** of the system, or the willingness of the user (data entry through analyst) to actually use the system.

Data quality refers to the completeness and validity of the data recoded in the system. There are many subcategories that may be used to evaluate a system's **data quality**; all depend quite a bit on the purpose of the surveillance activity. Mode of data entry (automated vs. manual) impacts correctness of the data, and in combination with an understanding of the completeness and validity of the records, contributes to the overall accuracy of the data. Case definition and diagnostic criteria impact the sensitivity and specificity of the results. S**ensitivity** refers to the proportion of cases of a disease (or other health-related event) detected by the system, or the ability of the system to detect outbreaks or monitor changes in the number of cases over time. **Specificity** refers to

the ability of the system to accurately identify as non-cases those without the disease or health-related event. The **positive predictive value** (PPV) of a system refers to the proportion of reported cases that actually have the disease or health-related event of concern. The **representativeness** of the system pertains to the ability of the system to describe the occurrence of the disease or health-event in the population of interest, and requires an understanding of how the data was collected (population sampled, active vs. passive surveillance, voluntary vs. directed participation).

Timeliness of the information refers to the time between the occurrence of the health related event and its appearance in a functional surveillance system. Both the lag time between the detection of the medical event and the time required to generate a system record impacts a system's timeliness. The **stability** of the system is affected by security issues, system downtimes, and manpower constraints, all of which directly impact the overall reliability and availability of data generated by that system.

Applications of System Attributes in Surveillance Programs

Public Health Awareness and Risk Assessment Programs

Programs aimed at increasing the public's awareness or determining the risk of a disease or health-related event are generally designed to capture a broad snapshot of a particular condition in the population. Determining the overall disease burden or prevalence is often the goal, therefore these programs often tolerate longer lag times between data collection and analysis (less timely surveillance systems). Data collection may be actively collected to meet the specific goals of the program (active surveillance), or may be through data mining of existing system (passive surveillance). Because these programs hope to determine the potential existence of a disease or condition, the data collection system employed aims to be as sensitive as possible, with less emphasis on specificity. Surveillance done in support of these programs tends to lead to broad intervention or prevention programs and often policy changes. Depending on the specific goals of the programs, often dictated by the stakeholders and associated funding, surveillance efforts may be focused on a very specific population or demographic group. In these cases the representativeness of the data collected is critically important.

In 2013, Argaw et al conducted a study to determine the risk factors for Visceral Leishmaniosis (VL) among resident and migrant populations in Kafta-Humera, Ethiopia (36). Using hospital records to identify cases, they collected risk factor data through the administration of questionnaires. The study showed sleeping under an acacia tree at night, poverty, and lower educational status to be associated with increased risk in both populations. Among the resident population they also found living in a house with thatched walls and sleeping on the ground to be significant risk factors; whereas HIV status and sleeping near dogs added risk for migrants. Bed net use, especially during the rainy season was found to be protective. Case definitions followed local hospital diagnostic algorithms, where patients had at least 2 weeks of fever plus weight loss and/or splenomegaly and a laboratory confirmed diagnosis of VL. Laboratory tests used had sensitivities ranging from 79% to 93% and specificities ranging from 85% to 96% (36). The high variability of diagnostic test capability and quality is one limitation of this study. Also, the findings of this study are limited to a very specific population in a very specific location. However, the risk factors and preventive measures identified are representative of this specific populous, and therefore the development of practical and relevant intervention practices are possible.

In order to better understand the risks factors for severe H1N1 infections associated with the 2009 global outbreak, Van Kerkhove et al analyzed data obtained primarily from surveillance programs of the Ministries of Health or National Public Health Institutes of 19 countries from the period 1 April 2009 to 1 January 2010 (37). Data sources included 70,000 laboratory-confirmed hospitalized H1N1 patients, 9,700 patients admitted to intensive care units (ICUs), and 2,500 deaths reports. The highest risk of hospitalization was among 5–14 year old patients; while the maximum risk of death was in those 50 years old or older. Although the infection rate was low in the over 65 year old group, the ratio of H1N1deaths to hospitalizations increased with age. The proportion of patients with one or more reported chronic conditions increased with severity of H1N1 infection (37). Overall, the findings demonstrated that risk factors for severe H1N1 infection are similar to those for seasonal influenza; notable differences included younger age groups and obesity. Although the study provides a broad overview of the risk factors associated with the H1N1 outbreak on a global scale, it does not provide enough detail to support the implementation of targeted preventive strategies on a local level. Another significant limitation is the wide differences in surveillance systems employed by each country, to include potential differences in H1N1 case definitions and the completeness and quality of risk factor data. This variability could have led to clustering of data

by country, making direct comparisons across countries inaccurate. The country level data lacked individual-level details which eliminated the ability to control for confounding and the subsequent identification of the independent contribution of individual risk factors.

Disease Control Programs

Disease Control Programs (DCP) are directed at "reducing the frequency of existing disease to levels biologically or economically justifiable or otherwise of little consequence" (6). Such programs require a more accurate determination of the particular disease or condition in the population. Determining the overall disease burden or prevalence is still important, but in addition the incidence (occurrence of new cases) of disease is also important. For this reason the surveillance system used must be timelier than that tolerated in awareness and risk assessment programs. In general data collection is active, targeted specifically to the population of concern. In addition, the methods used to determine the presence of disease must be sensitive enough to capture the disease (some false positives may be tolerated) but specific enough to limit the number of false positives. The tolerated levels of sensitivity and specificity are dictated by the proposed control practices as well as the impact of not achieving control of the disease. For example, resource intensive control practices such as central quarantines may warrant more specific (and potentially costly) diagnostic testing; whereas quarantines conducted at home residencies may be initiated with less specific testing. A high rate of false positives (high sensitivity with low specificity) may be tolerated in order to control a particularly virulent disease; whereas the same would not be tolerated with a less virulent disease, especially if treatment is costly or invasive.

In 1995 the CDC established Active Bacterial Core surveillance (ABCs) as a core component of their Emerging Infections Programs network (EIP) (38). The collaboration between the CDC, state health departments, and universities is an active laboratory- and population-based surveillance system for invasive bacterial pathogens of public health importance (38). The ABCs data have been used in disease control programs to track disease trends in response to public health interventions, including the decline in pneumococcal disease following the introduction of the pediatric pneumococcal conjugate vaccine (38, 39). Using very specific case definitions and reporting criteria, the system is designed to determine the incidence and epidemiologic characteristics of invasive disease due to

Haemophilus influenzae, *Neisseria meningitidis*, group A Streptococcus, group B Streptococcus, *Streptococcus pneumoniae* and methicillin-resistant *Staphylococcus aureus* (MRSA) (38, 39). Programs to assist state and local health departments with surveillance and control of MRSA and drug-resistant *Streptococcus pneumoniae* have been developed based primarily on lessons learned from ABCs (38, 39). A comparison of ABC surveillance data for invasive MRSA infections in 2011 to the baseline incidence of 2007 identified a decrease of nearly 26% (39). The decrease was attributed to the Department of Health and Human Services Action Plan to Prevent Healthcare-Associated Infections in support of the National Metric for Healthy People 2020 (39). The strengths of the CDC ABCs program include strictly defined case definitions, standardized report submission templates and protocols, and annual audits of all participating laboratories. One of the most significant limitations is that only 10 U.S. population areas are represented in these surveillance activities (as of April 2014).

In an attempt to better understand an outbreak of human encephalitis that was associated with concurrent extensive bird mortality in the northeastern United States in 1999, scientists conducted complete genome sequencing of the virus isolated from the brain of a dead Chilean flamingo (40). Together with gene amplification from several other species including mosquitoes and two fatal human cases, they were able to determine that West Nile virus (WNV), previously never detected in the U.S. was responsible for the human disease (40). The rapid identification of the causative agent by these researchers was the first step in the development of systems and procedures to implement effective WNV control programs. The CDC Division of Vector-borne Disease leads the WNV control program for the United States (41). Their published guidelines provide details on how to conduct surveillance efforts as well as how to implement prevention and control strategies for the virus (41). Because laboratory confirmation of disease in humans is not only unreliable but also lags several weeks behind active infection, effective WNV surveillance must also include efforts to monitor the virus in the mosquito vectors and non-human vertebrate hosts. This comprehensive approach spurred the 2000 creation of ArboNET, the national arbovirus surveillance system which serves as a surveillance capture platform to monitor WNV infections in humans, mosquitoes, birds, and other animal (42). Human data includes both disease cases and information collected from donor screening programs (presumptive viremic donors) (42). Human data elements include (but are not limited to) age and sex, county of residence, date of symptom onset or date of donation, case status, and laboratory of diagnosis (42). Non-human data elements include species, state and county, and date of symptom onset or collection. As a notifiable disease, health

care providers are required to report data to their state/local health department, who then enters it into an electronic database which is uploaded to the CDC database weekly (42). The CDC analyzes the data and disseminates information regularly, weekly during the WNV season. Information includes both human and animal disease cases and deaths by state, human demographics, mosquito cases/pools, and a comparison of current to previous year at same week (42). Most of the ArboNET's limitations are due to it being a passive surveillance system: there is minimal clinical and laboratory data available to permit the confirming of cases, delays between case occurrence and reporting may exist, and variability in reporting practices by location may lead to gaps in data from some areas. One of the systems greatest strengths is its comprehensiveness, merging data from human, animal, and vector sources. In addition, the data within the system is standardized by defined case definitions and reporting protocols. The timeliness of both the data reporting and the report dissemination provides incidence and geographic and temporal trend analysis, enabling the implementation of control strategies.

Disease Eradication Programs

The goal of a Disease Eradication Program (DEP) is to eliminate selected organisms from a populous which is specifically defined by population characteristics (e.g. species, race, gender) and location (e.g. herd, farm, region, state, city) (6). Such programs require exact determination of the particular disease or condition in the population. The accurate and timely diagnosis of both prevalent and incident cases is critical and therefore active data collection, targeted specifically to the population of concern is used (6). Eradication programs generally involve strict intervention and control strategies. In animal populations these range from slaughtering affected animals to slaughtering the entire herd. For human populations it may include an aggressive quarantine and vaccination program. Both scenarios are resource intensive and very costly. For this reason the methods used to determine presence of disease should not produce a high level of false positives. However, complete eradication of disease assumes that all positive cases are correctly identified; therefore there cannot be any false negatives, either. Directors of DEP must have a deep understanding of not only the pathogen/organism being eradicated and the strengths and limitations of the surveillance system being used to determine disease presence or absence.

To date, only two diseases have ever been given the status of eradication: smallpox and rinderpest. Although widespread rinderpest eradication efforts started as early as the 1900s (the OIE was formed in response to the disease) most took place on an individual country basis (43). With the support of the FAO, the OIE, and the International Atomic Energy Agency, the Global Rinderpest Eradication Programme (GREP) was created (43, 44). Initiated in 1994 as an international coordination mechanism to promote the global eradication of rinderpest by 2010, the GREP established technical guidelines and standards for achieving the goal (43, 44). The eradication of the disease was confirmed by the World Organization for Animal Health on 25 May 2011 (43). The monumental accomplishment may be attributed to multiple factors. The creation of GREP was paramount as it coordinated eradication efforts internationally and established technical guidelines and standards. Specifically, it established appropriate disease surveillance techniques such as uniform sampling strategies, supported national laboratory services in organizing standardized surveillance activities, assisted national veterinary services to conform to OIE guidelines for declaration of freedom from disease and infection, and stockpiled high quality vaccines (43). In addition the GREP articulated strategies for the prevention of or response to the re-introduction of the rinderpest virus to include effective national/regional emergency plans (43). Initially, for a country to be recognized as free from rinderpest they were required by the OIE Terrestrial Animal Health Code to go through three phases: 1) provisionally-free status of self-declaration, pending official recognition of freedom from disease 2) disease-free status, and 3) infection-free status. Later they were only required to achieve infection-free status. Doing so required the country to 1) provide a letter declaring historical freedom from rinderpest according to the 1999 declaration made by OIE Delegates; 2) provide a dossier for historical freedom, attesting to never have experienced the disease or that the disease had been eradicated from the country for more than 25 years, and/or that the country stopped vaccination more than 10 years prior; 3) submit field and laboratory rinderpest data required by the dossier (44). Several features specific to the pathogenesis of the disease were also important to the eradication of the disease. For example, there is no carrier state where individual animals could act as long-term reservoirs of infection; animals exposed to rinderpest in nature develop clinical signs that either lead to death or recovery with life-long immunity (45). Because, some animals only develop mild disease the existence of serological methods to aid in definitive diagnosis is also critical (under OIE protocol, even these mild cases of disease had to be fully investigated) (45). Lastly, an effective vaccine that confers life-long immunity was absolutely essential to the program's control efforts and the eventual eradication of the disease (45).

Comparing Surveillance System Data

Until now, this review has focused on the evaluation of public health surveillance systems as separate entities and on an individual basis. In the next section the concentration shifts to evaluating surveillance systems by comparing similar systems to each other, specifically through assessing each one's data. Such comparisons may be used to validate overlapping data, estimate underlaps in data, identify redundancies in order to improve efficiencies and maximize resource management, and/or to aid in the selection of the appropriate system for a given project/incentive. The data used in surveillance systems can come from a many different types of sources and can be in a variety of formats. These differences can significantly impair the ability to compare the data, making subsequent surveillance system comparisons based on the data challenging. This section will introduce potential differences in data; discuss issues that arise when attempting to compare data from different systems, and review approaches used to conduct these comparisons. The discussion assumes the data being compared are collected in order to answer related if not the same overarching question, such as the occurrence of a disease or event.

Potential Differences in Data

Data variability is a very broad term used to describe the difference data due to factors, such as data source, data type, and/or the method of data collection/system used to compile data. **Data source** refers to the origin of the data, where it came from (e.g. hospitals, health departments, public records, media). **Data type** refers to the category of data collected (e.g. syndromic, test result, fatalities, armed conflict). In addition, data variability may be influenced by the mode of **data collection** (manual versus automated entry), which is further influenced by differences in **data format** (free-text, menu-driven, multiple choice). All of these factors may contribute to differences in naming conventions for the event or disease of interest, which ultimately makes the direct comparison of data difficult.

Data quality is affected by many different factors and may be discussed in terms of its validity and completeness. **Validity** can be measured by the sensitivity and specificity of the data. **Sensitivity** is calculated by determining the proportion the total population's cases that the dataset correctly identifies as cases or event or the probability of the dataset correctly identifying as cases or events those that are really are. **Specificity** is the proportion of the total

population's non-cases or non-events that the dataset correctly identifies as non-cases or non-events or the probability of the dataset correctly identifying as non-cases or non-events those that really are. **Completeness** refers to the ability to capture or ascertain the information needed to accurately assess the objectives of the dataset. Incomplete data may lead to biased conclusions. For example, if data collection is not emphasized as important, or if data entry is cumbersome, there may be a large amount of partial or missing data. Differences in data quality directly contribute to observed differences in data.

Approaches to Comparing Data

Data Harmonization Techniques

One of the barriers to comparing data from different systems is the lack of standardized naming conventions; therefore it has been proposed that the first phase of dataset comparison should be the creation of a combined dataset. Nadrowski et al. propose the use of an open source web platform for the upload, validation, and storage of data from different sources (46). The system is designed to allow for harmonization of naming conventions through a "bottom-up approach" to generating category lists from the primary data sources. Although this method offers a potential solution to the heterogeneity of naming conventions, there are several concerns about its practicality. The system is labor intensive, requiring manual entry of the primary source data into an 'import workbook' to facilitate its incorporation into the merged database (46). In addition, there are multiple layers of data management involved in the data validation process which make this technique cumbersome and potentially impractical (46). Another approach is to create a consortium of representatives for each data system being compared before attempting to merge the datasets (47.) This approach permits the collaborative creation of common naming conventions and case or event definitions (47). Unfortunately, this approach is not practical in many if not most instances. On a case-by-case basis or for a particular joint project it may be feasible, but in most instances large scale meetings are costly and potentially non-productive.

Metadata Coverage Index (MCI)

Liolios et al. propose a record scoring system based on the richness of data sources as a proxy for data quality (48). Their Metadata Coverage Index (MCI) calculates the number of fields in a record or database for which information is provided as a percentage (score) of the total fields available (48). Although the MCI may have utility in determining data entry compliance within specific databases or by specific record users, there are several limitations to its use to compare data quality across data sources. The MCI score does not correlate to the validity of the data; a complete but inaccurate dataset will receive a higher score than a partially complete but 100% accurate one. Also, there is no mechanism to weight fields that are deemed to be of higher importance; for example the outcome variable of interest versus an administrative code. Lastly, because the MCI score is a proportion, it is highly dependent on the denominator or number of available fields of data. This means even datasets that collected the exact same data will have different MCI scores when the data derives from sources with varying available fields. In fact, it would assign a lower score to the source with a larger collection capability, which may actually have better future use (a larger potential to expand data collection) than the more limited one.

Modeling Techniques

More complicated approaches have been proposed as well. In an attempt to create an efficient and effective way to merge data from different distributed information retrieval systems, Wu and Li review existing ways to assess the degree of overlap between data systems and propose some new approaches (49). When overlap is expected to be light and all data is of a similar quality, they suggest the simple Round-robin evaluation of each document or entry separately (49). With excessively large data sets this method would quickly become labor and resource intensive. The zero-one score normalization method calculates a raw score for each database based on the number of documents retrieved then uses a mathematical function to normalizes each score into a range of [0,1] linearly. This permits the ranking of databases through these now-normalized scores (49). Wu and Li, however, prefer to use the cubic and logistic regression to convert document ranks and normalized scores to actual estimates of the relation between rank and probability of relevance (49). Another approach they reviewed involved using weight assignments; where databases expected to have the least overlap are assigned heavier weights than those with higher

expected overlaps. They then propose their own methods, the shadow document method (SDM) and the Multi-Evidence Method (MEM). The SDM looks directly at the amount of overlap between two component databases through the use of queries and calculates a score based on the probability of a document not being found in both databases. The multi-evidence method (MEM) requires the averaging the score of every document retrieved from the databases, then multiplication of the average by a factor which is a function of the number of databases that include the document in their results. Although these methods might be effective tools for determining the extent of overlap between data sources for the purpose of merging data from distributed information retrieval systems, they may be too complicated to be practical. The use of these modeling techniques would require extensive training and may become labor and resource intensive.

Capture-Recapture Techniques

Capture-recapture (CR) techniques have long been used in wildlife biology to estimate total population numbers (50). The approach involves two trapping events. The first trapping produces the "capture" sample; these animals are counted, tagged, and released back into the population. The second trapping produces the "recapture" sample; in addition to counting the total number captured in this second sample, the number of previously tagged animals (recaptured animals) is also counted (50). The proportion of animals in the second trapping that have never been captured before represents the percent of the wildlife population that was not captured by either trapping event. The total population size is then estimated by multiplying the sample size by this percentage. In a similar way, CR techniques can be used to estimate the total number of cases or events expected when incomplete data exists. In doing so, CR techniques facilitate the comparison of data by providing an estimate of the denominator data required for completeness calculations (50). CR techniques were used to estimate the total number of people infected with human immunodeficiency virus type 1 (HIV-1) in Lazio, Italy, during 1990s (51). The resulting population estimates were consistent with those using a more complex compartmental mathematical model. The study concluded that CR techniques provide a simple and inexpensive means of obtaining accurate estimates of the total number of those infected with HIV when incomplete data from multiple sources must be used (51). In 2007, the Navy and Marine Corps Public Health Center (NMCPHC) utilized CR techniques to conduct a small pilot study of Leishmaniosis cases in the San Diego area naval medical system (52). Initially implemented to estimate the total

burden of disease, the resulting denominator data was then used in completeness calculations, enabling the comparison of data sources and a determination of the effectiveness of the three surveillance systems used (52). The limitations of using CR to compare data from different sources are directly related this technique's many underlying assumptions. For example, it is implicit that each case has been diagnosed accurately and that matching between sources has been done appropriately. Also, it is assumed that for any single source, each case in the population has the same "catchability" or probability of ascertainment, that selection bias does not exist between sources (50). Violation of these assumptions may lead to inaccurate estimations of total population data and therefore incorrect conclusions of data completeness.

Descriptive Comparisons

Sometimes it may not be appropriate or desirable to use the above mentioned methods for dataset comparisons. In such cases detailed descriptive comparisons of observed differences may be more appropriate. In such circumstances it is still important to conduct the comparison systematically in order to ensure each dataset or system is being evaluated fairly and equally. Doing this involves first determining which characteristics to evaluate. These generally are driven by the reason for the data comparison, and may include characteristics such as the number of variables available to define the event of interest (e.g. patient demographics, disease status, category of event, unique identifiers), the methods used to validate the data, and the sources used to gather the data. In 2012, in support of research studying the dynamics of civil war, a comparison of two leading conflict events datasets, the Uppsala Conflict Data Program (UCDP) and the Armed Conflict Location Events Dataset (ACLED), was conducted (53). The goal of the comparison was to describe the differences between the two datasets (including strengths and weaknesses) and provide guidelines as to the most appropriate use of each (52). Characteristics used to make these determinations include data source, conflict event definition, and quality-control practices. The study found that both datasets shared a heavy reliance on media as their primary data source; however the two used very different conflict event definitions (52). The UCDP restricts it domain to events which resulted in a fatality, thereby limiting the definition of a conflict event to those with fatalities. The ACLED dataset, on the other hand, has a domain that includes not only events with fatalities and without fatalities, but also events that are characterized as non-violent (52). In addition, it was discovered that the ACLED dataset had significant quality-control issues that could result in

biased results if left unchecked by the researcher. In conclusion, it was discovered that those interested in non-violent events such as troop movement only had the ACLED dataset to choose from, but researchers were told to be wary of the data due to quality-control issues (52). Hedden et al conducted detailed descriptive analyses of multiple adult mental health data sources in order to facilitate the comparison of adult mental health prevalence estimates generated in the United States (54). In specific, the study aimed to compare the prevalence estimates generated from the 2009 National Survey on Drug Use and Health (NSDUH) to estimates generated from other national data sources (54). Descriptions included details on each source's survey design and instrumentation, mode of data collection, and methods used to produce the estimates. The comparison focused on the specific mental health indicators of serious mental illness (SMI), any mental illness (AMI), serious psychological distress (SPD), major depressive episodes (MDE), and suicidality (54). The study concluded that the substantial differences in methodologies between the data sources explained the observed differences in adult mental health prevalence estimates, and recommended using the data from multiple sources in order to get the most comprehensive picture of mental health in the United States.

The Use of Animals as Sentinels in Zoonotic Disease Surveillance

The use of animal disease surveillance in public health was briefly mentioned earlier in this chapter as a valuable public health surveillance strategy. The purpose of the current section is to further discuss the role of animals in public health, focusing on their use as sentinels in zoonotic disease surveillance. This section will introduce the concept of "One Health", discuss the application of animals as sentinels, and review research demonstrating animals as sentinel surveillance in public health strategy.

One Health

The One Health concept recognizes that human, animal, and ecosystem health are inextricably linked. It seeks to promote, improve, and defend the health and well-being of all species by enhancing cooperation and collaboration between physicians, veterinarians, other scientific health and environmental professionals. One of the stated

objectives of One Health is to achieve joint cross-species disease surveillance and control efforts in public health (55). In 2008 the American Veterinary Medical Association (AVMA) published the "One Health: A New Professional Imperative" dedicated to finding a holistic, collaborative approach to contemporary health issues through the convergence of human, animal, and environmental domains (56). The document highlights that of the approximately 1,500 diseases recognized in humans; almost 60% are due to multi-host pathogens characterized by their movement across species lines. While multi-host pathogens may present considerable challenges implementing disease control practices, the existence of these "other" hosts may provide surveillance opportunities that subsequently contributed to enhanced prevention strategies.

Animals as Sentinels

Sentinel surveillance focuses on specific subpopulations to enhance detection of disease and/or improve the cost-effectiveness of surveillance activities (57). The goal of using animals as sentinels for human disease surveillance is to estimate disease risk in order to make informed recommendations on preventive practices and therapeutic protocols (57). The use of sentinels may answer a variety of surveillance questions such as 1) the presence of a pathogen in a new area, 2) changes in the prevalence or incidence of a pathogen or disease over time, 3) the rate and/or direction of pathogen spread, 4) the ecology of the pathogen, and 5) the effectiveness of pathogen control practices (57).

There are many considerations when determining if an animal is a suitable sentinel for disease surveillance. Compiling a definitive list of characteristics is not practical, however, because the criteria against which the usefulness of a given sentinel population is assessed are influenced by the aim of the surveillance and the context of its use (58). Halliday et al., instead propose the use of a **"framework" for evaluating animals as sentinels** for infectious disease. Key to this framework is an understanding of how three fundamental components of the surveillance program interact: the **sentinel population** being used to conduct the surveillance, the **target population** for which the surveillance is being conducted, and the **pathogen** under surveillance (58).

In order to serve as sentinels, the animal models must be susceptible to the pathogen and must produce a measurable symptom or response that indicates exposure or infection (57). How the sentinel responds to the pathogen determines how pathogen presence will be measured as well as the application of the data once it is obtained. For example, disease presence may be easily observed through morbidity or mortality or it may involve active diagnostic testing through sampling such as sero-conversion or pathogen presence. The **sentinel response** can be viewed as a test for the presence of the pathogen within the target population, and as such has properties that are analogous to test sensitivity and specificity. **Sentinel sensitivity** is then the susceptibility of the sentinel to the pathogen and the **sentinel specificity** is the ease with which a sentinel response can be interpreted and attributed to a particular pathogen (58). In addition, the **temporal characteristics** of the sentinel response may dictate the limits of the sentinel's use in public health strategies; rapidly detectable responses may be used for early warning, whereas slower but long duration responses may be useful in prevalence studies. The **target population's response** to the pathogen influences the application of the sentinel data as well. For example, if the target population tends to have a more severe response to the pathogen, then the surveillance system may tolerate lower sentinel specificity in favor of higher sentinel sensitivity. Awareness of the relationship between the sentinel and the target population is not required to address all questions, but will allow for the investigation of more complex epidemiological questions and better informed interpretation of the data collected using that sentinel. At a minimum, the sentinel and target population should be spatially associated (58). Knowledge about the pathogen further facilitates an understanding of the other components: the sentinel response, the target population's response, and the potential for transmission between the two populations. Specifically **pathogen factors** that should be understood include, but are not limited to: **infectivity**, the ability to establish an infection; **pathogenicity**, the ability to produce clinical or detectable disease; and **virulence**, the ability to produce severe disease.

Research Demonstrating Animals as Sentinels

A survey of Washington state veterinarians in 2008 showed that more than three quarters of respondents believed a veterinarian's role in educating clients on zoonotic disease was very important, but only 43% reported doing so on a daily basis (59). Seventy-four percent of the population indicated wanting more public health agency assistance with issues involving zoonotic disease; specifically a web-based data dissemination system (59). Despite the

recognition that animal and human health are intrinsically linked, and the apparent willingness of veterinarians to play their part, there still are no organizations or agencies dedicated to conducting zoonotic disease surveillance in animal and human populations simultaneously. Instead, the value of animal disease surveillance in public health is primarily demonstrated in a series of independent studies.

Anthrax

Anthrax is a zoonotic pathogen that has also been identified as a potential bioterrorism agent. A systematic review of the scientific literature from 1966 to 2005 determined that sentinel surveillance for sporadic cases of livestock in non-endemic areas may serve as an indicator for a potential attack (60). In addition, although dogs and cats are less susceptible to the pathogen than ruminants, the review recommended including them in surveillance activities due to their proximity to humans (60). The review did not indicate the methods of surveillance that should be used. Anthrax continues to be an issue in the Serengeti ecosystem of Tanzania (61). Despite the recognized value of serologic data for disease surveillance, it is rarely used in studies of anthrax in the area. Reasons include the perception that diagnosis can be based solely on the syndrome of sudden hemorrhagic death in herbivores, the rapid deterioration of effected carcasses which prevent productive sampling, the risk of exposure to those collecting the samples, or the lack of outbreak reporting due to inaccessibility (61). In an effort to find a better way to monitor Anthrax in the region, Lembo et al. conducted a study investigating to potential to use sero-surveillance on native wildlife species, local livestock, and domestic dogs. Using case reports and a serologic assay that enabled multispecies comparisons, researchers conducted a sero-surveillance study in the Serengeti ecosystem in Tanzania from 1996–2009 (61). High seroprevalence among carnivores suggested regular nonfatal exposure in this species and reflected known patterns of endemicity for the region. Also, because of the existence of robust preventive medicine campaigns in the area, the authors felt that routine surveillance of these domesticated dogs would be relatively easy to maintain. The study concluded that dogs could serve as indicator species in sentinel surveillance for Anthrax, providing new information about the pathogen in the Serengeti ecosystem (61).

Intestinal Parasites

In 2009 a study out of the Netherlands investigated the role of pets in the transmission of zoonotic parasites by analyzing fecal and hair samples from healthy household dogs and cats (62). The study found *Toxocara* eggs in 4.4% of dogs and 4.6% of cat fecal samples and in 12.2% of dogs and 3.4% of cat fur samples. In addition *Giardia* was present in from 15.2% of the dog and 13.6% of the cat feces. *Cryptosporidium* sp. were present in 8.7% of dog and 4.6% of cat feces (62). In addition, the study evaluated the close physical contact and other behaviors between owners and their pets that further increase the risk of transmission of these zoonotic parasites. Although not stated explicitly, the study demonstrates the value in routine monitoring of these parasites in even healthy animals as a mode of zoonotic disease surveillance for public health purposes (62). In a 2012 study by Schurer et al. companion dogs were used to estimate the risk of internal parasites in the indigenous people of Saskatchewan (63). Specifically they looked at the following intestinal parasites; *Echinococcus granulosus*, *Toxocara canis*, *Toxoplasma gondii*, *Diphyllobothrium spp*., and *Giardia duodenalis*. The authors cited many reasons these dogs are suitable sentinels for internal parasites in these communities. The dogs are highly susceptible and highly exposed to these parasites; dogs in remote and indigenous communities are often free-ranging and have access to human food, garbage, and carcasses of local fish and wildlife (63). Sample collection is relatively non-invasive, diagnostic tests can be run on either fresh feces found in the environment or that taken directly from the pet (63). Diagnostic testing for these parasites includes routine laboratory tests such as fecal floatation, microscopy, and commercial assays; all of which can be performed at the local veterinary facility (63). Overall, the prevalence of canine intestinal parasitic infection in the indigenous communities was 20–71%, which was 5–16 times higher than a nearby urban center in Saskatchewan (63). This information was interpreted to mean that the indigenous human population was at a 5-16 higher risk of acquiring these parasites than their urban counterparts (63).

Influenza

Historically sentinel surveillance for influenza viruses has been done on either avian or porcine species. Recent work, however, has challenged us to think outside the box. During the 2003-2004 avian influenza A (H5N1) virus outbreak in Asia there were anecdotal reports of domestic and zoo cats being infected with fatal H5N1 virus from

infected chickens (64). Previously this species was through to be resistant to infection. Experimental infection of 4- to 6-month-old European shorthair cats showed that cats can be infected with H5N1 virus both by horizontal transmission and by feeding on virus infected birds (64). The implications of the findings are that domestic cats may be playing a role in the spread of H5N1 virus between poultry farms, and from poultry to humans (64). Another study on experimentally infected cats showed the virus replicated not only in the respiratory tract but also in multiple extra-respiratory tissues causing systemic disease (65). All infected cats subsequently excreted virus not only via the respiratory tract but also via the digestive tract (65). The findings of these studies indicate that domestic cats may serve as sentinels in H5N1 surveillance; they are susceptible to infection and once infected demonstrate outwards signs of disease (death or multi-system disease) (65). In addition, domestic cats may play an important role in the epidemiology of the disease, serving as a source of infection to poultry, other cats, or even humans.

Lyme Disease

One of the aims of this project is to demonstrate the use of military pet dogs as sentinel surveillance for military populations. In support of this, a pilot study investigating the effectiveness of pet dog serology in sentinel surveillance for human Lyme disease was conducted for these populations. To facilitate the pilot study, a detailed review of similar studies was conducted, resulting in a much more extensive review of research demonstrating animals as sentinel surveillance in this disease.

Human Lyme disease is the most commonly reported vector-borne illness in U.S. and the seventh most common Nationally Notifiable disease (66). The causative agent is a spirochetal bacteria from the genus *Borrelia,* specifically *Borrelia burgdorferi* sensu lato (67). Within this group there are several relevant disease causing genospecies, this discussion will focus on the major *Borreliae* known to cause disease in humans, their vectors, and the geographical location of each. The vectors of *B. burgdorferi* sensu lato are various species of *Ixodes* ticks, whose distribution is associated with the specific geographic distribution of the *Borreliae* genospecies (67). In Europe both *I. ricinus* and *I. persulcatus* serve as vectors, carrying the three primary genospecies of concern for the region; *B. burgdorferi* sensu stricto (*I. ricinus* only), *B. garini*i, *B. afzelli*, and *B. spielmanii* (*I. ricinus* only). *B. garinii* and *B. afzelii*, carried by the associated ticks, also cause human borreliosis in Asia (67). In the United States, *I.scapularis* (northeastern, midwestern, and southeastern states down to northeastern Mexico, and up into southern

Canada) and *I. pacificus* together with *I. neotomae* (western states) carry the genospecies of concern in the United States, *B. burgdorferi* sensu sticto (Bb) (67). Studies using dog serology in sentinel surveillance for human Lyme disease in the United States were most likely detecting Bb, not one of the other *Borreilae* genospecies. Unless otherwise indicated, for the remainder of this document, the *Borrelia* pathogen will be referred to as Bb, the most common disease causing genospecies in the United States.

Typical symptoms of human Lyme disease include fever, headache, fatigue, and a characteristic skin rash called erythema migrans (66). Diagnosis can be difficult due to the vagueness of symptoms and the ambiguity of laboratory testing. If left untreated, infection can spread to joints, the heart, and the nervous system. Sixty percent of those infected develop chronic, severe arthritis (66). Even with treatment, 10-20% may develop Post-Treatment Lyme Disease Syndrome (PTLDS) (66). There is no current human vaccine. In 1998, direct medical costs associated with Lyme disease were estimated to be $2,970 per case (68). Indirect costs, including associated non-medical costs and projected losses in productivity, were estimated to be $5,202 per case. For the same time period mean productivity loss per clinically defined late stage patient was $9,108 (68).

Due to the high impact of the disease, difficulties in definitive diagnoses and treatment, and the lack of an existing human vaccine, the best way to combat the disease is through preventive measures. The cornerstone of preventive strategies, however, is to understand the risk of disease. Over the last several decades, there has been a plethora of research assessing the use of different surveillance strategies for the purpose of assessing the risk of Lyme disease in human populations. Such activities range from the observation of climatic factors and forest growth patterns (69 - 72) to the monitoring of relevant wildlife and tick populations (73 - 78). There have also been many studies and reviews assessing the effectiveness of dogs as sentinels for the risk of Lyme disease in humans (79-87). Dogs are appropriate sentinels for several reasons. Not only are they susceptible to infection but because of their increased likelihood of exposure to ticks, they appear to be at greater risk of infection than people. In addition, attached ticks are less likely to be detected on dogs due to the animal's fur, which allows the tick the full 36-48 hours needed for pathogen transmission. Also, dogs mount an immunological response that is easily detected through routine commercial serological assays. Lastly, because they share the same environment and visit the same outdoors areas, dogs give a good indication of the potential exposure to infected ticks for humans.

A 1994 study investigated the use of dogs as sentinels for Lyme disease in humans in Massachusetts (81). The researchers used regression analysis to determine the relationship between incident human cases of Lyme disease and canine seroprevalence of antibodies to Bb. The study showed companion dogs to be highly predictive of the incidence of Lyme disease in humans for the state (81). There were some limitations to this study, however. Seroprevalence estimates for this study were based on the assumption that seropositive dogs received their exposure in the towns where they resided. In addition, details about the specificity of the diagnostic test used were not provided; most diagnostic testing available at the time could not discern between antibodies to Bb or those of related spirochetal bacteria such as Leptospires (the bacterial agent of Leptospirosis). Olson et al. used canine sera solicited from local veterinarians, trappers, animal shelters, and humane societies to determine the utility of dogs as sentinel surveillance for Lyme disease in San Diego County, California (82). Although samples were initially screened for sero-positivity using standard ELISA antibody testing, positive samples were confirmed through the use of split-sample Western blot and indirect fluorescent antibody (IFA) testing (82). No regional clustering of sero-positive animals was detected, which reflected the observed low incidence of human disease expected for the area (82). The more strict diagnostic testing modalities (confirmatory antigen testing) in combination of the use of wild and stray dog populations (more representative of area prevalence) make this study design more accurate for determining the true risk of Lyme disease in the area. Unfortunately, the impact of this study's findings appears less significant given the low prevalence of the area. Aimed at determining whether dogs prone to *Borrelia* sero-positivity pose a threat to their owners, researchers in the Netherlands conducted a seroprevalence study of hunters with hunting dogs and hunters without hunting dogs (83). The study found that although hunting dogs (18%) had slightly higher seroprevalance than non-hunting dogs (17%) and hunters with hunting dogs (15%) had slightly higher seroprevalence than those without hunting dogs (13%), only 12% of the sero-positive hunters had hunting dogs which were also sero-positive. On the basis of these findings the researchers concluded that ownership of dogs with increased risk of infection is not associated with higher risk of human Lyme disease in people, indicating humans and dogs are infected independently and further supporting the use of dogs in sentinel surveillance for Lyme disease (83). Once again, this study was limited by the abilities of the diagnostic tests available at the time. Although methods were used to account for the cross-reactivity of dogs vaccinated against the four most common canine leptospiral serogroups, the lack of confirmatory testing limited the specificity of the results.

In 2004 Duncan et al. employed the use of a new highly sensitive and specific in-house serological test to assess the use of dogs in sentinel surveillance for human Lyme disease in the Southeastern and Mid-Atlantic United States (84). The new test was the SNAP® 3Dx assay. Unlike IFA testing and other antigen-based ELISAs, the SNAP® 3Dx assay detects antibodies to a C6-peptide antigen which is only present when an animal is naturally infected with the Bb pathogen (84). According to the manufacturer, the test sensitivity and specificity were 92% and 100% respectively (84). Comparisons among the sampled states indicated a trend of increasing sero-prevelence with northern movement up the eastern coastline. These findings reflect the distribution of human Lyme disease reported to the CDC, and where interpreted by the authors to further support the use of dogs in sentinel surveillance for the disease (84). In 2011, researchers from the CDC compared U.S. human Lyme disease data with recently published data on national canine Bb seroprevalence data (85). The canine data came from a national database of SNAP® 3Dx and 4Dx assay results that were collected from voluntarily participating veterinary clinics nation-wide. Canine seroprevalence was calculated as the number of sero-positive dogs divided by the total number of tested dogs for each state (88). Using linear regression analysis, the authors showed that state canine sero-prevelence and human Lyme disease incidence were positively correlated (85). In particular, the association was highest among states with very low (<2%) or very high (>5%) canine sero-prevelence. Although the use of the C6 peptide assay may have increased the specificity of these serological findings, there were still limitations to this approach. As with any test, manufacturer claims of accuracy in terms of sensitivity and specificity are limited to the diagnostic methods employed in their laboratory; no test is perfect, and all results should be viewed with healthy skepticism. For example, the sensitivity and specificity is limited the *Borrelia* species and non-*Borrelia* pathogens tested for by the company. Also, the authors of both studies assumed the locations where the test results were obtained represent the true exposure locality. Due to the duration of the canine antibody response (discussed in a later section), dogs with a history of travel may not accurately represent the risk of disease for the test area.

Smith et al. opted for an alternative approach to dogs as sentinels in Lyme disease surveillance, using them as a means of sampling ticks (86). This approach still capitalized on several of the features that make dogs good sentinels for Lyme surveillance, but is not inhibited by some of the issues associated with sero-surveillance studies. The use of dogs to collect ticks permits the researchers to target the sampling to areas where humans are most likely to be exposed because companion dogs tend to share the environment with their human counterparts. Also, because

dogs typically spend more time in these outdoor environments and the time spend tends to be directly in the tick habitat (tall grasses, shrubs, bushes), dogs are more likely to be exposed to tick. Lastly, the dogs' hair coat contributes an increased likelihood of tick bites and sustained attachment, further supporting the dog's usefulness in collecting tick samples. Participating owners were issued a questionnaire that included questions on their pet's travel history and acaricide use (86). The collected ticks were analyzed using PCR technology to determine the presence of the *Borrelia* pathogen. Overall, the United Kingdom-based study found an overall prevalence of *Borrelia* in 2.3% of all ticks sampled, and a 0.5% prevalence of infected ticks on all dogs sampled (86). None of the infected ticks came from dogs that had travelled abroad. The researchers conclude that using dogs to collect ticks is an effective alternative approach to their use as sentinels for Lyme disease surveillance in humans. There are limitations to this approach, however. Most noteworthy, is that veterinary clinics do not have the ability to run PCR on tick samples, making this mode of surveillance less practical. Another relates to the particular animal population surveyed, dogs seen at veterinary clinics. This population of animals is likely on a high plane of preventive medical care, and although analysis did not show a significant association between a history of acaricide use and the presence of ticks in this population, there may be other owner behaviors or pet characteristics that may have influenced the study findings.

Overall, these studies successfully demonstrate the potential to use dogs in sentinel surveillance for Lyme disease in human population. Their review, however, reveals some common limitations. First and foremost, as discussed briefly above, test results based on seroprevalence only indicated the location the animals were tested in and potentially not the true location of exposure. Dogs testing positive may well have been exposed elsewhere or dogs testing negative may spend a significant amount of time in low risk areas. Also, although some studies did collect data pertaining to known canine risk factors for Bb sero-positivity, none analyzed the potential impact of these factors on sero-prevelence estimates by location. For example, locations with a high percent of toy breeds may predict lower prevalence simply because the decreased likelihood of tick exposure in these breeds (less likely to spend time outdoors or recreating in tick habitats). In addition, most of the studies used dogs seen at veterinary clinics, which generally represent animals with a high level of preventive care and subsequently a lower probability of tick exposure. This approach may underestimate the true prevalence in the area and brings question to the basic generalizability of the findings. A more effective surveillance approach might be to include stray and wild dog

populations in these activities. Most studies occurred in areas where Lyme disease is known to be endemic and therefore serological screening for *Borrelia* antibodies is part of the dog's routine health care. However, in regions where human Borreliosis is less common or just now emerging, routine serological screening is less likely to occur. Coupled with the fact that 95% of infected dogs do not produce outwardly visible signs, dogs may not be a very sensitive indicator of Lyme disease in these low prevalent areas (67).

The inability of the reviewed studies to be able to distinguish test location from exposure location is directly related to the methods used to determine canine exposure. For most of the studies exposure was determined using serology and seroprevalence. The limitation of these seroprevalence estimates is their inability to differentiate between recent versus chronic exposures. In 2012, scientists from Cornell University's College of Veterinary Medicine published research detailing the different antibody dynamics that occur throughout the various stages of natural *Borrelia* infection (89). The findings lead to the development of a novel multiplex assay capable of detecting these antibodies, resulting in the ability to discern acute from chronic infections (89). Later that year the assay was used to determine the incidence of Bb in both dogs and horses in New York State (87). The study recommends the use of the assay as a more sensitive tool in recognizing the risk of Lyme disease in human populations (87). As of the data of this publication, however, the assay is not yet commercially available.

References

1. Salman,MD. Surveillance and monitoring systems for animal health programs and disease surveys. In: Salman MD eds. Animal disease surveillance and survey systems, methods and applications. 1st ed. Ames, Iowa: Blackwell Publishing, 2003; 3-13.

2. World Health Organization (WHO) website. World Health Statistics, 2013. Available at: www.who.int/gho/publications/world_health_statistics/en. Accessed Mar 10, 2014.

3. Centers for Disease Control and Prevention. Updated guidelines for evaluating public health surveillance systems: recommendations from the guidelines working group. MMWR 2001; 50(No.RR-13)Atlanta, Georgia, Centers for Disease Control and Prevention.

4. Centers for Disease Control and Prevention: Framework for evaluating public health surveillance systems for early detection of outbreaks. MMWR 2004; 53(No.RR-5)Atlanta, Georgia, Centers for Disease Control and Prevention.

5. Animal and Plant Health Inspection Service (APHIS) website. About APHIS. Available at: www.aphis.usda.gov. Accessed Mar 17, 2014.

6. Christensen J. Application of surveillance and monitoring systems in disease control programs. In: Salman MD, eds. Animal disease surveillance and survey systems, methods and applications. 1st ed. Ames, Iowa: Blackwell Publishing, 2003; 15-34.

7. Institute of Medicine. 2007. "Global infectious disease surveillance and detection: assessing the challenges—finding solutions", Workshop Summary. Washington DC: The National Academies Press.

8. World Organisation for Animal Health (OIE) website. Available at: www.oie.int. Accessed Mar, 15, 2014.

9. World Health Organization (WHO) website. The Global Health Observatory (GHO) portal. Available at: www.who.int/gho. Accessed Mar 15, 2014.

10. Centers for Disease Control and Prevention (CDC) website. Available at: www.cdc.gov/nchs/surveys. Accessed Mar 10 2012.

11. Department of Defense Directive (DODD) 6490.02E. Washington DC Headquarters, 8 Feb 2012.

12. World Health Organization (WHO) website. World Health Statistics, 2013. Available at: www.who.int/gho/publications/world_health_statistics/en. Accessed Mar 10, 2014.

13. DRSi for HPOs Information Briefing by U.S. Army Public Health Command, Army Institute of Public Health, Epidemiology and Disease Surveillance Portfolio, updated April 2012.

14. Personal interview with Zachary McCormic, Epidemiologist, U.S. Army Public Health Command, Army Institute of Public Health, Epidemiology and Disease Surveillance Portfolio, email: zachary.mccormic@us.army.mil.

15. Armed Forces Reportable Medical Events Guidelines and Case Definition. Armed Forces Health Surveillance Center (AFHSC), March 2012.

16. Crump JA, Luby SP, Mintz ED. The global burden of typhoid fever. Bull World Health Organ . 2004; 82(5): 346-353.

17. Majowicz SE, Musto J, Scallan E, et al. The global burden of nontyphoidal salmonella gastroenteritis. Clinical Infectious Diseases 2010; 50:882–889.
18. Josseran L, Nicolau J, Caillère N, et al. Syndromic surveillance based on emergency department activity and crude mortality: two examples. Euro Surveill. 2006; 11(12):225-229.

19. Smith GE, Cooper DL, Loveridge P, et al. A national syndromic surveillance system for England and Wales using calls to a telephone helpline. Euro Surveill. 2006; 11(12):220-224.

20. AnderssonT, Bjelkmar P, Hulth A, et al. Syndromic surveillance for local outbreak detection and awareness: evaluating outbreak signals of acute gastroenteritis in telephone triage, web-based queries and over-the-counter pharmacy sales. Epidemiol. Infect. 2014; 142: 303–313.

21. Fu-Chiang T, Espino JU, Dato VM, et al. Technical description of RODS: a real-time public health surveillance system. J Am Med Inform Assoc. 2003; 10:399–408.

22. Muscatello DJ, Churches T, Kaldor J, et al. An automated, broad-based, near real-time public health surveillance system using presentations to hospital Emergency Departments in New South Wales, Australia. BMC Public Health 2005; 5:141. Available at: www.biomedcentral.com (doi:10.1186/1471-2458-5-141). Accessed April 4, 2014.

23. Wu JT, Ho A, Ma ESK, et al. Estimating Infection Attack Rates and Severity in Real Time during an Influenza Pandemic: Analysis of Serial Cross-Sectional Serologic Surveillance Data. PLoS Med. 2011;8(10). Available at: http://www.plosmedicine.org/article/info (doi:10.1371/journal.pmed.100110). Accessed April, 5 2014.

24. Food and Agriculture Organization of the United Nations website. Available at: www.fao.org. Accessed Mar 10, 2014.

25. Taylor LH, Latham SM, Woolhouse ME. Risk factors for disease emergence. Philos Trans R Soc Lond B Biol Sci. 2001; 356:983-989.

26. Day MJ, Breitschwerdt E, Cleaveland S, et al. Surveillance of zoonotic infectious disease transmitted by small companion animals. Emerg Infect Dis. 2012; 18(12): e1. doi: 10.3201/eid1812.120664

27. Vrbova L, Stephen C, Kasman N, et al. Systematic review of surveillance systems for emerging. Transbound Emerg Dis. 2010; 57:154-161.

28. American Veterinary Medical Association website. Market Research Statistics. U.S. Pet Ownership. Available at: www.avma.org/KB/Resources/Statistics/Pages/default. Accessed April 4, 2014.

29. Madoff L. Cooperation between animal and human health sectors is key to the detection, surveillance, and control of emerging disease, IMED 2007 meeting in Vienna, February 2007. Euro Surveill. 2006; 11(12):E061221.4.

30. ProMED-mail website. Available at: http://www.promedmail.org. Accessed April 1, 2014.

31. Cowen P, Garland T, Hugh-Jones ME, et al. Evaluation of ProMED-mail as an electronic early warning system for emerging animal diseases: 1996 to 2004. JAVMA. 2006; 229(7):1090-1099.

32. Mawudeku A, Michael B. Global Public Health Intelligence Network (GPHIN). Available at: www.mt-archive.info/MTS-2005-Mawudeku.pdf. Accessed April 2, 2014.

33. Mykhalovskiy E, Weir L. The Global Public Health Intelligence Network and early warning outbreak detection: a Canadian contribution to global public health. Can J Public Health. 2006; 97(1):42-44.

34. GOARN website. Global Outbreak Alert and Response Network- GOARN. Available at: www.who.int/csr/outbreaknetwork/en. Accessed April 2, 2014.

35. HealthMap website. Available at: www.healthmap.org/site/about. Accessed April 4, 2014.
36. Argaw D, Mulugeta A, Herrero M, et al. Risk Factors for Visceral Leishmaniasis among Residents and Migrants in Kafta-Humera, Ethiopia. PLoS Negl Trop Dis. 2013; 7(11): e2543. doi:10.1371/journal.pntd.0002543.

37. Van Kerkhove MD, Vandemaele KAH, Shinde V, et al. Risk factors for severe outcomes following 2009 Influenza A (H1N1) infection: a global pooled analysis. PLoS Med. 2011; 8(7): e1001053. doi:10.1371/journal.pmed.1001053.

38. Centers for Disease Control and Prevention website. Active Bacterial Core Surveillance. Background. Available at: http://www.cdc.gov/abcs/overview/background.html. Accessed April 8, 2014.

39. Centers for Disease Control and Prevention website. Active Bacterial Core Surveillance Report, Emerging Infections Program Network, Methicillin-Resistant Staphylococcus aureus, 2011. Available at: www.cdc.gov/abcs/reports-findings/survreports/mrsa11.pdf. Accessed April 8, 2014.

40. Lanciotti RS, Roehrig JT, Deubel V, et al. Origin of the West Nile Virus responsible for an outbreak of encephalitis in the Northeastern United States. Science. 199; 286:2333-23336.

41. Centers for Disease Control and Prevention. Division of Vector-Borne Diseases. West Nile Virus in the United States: Guidelines for Surveillance, Prevention, and Control. 4th Revision, June 14, 2013. Available at: www.cdc.gov/westnile/resources/pdfs/wnvGuidelines.pdf. Accessed 9 April 2014.

42. Fischer M. ArboNET: National surveillance system for arboviraldiseases in the United States (powerpoint). Arboviral Diseases Branch Centers for Disease Control and Prevention Fort Collins, Colorado June 25, 2008. Available at: www.iom.edu/~/media/Files/Activity%20Files/Global/ZoonoticDisease/Presentation4Fischer.pdf. Accessed: April 9, 2014.

43. The FAO website. The Global Rinderpest Eradication Programme Progress report on rinderpest eradication: Success stories and actions leading to the June 2011 Global Declaration. Available at: www.fao.org/ag/againfo/resources/documents/AH/GREP_flyer.pdf. Accessed April 9, 2014.

44. FAO website. GREPIN ACTION. Required steps for recognition of freedom from rinderpest Available at: www.fao.org/ag/againfo/programmes/documents/grep/country_Action.pdf. Accessed April 9, 2014.

45. James AD. Guide to epidemiological surveillance for rinderpest. Rev. sci. tech. Off. int. Epiz., 199;17 (3):796-809. www.oie.int/doc/ged/D9478.PDF. Accessed April 10, 2014.

46. Nadrowski K, Ratcliffe S, Bonisch G, et al. Harmonizing, annotating and sharing data in biodiversity-ecosystem functioning research. Methods in Ecology and Evolution 2013; 4:201-205.

47. Hartwell SW, Deng X, Fischer W, et al. Harmonizing databases? Developing a quasi-experimental design to evaluate a public health re-entry program. Evaluation and Program Planning 2012; 35: 461-472.

48. Liolios K, Schriml L, Hirschman L, et al. The Metadata Coverage Index (MCI): A standardized metric for quantifying database metadata richness. Standards in Genomic Sciences 2012; 6:444-453.

49. Wu S, Li J. Merging Results from Overlapping Databases in Distributed Information Retrieval. In Proceedings. 21st Euromicro International Conference on Parallel, Distributed, and Network-Based Proceedings 2013;102-107.

50. Hook EB, Regal RR. Capture-Recapture Methods in Epidemiology: Methods and Limitations. Epidemiologic Reviews 1995; 17:243-64.

51. Abeni, DD, Brancato, G, Perucci, CA. Capture-recapture to estimate the size of the population with Human Immunodeficiency Virus Type 1 infection." Epidemiology 1994; 5: 410-414.

52. Navy and Marine Corps Public Health Center. Capture-Recapture Methods: Technical Report. September 19, 2007.
53. Eck K. In data we trust? A comparision of USDP GED an ACLED conflict events datasets. Cooperation and Conflict. 2012; 47:124-141.

54. Hedden S, Gfroerer J, Barker P, et al. Comparison of NSDUH mental health data and methods with other data sources. Center for Behavioral Health Statistics and Quality, Substance Abuse and Mental Health Services Administration (SAMHSA). Available at: www. samhsa. gov/data/NSDUH/2k12Findings/CBHS QDataReviewC2MentalHealth2012.htm. Accessed May 10, 2013.

55. One Health Initiative website. Available at www.onehealthinitiative.com. Accessed April 10, 2014.

56. American Veterinary Medical Association One Health: A New Professional Imperative, One Health Initial Task Force: Final Report, July 15, 2008. Available at www.avma.org/KB/Resources/Reports/Pages/One-Health. Accessed April 10, 2014.

57. McCluskey B. Use of sentinel herds in monitoring and surveillance systems. In: Salman MD eds. Animal disease surveillance and survey systems, methods and applications. 1st ed. Ames, Iowa: Blackwell Publishing, 2003; 119-123.

58. Halliday JEB, Meredith AL, Knobe DL, et al. A framework for evaluating animals as sentinels for infectious disease surveillance. Journal of The Royal Society Interface. 2007; 4(16): 973-984.

59. Lipton BA, Hopkins SG, Koehler JE, et al. A survey of veterinarian involvement in zoonotic disease prevention practices. JAVMA, 2008; 233(8):1242-1249.

60. Rabinowitz, P, Gordon ., Chudnov,D, et al. Animals as sentinels of bioterrorism agents. Emerging infectious diseases. 2006; 12(4):647.

61. Lembo T, Hampson K, Auty H, et al. Serologic Surveillance of Anthrax in the Serengeti Ecosystem,Tanzania, 1996–2009. Emerging Infectious Diseases. 2011; 17(3): 387-394.

62. Overgaauw PA, van Zutphen L, Hoek D, et al. Zoonotic parasites in fecal samples and fur from dogs and cats in The Netherlands. Veterinary parasitology, 2009; 163(1):115-122.

63. Schurer JM, Hill JE, Fernando C, et al. Sentinel surveillance for zoonotic parasites in companion animals in indigenous communities of Saskatchewan. Am. J. Trop. Med. Hyg. 2012; 87(3):495–498.

64. Kuiken T, Rimmelzwaan G, van Riel D, et al. Avian H5N1 influenza in cats. Science, 2004; 306(5694):241-241.

65. Rimmelzwaan, GF. Influenza A virus (H5N1) infection in cats causes systemic disease with potential novel routes of virus spread within and between hosts. The American journal of pathology. 2006; 168(1): 176-183.

66. Center for Disease Control and Prevention website. Lyme Disease Home Page. Available at: www.cdc.gov/lyme/index/html. Accessed: Dec 17, 2013.

67. Greene CE, Straubinger RK, Levy SA. Borreliosis. In: Infectious Diseases of the Dog and Cat. 4th ed. St. Louis, Missouri: Elsevier Sauders Co, 2012; 447-465.

68. Zhang, X, Meltzer MI, Pena CA, et al. Economic Impact of Lyme Disease. Emerging Infectious Diseases. 2006; 12(4):653-660.

69. Gatewood AG, Liebman, KA, Vourc'h, G, et al. Climate and tick seasonality are predictors of Borrelia burgdorferi genotype distribution. Applied and environmental microbiology. 2009; 75(8):2476-2483.

70. Subak S. Effects of climate on variability in Lyme disease incidence in the northeastern United States. American Journal of Epidemiology. 2003; 157(6):531-538.
71. McCabe GJ, Bunnell JE. Precipitation and the occurrence of Lyme disease in the northeastern United States. Vector-Borne and Zoonotic Diseases. 2004; 4(2):143-148.

72. Ostfeld RS, Schauber EM, Canham CD, et al. Effects of acorn production and mouse abundance on abundance and Borrelia burgdorferi infection prevalence of nymphal *Ixodes scapularis* ticks. Vector Borne and Zoonotic Diseases. 2001; 1(1):55-63.

73. Richter D, Matuschka FR. Differential risk for Lyme disease along hiking trail, Germany. Emerging Infectious Diseases. 2011; 17(9)1;1704-1706.

74. Diuk-Wasser, MA, Gatewood Hoen A, Cislo P, et al. Human risk of infection with Borrelia burgdorferi, the Lyme disease agent, in eastern United States. American Journal of Tropical Medicine and Hygiene. 2012; 86(2): 320.

75. Rand, P. W., Lacombe, E. H., Dearborn, R., et al. Passive surveillance in Maine, an area emergent for tick-borne diseases. Journal of Medical Entomology. 2007; 44(6):1118-1129.

76. Levi T, Kilpatrick AM, Mangel M,et al. Deer, predators, and the emergence of Lyme disease. Proceedings of the National Academy of Sciences. 2012; 109(27):10942-10947.

77. Bhide MR, Travnicek M, Levkutova M, et al. Sensitivity of Borrelia genospecies to serum complement from different animals and human: a host-pathogen relationship. Fems Immunology and Medical Microbiology 2005; 43(2):165-172.

78. Gill JS, McLean RG, Shriner RB, et al. Serologic surveillance for the Lyme disease spirochete, Borrelia burgdorferi, in Minnesota by using white-tailed deer as sentinel animals. Journal of Clinical Microbiology. 1994; 32(2): 444-451.

79. Little SE, Heise SR, Blagburn BL,et al. Lyme borreliosis in dogs and humans in the USA. Trends in Parasitology, 2010; 26(4):213-218.

80. Eng TR, Wilson ML, Spielman A et al. Greater risk of Borellia burgdorferi infection in dogs than in people. J of Inf Disease 1988; 158:1410-1411.

81. Lindenmayer JM, Marshall D, Onderdonk AB. Dogs as sentinels for Lyme Disease in Massachusetts. Am J of Public Health 1991; 81:1448-1455.

82. Olson PE, Kallen AJ, Bjorneby JM, et al. Canines as sentinels for Lyme Disease in San Diego County, California. J VET Diagn Invest 2000; 12:126-129.

83. Goossens HAT, van den Bogaard AE, Nohlmans MKE. Dogs as sentinels for human Lyme Borreliosis in the Netherlands. J Clin Microbiol 2001; 39:844-848.

84. Duncan AW, Correa MT, Levine JF, et al. The dog as a sentinel for human infection: prevalence of Borrelia burgdorferi C6 antibodies in dogs from Southeastern and Mid-Atlantic states. Vector-Borne and Zoonotic Diseases 2004; 4:221-229.

85. Mead, P, Goel R., Kugeler K. Canine serology as adjunct to human Lyme disease surveillance. Emerging infectious diseases, 2011;17(9):1710-1712.

86. Smith FD, Ballantyne R., Morgan ER., et al. Estimating Lyme disease risk using pet dogs as sentinels. Comparative immunology, microbiology and infectious diseases. 2012; 35(2):163-167.

87. Wagner B, Erb H N. Dogs and horses with antibodies to outer-surface protein C as on-time sentinels for ticks infected with Borrelia burgdorferi in New York State in 2011. Preventive Veterinary Medicine, 2012; 107(3):275-279.
88. Bowman D, Little SE, Lorentzen L, et al. Prevalence and geographic distribution of Dirofilaria immitis, Borrelia burgdorferi, Ehrlichia canis, and Anaplasma phagocytophilum in dogs in the United States: Results of a national clinic-based serologic survey. Veterinary Parasitology 2009;160:138-148.

89. Wagner B, Freer H, Rollins A, et al. Antibodies to Borrelia burgdorferi OspA, OspC, OspF, and C6 antigens as markers for early and late infection in dogs. Clinical and Vaccine Immunology, 2012; 19(4):527-535.

Chapter 3: Review of Data Systems used in Public Health Surveillance

Introduction

Public Health Surveillance

According to the World Health Organization (WHO), public health surveillance is the continuous, systematic collection, analysis and interpretation of health-related data needed for the planning, implementation, and evaluation of public health practice. Such surveillance can serve as an early warning system for impending public health emergencies, document the impact of an intervention, track progress towards specified goals, and monitor and clarify the epidemiology of health problems. The data from public health surveillance inform public health policy and strategies (1). Inherent in this definition is the implication that action will be taken based on the collected public health data.

Most developed countries have the ability to detect and diagnose human, animal, and plant disease, but many developing countries- where the majority of the global population resides- do not have the resources or infrastructure to support such activities. In these regions of the world surveillance activities are often supported by outside agencies and organizations such as the World Health Organization (WHO), the Food and Agriculture Organization of the United Nations (FAO), and the World Organization for Animal Health (OIE). Within the United States, these activities fall under the Centers for Disease Control and Prevention (CDC), state and local health departments, state agricultural departments, and the United States Department of Agriculture Animal and Plant Health Inspection Service (USDA APHIS). Although, each of these organizations is unique in its specific purpose and situation, their long history of existence offers an excellent opportunity for learning which factors lead to both successes and failures in public health surveillance efforts. Because each of these organizations is unique in their specific purpose and objectives, a review of each organization's surveillance systems offers an excellent opportunity to delineate the strengths and limitations of each approach.

Public Health Surveillance in the Military

Public health surveillance in the military also involves the continuous systematic collection, analysis, interpretation, and dissemination of population based health-related events data. Much like other surveillance programs, the data are used to reduce morbidity and mortality, to estimate the distribution, trends, and risks associated with significant medical events, and to aid in the development and assessment of policy and resource allocation (2). In addition, health-related event data in the military are used by commanders and public health officials to enable early intervention and control strategies in order to prevent adverse military health outcomes and ensure overall strategic mission accomplishment (2). Currently the military has several human medical data systems available for use in public health surveillance. Because each system has different capabilities and attributes, an understanding of each one is critical for military leaders to make appropriate decisions based on the data.

Describing Data Systems Used in Public Health Surveillance

Surveillance systems can vary greatly depending on purpose, priority, infrastructure and resources. Successful surveillance programs undergo periodic evaluations to ensure that problems of public health importance are being monitored efficiently and effectively. In 2001 the Centers for Disease Control and Prevention (CDC) published guidelines for evaluating public health surveillance programs (3). These guidelines outline the critical components of successful systems, providing a standardized approach to describe and subsequently compare surveillance systems (3).

This study will apply these guidelines to systematically describe the five military medical data systems most commonly used for public health surveillance purposes, as well as briefly discus the capabilities and attributes of data systems used in selected national and international public health surveillance programs. In combination, this study aims to produce a document that can be used by military commanders and public health officials as a tool to understand differences in military medical data systems and subsequently inform optimal decision making and contribute to mission accomplishment.

Background

A Systematic Approach to Describing Data Systems Used in Public Health Surveillance

The following discussion is based on the CDC Updated guidelines for evaluating public health surveillance systems (1). Unless otherwise specified, this document serves as the primary reference.

When evaluating public health systems, the focus should be on how well the system operates to meet the stated purpose and objectives of the given surveillance program. In general, the evaluation should involve an assessment of system attributes such as simplicity, flexibility, data quality, acceptability, sensitivity, positive predictive value, representativeness, timeliness, and stability. It should also attempt to incorporate informatics concerns such as hardware/software comparability, user interface, standard data format and coding, appropriate quality checks, and adherence to confidentiality and security standards. As public health surveillance systems vary in methods, scope, purpose, and objectives, attributes that are important to one system might be less so or completely irrelevant to another. For this reason, it may not be possible to assess each attribute in each system. Below is a discussion of the specific system attributes used in this review. The template used to collect the information from each military medical data system can be found in **Appendix A-3**.

As stated earlier, the specific attributes and capabilities of a disease reporting/surveillance system are directly related to the stated purpose of that system. For this reason it is imperative that any standardized system description starts with a listing of the system's **purpose**, **objectives**, and **justification**. The purpose is why the system exists, whereas its objectives state how the data will be used for public health action. The justification explains how this system alone uniquely supports the stated purpose and specific objectives.

Next the review should determine how the system defines a health-related event, essentially, what the system uses as an **operating case definition**. A system's case definition is dependent on the type of data the system accesses. Included in this description should be the system's inclusion/exclusion criteria, whether the system is one based on syndromes, laboratory reports, International Classification of Disease (ICD) codes, or epidemiological reports. In

addition, if possible, an understanding of the reported level of certainty (probable, suspect, confirmed) should be included.

Equally important is an understanding of the **system components** such as population included, data sources, and level of integration with other systems. These three components are critical to understanding the representativeness of the data; without a clear understanding of who the data represents it is not possible to implement a successful public health policy or strategy. An understanding of the system's purpose aids in understanding the **target population**. Population variables may include categories such as gender, age, race, and occupation. In the military, it also generally includes (but is not limited to) service component, branch of service, and rank. **Data source** refers to where the data in the system is collected from, both medical and other (population data). This should include automated and non-automated record sources. Data type, on the other hand, refers to the category of data collected as discussed above (e.g. syndromic, laboratory based, medical, epidemiological reports). A system's **level of integration** with other data sources not only aids in identifying potential gaps and overlaps in data, but it also addresses the interoperability of the systems, and the ability of a system to be upgraded and its ability to respond to growing memory demands and queries. In addition to these rather fixed components are some that are more fluid, such as the system's ability to be managed (edited) and manipulated (analyzed, reported, disseminated). An understanding of these system components is also critical to understating the overall quality of the data and therefore the applicability of conclusions made.

The two remaining descriptive attributes relate to the resources (both human and informatics) required for the system to be operational. Criteria used to evaluate this category include **accessibility** to the data, **user-friendliness** of the system, and overall **efficiency** of the data system. Closely related to accessibility is **stability**. The stability of the system is affected by firewall issues, system downtimes, and manpower constraints, all of which directly impact the overall reliability and availability of data generated by that system.

These last three descriptive attributes are also related to the most critical aspect of health surveillance system evaluations; the assessment of the **system's performance** in the field. A perfectly designed system is still a failure if there are extenuating circumstances and/or issues that prevents its use. There are many factors that come into play

when evaluating a system's performance. **Simplicity** refers to the structure of the system and its ease of operation; specific categories include methods of data collection (number and type of reporting sources), level of integration with other systems, methods for data analysis and data dissemination, time required to maintain and update the system, and human resource/staffing requirements. Overly complicated systems may be prone to system failures or may have extremely high human resource requirements. **Flexibility** is important as it represents the ability of a system to accommodate changing query requests and adapt and evolve to meet new requirements. These two factors (simplicity and flexibility) contribute to the overall **acceptability** of the system, or the willingness of the user (data entry through analyst) to actually use the system.

In terms of the end-user or analysts, there are several factors that can be used to assess a system's performance. **Utility** refers to the level of usefulness of the retrieved data; the appropriateness of the estimators provided, and the ability to accurately detect disease trends. In order to generate useful information, the system should be able to stratify the data **temporally**, **spatially** and **categorically**. In addition, access to accurate and representative **denominator data** (from the same source, comparable across categories, representative over time) is important. **Timeliness** of the information refers to the time between the occurrence of the health related event and its appearance in a functional surveillance system. Both the lag time between the detection of the medical event and the time required to generate a system record impacts a system's timeliness.

The topic of **data quality** is rather complicated and depends quite a bit on the purpose of the surveillance activity. There are many subcategories that may be used to evaluate a system's data quality. **Mode of data** entry (automated vs. manual) impacts correctness of the data, and in combination with an understanding of the c**ompleteness** and **validity** of the records contributes to the overall **accuracy** of the data. Case definition and diagnostic criteria impact the **sensitivity** and **specificity** of the results, whereas an understanding of how the data was collected (population sampled, active vs. passive surveillance, voluntary vs. directed participation) pertains more to the **representativeness** of the reported results.

A comparative review of the performance of the five military human medical data systems most commonly used for public health surveillance was conducted in a separate study (Chapter 4). The remainder of this chapter will apply

the terms and definitions discussed above to a broad overview of existing global zoonotic disease programs (Section I), then a detailed description of the five human medical data systems most commonly used for public health surveillance in the military (Section II). The goal of this review is to create a resource for military leaders and public health officials to use when determining which data system is best suited for certain tasks or analyses.

Section I: US and Global Data Systems Used in Public Health Surveillance

National and Global Public Health Surveillance

Early detection is essential to the control of emerging, reemerging, and novel infectious disease, whether naturally occurring or intentionally introduced (4). Containing the spread of such diseases in a deeply interconnected world requires active vigilance for changes in disease trends that are the cornerstone of public health surveillance. Most developed countries have the ability to detect and diagnose human, animal, and plant disease, but many developing countries- where the majority of the global population resides- do not have the resources or infrastructure to support such activities (4). A major challenge to global disease surveillance and detection is the disparity in existing local and regional surveillance capabilities. These inequalities may range from differences in the ability to make diagnoses at the patient level to differences in reporting, data capture methods, and protocols. This section briefly reviews the data systems used in public health surveillance by four different organizations; three at a global level and one nationally. Although each of the organizations has a unique mission, much can be learned by studying their overall approach to surveillance.

The aim of this section is to apply knowledge of existing public health surveillance systems to the betterment of military public health surveillance programs.

Methods

Data Collection

The list of data systems reviewed in this section are not meant to be all inclusive, rather it was limited to the data systems used by some of the best known international and national public health organizations: The World Health Organization (WHO), the Food and Agriculture Organization (FAO), the World Health Organization for Animal Health (OIE), and the United States Department of Agriculture (USDA). In order to be included in the review, information on each organization's data systems had to be freely available for review on public domains.

Descriptions of each public health surveillance system were compiled using public domain material from organizational websites and when available pertinent documents from the literature. The limitations of this approach resulted in gaps in information at times, most notably in details pertaining to data entry, quality assurance, and data analysis.

GHO: WHO's Data and Statistics Portal

I. Background

The World Health Organization (WHO) is the directing and coordinating authority for health within the United Nations (UN) system which currently includes 194 Member States (1). It is responsible for providing leadership on global health matters, shaping the health research agenda, setting norms and standards, articulating evidence-based policy options, providing technical support to countries, and monitoring and assessing health trends. The WHO has six main global objectives which they refer to as agenda points: 1) promoting health development, 2) fostering health security, 4) harnessing research, information and evidence, 5) enhancing partnerships with UN agencies and other international organizations, donors, civil society and the private sector, and 6) improving health performance at country, regional and international levels (1).

II. Purpose, Justification, Objectives

The Global Health Observatory (GHO) is the WHO's portal to health-related statistics from around the world (5). The aim of the GHO portal is to provide easy access to country data and statistics with a focus on comparable estimates; enabling the WHO to analyze and monitor global, regional, and country trends. The GHO database provides access to an interactive repository of health statistics. Users are able to display data for selected indicators and health topics at the country and regional level (5).

III. Operations

A. Operating Case Definitions

The data within the GHO support global health priorities such as the health-related Millennium Development Goals, mortality and burden of disease, health systems, environmental health, non-communicable diseases, infectious diseases, health equity and violence and injuries (5). The system is based on “indicators” that are not only relevant to global public health, but are also readily available, of good data quality, and able to produce reliable and comparable health estimates across member states (6). Taken together, these indicators provide a comprehensive summary of the current status of national health and health systems in the following nine areas: life expectancy and mortality, cause-specific mortality and morbidity, selected infectious diseases, health service coverage, risk factors, health systems, health expenditure, health inequities, demographic and socioeconomic statistics (5).

B. Target Population

The data in the GHO is limited to the 194 Member States (5). Estimates based on this data may be used in discussion of global health issues, but are not intended to be the definitive representation of global health. Also, they should not be regarded as the nationally endorsed statistics for these Member States, which may use alternative methodologies to derive their estimates.

C. Data Sources and Integration with Other Systems

The WHO Department of Health Statistics and Information Systems compiles reports using publications and databases produced and maintained by WHO technical programs and regional offices (7). A number of demographic and socioeconomic statistics are also derived from databases maintained by a range of other organizations including: the United Nations International Telecommunication Union (ITU), the United Nations Department of Economic and Social Affairs (UNDESA), the United Nations Educational, Scientific and Cultural Organization (UNESCO), the United Nations Children's Fund (UNICEF) and the World Bank (7).

D. Data Management and Quality Assurance

Country data may differ in terms of definitions, data-collection methods, population coverage and estimation methods used. Estimates from GHO come from the Indicator and Measurement Registry (IMR) (6). The IMR is a central source of indicator and measurement definitions (text and computer-readable), metadata, and translations. It also ensures consistency across statistical domains, promotes interoperability of indicator exchange formats, provides internet access to indicator definitions, and incorporates other appropriate international standards required to support the compilation of metadata (6). Ideally data comes from the national level; the IMR includes 31 data source codes ranging from civil registries and government statistics to population-based surveys and health professional education institution reports. WHO defined data categories for health indicators and measurements include demographics, health equity monitor, health service coverage, health systems resources, morbidity, mortality, risk factors, and socioeconomics. Data types may include categorical, count, financial, percent, rate, ratio, score, and statistics (6).

The estimates from GHO are derived from multiple sources, depending on each indicator and on the availability and quality of data (5, 6). In many countries, statistical and health information systems are weak and the underlying empirical data may not be available or may be of poor quality. The use of standardized categories and methods enhances the cross-national comparability of the data, but may also lead to some distortion of estimates. For the sake of the reports generated by the WHO, statistical modelling and other techniques are used in an attempt to fill

data gaps; all estimates are cleared only following consultation with Member States. However, the estimates should not be regarded as the nationally endorsed statistics which may have been derived using alternative methodologies (6, 7).

E. Data Analysis

The data in GHO are arranged in a format that is intended to provide easy access to country data and statistics with a focus on comparable estimates to facilitate the monitoring of spatial trends (5). The GHO has "theme pages" that cover global health priorities such as the health-related Millennium Development Goals (MDGs), mortality and burden of disease, health systems, environmental health, non-communicable diseases, infectious diseases, health equity and violence and injuries (1, 6). The theme pages present: highlights showing the global situation and trends using regularly updated core indicators; data views customized for each theme, including country profiles and a map gallery; publications relevant to the theme; and links to relevant web pages within WHO and elsewhere (1,5,6). The GHO country data includes all country statistics and health profiles that are available within WHO (5, 6).

F. Data Dissemination and Reporting Options

The GHO database provides access to an interactive repository of health statistics (5). Users are able to display data for selected indicators, health topics, countries and regions, and download customized tables in Excel format. In addition, the GHO issues analytical reports on priority health issues (5)

User Generated Reports

The GHO database is freely accessible to anyone with internet access (5). User-generated reports are pull-down driven and may be searched by topic, country, or indicator (6). Report properties may include data type and unit of measure, data sources, and method of measurement, estimation, or aggregation (potential for disaggregation). There are also free text fields in the GHO generated reports where analysts can discuss the expected frequency of data

dissemination and collection, limitations of the presented data, and make recommendations for external links or points of contact (6).

World Health Statistics

The World Health Statistics annual publication, compiles statistics for key health indicators monitored by the WHO (7). The analytical report addresses cross-cutting topics such as women and health and a summary of the progress made towards achieving the health-related MDGs and associated targets for its 194 Member States (7). In 2013 it also included summaries on the topics of reducing the health gaps between the world's most-advantaged and least-advantaged countries, and on current trends in Official Development Assistance (ODA) for health. Every effort has been made to ensure the best use of country reported data – adjusted where necessary to deal with missing values, to correct for known biases, and to maximize the comparability of the statistics across countries and over time (7).

EMPRES: FAO's Emergency Prevention System

I. Background

In 1943 the United States, along with forty-four other governments, committed themselves to founding a permanent organization focused on food and agriculture security (8). Two years later, the first conference of the Food and Agriculture Organization of the United Nations (FAO) was held establishing the FAO as a specialized United Nations agency (8). Achieving food security for all is the focus of FAO's efforts; specifically, ensuring people have regular access to sufficient high-quality food to lead active, healthy lives (8). In order to accomplish this, the FAO has three main goals 1) the eradication of hunger, food insecurity and malnutrition, 2) the elimination of poverty and the driving forward of economic and social progress for all and, 3) the sustainable management and utilization of natural resources, including land, water, air, climate and genetic resources for the benefit of present and future generations (8, 9). The organization's specific objectives are to help eliminate hunger, food insecurity and malnutrition; make agriculture, forestry and fisheries more productive and sustainable; reduce rural poverty; enable inclusive and efficient agricultural and food systems; and increase the resilience of livelihoods to disasters (8). The

FAO believes that a key component of fighting hunger, malnutrition, and poverty is protection against animal and plant diseases and pests, and food safety threats (8,9).

The FAO Emergency Prevention System (EMPRES) has the mandate to address prevention and early warning across the entire food chain (9). This is done through the following three systems: EMPRES Animal, EMPRES Plant Protection, and EMPRES Food Safety (9, 10). The remainder of this discussion will only pertain to EMPRES's utility in Animal Health; therefore for the remainder of the manuscript the use of the term EMPRES pertains to the EMPRES Global Animal Disease Information System.

II. Purpose, Justification, Objectives

EMPRES Global Animal Disease Information System (EMPRES-i) is a web-based application that aims to clarify animal disease events internationally by facilitating the organization of, and access to regional and global animal disease information for veterinary services worldwide (9). The purpose of EMPRES-i is to capture and consolidate timely and reliable disease information in order to enhance early warning and response to transboundary and high impact animal diseases, including emergent zoonoses. The data in EMPRES-i are utilized to support prevention activities, improve management, and encourage progressive approaches to control practices (9, 10).

III. Operations

A. Operating Case Definitions

The extreme diversity among the sources (from non-medical media sources to diagnostic reference laboratories), prohibits the application of strict operating case definitions. Instead, the system offers a highly sensitive (but less specific) mode of disease status awareness.

B. Target Population

The EMPRES-i aims to represent global data on emerging zoonotic and transboundary animal disease. The data collected in the system can be linked to the specific data source and location of the disease event, and therefore can be utilized locally or globally.

C. Data Sources and Integration with Other Systems

The EMPRES-i is an internet platform that facilitates the sharing of animal health data through the use of a secured internet-based database (using Oracle) (9, 10). Data sources include both official and unofficial sources of information (9, 10). Sources include country or regional project reports, filed mission reports, partner Non-Governmental Organizations (NGOs), cooperating institutions, government Ministries of Agriculture and Health, FAO in country representations or other United Nations parties, public domains, the media and web-based health surveillance systems (10, 11). Users from these divers sources are granted differential access (guest, manager, administrator) with varying levels of privileges (view, edit, analyze) (11).

D. Data Management and Quality Assurance

Data is added to EMPRES-i by designated users for selected data sources dependent on user access and privilege levels (10). FAO officials use both official and unofficial sources such as in-country assistance projects and personal contacts with NGOs and other institutions in order to verify data (10). The data in EMPRES-i is intended to enhance early warning efforts, and as such is intended to be highly sensitive. Users can specifically query the database for details on disease status (confirmed or not), diagnostic criteria used, and may even contact public health officials or laboratories involved in the investigation if further details are required or desired (9, 10, 11).

E. Data Analysis

EMPRES-i provides up-to-date information on the global animal disease distribution and current threats at the national, regional and global level. Users can navigate through the system to perform specific searches of diseases or outbreaks by locality, period, or disease (see data dissemination and reporting options below).

F. Data Dissemination and Reporting Options

Depending on user's assigned privileges, different levels of data details and confidentiality may be accessed. In addition to disease data, EMPRES-i provides access to publications, manuals, and other resources such as details of chief veterinary officers (CVOs) and FAO/OIE (World Organization for Animal Health) reference laboratories (8, 9, 10). The system is constantly undergoing improvements; currently the public can access the following features:

Early Warning Messages

Consolidated information in EMPRES-i is used to generate and disseminate early warning messages (9, 10).

Disease Event Database

Under the disease event tab of the EMPRES-i website, users can access and retrieve information on wild or domestic animal disease outbreaks/cases throughout the world according to criteria defined by the user (10). Details include, but are not limited to disease (status, serotype), date (observed and reported), animal data (species, number of cases, number at risk, deaths, destroyed, slaughtered), laboratory (disease tested, test performed, results, date, reference laboratory), location (region, latitude/longitude), source (media, agency, etc.) (10). Searches can include both current and historical disease data. The data can be exported to PDF, CSV, or Excel formats for further analysis by the user (10).

Advanced Search/Mapping/Graphing Tools

EMPRES-i can be used to create graphical representations of specific outbreaks/cases by time or by location (9, 10). Maps can be tailored to include layers such as livestock populations, biophysical features, socioeconomics, animal health, trade, etc. (8, 9, 10). Maps and graphs can be exported into many different formats (10).

My EMPRES-i

The 'My EMPRES-i' tab enables users to log in and access a personalized page designed by the user (10). Display filters include, but are not limited to: disease, period, and geographical region of interest (10). Users can subscribe to various newsletters and bulletins through this tab as well.

WAHIS: OIE's Web Interface for Animal Health Data

I. Background

The need to fight animal diseases at global level was formally identified in 1924 and led to the creation of the Office International des Epizooties (OIE) (12). In May 1994 the World Assembly of Delegates of the OIE requested the Foot and Mouth Disease and Other Epizootics Commission (now called the Scientific Commission for Animal Diseases) to develop a procedure for the official recognition by the OIE of the foot and mouth disease free status of Member Countries (12). The procedure has since been expanded to include rinderpest, contagious bovine pleuropneumonia and bovine spongiform encephalopathy. In 1998, the official agreement between the World Trade Organization (WTO) and the OIE further confirmed the OIE's mandate to recognize disease and pest-free areas for trade purposes (12). By acquiring and maintaining its official status, a country demonstrates transparency, thereby gaining the trust of its trade partners, neighboring countries and the international community as a whole. In May 2003 the Office became the World Organization for Animal Health but kept its historical acronym OIE (12).

The OIE is the intergovernmental organization responsible for improving animal health worldwide and is recognized as a reference organization by the World Trade Organization (WTO) (12). The OIE has six stated objectives: 1) ensure transparency in the global animal disease situation, 2) collect, analyze and disseminate veterinary scientific information, 3) encourage international solidarity in the control of animal diseases, 4) safeguard world trade by publishing health standards for international trade in animals and animal products, 5) improve the legal framework and resources of national Veterinary Services, and 6) provide a better guarantee of food of animal origin and promote animal welfare through a science-based approach (12).

II. Purpose, Justification, Objectives (WAHIS)

The World Animal Health Information System (WAHIS) is the internet-based system that the OIE uses to collect and publish their animal disease data (12). The purpose of the system is to support transparency in the global animal disease situation by providing timely data sharing with the international community. The system consists of two components: an early warning system to inform the international community, by means of "alert messages", of relevant epidemiological events that occurred in OIE Member Countries, and a monitoring system that tracks OIE listed diseases (presence or absence) over time (12, 13). In addition, WAHIS plays a critical role in almost all of the stated OIE objectives, specifically those pertaining to data collection, analysis, and dissemination. Member Countries that conform to the obligations of the OIE have the trust of its trade partners, neighboring countries and the international community.

III. Operations

A. Operating Case Definitions

The diseases monitored in the WAHIS are limited to those included on the OIE list of notifiable terrestrial and aquatic animal diseases. The list was generated to support Member Countries' efforts to prevent the transboundary spread of important animal diseases, including zoonoses, through transparent and consistent reporting (12, 13). The list is reviewed annually. The OIE - Terrestrial Animal Health Code provides Member Countries with guidelines for

disease detection. The document also details the requirements for notification (through WAHIS) or, if not possible, by fax or e-mail) and discusses methods of prevention and control (13).

B. Target Population

The target population of the OIE is the international community as a whole; however, the data captured in WAHIS only represents animal disease data from Member Countries. In 2013 the OIE had a total of 178 Member Countries and maintains permanent relations with 45 other international and regional organizations and has Regional and sub-regional Offices on every continent (12).

C. Data Sources and Integration with Other Systems

One of the formal obligations of OIE Member Countries is the submission of information on the relevant animal disease situation – including on zoonoses present in their territory - in the most timely and transparent way. The World Animal Health Information System (WAHIS) is the internet-based computer system used by Member Countries to submit the animal disease data (12). Access to this secure site is only available to authorized users, namely the Delegates of OIE Member Countries and their authorized representatives (12). The OIE also uses the system to disseminate the information to other countries, permitting affected regions to implement necessary preventive actions. Information is sent out immediately or periodically depending on the seriousness of the disease. This objective applies to disease occurrences both naturally occurring and deliberately caused.

D. Data Management and Quality Assurance

Member Countries submit animal health data through WAHIS to the OIE where it is verified by OIE officials. Once verified by the OIE the data makes it to WAHID (World Animal Health Information Database) for public viewing (12, 13).

Epidemiological Event Data

Whenever an important epidemiological event occurs in a Member Country, the Member Country must inform the OIE by sending an Immediate Notification which includes the reason for the notification, the name of the disease, the affected species, the geographical area affected, the control measures applied and any laboratory tests carried out or in progress (12). Once they have been received, verified and validated by the OIE, the immediate notifications are published in the OIE's three official working languages (English, French and Spanish) under the heading "Alerts" and sent to everyone on the OIE-Info Distribution List (12). This list is open to Member Country Delegates, the OIE Reference Laboratories and Collaborating Centers, and international and regional organizations. In addition, institutions and individuals may directly receive the information if they apply for a subscription. Members must then continue to send weekly Follow-up Reports so that the event can be monitored as it evolves (12, 13). In all cases, the country must submit a final report to notify that the event has been resolved or that the disease has become endemic. In both cases, the country will continue to submit information in its biannual reports if the disease is on the OIE List (12).

Recurring Report Submission

Biannual Reports provide information on the presence or absence of diseases on the OIE List and the prevention and control measures applied. For diseases reported as being present in a country/territory during a given six-month period, the country/territory in question must provide quantitative data on the number of outbreaks, susceptible animals, cases, deaths, animals destroyed and animals vaccinated (12, 13). For diseases that are present and are notifiable in the country, the OIE recommends that countries provide quantitative data by month and by first administrative division. Countries/territories that so wish can enter their data in WAHIS each month during a given six-month period (i.e. without waiting until the end of the six-month period), thereby providing the international community with the most recent information on the diseases that are present and which Member Countries consider are the most important. The selection of which option will depend on the national surveillance and monitoring systems in the country/territory in question and the type of information generated by these systems. The two biannual reports of a given year are combined as part of the Annual Report for OIE-listed diseases (12, 13).

E. Data Analysis

The WAHIS is a secured web site; access is restricted to authorized users. The information in WAHIS is limited to what is submitted by Member Countries and includes immediate notification and follow-up reports, six-monthly reports and annual reports. The system consists of an early warning system to inform the international community, by means of "alert messages", of relevant epidemiological events that occurred in OIE Member Countries and a monitoring system in order to monitor OIE Listed diseases (presence or absence) over time. In addition, according to the website, "the system not only provides countries with a simpler and quicker method of sending notifications and reports on disease information, but also allows them to utilize analysis capabilities to produce essential and useful information without delays." The details of these analytical capabilities were not accessible.

F. Data Dissemination and Reporting Options

The OIE works to make the animal disease data (to include diseases transmissible to humans and intentional introduction of pathogens) it receives from its Member Countries available in real-time. Multiple dissemination and reporting options exist, with information sent out immediately or periodically depending on the seriousness of the disease.

WAHIS

The World Animal Health Information System (WAHIS) is only accessible to authorized users, namely Delegates of OIE Member Countries and their authorized representatives. Members receive "alert messages" when relevant epidemiological events occur in OIE Member Countries (12).

WAHID

The World Animal Health Information Database (WAHID) Interface provides public access to all data held within OIE's new World Animal Health Information System (WAHIS). The data in WAHID is compiled into three main

categories: 1) disease information by country/territory, 2) overall disease information, and 3) control measures. It is the cornerstone in OIE efforts to improve the transparency, efficiency and speed with which animal health information is disseminated throughout the world (12).

World Animal Health: OIE Publication

In cooperation with the WHO and the FAO, OIE Member Countries are asked once a year to complete their annual report with information on non OIE-listed diseases. In this report they are asked to specifically discuss the impact of any zoonoses on humans and animal populations. Member countries are encouraged to provide information from their national reference laboratories, including diagnostic tests as well as vaccine data including manufacturers and vaccine production. All of the information is then gathered into the OIE *World Animal Health* publication (12).

NAHRS: USDA APHIS's Animal Health Reporting System

I. Background

The foundation for the Animal and Plant Health Inspection Service (APHIS) was built in 1883 when the United States Department of Agriculture (USDA) established the Department's first regulatory program, the Veterinary Division (14). Through a multitude of reorganizations in response to expanding missions, the USDA consolidated the previously independent animal and plant health bureaus in 1972, formally establishing the APHIS (14).

The Animal and Plant Health Inspection Service is a multi-faceted agency with a broad mission that includes protecting and promoting U.S. agricultural health, regulating genetically engineered organisms, administering the Animal Welfare Act and carrying out wildlife damage management activities (14). These efforts support the overall mission of the USDA, which is to protect and promote food, agriculture, natural resources and related issues. In the event that a pest or disease of concern is detected, APHIS implements emergency protocols and partners with affected States to quickly manage or eradicate the outbreak (14). This aggressive approach has enabled APHIS to successfully prevent and respond to potential pest and disease threats to U.S. agriculture (14).

II. Purpose, Justification, Objectives

The National Animal Health Reporting System (NAHRS) was designed to collect data from chief animal health officials from each state on the presence of confirmed OIE listed diseases in the United States (15). Although it is managed by the USDA APHIS (Veterinary Services Centers for Epidemiology and Animal Health National Surveillance Unit), it is a result of the joint efforts of the United States Animal Health Association (USAHA), American Association of Veterinary Laboratory Diagnosticians (AAVLD), and USDA-APHIS (14, 15). The specific objectives of NAHRS are: 1) demonstrate the integrated and transparent nature of disease surveillance and reporting in the United States and ultimately help protect the global market share of America's animals and animal products sold, 2) provide the primary source of information used in the completion of OIE reports by USDA-APHIS-VS, 3) provide reporting that reflects summary-level disease data for individual states as well as for the nation (14).

III. Operations

A. Operating Case Definitions

The animal diseases monitored in NAHRS are limited to those on the OIE-reportable disease list including diseases for cattle, small ruminants, horses, swine, poultry, and aquaculture; it does not include OIE-reportable diseases for amphibians, bees, lagomorphs, and diseases OIE categorizes as ‘other' (e.g., camelpox and leishmaniasis) (15). The NAHRS reportable disease list is modified to keep up with changes to the OIE list (at the beginning of the next calendar year); only confirmed occurrences of NAHRS reportable diseases are to be reported (14). Defined laboratory testing standards as well as the specific reporting criteria are detailed in the NAHRS Operation Manual and Uniform Methods & Rules; these criteria are considered minimum guidelines (14). The NAHRS Steering Committee and National Surveillance Unit are developing a U.S. National List of Reportable Diseases (NLRAD) that will provide a single, standardized list of diseases that must be reported in the U.S. (14, 16).

B. Target Population

The NAHRS aims to collect data on reportable animal diseases for all cattle, small ruminants, horses, swine, poultry, and aquaculture in the United States (16). Participation in NAHRS reporting, however, is completely voluntary. In federal fiscal year 2010, 48 states participated in NAHRS by providing monthly disease status reports; only Missouri and Louisiana did not participate (14).

C. Data Sources and Integration with Other Systems

State animal health officials collect and collate the confirmed occurrence of NAHRS reportable diseases. After approval by a designated state official, the state monthly report is entered directly (manually) into the system and then submitted to NAHRS via the secure NAHRS online reporting system, or by fax or e-mail (14, 15). NAHRS reporting is coordinated by the USDA-APHIS-VS-CEAH-National Surveillance Unit (14, 15). The NAHRS is not currently integrated with any other data system.

D. Data Management and Quality Assurance

All reported disease data is submitted to NAHRS via an online reporting form which specifically lists the diseases of concern and limits the participant responses to yes ('Y') or no ('N') (14). Participants must report on ALL listed diseases for livestock, poultry, and aquaculture species in their state. A 'yes' response indicates that at least one case of that particular disease was confirmed during the specific month; a "no" response indicates that, as far as state animal health officials are aware, no new cases of disease were confirmed in the state during the specific month. Confirmed occurrences of NAHRS reportable diseases are those diseases that meet NAHRS reporting criteria and/or are confirmed by using additional information (i.e., other testing methods or additional epidemiological information) (14, 15).

State animal health officials determine if a case is a valid one that should be reported to NAHRS (14, 15). State animal health officials are encouraged to use information from as many verifiable sources as possible to complete

their monthly NAHRS reports. This may include information from state or university animal disease diagnostic laboratories, private or industry laboratories, private veterinarians, state fish or wildlife agencies, extension veterinarians, the National Veterinary Services Laboratory (NVSL), the National Animal Health Laboratory Network (NAHLN), public health agencies, and other reliable sources (14, 16). The State chief animal health official makes the final determination on data reported to NAHRS for their state (14).

E. Data Analysis

The information in NAHRS is limited to disease status data (disease, present vs. free, data of last occurrence) and is collected for the intent of compiling the national summary-level data on disease occurrences that is provided to OIE (14). The USDA also uses NAHRS information in U.S. trade negotiations and to respond to inquiries or audits from trade partners (15). The NAHRS information is also used by the USDA-APHIS-VS-Centers for Epidemiology and Animal Health to provide historic information on occurrences of reportable diseases in the U.S.

F. Data Dissemination and Reporting Options

World Organization for Animal Health (OIE)

The primary reason for the NAHRS is to collect information to complete the biannual and annual reports that the U.S. is required to submit to the World Organization for Animal Health (OIE).

Local and State Level Utilization

Individual states benefit from participating in NAHRS by strengthening their internal animal health surveillance abilities through enhancement of disease reporting infrastructure and by using standardized disease reporting criteria (16). NAHRS reporting information enables states and animal industries to provide timely and accurate information for maintaining and developing international trade (14, 16). Further, NAHRS reporting provides states with important current and historic records of occurrences of reportable diseases; this information may be used to inform decision-making related to animal health issues including emerging animal health situations (16).

NAHSS

The National Animal Health Surveillance System (NAHSS) is a network of federal, state, university, industry, laboratory, and other partners aimed at collaborating surveillance efforts in order to protect animal health (16). The goal of the NAHSS is to integrate animal health monitoring and surveillance activities conducted by many federal and state government agencies into a comprehensive and coordinated system. The information in NAHRS in one of the data sources the NAHSS hopes to utilize (16). In addition the NAHSS hopes to incorporate information from the National Animal Health Monitoring System Reports and Information Sheets (NAHMS), National Surveillance Unit (NSU), and the National Animal Health Laboratory Network (NAHLN) (16).

SCS: USDA APHIS's Surveillance Collaboration System

I. Background

Veterinary Services is part of the Animal and Plant Health Inspection Service (APHIS). The core services provided by APHIS VS include surveying animals in order to estimate prevalence of disease, investigating the health status of animals in order to detect disease, and applying disease control measures (14, 17). Additional core functions include enforcing laws and regulations that protect animal health and commercial agriculture interests, developing policies and plans that support animal health efforts, informing people about animal health issues, and mobilizing and linking animal health service providers to identify and solve health problems (17).

II. Purpose, Justification, Objectives

The APHIS-VS conducts disease surveillance and monitoring for several diseases across multiple species. Because of differences in existing industry standards, disease pathogenesis, and potential prevention and control measures; each disease is typically managed by a separate APHIS activity unit. As a result, animal disease data has historically existed in several separate databases, preventing a complete understanding of disease status at a national level. The Animal Health and Surveillance Management (AHSM) System is VS's enterprise-level animal health and

surveillance electronic information management system which aims to manage surveillance data for numerous species and diseases; however, AHSM identified the need for a comprehensive and integrated animal health management commercial-off-the-shelf (COTS) surveillance software product (17).

In December 2011 the USDA APHIS VS began an initiative aimed at integrating all state level animal disease data into one comprehensive system; the Surveillance Collaboration System (SCS) (17, 18). Using a commercial off-the-shelf product (CoreOne), the purpose of the system is to provide a platform for the integration of field level activities related to surveillance across species (17, 18). Specifically, the objectives of the system are 1) provide a repository for core animal health data and manage regulatory work, 2) increase the speed, accuracy, and efficiency of laboratory data through the use of Mobile Information Management (MIM) which captures specimen collection data as it is collected from the field, 3) track laboratory diagnostics electronically from submission through results reporting, 4) provide e-mail alerts of laboratory results to appropriate animal health officials, 5) support data analysis such as pattern recognition, spread rate, emergence of threats for animal populations, and 6) provide critical data collection and support to investigations such as recording responses taken and scheduling future response activities (17).

III. Operations

A. Operating Case Definitions

The SCS serves as a central repository for animal disease data collected at the state level, to include the granularity and details of the actual filed report (17, 18, 19). The system does not impose specific case definitions, but rather provides a standardized template for data collection (19). The specific diagnostic modality used or control measures employed remain at the discretion of the operating veterinarian and animal health officials, but the system's drop-down menus ensure consistency of data-entry and therefore facilitates a more accurate assessment of overall disease statuses (18, 19).

B. Target Population

The aim of the SCS is to collect all data pertinent to the monitoring of animal diseases throughout the United States; however, utilization of the SCS at the state level is completely voluntary (18). In an attempt to encourage its use, the USDA offered the software at no cost to all relevant veterinary activities throughout the United States. Currently thirty-five states are using the system.

C. Data Sources and Integration with Other Systems

Data sources for SCS can be divided into two broad categories: test charts, forms, or other documents without animal details and reports with animal details (19). Examples of charts, forms, or other documents without animal details include herd level testing summaries, herd level vaccination summaries, and premise history summary reports. A case report tracking an individual animal is an example of a test chart, form, or other document with animal details. Access to SCS is granted by the USDA APHIS-VS Chief Information Officer. There are several levels of access. Listed hierarchically they are: read only, data entry, data extractor, and data manager. The data manager is the only one who can change data already entered. Data is generally entered into the SCS by animal health officials at the state level (17, 18). The system also has the capability to use MIM to receive near real-time data from the mobile devices used by animal health officials in the field (18). In addition, the SCS system is (or hopes to become) integrated with the following VS data systems: Animal Health and Surveillance Management System (AHSM), Emergency Management Response System (EMRS), National Animal Health Laboratory Network (NAHLN), Animal Disease Traceability, National Veterinary Services Logistic System (NVLS), National Veterinary Services Automated Information Management System (AIMS), National Veterinary Services Laboratory Information Management System (NVSL-LIMS), User Fee System (UFS), and Veterinary Services Process Streamlining (VSPS) (17). States that opt to employ the system are able to migrate data from their earlier systems into the SCS (18).

D. Data Management and Quality Assurance

Data entry into the SCS is menu-driven, which helps ensure consistency and overall uniformity of the data (18, 19). While SCS has many drop-down windows, there are free text and numeric fields available too. Data categories are dictated by the specific report being entered (19). For example, a Herd Brucellosis Test Summary Charts includes premise name, identification, state, herd number, species, submission reason, submitter's name/identification, test accession number, sample collection date, and total number tested. By selecting the specific accession number one may further review the details of the individual animals tested. In addition, SCS has the capability of maintaining scanned copies of original test charts, forms, or other documents, which permits users to further investigate the details of the data (18, 19).

The data in the SCS is an up-to-date representation of the surveillance activities of those using the system (17, 19). Data may include premise visits, laboratory testing, and disease outbreak investigations (17, 19). There are no externally imposed validations or quality assurance processes inherent in the SCS; the data captured is in the raw, user submitted form. The advantage of this approach is the depth of the data captured; reports are not limited to those with positive test results or specific case definitions. Another advantage is that a complete picture of surveillance efforts and activities in a state or related to a species or specific disease can be observed as a whole. Both of these support the aim of the SCS; to serve as a surveillance collaboration system for the purpose of disease monitoring, investigation and management (17, 18).

E. Data Analysis

The data collected in SCS is used to describe the workflow associated with animal health surveillance within the United States (18). The organizational units used to describe these workflows range from the lowest granularity (Task) to the highest (Service) (18). A 'task' is a discrete action or piece of work such as drawing a blood sample or assessing a caudal fold test. A group of tasks which achieve an objective are an 'activity'; a tuberculosis caudal fold herd test includes the tasks of scheduling the test, coordinating the resources, injecting the animals with the tuberculin, recording the premises and animal information, assessing the response, and recording and reporting the

results. An ‘activity’ is equivalent to a reportable event and is often recorded on an official form. An ‘activity group’ is a set of related activities for the purposes of either evaluation or management of a place (premise or zone) or thing (animal group, individual animal, product, fomite, or vector) (18). The service, activity and task workflow units are intended to provide the proper context for accurate interpretation of the data collected in SCS.

F. Data Dissemination and Reporting Options

The SCS software provides the infrastructure needed to operate a comprehensive and integrated animal information platform. The broad functionality of the system supports the collection of data that can be used by all levels of animal health surveillance activities; ranging from the routine management of herd health practices to the investigation of an outbreak. There are no recurring reports created or routine data disseminated directly from the SCS; rather, all reports are user-driven. The data in SCS is maintained in “sessions”, individual groupings of data by state. Those with access can see the data in their own state but not from other states unless they are in a multiple state area. Queries of SCS are drop-down menu driven and include (but are not limited to) details on premises, herds and animals, tests, treatments, or vaccination and other activities (18, 19). User-defined reports and other analyses can be output into Microsoft Excel format, as overlays on Google Maps, or as thematic maps (14, 18, 19).

To get access you must submit a form signed by your supervisor stating what kind of access you need. That goes to the Vet Services Chief Information Officer for final approval. Data entry people can see the data in their own state but not from other states unless they are in a multiple state area. There are several levels of access. Read Only, Data Entry, Data Extractor, and Data Manager. Those are in hierarchical order. The Data Manager is the only one who can change data already entered. Each state has a session of SCS so each approval is limited to a state or group depending on the need.

The information in Table 3.1 is not meant to be all inclusive, rather the review was limited to data systems used by some of the best known international and national public health organizations: The World Health Organization (WHO), the Food and Agriculture Organization (FAO), the Scientific Commission for Animal Disease (OIE), and the United States Department of Agriculture (USDA). In order to be included in the review, information on each organization's data systems had to be freely available for review on public domains.

Table 3.1: Comparison of Key Aspects of Global and US Data Systems Used in Public Health Surveillance

Organization	Data System	Purpose	Data Sources	Target Population	Limitations	Strengths
WHO (World Health Organization)	**GHO** (Global Health Observatory)	Monitor global health trends	Country-level data pulled from existing data sources: United Nations databases[1] (ITU, UNDESA, UNESCO, UNICEF) World Bank	Member States (194 in 2013)	Variability of country-level data	Breadth of data sources (31 potential data sources) Standardized data entry makes data comparable across countries/regions
FAO (Food and Agriculture Organization of the United Nations)	**EMPRES-i** (Emergency Prevention System-internet)	Early warning for zoonotic and transboundary animal diseases globally	Country/regional data from official and unofficial sources: Country/regional project and mission reports, NGOs[2], institutions, Ministries of Agriculture and health, in country FAO or United Nations reps, public domains, media and web-based health surveillance systems	Animal populations globally	Lack of case definition (low specificity)	Breadth of data sources Contains both confirmed and non-confirmed animal data (high sensitivity)

Table 3.1: Comparison of Key Aspects of Global and US Data Systems Used in Public Health Surveillance (continued)						
Organization	**Data System**	**Purpose**	**Data Sources**	**Target Population**	**Limitations**	**Strengths**
OIE (Office International des Epizooties)	**WAHIS** (World Animal Health Information System)	Transparency in global animal diseases that may affect trade	Country-level data entered by OIE delegates	Member Countries (178 in 2013)	Limited to OIE specified animal diseases	Standardized guidelines for disease detection and notification processes
USDA (United States Department of Agriculture)	**NAHRS** (National Animal Health Surveillance System)	Collect US OIE-reportable disease data	State-level veterinary data entered by state animal health officials	United States (48 participate)	Limited to present/absent status of OIE specified animal diseases by state Reporting is not mandatory and case validation is at the discretion of the state	Facilitates OIE-mandated report submission Provides summary data international trade negotiations
USDA (United States Department of Agriculture)	**SCS** (Surveillance Collaboration System)	Integration of US field level animal health surveillance related activates across species	Field-level animal health reports USDA VS data systems[3]: AHSM, EMRS, GDB, NAHLN, Animal Disease Traceability, NVLS, NAIMS, NVSL-LIMS, UFS, and VSPS	United States (participation varies)	Lack of uniform participation contributes to gaps in data	Standardized data entry makes data comparable Data maintains granularity to support detailed investigations

Section II: Military Data Systems Used in Public Health Surveillance

Public Health Surveillance in the Military

Prior to the Civil War, medical care in the military was provided mainly by regimental surgeon and surgeons mates (20). Although there were attempts to establish a centralized medical system, primitive communication and transportation efforts made this difficult (21). In World War I, the U.S. Army Medical Department expanded and developed greater organization and structure (20). Medical care that began on the battlefield was successfully transferred to facilities of greater medical capabilities away from the front. After World War II, the Executive Branch of the U.S. Government underwent significant reorganization (20, 21). The separate Department of War and Department of the Navy were re-aligned under a single Department of Defense (DoD) (20, 21). There were significant growing pains under this reorganization as both the Army and the Navy had independently developed their own medical systems. In addition, the Air Force, part of the Army until 1947, also had its own medical system (21).

Post-Cold War downsizing of the U.S. military, with a simultaneous increase in geographically dispersed overseas operations forced the military to focus on maximizing the health, fitness, and medical preparedness of the forces being deployed (22). This shift in focus from reactionary medicine to preventive medicine was the corner stone of the post-Cold War military medical support strategy (22). Following the Persian Gulf War there were a number of medical complaints from Gulf war veterans, but investigation into potential causes was hindered because relevant records were often inaccessible or nonexistent, those that were accessible lacked uniformity and accuracy, and those that were uniform and accurate were generally not automated, making investigation and analysis cumbersome. These short-comings made it evident that “deployment medical surveillance” needed to become a DoD priority (22).

According to Department of Defense Directive 6490.20E (2), public health surveillance in the military shall be conducted to enable early intervention and control strategies to prevent adverse effects on mission accomplishment and to provide information for shaping commanders’ decision making. The primary objectives of these systems in the military are timely and proficient public health response to medical events that ultimately reduce morbidity and mortality, estimation of the distribution, trends, and risks associated with significant medical events, and the

development and assessment of policy and resource allocation for the prevention and control of communicable diseases (2). Currently, the military has many human disease reporting/surveillance systems.

Methods

Data Collection

Descriptions of each medical data systems were compiled utilizing a standardized outline adapted from the CDC Guidelines for Evaluating a Public Health Surveillance System (**Appendix A-3**). Information on each system was collected through a combination of pertinent literature reviews, online user training modules, as well as interviews with subject matter experts or end users of each system. When authorized, live demonstrations of how to use the system were conducted. All information was then entered into the standardized outline and forwarded back to the subject matter experts for content review, clarification, and eventual approval.

MDR: Military Health System Data Repository

I. Background

In the DoD, the Military Health System (MHS) is the framework for global health information exchange (23). Within the MHS are the DoD global health information network and common services utilities, some of which include medical, personnel, deployment, and financial data from all branches of military service (24). Specific hospital and personnel data is then further consolidated into the MHS Data Repository (MDR) (25). The MDR serves as the central repository for medical data such as outpatient visits, inpatient admissions, and laboratory orders (and some results) (25). It also compiles population and personnel data for Service members, dependents, retirees, and other beneficiaries (25). Lastly, the MDR captures financial (claims) data (25). This central repository of data can be used as an electronic data stream for disease surveillance purposes. Several of the military human disease data systems utilize the data available in the MDR, however, how the data is used and what data it is combined with varies greatly.

II. Purpose, Justification, Objectives

The purpose of the MDR is to serve as the centralized data repository for the Department of Defense's MHS (25). The justification for the system is its utility in health related financial and administrative management. As the single point for health data integration, data quality edits, online data storage and healthcare data transfers, the MDR captures and validates data from more than 260 DoD health data networks worldwide giving it the ability to quantify resource needs and associated costs (23, 25). The system has four main objectives 1) preserve historical "raw" data by type and by year, 2) normalize and process all data types against specific business rules, 3) assure controlled access to privacy protected data, and 4) supply interagency data sharing with the Department of Veterans Affairs and Health and Human Services (23, 24).

III. Operations

A. Operating Case Definitions

Cases in the MDR may be identified using International Classification of Disease (ICD) codes entered in the medical record (25, 26). Because each time an ICD code is associated with a record it has the potential to be counted as a case, it is very important for analysts to define inclusion criteria when conducting system queries based on these codes. Examples of inclusion criteria are incidence rules (i.e. a person can be considered a case once in their lifetime) and health encounter criteria (outpatient vs. inpatient, number of visits, time between visits) (26). The MDR contains "raw" unedited health event data; data quality assurance is the responsibility of the individual conducting the query. The system does not automatically incorporate Reportable Medical Event reports.

B. Target Population

The MDR population includes Service members (SM) from all branches of Service to include active duty, Reserve/Guard Component (while on orders), and retirees; SM dependents (military family members) and other eligible beneficiaries (23, 25, 26). Complete medical and personnel data from both military Medical Treatment

Facilities (MTFs) and civilian medical facilities who accept/file Tricare Managed Activities (TMA) health insurance claims are available for this population (23, 25, 26). Data from Veteran Affairs hospitals is currently not available.

C. Data Sources and Integration with Other Systems

MDR serves as the central repository for the Military Health System (MHS) receiving information from more than 260 DoD health data network systems worldwide (23, 24). Briefly, data types include medical, population, and fiscal (24).

The MDR is a fully integrated database with tools for accessing and managing health related data from both MTFs and MHS approved purchased care from civilian medical facilities; hereafter referred to as civilian medical facilities (23, 25). In-patient data is updated monthly from Standard In-patient Data Record & Tricare Encounter Data/Health Care Service Record non-institutional files (SIDR & TED/HCSRI). Out-patient (ambulatory) data is updated daily to monthly from Comprehensive Ambulatory/Professional Encounter Record & Tricare Encounter Data/Health Care Service Record non-institutional files (CAPER & TED/HCSRI) (23, 25, 26). Data from ancillary laboratory and radiology files come from Composite Health Care System (CHCS) on a monthly basis (25, 26). The MDR does not receive Reportable Medical Events Reports (26).

Population data in the MDR comes from the Defense Eligibility Enrollment Records System (DEERS), allowing it to capture complete personnel records for anyone eligible to receive medical care in the military system (MTFs and civilian medical facilities), to include full denominator data for active duty, Reserve/Guard Component (while on orders), retirees, SM dependents (military family members) and other eligible beneficiaries (24, 25, 26).

D. Data Management and Quality Assurance

The Defense Health Services Systems (DHSS) manages the MDR (23). All data input for the MDR is automated; as the central repository for the MHS there is an extensive list of external interfaces (24, 25). The data in the MDR is

cleaned by the DHSS, but, as with any passive surveillance system there is still the potential for incomplete records or inconsistent entries (25, 26).

E. Data Analysis

The data in the MDR are extracted by experienced SAS programmers for the production of user-friendly formats such as frequencies, counts, and person-years (25, 26). Because the MDR has complete denominator data for military beneficiaries, it is able to produce estimates of disease rates for the entire military population. Population query filters are extensive to include, but can incorporate gender, age, race, grade, marital status, service, occupation, location, and calendar year (25, 26). In addition medical data source and method of diagnostic validation can be specified. All records are maintained for person, place, and time (25, 26). The data are routinely used to analyze population-based morbidity, conduct risk assessments, determine emerging threats, and monitor the effects of health policy (25, 26). In addition, the data are used in health care management such as business case studies, inventory/supply management, and reimbursements justifications (23).

F. Data Dissemination and Reporting Options

MDR Ad Hoc Queries

The DHSS issues MDR accounts to qualified individuals. Applicants are limited to those with Common Access Card (CAC) identification cards and are therefore limited to active-duty military personnel, selected Reserve personnel, DoD civilian employees, and eligible contractor personnel (25). Access is password protected with extensive physical and electronic security measures. Because of the nature of the data contained in the MDR, access is further restricted to those with SAS programming expertise and functionality (25). Local protocols dictate query processes, information accessed and information management storage. Any data requests generating more than 10,000 person level records containing Protected Health Information/Personally Identifiable Information (PHI/PII) must be requested directly through the DHSS and cannot be obtained through ad hoc queries. The DHSS fulfills the

data requests by compiling and shipping the data on a disc (26). The MDR itself does not generate any recurring reports.

MDR Queries through DHSS

Ad hoc requests can be submitted by system end users or directly to the DHSS; submission requirements, request channels, and accessibility limitations are subject to DHSS protocol. Any data requests generating more than 10,000 records will be compiled and shipped from the DHSS on a disc (26). In these instances only final results are downloaded which eliminates the ability of the analyst to perform quality assurance or debugging checks without addition resubmissions.

M2

The Military Health System Management Analysis and Reporting Tool (M2) is a user friendly interface where end-users can extract data through menu-based queries and create data outputs (27, 28). Access is password protected and limited to those with CAC ID cards; physical and electronic security measures are the same as with MDR. The system merges several stand-alone applications into a single platform designed for the purpose of improved business intelligence reporting (27, 28). The ad-hoc query tool enables proactive healthcare management and trend analysis for MHS operations worldwide. It permits the monitoring of patients' use of services, runs patient and provider profiling studies, identifies opportunities to transfer health care to and from military facilities, and delivers summary and detailed financial data (27, 28). In addition, it has the capability of providing summary and detailed clinical data on diseases specified by condition (ICD code), health related event (in-patient, out-patient, laboratory order), and population (26, 27, 28). Population filters are the same as the MDR (26, 27, 28). Also like the MDR, the M2 enables reports of not only count data but also the calculation of rates.

Defense Medical Surveillance System (DMSS)

I. Background

In 1986, the Army established a data center to support its HIV-related screening, clinical care, and epidemiological research programs. In 1993, the system transitioned to the Army Medical Surveillance System, expanding its scope to include all illnesses and injuries of public health or military operational importance. In 1997, the Army Medical Surveillance System transitioned to the Defense Medical Surveillance System (DMSS), and the Army Medical Surveillance Activity (AMSA) was assigned responsibility for its operation (22). The Deputy Secretary of Defense established the Armed Forces Health Surveillance Center (AFHSC) in 2008, by partnering AMSA, the DoD Global Emerging Infectious Disease Surveillance and Response System (DoD-GEIS), and the Global Health Surveillance Activity (29). Today, the AFHSC has the responsibility for comprehensive health surveillance for the DoD, which includes the monitoring and maintaining of the information in DMSS.

II. Purpose, Justification, Objectives

The purpose of DMSS is to serve as the central repository of medical surveillance data for the U.S. Armed Forces in support of population-based surveillance (22). The justification for the system is to maximize health, fitness, and medical preparedness of forces being deployed and to minimize disease and injury risks during deployments. The system has three main objectives 1) create longitudinal records on Service members (SMs) by documenting their military and medical experiences throughout their military careers, 2) document statuses of and changes in demographic and military characteristics of the SM population, and 3) permit the timely assessment of the morbidity experiences of SMs with shared characteristics, who were in specific locations, or had similar experiences on specific time scales (22).

III. Operations

A. Operating Case Definition

Cases in DMSS are identified using ICD codes, based on either ICD codes entered in a medical records or the existence of a Reportable Medical Event (RME) report (30, 31). Because each time an ICD code is associated with a record or RME report it has the potential to be counted as a case, it is very important to specify inclusion criteria when requesting case count data. Examples of inclusion criteria are incidence rules (i.e. a person can be considered a case once in their lifetime) and health encounter criteria (outpatient vs. inpatient, number of visits, time between visits). If the requester does not provide specific inclusion guidelines, AFHSC uses the Tri-Service Reportable Event Guidelines and Case Definitions or their own published case definitions (30, 31). Records that are deemed invalid via DMSS's "edit check" program process are excluded from entry into the DMSS database.

B. Target Population

The target population for DMSS is Service members from all branches of service to include active duty, Reserve Component (while on orders), deployed contractors, and eligible retirees (22, 32). Complete medical and personnel data from both military MTFs and civilian-TMAs are available for this population; denominator data comes from the number of SM records associated with each installation (22, 31). The system also captures data for dependents (military family members) and other eligible beneficiaries, however it is limited to the medical data associated with health encounters and lacks the corresponding denominator data or personnel details for this population (31). Data from the Veteran Affairs hospitals or from non-deployed contractors is not available.

C. Data Sources and Integration with Other Systems

DMSS contains data relevant to more than 7 million individuals who have served in the armed forces since 1990 (22). Briefly, data types include medical, population, deployment, and serological.

The medical data from both MTFs and civilian-TMA facilities comes from the MDR and includes both in-patient and out-patient records (22, 31, 32). Although not yet routinely used, the DMSS has had direct access to the laboratory data (HL7) in the MDR since 2010 (31). The DMSS also receives RMEs from all branches of service (22, 31, 32).

Population data in the DMSS comes from the Defense Manpower Data Center (DMDC) which includes complete personnel records documenting the entire "military experience" of the Service member (22). Information includes demographic data, occupational history, training records, and assignment history. In addition, DMSS incorporates data from the Armed Forces Medical Examiners System (AFMES) which provides casualty data for Service members that are deceased. The combined data from these two systems allow DMSS to have complete records for all Service members from date of entry through career completion.

Deployment data in DMSS is primarily in the form of deployment rosters (DMDC) and pre and post deployment health assessments Medical Protection System/Remote Data Entry System (MEDPROS/RIDES) (22, 31). Disease and non-battle injury data from the Theater Medical Data Store (TMDS) does not feed directly into DMSS's case count data, but rather must be queried separately and then is only available as aggregate data (31).

The Armed Forces Health Surveillance Center has operational oversight of both DMSS and the Department of Defense Serum Repository (DoDSR) (22, 32). Since 1989, sera remaining after routine HIV-1 antibody testing and sera collected before and after major deployments have been maintained at the DoDSR for each Service member (22, 32). As of this writing, DoDSR currently houses over 50 million specimens, and continues to grow by approximately 2.3 million specimens per year (32). Specimens contained in the DoDSR are available to researchers and other investigators within the Department of Defense for the purposes of conducting militarily relevant investigations. Access to the specimens is regulated by AFHSC (32).

D. Data Management and Quality Assurance

The data in DMSS is managed by the AFHSC. All data input for DMSS is automated; external interfaces include, but are not limited to those mentioned above (MDR, DMDC, and AFMES). DMSS does not have the capability to edit any of the data directly, but does have "edit check" programs in place to ensure data quality before it becomes integrated into the system. Such checks include completeness (ensure all essential fields are filled in), consistency (birthdates remain the same for different entries), and accuracy (compliance with specified formats and within acceptable ranges) (31). After processing, errors may still remain, including differential diagnoses ICD codes that get counted as cases or vaccinations that are counted as disease because of incorrect disease designation selection in the electronic medical record. These errors, however, are not unique to DMSS, and are found in all surveillance systems based on ICD codes.

E. Data Analysis

The data in DMSS are in table format (SQL); SAS is used to access, collect, and analyze the data requested and can produce frequencies, counts, person-years, rates, and various charts. Denominator data exists, but does not include beneficiary data (complete denominator data exists for DMED, see below). Population queries can include variables such as gender, age, race, marital status, service, occupation, location, and calendar year (31). In addition medical data source can be specified. Complete longitudinal records are created for all Service Members (22, 31). Examples of historical studies that have been conducted with DMSS data include: population-based mortality (trend analysis by sex and age), risk assessments (respiratory illness hospitalization compared among exposed vs. unexposed cohorts), emerging threats (track cases to common exposures, deployment surveillance (hospitalization rates between those deployed and those not), policy effects (evaluation of pre vs. post policy impacts), serological surveys (nature, etiology, distribution, magnitude of disease), sero-epidemiological research (linkages of data relevant to individual characteristics, exposure states, medical events, an serum specimens), and routine reports and summaries (annual population-based morbidity reports, for example on Lyme disease) (22, 32).

F. Data Dissemination and Reporting Options

DMSS Ad Hoc Queries

All data requests must be submitted through AFHSC and will be categorized as either Operational/Public Health Practice or Research Support (31, 32). Only queries categorized as Research requires additional Internal Review Board approval, which generally takes at least two weeks. Requestors must specify, as a bare minimum, the disease of interest including a specific case definition, inclusion and exclusion criteria, incidence rules, target population, other applicable count clarifications (time between outpatient visits). Data is normally provided in spreadsheet format and generally does not include unique identifiers unless prior approval was granted. Because complete denominator data does not exist for the entire population captured in DMSS, data is generally case count in nature. Additional variables per case include gender, age, race, grade, marital status, service, occupation, location, and calendar year.

DMED

The Defense Medical Epidemiological Database (DMED) is a web-based application that provides remote and rapid (within seconds) access to data summaries in response to user-defined queries (31, 32). The data available in DMED are synchronized with DMSS on a monthly basis, but is only a subset of the DMSS data that represents the active duty Service member population only (31, 32). Also, data are only aggregate or summary in nature with queries limited to primary diagnoses only. Population filters including gender, age, race, grade, marital status, service, occupation, location, and calendar year. The data formats include frequencies, counts, person-years, rates, line charts, and bar charts (31, 32). Access is limited to those with CAC ID cards and is therefore restricted to active-duty military personnel, selected Reserve personnel, DoD civilian employees, and eligible contractor personnel (32).

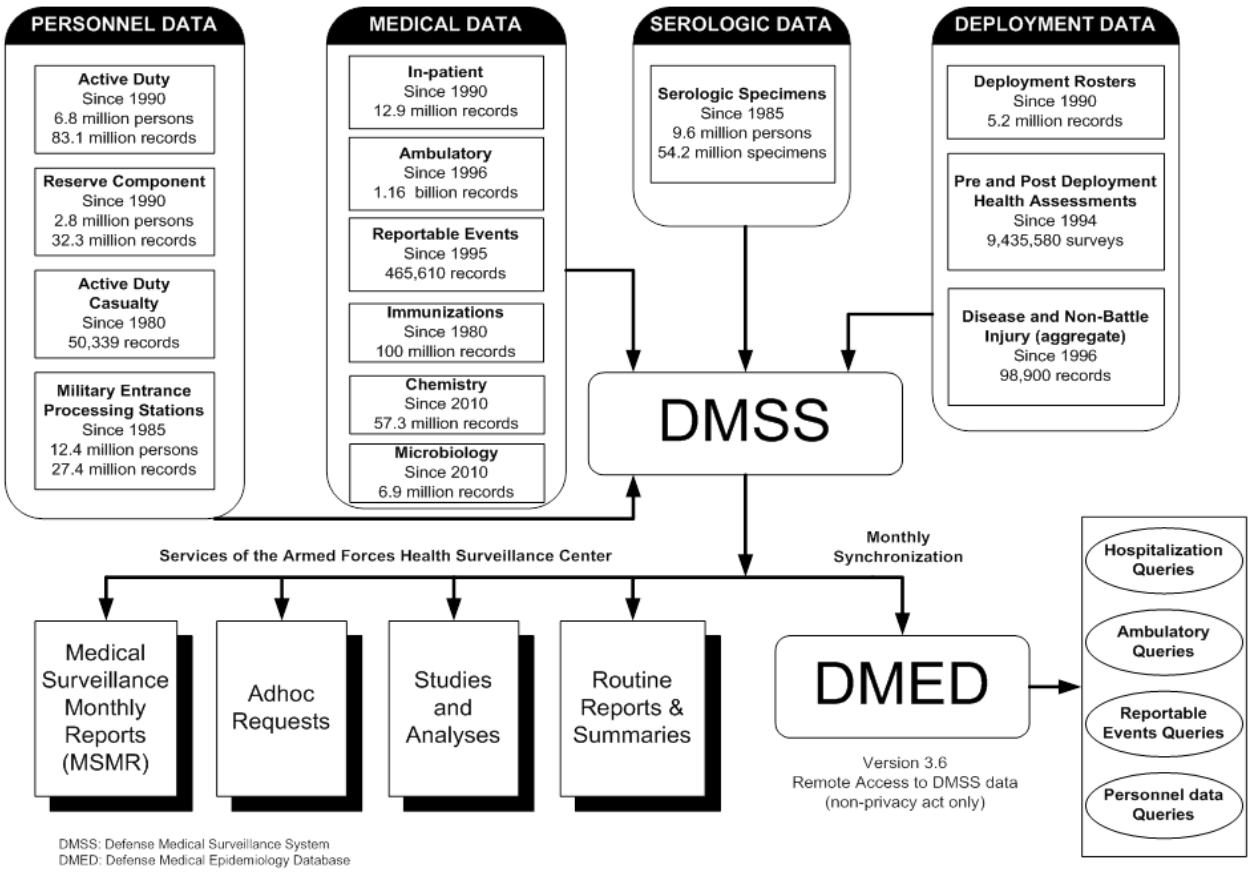

Figure 3.1. Overview of Defense Medical Surveillance System architecture, with key data sources and functional relationships (22)

Medical Surveillance Monthly Reports (MSMR)

The Medical Surveillance Monthly Reports (MSMR) is a recurring publication produced by the AFHSC (32). The aim of the publication is to report frequencies, rates, and trends of ambulatory visits, hospitalization, and RMEs among active SMs as well as share information on cases and outbreaks of illnesses, injuries, and exposures with broad military or medical relevance. The publication can be accessed at the AFHSC website or is mailed to designated recipients.

HL7

I. Background

HL7 (Health Level 7) technically refers to the standardized format used by electronic hospital information systems that facilitates the exchange, integration, sharing, and retrieval of electronic health data (4); the data from these information systems is referred to as HL7 data. HL7 data in the military system comes from services that occur during or related to patient encounters such as outpatient visits, clinic appointments, pharmacy orders and fills, radiology exams, laboratory test orders and results, and inpatient admissions and discharges (33). Algorithms and tools developed by the Navy and Marine Corps Public Health Center permit the interpretation of HL7 data for the purpose of augmenting and improving military public health surveillance activities (34). Although all service branches are represented in HL7 data, it only represents health care events that occurred at fixed facility MTFs (no ship or temporary medical facilities) (35).

II. Purpose, Justification, Objectives

The purpose of HL7 is the management of electronic health data resulting from ancillary medical services and diagnostics such as laboratory, radiology, and pharmacology (4, 33, 36). The justification of HL7 use in the military is to conform to international standards for the exchange, management and integration of electronic healthcare information, increasing the effectiveness and efficiency of healthcare information delivery within the military health care system. The data can be utilized to augment and improve military public health surveillance activities by providing case validation through confirmation of medical encounter data with laboratory, pharmacological, and/or radiological data. Objectives that support health surveillance include: 1) identification of cases of reportable diseases (case finding of RMEs), 2) analysis of case burden (health reports and inquiry), 3) identifying illness trend, 4) investigation into the types and subtypes of diseases (vaccine development, multidrug resistance), and 5) estimation of occupational exposures as indicated by organic biomarkers (33, 35, 36).

III. Operations

A. Operating Case Definitions

It is difficult to apply the term "case definition" to HL7 data; HL7 itself does not define cases, rather it may be used as a tool to augment the defining of a case. For example, HL7 does not generate ICD codes but the data message may aid physicians in making a diagnosis that leads to the selection of that ICD code (36). In this way, the data in HL7 can be used for identification of diseases based on Reportable Medical Events criteria listed in the Armed Forces Reportable Medical Events Guidelines and Case Definitions, but is not limited to this application (30). The data can also be used for the identification of medical events as established by professional standards from outside the DoD (e.g., CDC, professional subject matter experts).

B. Target Population

The HL7 captures all of the ancillary medical services that occur at fixed MTFs, therefore it represents Service members (SM) from all branches of service to include active duty, Reserve Component (while on orders), trainees, contractors, and retirees, as well as SM dependents (military family members) and other eligible beneficiaries. It does not incorporate data from deployed personnel (aboard ship, in-theater) or personnel receiving care in other than MTF facilities (field unit, Battalion Aid Stations), nor does it include civilian-TMA data (36).

C. Data Sources and Integration with Other Systems

The HL7 data are derived from the military's CHCS, which manages all hospital orders, patient administration, and appointment scheduling. Data types include the ancillary medical services of pharmacy, radiology, and laboratory departments (36).

The dependency on CHCS explains why procedures performed at locations without CHCS are not automatically captured in the system (civilian facilities, field units, etc.). Data on procedures performed outside of a fixed MTF

can be entered manually if appropriate documentation is provided (36). Both the Disease Reporting System-internet (DRSi) and the Electronic Surveillance System for the Early Notification of Community-based Epidemics (ESSENCE) routinely incorporate HL7 data in their case finding modules; however this is not done automatically, rather through the application of appropriate algorithms. The DMSS system has a similar capability, but it is rarely used (31, 36).

D. Data Management and Quality Assurance

Medical services data are either entered into CHCS manually by hospital staff ordering the service or automatically via the laboratory device processing the medical test. The data are not transmitted as an HL7 message into the CHCS central server until they are certified; pharmacy data are certified when the transaction is completed (order filled and picked up), radiology data are certified when the report is filed (results/interpretations), and laboratory data are certified when tests are finalized. Once data are certified an HL7 message is sent to the CHCS host where it is archived and batched (specific triggers are set up to send messages). There is a one to two day lag between certification at the MTF level and the data's availability to be accessed in HL7 messages (36).

Data in the HL7 database are raw and unedited; duplicate records may exist. For example the reordering of a test will generate more than one data entry. Methods exist to identify and account for these duplications. When a test is cancelled, inaccurate, or misread, the details are recorded in the "Test Result", "Clinical Comments", or "Result Notes" fields. Differences in MTF capabilities (laboratory, radiology, and pharmacy) and naming conventions contribute to a lack of uniformity between locations and can lead to issues with data consistency. There is also a potential for missing data. As stated earlier, HL7 does not capture data from non-fixed MTF facilities, but it also does not capture conditions that are diagnosed without laboratory confirmation or pharmacological treatment. Data entry errors can occur; for example, data from outsourced clinics may be recorded in inappropriate data fields, and the existence of free-text fields permit differences in formatting, abbreviations, and allow for spelling errors. However, algorithms have been developed to account for variation in free-text fields, including the identification of new tests and results, quality assurance test entries (not used for clinical purposes), and provider indications (36).

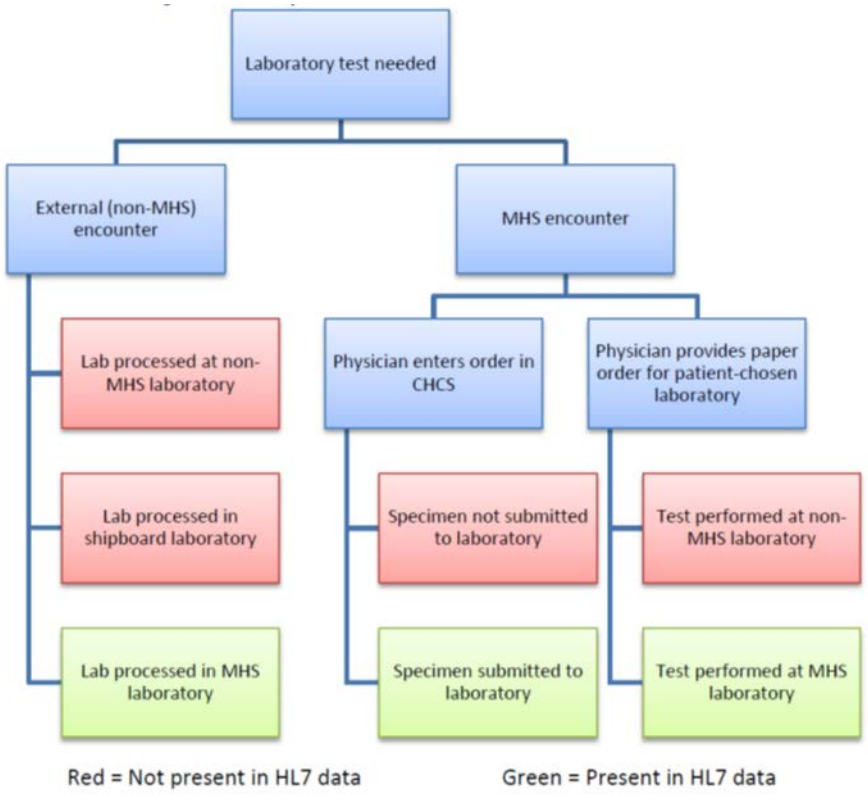

Figure 3.2. Example of HL7 Data Flow Process (35)

E. Data Analysis

The data in HL7 are in the raw form; extraction and compilation of data into formats suitable for public health analysis requires the expertise of experienced users. For the military these personnel are located at the EpiData Center Department at the Navy and Marine Corps Public Health Center (EDC NMCPHC). All analyses are based on count data, as complete population data does not exist within the system. Data are linked by social security number, therefore comorbidity information (based on laboratory, radiology, or pharmacology) and extensive demographic data can be accessed in order to perform trend analysis for known cases. Data are collected at the MTF level but can be broken out by patient or branch of service; grouping of data above the MTF level (region, state, country) must be performed by the user. Data can also be analyzed temporally either daily, monthly, quarterly, or annually. Observing changes in case counts over time requires careful interpretation as they may actually reflect changes in diagnostic capabilities or physician diagnostic preferences. Existing time fields include: order date, effective date, collection date, certification date, and message date. All data can be analyzed by levels of

diagnostic certainty (suspect, probable, confirmed, pending). Caution must be employed if data from HL7 are used alone to estimate true burden of disease; as the data are greatly influenced by provider practices, MTF specific capabilities, and naming conventions (35, 36).

F. Data Dissemination and Reporting Options

The batched HL7 messages are forwarded to DHSS main servers at least once daily. On a daily basis, DHSS forwards raw parsed data to EDC NMCPHC. The EDC NMCPHC manages all HL7 related reports and query requests (35, 36).

HL7 Ad Hoc Queries

All data requests must be submitted in writing to EDC NMCPHC using a specific data request form; data can only be provided based on established Memorandums of Understanding (MOUs) or equivalent agreements (36). Accurate data acquisition requires knowledge both of the disease or condition of interest, as well as working knowledge of MTF practices and capabilities. Therefore fairly detailed requests must be provided by the requestor to EDC NMCPHC. The staff at EDC NMCPHC is skilled in searching the different datasets, querying the free-text windows, and searching using multiple naming conventions. Data transfer to the requestor requires a secure CAC enabled connection (36).

Recurring Reportable Medical Events Reports

Daily Reportable Medical Events Case Finding Reports are automatically generated and sent to Air Force, Army, and Navy Preventive Medicine personnel. The reports include case count data for 53 Reportable Medical Events including details on the MTFs reporting the condition, patient demographics, disease name and diagnostic classification (suspect, positive) (36).

ESSENCE

I. Background

In the late 1990s, the DoD identified the need to create a sensitive, specific, standardized, real-time surveillance system for the U.S. National Capital region (4, 37). The specific goal was to have a system that was capable of detecting outbreaks of emerging infections and bioterrorist attacks (37). In 1999, a partnership between the Walter Reed Army Institute of Research/Department of Defense Global Emerging Infections Surveillance and Response System (WRAIR/DoD-GEIS) and Johns Hopkins University Applied Physics Laboratory (JHU/APL) developed ESSENCE (37). ESSENCE (Electronic Surveillance System for the Early Notification of Community-based Epidemics) is a web-based medical surveillance system developed to support outbreak detection and medical situational awareness (38). After 11 September 2001 it was expanded to include all DoD MTFs, becoming a key component of the US national biosurveilalnce network in support of the CDC BioSense application (http://www.cdc.gov/biossense) (37).

II. Purpose, Justification, Objectives

The purpose of ESSENCE is to promote detection and monitoring of disease clusters and outbreaks. The justification for the system is its unique potential for early disease detection based on surveillance of syndromic groups, which allows it to deliver actionable health information to decision makers before an outbreak occurs. The system has four main objectives: 1) sentinel event detection, 2) disease outbreak detection 3) forecasting or predicting health related events before they happen, and 4) health event monitoring (e.g., injury prevention, chronic disease burden, exposure monitoring, etc.) (37, 38).

III. Operations

A. Operating Case Definitions

Although ESSENCE has the ability to mine data based exclusively on ICD codes, the system's strength comes from its ability to monitor unusual changes in groups of diagnostic categories that represent specific syndromes or disease categories of concern, these are referred to as syndromic groups (4, 37, 38). There are ten broad syndromic groups already established in ESSENCE: neurological, botulism-like, shock/coma/death, rash, localized cutaneous lesion, gastrointestinal, influenza specific, influenza-like illness (ILI), ILI- alternative case definition, fever-unexplained, and hemorrhagic illness (37, 39). These syndrome groups are made up of various de-identified data fields including laboratory/radiology order codes (Current Procedural Terminology or CPT), pharmaceutical order codes (Categories of Pharmaceuticals Ordered or GC3), disease and non-battle injury data (by ICD), potential reportable medical event data (based on ICD), as well as patient disposition (admitted, outpatient, home) and diagnostic codes (ICD) (37). ESSENCE also has the unique ability to incorporate data from the free text field of patient's chief complaint entry from electronic medical record. In addition, users can create "user-defined" syndromic groups based on specific criteria (selected ICD codes, pharmacy codes, and radiology orders)(37). Data are processed through algorithms designed to identify records that meet the appropriate criteria and are then grouped together according to the disease/syndrome group they fall within, by date and location. Users can then define the acceptable thresholds and set alerts for the detection of the syndrome above expected levels within a specified user defined location (37, 39).

B. Target Population

The ESSENCE captures all beneficiaries seeking outpatient care at fixed MTFs. This population comprises Service members (SM) from all branches of service to include active duty, Reserve Component (while on orders), and retirees; SM dependents (military family members) and other eligible beneficiaries. The medical data, however, is primarily limited to outpatient or ambulatory encounters at MTF, although outpatient cases that result in admissions are also captured. Data from the Veteran Affairs hospitals or from the deployed environment is not available (39).

C. Data Sources and Integration with Other Systems

The data in ESSENCE are extracted from the military's electronic medical record (EMR) system, Armed Forces Health Longitudinal Technology Application (AHLTA) (37, 39). Specifically, the majority of the data comes from the Ambulatory Data Module which is limited to out-patient encounters seen at MTFs (39). The uniqueness of ESSENCE is its ability to monitor syndrome groups. In addition, ESSENCE has the potential to incorporate information from various data fields including laboratory/radiology order codes (Current Procedural Terminology or CPT), pharmaceutical order codes (Categories of Pharmaceuticals Ordered or GC3), disease and non-battle injury data (by ICD), potential reportable medical event data (based on ICD), as well as free-text fields from patient records (chief complaint), and ICD based patient disposition and diagnostic codes. Demographic data in ESSENCE is limited to that available in the medical record (39).

D. Data Management and Quality Assurance

ESSENCE is managed by the DHSS (37). All data input for ESSENCE is automated; most are updated on a daily basis (39). The data in ESSENCE are clustered at the installation level, allowing Public Health (PH) and Preventive Medicine (PM) personnel to detect trends for patients seen at their MTF (39). Although unique identifiers exist for each record, they are generally not available to the user unless they have a restricted access account (39).

The discussion of quality assurance for a data system like ESSENCE is complicated. Unlike surveillance systems that are designed to capture true cases (high specificity), the value of ESSENCE is its ability to focus attention on statistical anomalies that may indicate the need for further investigation by identifying possible outbreaks in very early stages (sacrificing specificity in order to enhance sensitivity). By design ESSENCE has a high potential for over-reporting because false positives are expected with syndromic surveillance. The determination of the epidemiological significance of detected trends, clusters, or alerts is the task of the trained user.

By incorporating ICD data, ESSENCE is subject to the same errors as other surveillance systems that use the codes (differential diagnoses ICD codes counted as cases, vaccinations counted as disease).

E. Data Analysis

ESSENCE uses statistical algorithms to generate alerts for the ten broad syndromic groups based on daily health encounter data for a specific MTF or user-defined geographic location (37). Using historical data, a prediction of expected daily ranges (threshold) for each of the syndrome groups is generated. Event detection algorithms alert local users to aberrant values (potential outbreaks) when the observed number of cases is statistically significantly higher than the established threshold (37). Tests of significance are based on one-sided 95% confidence intervals where no alert is generated if the p-value is > 0.05, an amber alert is generated if the p-value is between 0.01-0.05, and a red alert is generated if the p-value < 0.01 (37). Algorithms are designed to account for changes in patient visits associated with days of the week; historical count information relies on data from the last four to twelve weeks (39). Because of incomplete population data, ESSENCE cannot be used to generate information about disease or syndrome rates. When "user-defined" syndromic groups are used instead, however, the user must define the acceptable thresholds in order to set alerts for detecting the syndrome of interest at the specific location.

F. Data Dissemination and Reporting Options

All access to ESSENCE requires CAC ID cards and is therefore limited to active-duty military personnel, Selected Reserve, DoD civilian employees, and eligible contractor personnel. There are two possible access levels. Level I provides a general over-view of the data, while Level II access permits detailed investigation of patient information at the MTF level by allowing access to personal identifiers (37).

E-mail Alerts to Users

ESSENCE generates e-mail alerts to users based on user-selected syndromic group or user-defined syndromes (37). The data generated includes graphical descriptions of the alert showing historical data, outpatient/inpatient visit counts, and new visit counts (14-day incidence rule) (37). The graphs also display increasing or decreasing patterns, and whether peaks are remarkably different from the distribution observed over the past 12 weeks (37). The user

can then choose to further investigate the data by patient category, clinic type, age range, and branch of service. The epidemiological significance of unusual findings and the actions needed are ultimately determined by the user.

Reportable Medical Event Query based on ICD Codes

ESSENCE permits users to investigate reportable disease events through ad hoc queries based on ICD codes. There are no recurring reports generated centrally from ESSENCE.

DRSi

I. Background

For the later part of the 1990s, the US Army reported diseases and medical conditions of public health importance with a system called Reportable Medical Events System (RMES) (40). The Disease Epidemiology Program of the U.S. Army Public Health Command (USAPHC), responsible for conducting surveillance and follow-up of all reportable medical conditions within the U.S. Army, identified during the H1N1 influenza pandemic and other outbreaks in the early 2000s that the RMES lacks the desired flexibility, completeness, and timeliness to effectively monitor and report disease information (40). The operating characteristics and cost effectiveness of the Disease Reporting System-internet "DRSi'"; Congressional and DoD requirements to employ joint/shared technology solutions (Navy's use of DRSi); and the decommissioning of RMES as part of the upgrade of the Defense Medical Surveillance System (DMSS), led to the adoption of DRSi Army-wide in October of 2010 and by the Air Force in January 2014 (40).

II. Purpose, Justification, Objectives

The purpose of DRSi is to provide a user friendly web-based reporting system for Reportable Medical Events. The system is justified by the need for a more flexible, complete, and timely system; Congressional and DoD mandate to employ joint/shared technology (Navy DRSi); and the decommissioning of previous system (RMES). In addition,

the ease of operation of the new system was believed to increase the number of reports submitted and number of facilities reporting cases, thereby improve surveillance efforts. The system has three main objectives 1) web-based reporting of reportable medical events (RMEs), 2) tracking of outbreaks, and 3) tracking of sexually transmitted infections (40).

III. Operations

A. Operating Case Definitions

Cases in DRSi are specifically defined by the Armed Forces Reportable Medical Events Guidelines and Case Definitions. Currently DRSi tracks 66 reportable conditions (41, 42). Military medical facilities have the capability to augment the list with diseases of local concern for additional surveillance if needed. Initially, the data in DRSi were driven by the ICD diagnostic codes from Medical Event Reports (MERs) (41, 42). Since 2011, the disease name as specified in the Guidelines is used instead. In addition DRSi has a case finding (CF) module that can pull cases based on laboratory data indicative of a reportable event from the HL7 data feeds. DRSi data includes levels of certainty for each case; suspect, probable, confirmed, pending, not confirmed (41, 42).

B. Target Population

The DRSi aims to collect reportable medical event data for all individuals who receive care at a MTF including all branches of Service, active duty, Reserve Component (while on orders), dependents, and eligible retirees. No information is collected for health encounters that occur outside the military system; there is no data for reportable events that occur at civilian-TMA or VA hospitals (41, 42).

C. Data Sources and Integration with Other Systems

DRSi does not receive any automated data from electronic medical records; all of the data is manually entered and uploaded (42). There are three ways in which a health encounter can become a case in DRSi. One way is for MTF

staff to submit the potential RME to the Preventive Medicine department for determination of medical event report status; if the case meets the criteria the PM staff converts the entry to a MER. Another way is when the patient record is directly entered as a MER; PM staff typically still reviews the data before it becomes final. Lastly, DRSi's case finding module allows for the querying of laboratory data from HL7. Then cases that meet the Armed Forces RME guidelines can be pulled into the case finding module in order to facilitate the creation of a MER if the case is determined to meet the RME case definition. At the time of this printing, DRSi did not yet have the capability to automatically pull demographic data directly from DEERs (cross-referenced by patient social security number) (42).

D. Data Management and Quality Assurance

The Disease Epidemiology Program of the U.S. Army Public Health Command (USAPHC) is responsible for conducting surveillance and follow-up of all reportable medical conditions within the U.S. Army. In order to ensure the appropriate capture of RMEs, all MTFs are required to have at least two trained DRSi users within their Preventive Medicine departments (41). These PM DRSi users are responsible for reviewing and approving all RMEs before they are submitted as a MER, contacting the author of the report or consulting patient records when data quality concerns arise (42). Data is hand-entered into the system and entry errors are possible; however, as long as the report has not been converted to a MER, editing is fairly easy and can be done at all levels. In addition, all non-sexually transmitted infection cases are reviewed by Disease Epidemiology staff to ensure data quality and integrity. The RMEs are required to be submitted within 48 hours of diagnosis; there is a potential 1-7 day lag time from submission to appearance in DRSi (42). In order to prevent double counting of cases, the MERs are linked to patient social security numbers (42).

E. Data Analysis

The data in DRSi are only available as case counts; the number of MERs per identified condition. There is no denominator data in DRSi, therefore it is not possible to calculate person-years and/or rates (42). Because DRSi has access to all of the demographic data available in DEERS, extensive trend analysis of the data by demographic categories is possible. All data can also be analyzed temporally (by year and month), spatially (by Service

component, MTF, or installation), and by level of diagnostic certainty (probable, suspect, confirmed)(42). Submission trends can also be monitored, serving as a way to track conformity to RME submission guidelines.

F. Data Dissemination and Reporting Options

All access to DRSi requires Common Access Cards (CAC) and is therefore limited to active-duty military personnel, selected Reserve personnel, DoD civilian employees, and eligible contractor personnel. There are two possible access levels; recorder access and supervisory access. Clinic and hospital staff generally has recorder access, while the Preventive Medicine staff has supervisory access. The system is primarily driven by pull-down windows as opposed to free-text fields, therefore user data entry errors are minimized (41, 42).

DRSi Ad Hoc Queries

Those with access (generally MTF PM and Regional Command staff) have the ability to perform queries and view summary reports (42). Queries are not limited to the RMEs defined by the Tri-Service Guidelines; users can expand their queries to incorporate diseases of particular interest in their region or as directed by command staff. Summary reports generated by DRSi can be in the form of excel spreadsheets, bar graphs, and pie-charts (42).

Reportable Medical Events Summary Reports

Disease Epidemiology Program staff from the U.S. Army Public Health Command (USAPHC) generate both daily and monthly Reportable Medical Event Reports (42). Report recipients include all registered DRSi users. The reports are limited to the RMEs as defined in the Armed Forces Reportable Medical Events Guidelines and Case Definitions. Daily reports include details on the reporting location, condition and diagnostic status (suspect, confirmed, pending, etc.) visit date, associated branch of service, service component (active duty, Reserve, retiree, dependent), as well as patient age and sex. Monthly reports include year- to-date total case count comparisons for each RME, trend analyses on selected diseases, and a break-out of the number of RMEs by condition per installation for the entire month (42).

Table 3.2: Comparison of Key Aspects of Military Data Systems

Data System[1]	Surveillance Category	Medical Facility Sources[2]	Target Population[3]	Medical Data Sources[4]	Non-Medical Data Sources[5]	Limitations	Strengths
MDR	Medical	MTF Civilian facilities	All DEERS eligible personnel	In-Patient Out-Patient Ancillary medical services	DEERS	Lacks RME data	Complete population and medical data M2
DMSS	Medical	MTF Civilian facilities Deployed medical facilities	Service Members	In-Patient Out-Patient RMEs TMDS DODSR	DMDC Pre and post deployment health assessment surveys	Incomplete non-SM population data Requires 3rd party (no direct access)	Longitudinal data on SMs (occupational, deployed, medical) DODSR DMED
HL7	Ancillary medical services (laboratory)	MTF	All DEERS eligible personnel	CHCS/AHLTA Laboratory Radiology Pharmacy		Potential for missed cases (low sensitivity)	Timely data Ability to confirm cases (high specificity)
ESSENCE	Syndromic	MTF	All DEERS eligible personnel	Out-Patient +/- HL7		Potential for false positives (low specificity)	Potential to detect trend before outbreak (high sensitivity) Ability to generate e-mail alerts to users
DRSi	Medical Event Reports (MERs; passive surveillance)	MTF (Army, Navy +/- Air Force)	All DEERS eligible personnel	Medical Event Reports +/- HL7	DEERS	Passive surveillance (relies on manual input of RMEs) RMEs only Army data only	User-friendly

1. Data Systems: MDR (Medical Data Repository); DMSS (Defense Medial Surveillance System); HL7 (Health Level 7); ESSENCE (Electronic Surveillance System for the Early Notification of Community-based Epidemics); DRSi (Disease Reporting System Internet)
2. Medical Facility Sources: MTF (Military Treatment Facility); Civilian Facility are those approved by the Military Health System as purchased care (claims data)
3. Target Population: DEERS eligible includes all Service Members and Retirees as well as their dependents

4. Medical Data Sources: RMEs (Reportable Medical Reports); TMDS (Theater Medical Data Store); DODSR (Department of Defense Serum Repository); CHCS/AHLTA (Composite Health Care System/Armed Forces Health Longitudinal Technology Application), the military's electronic medical record system)
5. Non-Medical Data Sources: DMDC (Defense Manpower Data Center); Neither HL7 nor ESSENCE have access to non-medical data sources, non-medical data is limited to that contained in the medical record.

References

1. Health Topics: Public Health Surveillance posted at the World Health Organization (WHO) website accessed January 23, 2014 at http://www.who.int/topics/public_health_surveillance/en/

2. Department of Defense Directive (DODD) 6490.02E. Washington DC Headquarters, (8 Feb 2012).

3. Centers for Disease Control and Prevention. Updated guidelines for evaluating public health surveillance systems: recommendations from the guidelines working group. MMWR 2001;50 (No.RR-13) Atlanta, Georgia, Centers for Disease Control and Prevention.

4. Institute of Medicine. 2007. "Global Infectious Disease Surveillance and Detection: Assessing the Challenges—finding solutions", Workshop Summary. Washington DC: The National Academies Press.

5. The Global Health Observatory (GHO) portal: http://www.who.int/gho, accessed 15 March 2014.

6. See: http://apps.who.int/gho/indicatorregistry/App_Main/browse_indicators.aspx, accessed 10 March 2013.

7. World Health Statistics, 2013. World Health Press. World Health Organization. www.who.int accessed 10 March 2014.

8. http://www.fao.org/; accessMarch 2014.

9. EMPRES-i flier, http://www.fao.org/; access March 2014.

10. EMPRES-I, Global Animal Disease Information System User Guidelines, Version 3.0, October 2012, http://EMPRES-i.fao.org; access 19 March 2014.

11. Martin, V. et al. Early Warning, Database, and Infection Surveillance for Avian Influenza Surveillance. J of Wildlife Diseases. 43(3): Supplemental 2007:71-76.

12. http://www.oie.int/ accessed March 2014.

13. 2013 © OIE - *Terrestrial Animal Health Code*

14. http://www.aphis.usda.gov/

15. NAHSS Strategic Plan, 2005;
http://www.aphis.usda.gov/vs/nahss/docs/NAHSS_Strategic_Plan_2005_0216.pdf

16. NAHSS Brochure, 2008; http://www.aphis.usda.gov/vs/nahss/docs/nahss_brochure_2008.pdf

17. extract from SCS Statement of Work for USDA APHIS VS
18. USDA VS SCS News Issue 2, July 2012

19. USDA APHIS VS Information briefing "Getting Started with Cognos Reports for Surveillance Collaboration Services (SCS). Document Version Jan 18, 2013.

20. Gabriel RA and Metz K. A History of Military Medicine, Vol. II, From the Renaissance Through Modern Times. Westport, Connecticut: Greenwood Press, 1992.

21. Jones, C. Unification of the Military Health System: A Half-Century of Unresolved Debate. Strategic Research Project for Senior Service College. U.S, Army War College, Carlisle Barracks, PA. 20 July 2001.

22. Ruberton MV and Brundage JF (2002). The Defense Medical Surveillance System and Department of Defense Serum Repository: Glimpses of the Future of Public Health Surveillance. Am. J. of Public Health; 92(12):1900-1904.

23. About MHS from the Health.mil Official Website for Military Health Medicine accessed on January 10 2014 at http://health.mil/About-MHS.

24. The Military Health System: Framework for National Health Information Exchange Poster, Developed by the Office of the MHS Chief Information Officer, February, 2007

25. MDR Information Briefing 2012 accessed from the Health Military System website on 20 December 2012 at http://www.health.mil/MHSCIO/programs_products/DHSS/DHSS-Products/MDR.aspx

26. Personal interviews with Nikki Jordan, December 2012, Epidemiologist, U.S. Army Public Health Command, Army Institute of Public Health, Epidemiology and Disease Surveillance Portfolio, email: nikki.jordan@us.army.mil.

27. M2 Factsheet posted on the Military Health System website accessed January 13, 2013 at http://www.health.mil/MHSCIO/programs_products/DHSS/DHSS-Products/M2.aspx

28. M2 Information Briefing posted on the Military Health System website accessed January 13, 2013 at http://www.health.mil/MHSCIO/programs_products/DHSS/DHSS-Products/M2.aspx

29. DeFraites, Robert F. The Armed Forces Health Surveillance Center: Enhancing the Military Health System's Public Health Capabilities. BioMed Centrtal Public Health. 2011; 11(Suppl 2)S1 (http://www.biomedcentral.com/1471-2458/11/S2/S1).

30. Armed Forces Reportable Medical Events Guidelines and Case Definitions. Armed Forces Health Surveillance Center (AFHSC), March 2012.

31. Personal interview with LCDR Rhonda Lizewski, Armed Forces Health Surveillance Center, March 2013, e-mail: rhonda.a.lizewski.mil@mail.mil

32. AFHSC website http://www.afhsc.mil/dodsr accessed April 26, 2013.

33. Navy and Marine Corps Public Health Center. HL7 Laboratory — Chemistry Technical Document. EpiData Center Department. September 2012.

34. Meyer et al.: Antimicrobial resistance surveillance in the AFHSC-GEIS network. BMC Public Health 2011 11(Suppl 2):S8.

35 . Nowak, G. Description of the MHS Health Level 7 Chemistry Laboratory for Public Health Surveillance. Defense Technical Information Center, 22 Jan 2014)

36. Personal interview with Gosia Nowak, Infectious Disease Epidemiologist, Navy and Marine Corps Public Health Center, Epi Data Center, e-mail: gosia.nowak@med.navy.mil

37. ESSENCE Information Briefing by U.S. Army Public Health Command, Army Institute of Public Health, Epidemiology and Disease Surveillance Portfolio, updated September 12, 2013.

38. ESSENCE Factsheet posted on the Military Health System website accessed December 10, 2013 at http://www.health.mil/MHSCIO/programs_products/DHSS/DHSS-Products/M2.aspx

39. Personal interview with Paul Pietrusiak, January 2013, Epidemiologist, U.S. Army Public Health Command, Army Institute of Public Health, Epidemiology and Disease Surveillance Portfolio, email: paul.pietrusiak@us.army.mil.

40. OPERATION ORDER 10-78 (DISEASE REPORTING SYSTEM-INTERNET). HEADQUARTERS, US ARMY MEDICAL COMMAND. Fort Sam Houston, TX 78234-6007. 81600Q September 2010

41. DRSi for HPOs Information Briefing by U.S. Army Public Health Command, Army Institute of Public Health, Epidemiology and Disease Surveillance Portfolio, updated April 2012.

42. Personal interview with Zachary McCormic, Epidemiologist, U.S. Army Public Health Command, Army Institute of Public Health, Epidemiology and Disease Surveillance Portfolio, email: zachary.mccormic@us.army.mil.

Appendices

Appendix A-3:

Standardized Outline for Evaluating Military Surveillance Systems

Surveillance Database Description
(Modified from CDC Guidelines)

I. Purpose, Justification, Objectives
   - A. Purpose
   - B. Justification
   - C. Objectives

II. Operations
   - A. Operating Case Definitions (inclusion/exclusion criteria, syndromic, laboratory, epidemiological, levels of certainty- probably, suspect, confirmed)
   - B. System Components
     1. Population
     2. Data Sources
     3. Integration with Other Systems
     4. Data Management (Input, Editing, Data Quality)
     5. Data Analysis
     6. Data Dissemination and Reporting Options
   - C. Flow Chart of System (Collection through Reporting to include time required)

III. Resource Requirements
   - A. Informatics Requirements (Accessibility)
   - B. Human Resource Requirement = User-friendliness, efficiency of system

Surveillance Database Performance Evaluation
(Modified from CDC Guidelines)

I. Utility: Level of Usefulness (timely, appropriate estimators, trend detection)
   A. As related to stated objectives
   B. As determined by users
II. Simplicity: Structure and Ease of Operation
   A. Method of collecting data including number and type of reporting sources
   B. Level of integration with other systems
   C. Method for data analysis
   D. Method for data dissemination
   E. Time required to maintain and update system
   F. Staff training requirement
III. Flexibility: Ability of System to Adapt and Upgrade
IV. Accuracy: Data Quality: Completeness and Validity of Recorded Data
   A. Refer to Case Definition (Specificity vs. Sensitivity)
   B. Representativeness of Data (Population Sampled, Active/Passive Surveillance, Voluntary vs. Directed data collection)
   C. Data entry (potential to make mistakes)
V. Acceptability: Willingness to use the System
VI. Sensitivity
   A. Ability to detect cases
   B. Ability to monitor changes in number of cases over time
   C. Sensitivity Measurement: compared to area prevalence data and/or validation of data collected
   D. Note: Increase sensitivity artifacts include heightened awareness of potential for disease, improved diagnostic testing, broadened case definitions
VII. Positive Predictive Value: Proportion of reported cases with actual case defined event
   A. Historical data on accurate identification of outbreaks, trends (high PPV)
VIII. Representativeness: Ability to accurately describe health related event temporally, spatially and demographically.
   A. Demographic breakout
   B. Temporal breakout
   C. Spatial breakout
   D. Denominator Information (same source, comparable across categories, consistent over time)
IX. Timeliness: speed in between steps SPEED TO GENERATING USEFUL REPORT
X. Stability: Reliability and Availability of System- Accessability
   A. System Downtime, Firewall Issues
   B. Manpower constraints

Chapter 4: Comparison of the Human Medical Data Systems Most Commonly Used for Public Health Surveillance in the U.S. Army

Introduction

Effective 22 July 2011, under the direction of the U.S. Army Surgeon General, the former U.S. Army Veterinary Command (VETCOM) joined with the former U.S. Army Center for Health Promotion and Preventive Medicine (USACHPPM) to create the new U.S. Army Public Health Command (USAPHC) (1). The mission statement for this new Command is: "Promote health and prevent disease, injury, and disability of Soldiers and military retirees, their Families, and Department of the Army civilian employees; and assure effective execution of full spectrum veterinary service for Army and Department of Defense Veterinary missions"(2). The command combines the technical skills of multiple disciplines, including veterinarians, physicians, entomologists, laboratory specialists, epidemiologists, and many others in order to achieve the USAPHC mission.

In 2011, the Epidemiology and Disease Surveillance Portfolio, Disease Epidemiology Program within the USAPHC initiated development of a Zoonotic Disease Report (ZDR) to provide U.S. Army public health personnel with critical health information regarding the presence and spread of zoonotic pathogens and to create opportunities for improved preventive medicine strategies. The audience of the ZDR is U.S. Army Public Health Command Regional and District level Commanders, as well as veterinarians and public health and preventive medicine professionals throughout the Army. The report combines zoonotic disease information from diverse sources, including: extant Army/DoD human disease databases, data generated by the Laboratory Services Portfolio for arthropod borne disease surveillance as well as rabies specimen testing, and animal data from intergovernmental organization's public access databases. The goal is that the animal data will eventually include zoonotic diseases diagnosed at U.S. Army Veterinary Treatment Facilities once the web-based electronic medical record known as the Remote Online Veterinary Record (ROVR) is fully deployed. Because of the existence of multiple human health and disease databases, the first step in creating the ZDR was to determine which database(s) best fit the needs of the ZDR.

The aim of this study was to conduct a systematic comparison of the human medical data systems most commonly used for public health surveillance in the U.S. Army in order to identify the one(s) most suited for use in the USAPHC ZDR.

Background

In accordance with Department of Defense (DoD) Directive 6490.02E (8 Feb 2012) military public health surveillance shall be conducted to enable early intervention and control strategies to prevent adverse effects on mission accomplishment and to provide information for shaping commanders' decision making (3). The primary objectives of these military medical surveillance systems are timely and proficient public health response to medical events to reduce morbidity and mortality, to estimate the distribution, trends, and risks associated with significant medical events, and to aid in the development and assessment of policy and resource allocation for the prevention and control of communicable diseases (3).

The collection of medical event data is a cornerstone of public health surveillance. The use of existing electronic data streams associated with hospital and provider systems is one option for the collection of such data (4). In the DoD, the Defense Health Agency (DHA), previously the Military Health System (MHS), is the overarching enterprise that provides health care to all military beneficiaries (5). Within the DHA are multiple electronic data streams, to include medical, personnel, deployment, and financial data for all branches of the military (6). Specific hospital and personnel data are consolidated into the MHS Data Repository (MDR). MDR serves as the centralized data repository for DoD beneficiaries, containing both personnel and medical data such as outpatient visits, inpatient admissions, and laboratory orders (6). The MDR also integrates business, customer, and financial data, to include purchased care claims for services received outside of the DHA. Several military human medical data systems utilize the available data in MDR; however, exactly which data are used and how the collected data are then analyzed varies greatly between these systems. For example, the focus of Defense Medical Surveillance System (DMSS) is medical surveillance in Service members and therefore the system compiles data on their medical events, medical readiness , and overall "military experience" (7). In contrast, the Military Health System Management and Reporting Tool (M2) was designed as a health system management tool and therefore draws from population,

clinical, and financial data (specifically claims, eligibility, and enrollment data) (8, 9, 10). Because the goal of the Electronic Surveillance System for the Early Notification of Community-Based Epidemics (ESSENCE) is the automated detection of epidemic outbreaks based on syndromes, the system uses outpatient and inpatient encounter data from MDR (11, 12, 13).

Evaluating Military Human Medical Data Systems

According to the Centers for Disease Control and Prevention (CDC), the purpose of evaluating public health surveillance systems is to ensure that problems of public health importance are being monitored efficiently and effectively (14). In 2001, the CDC published updated guidelines for evaluating public health surveillance systems and in 2004 a framework for evaluating public health surveillance systems for early detection of outbreaks (3, 14). Although the current study is evaluating medical data systems rather than health surveillance systems, the tasks detailed in these reports are still applicable and guided the military human medical data system comparison. The CDC reports emphasized the importance of establishing a standardized framework when conducting an evaluation comparing multiple data systems. The framework used should include the following three components: 1) detailed system descriptions, 2) a comparison of specific data systems attributes, and 3) an evaluation of each system's data completeness.

Chapter three of this document provides detailed descriptions of the five most commonly used military human medical data systems. The reader is referred to this chapter for more details. Each medical data system is described briefly below.

A Brief Description of the Human Medical Data Systems Most Commonly Used by the U.S. Army

The Electronic Surveillance System for the Early Notification of Community-Based Epidemics (ESSENCE) system is a syndromic web-based surveillance system that uses clinical data from the DHA to monitor and generate alerts for rapid or unusual increases in groups of diagnostic categories that represent specific syndromes or disease

categories of concern such as Influenza-like illness, hemorrhagic illness, and botulism-like (11, 13). The ESSENCE system is designed to utilize not only health encounter diagnoses, but also chief complaint information entered as free text, and ancillary hospital data from pharmacy, laboratory, and radiology departments (11, 12, 13). All of the data in ESSENCE come from electronic medical records from fixed military treatment facilities (MTF). Information is clustered at the installation level; all branches of Service are represented (11, 13).

The Disease Reporting System-internet (DRSi) system is the Army's web-based reporting system for Reportable Medical Events (RMEs) (15, 16). The RME case definitions are described in the Armed Forces Reportable Medical Events Guidelines and Case Definitions; currently 66 RMEs are reportable within DRSi. All RME case reports are generated by personnel at (MTFs); therefore cases diagnosed outside of the military health system (e.g., purchased care) are typically not captured (15, 16). Once initiated, the case report can be reviewed by hospital preventive medicine personnel and personnel from the Disease Epidemiology Program at the USAPHC Army Institute of Public Health (15). Data in DRSi are used to identify and track disease outbreaks and perform RME trend analysis at the installation and regional level (15, 16).

Health Level 7 (HL7) technically refers to the standardized format used by electronic hospital information systems that facilitates the exchange, integration, sharing, and retrieval of electronic health data (17); the data from these information systems are referred to as HL7 data. The HL7 data in the military system come from services related to patient encounters such as outpatient visits, clinic appointments, pharmacy orders and fills, radiology exams, laboratory test orders and results, and inpatient admissions and discharges (18). Algorithms and tools developed by the Navy and Marine Corps Public Health Center (NMCPHC) permit the interpretation of this text string HL7 data to augment and improve military public health surveillance activities (19, 20). Although all Service branches are represented in HL7 data, it only includes health care events that occurred at fixed facility; off-shore (Navy fleet hospital) or temporary MTF data are not captured (20).

The Defense Medical Surveillance System (DMSS) is the designated central repository of medical surveillance data for the U.S. Armed Forces (7). The target population for DMSS is Service members; data for dependents (military family members) and other eligible beneficiaries is limited to medical data only (7, 21). One of its primary

objectives is to create longitudinal records on service members, documenting their military and medical experience throughout their military career (7). In addition to direct data feeds from MDR, DMSS receives Reportable Medical Event data from all Service branches and is linked to the Department of Defense Serum Repository (DoDSR), a central archive of sera drawn from service members specifically for medical surveillance purposes (21). Although not yet routinely used, DMSS has had direct access to laboratory data (HL7) since 2010 (7, 21).

The Military Health System Management Analysis and Reporting Tool (M2) merges several stand-alone health care management applications into a single platform (8, 9). Although the purpose of M2 was improved business intelligence reporting, the completeness of the system's medical data feeds and sources makes it well-suited for health surveillance purposes as well (10). The M2 captures data equally for service member and beneficiary populations from both military and purchased care from civilian health care facilities (8, 9, 10).

When comparing specific medical data systems, the focus should be on how well the data systems operate to meet the stated purpose and objectives of the given public health program. The remainder of this chapter is dedicated to the last two components of the medical data system evaluation framework: a comparison of specific data systems attributes (Section I), and an evaluation of each system's data completeness (Section II). The evaluations done in each of these sections are sequential; based on the findings of Section I, some data systems did not proceed to the evaluation conducted in Section II.

Section I: A Comparison of Military Medical Data System Attributes

Medical data systems vary in their purpose and objectives, thus their attributes of each system vary. To conduct this evaluation, the specific goals of the ZDR were used to determine which data system attributes to focus on for comparison purposes. The goal of the ZDR is to provide commanders and preventive medicine and public health personnel with timely and accurate zoonotic disease data relevant to their military communities that can be refined both spatially and temporally. The ability to meet this goal was evaluated by determining 1) how well each system

matched the specific ZDR criteria of population source, data source, and data types, 2) how well each system met the specific ZDR query requirements, and 3) how well each system truly performs when queried.

Zoonotic Disease Report Criteria

The attributes of medical data systems are a reflection the particular data system's purpose, objectives, or goals; therefore an understanding of each system's goals is important in determining its suitability for use in the ZDR. Systems with goals similar to that of the ZDR, as stated above, may be most appropriate. To best represent the full military community, the target population should include Service members, dependents, and other beneficiaries. To accurately reflect the burden of disease, the data should be as comprehensive and complete as possible with data sources including both military medical treatment facilities and civilian care facilities (purchased care), and data types including those of inpatient, outpatient, ambulatory, laboratory, and demographic records.

Zoonotic Disease Report Query Requirements

The ZDR will be generated from the portfolio level, the level that advises regional commands, USAPHC Epidemiology and Disease Surveillance (EDS) Portfolio personnel should have direct access to the data. Being able to conduct an agreed upon query themselves ensures the quality, reliability, accuracy, and timeliness of each data extract. In addition, this access must include visibility of unique identifiers for each case. This level of access ensures the quality of each query, as it provides USAPHC with the ability to validate the data. In order to provide accurate health information for the target population, the population data should include details on Service members, dependents, and beneficiaries. Lastly, in order to be as representative and accurate as possible assess disease burden, the system should receive medical data from data sources and types as described above.

System Performance

The medical data system used must meet certain performance criteria in order to provide the ZDR stakeholders with timely and accurate zoonotic disease information. . For example, the system's accessibility must be reliable, with

minimal “down time” or system glitches that serve as barriers to the system’s ability to fill query requests. In order for the data to be timely, data retrieval should be quick and relatively simple to operate (if accessed directly) or require minimal communication with those conducting the query (if direct access is not possible).

Methods

Ability to Meet ZDR Criteria

A standardized template adapted from the Centers for Disease Control and Prevention (CDC) Guidelines for Evaluating a Public Health Surveillance System was created in order to compile information gathered about each system. The template emphasizes the objectives and criteria of the U.S. Army Public Health Command Zoonotic Disease Report.

Specific categories include: system goals, target population, data sources, and data types.

1. System goals refers to the stated purpose, objectives, and goals of the particular medical data system.
    - The goals of the ZDR are timely and accurate zoonotic disease data for the desired military community that can be refined both spatially and temporally.
2. Target population refers who the particular medical data system is able to gather complete data from.
    - The target population for the ZDR includes Service members, their dependents, and other beneficiaries for all branches of service.
3. Data sources refers to the electronic data streams the particular medical data system has accesses.
    - To be a comprehensive as possible, the medical data system for the ZDR should include medical and personnel data for the full target population.
4. Data types refers to the specific data product the medical data system uses; for example ICD codes, laboratory results, hospital orders, or medical event reports.
    - To be a comprehensive as possible, the medical data system for the ZDR should use a variety of data types.

In order to construct a balanced and reliable description of each system, the following sources of information were used: interviews with system end-users, published literature, and personal database exploration.

Ability to Meet ZDR Query Requirements

System queries were conducted in order to objectively assess each system's ability to meet the query requirement of the ZDR. In order to assure equal comparability each system was queried for the same time period (1 April 2011 to 30 April 2012) and for the same four zoonotic diseases.

Zoonotic Disease Selection: Stakeholder Survey

In an effort to determine the disease information and reporting needs of both military and civilian organizations working to improve human health, the USAPHC Zoonotic Disease Monitoring Working Group created stakeholder survey. Of the 25 individuals that responded to the survey, 13 were military, 11 were civilian, 4 were from the CDC Zoonotic Disease Working Group, and 1 was of unknown affiliation. Sixteen of the 25 (64%) of the respondents were veterinarians, making up the majority of the population. The survey was distributed using the online survey using Vovici Survey Software, version 5.5, and was conducted from 11 May 2012 at 1500 until 31 May 2012 at 1700. Results were compiled by the USAPHC Zoonotic Disease Monitoring Working Group. Questions ranged from overall satisfaction with current zoonotic disease monitoring and reporting practices to recommendations for improved reporting methods and practices. The sections geared towards determining the zoonotic pathogens of highest concern both regionally and globally were used to aid in selecting the four zoonotic diseases for use in the systems queries conducted in this study.

Each of the five military human medical data systems were queried for case count data on the following four zoonotic diseases: leishmaniosis, hantavirus, leptospirosis, and borreliosis. All four diseases are considered reportable medical events by the Armed Forces Reportable Medical Events Guidelines (22). For the purposes of comparison, attempts were made to make the queries as standardized as possible within each system. However, because of differences inherent in each system, each system query process is outlined below.

Query Process

The USAPHC Epidemiology and Disease Surveillance Portfolio personnel have direct access to the DRSi, M2, and ESSENCE data systems. This permitted query requests to remain informal and internal to the study team. Obtaining the data contained within the DMSS and HL7 systems required formal request processes. Case definitions were based on the Armed Forces Reportable Medical Events Guidelines (22). A detailed listing of these case definitions is located in **Appendix A-4**.

In addition, the following parameters were applied to all data queries performed or requested:

1. All cases of hantavirus, leptospirosis, leishmaniosis, and borreliosis that met the case definition.
2. Population to include all Service members and other beneficiaries.
3. Limit cases counts to those associated with Army installations.
4. Limit to cases reported or diagnosed and laboratory tests from 1 April 2011 to 30 April 2012.

Query Traits

The information gathered from the query process was used to assess each system's ability to satisfy the ZDR's query requirements. Specifically, the following query traits were evaluated: accessibility, ability to validate, population data, and medical data. Accessibility refers to the ability of a system to support USAPHC personnel regularly successfully constructing and running queries, and exporting the resultant data. Unique Identifiers refers to the ability of USAPHC personnel to retrieve medical data retaining some form of a Personal Identifier (i.e., a form of Personally Identifiable Information (PII)) that is unique to the patient such as a social security number. Population data refers to the capability of the system to produce denominator data that represent the ZDR target population: Service members and other beneficiaries obtaining care at both military and civilian medical facilities. The last category, medical data, refers to the comprehensiveness of the medical data sources (laboratory, inpatient records, outpatient records, reportable medical event case reports) specifically in relation to the ZDR target population. Weighting of query trait categories was performed in order to appropriately emphasize traits that are critical to the success of the ZDR. Accessibility and unique identifiers were determined by USAPHC personnel to be critical and

therefore each assigned weights of six. Because the population data category is made up of four subcategories (Service members at military medical facilities, Service members at civilian medical facilities, other beneficiaries at military medical facilities, and other beneficiaries at civilian medical facilities), the number of points each system received in this category represents the number of subcategories the system represents. The medical data category also uses these four subcategories as they apply to laboratory, inpatient, outpatient, and reportable medical event case report data (4 medical data categories, each with a total of 4 subcategories, for a total of 16 points).

System Performance

The medical data systems queries were used to assess the performance of each system. Assessment criteria included ability of system to fill the query request, speed at which the query was able to be completed, and the amount of communication required to ensure the query was accurately completed,

Results

Ability to Meet ZDR Criteria

Each medical event data collection system has a unique purpose, justification, and objective. These differences impact how data are collected, compiled, analyzed, and extracted. However, in order to produce a successful Zoonotic Disease Report, each system was evaluated based on its ability to meet the specified system criteria of the ZDR (Table 4.1). None of the systems met all four of the ZDR criteria. ESSENCE was able to meet three of the four, having similar goals, capturing the target population and employing the largest range of data types. However, as discussed further below, ESSENCE was unable to perform the requested queries, which prevented its recommendation for inclusion in the ZDR. DMSS satisfied three of the four ZDR criteria; however, the target population for DMSS is the Service member (data may not be complete for other beneficiaries), which does not match the full ZDR target population, making its future use in the ZDR questionable. M2 also met three of the four ZDR criteria; however, it lacks reportable medical event case report data which is a necessary component of the data type category. The strength of the other three categories however, justified pursuing further evaluation of the M2

system. Although the data compiled from HL7 laboratory data sources can be used to augment and improve military public health surveillance activities, the data are not intended to serve as a stand-alone medical surveillance system. As such, it was not surprising that HL7 did not meet most of the criteria of the ZDR; however, the uniqueness of the data from HL7 still made it a viable option for the ZDR.

Table 4.1: Meets Zoonotic Disease Report (ZDR) Criteria

	HL7	DMSS	DRSi	M2	ESSENCE
System Goals[1]	No	Yes	No	Yes	Yes
Target Population[2]	Yes	No	Yes	Yes	Yes
Data Sources[3]	No	Yes	No	Yes	No
Data Types[4]	No	Yes	No	No	Yes

1. System goals refers to the stated purpose, objectives, and goals of the particular medical data system.
   The goals of the ZDR are timely and accurate zoonotic disease data for the desired military community that can be refined both spatially and temporally.
2. Target population refers who the particular medical data system is able to gather complete data from.
   The target population for the ZDR includes Service members, their dependents, and other beneficiaries for all branches of service.
3. Data sources refers to the electronic data streams the particular medical data system has accesses.
   To be a comprehensive as possible, the medical data system for the ZDR should include medical and personnel data for the full target population.
4. Data types refers to the specific data product the medical data system uses; for example ICD codes, laboratory results, hospital orders, or medical event reports.
   To be a comprehensive as possible, the medical data system for the ZDR should use a variety of data types.

Ability to Meet ZDR Query Requirements

The point allocation displayed in Table 4.2 mirrors the ability of each data system to meet ZDR query requirements. The inability of either DMSS or HL7 to be independently queried by USAPHC personnel greatly inhibits their usefulness in creating the ZDR. Although the DMSS data did eventually include unique identifiers, the initial lack of unique identifiers and the subsequent effort required to obtain them further reduced the system's utility for use in the ZDR. ESSENCE has the potential to meet several of the ZDR query criteria; however, the requested query was never completed because of system issues such as geographic limitations, data range limitations, and internal query errors. Although these issues eliminated it from further analysis and prevented its recommendation for inclusion in the ZDR, anticipated future updates to the ESSENCE system may render it a viable an option that could be reassessed at a later date. The M2 data system is noted to have the highest point allocation. Furthermore, the two

areas of deficiency with M2 (laboratory data and reportable medical event case report data) could be supplemented with uniquely identified data supplied by HL7 and DRSi.

Table 4.2: Ability to Meet Required Zoonotic Disease Report Query Traits

	HL7	DMSS	DRSi	M2	ESSENCE
Ability to Query Data Independently* (Accessibility)	0	0	6	6	6
Unique Identifier Data * (Ability to Validate Data)	6	0	6	6	6
Representativeness of Denominator Data+ (Population Data)	0	2	2	4	2
Laboratory Data^ (Medical Data)	2	0	0	0	0
Inpatient Data^ (Medical Data)	0	4	0	4	4
Outpatient Data^ (Medical Data)	0	4	0	4	4
Reportable Medical Event Data^ (Medical Data)	0	4	4	0	0
Total	**8**	**14**	**18**	**24**	**22**

*"Accessibility" and "Ability to Validate" categories, were assigned weights of 6 reflecting the relative importance of these system attributes; systems meeting the requirement received 6 points, those that did not received 0 points.
+ Population Data are comprised of 4 subcategories (Service members (SM) at military medical facilities, SM at civilian medical facilities, beneficiaries at military medical facilities, and beneficiaries at civilian medical facilities); each subcategory receives 1 point.
^ Medical Data are comprised of 4 subcategories (SM at military medical facilities, SM at civilian medical facilities, beneficiaries at military medical facilities, and beneficiaries at civilian medical facilities) as they apply to the laboratory, inpatient, outpatient, and reportable medical event report data sources (for a total of 16 points).

Assessment of System Performance

ESSENCE

Because ESSENCE was unable to successfully run the requested queries it was not included in this phase of the evaluation

HL7

HL7 cases are those that were confirmed positive via laboratory methods and identified via pre-existing case-finding algorithms; although the query request required formal submission to and approval by NMCPHC, the process was very easy and the data were provided within a timely manner. In addition, the amount of details provided in the data allowed for internal quality assurance checks and eliminated the need for any clarification or follow up

communication. Because of the uniqueness of the data provided by this system, HL7 was retained for further analysis.

DRSI

By design, DRSi cases only include reportable medical event case reports that are generated and submitted within the DRSi system. Data retrieval was performed internally by USAPHC and was supplied very quickly. Because of the ease in retrieving DRSi data and the uniqueness of this data source, DRSi was retained for further analysis.

DMSS

Although accessing data from DMSS requires a formal request process, the protocol is very clear and support staff is very helpful and informative. However, as issues arose, the formal process became cumbersome. One of the most significant issues noted was the incompleteness of data as identified by the large number of missing RMEs. This issue had been recognized earlier and was thought to have been rectified; however, this issue remained unresolved at the time of our query. This issue coupled with the initial lack of unique identifiers and the inability of USAPHC personnel to perform the system queries led to the elimination of the DMSS system from further consideration as a data source for the ZDR.

M2

The M2 data retrieval was performed internally by USAPHC personnel. The data available in M2 are extensive and the system is complex. Thus, the query process was fairly time consuming, though this may be reduced with further training and increased experience with the system. Because of the extensive nature of the data source and the ability of USAPHC personnel to retrieve the data themselves, M2 was retained for further analysis.

Discussion

The purpose of this section of the study was to determine which military human medical data system to include in the USAPHC ZDR. This focus of this section was to evaluate each system's ability to meet the goals of the ZDR by determining 1) how well each system matched the specific ZDR criteria of system goals, population source, data

source, and data types, 2) how well each system met the specific ZDR query requirements, and 3) how well each system performed when queried.

According to published CDC guidelines (14), a thorough evaluation of data systems should focus on how well the systems operate to meet the stated purpose and objectives of the given surveillance program, or in this case the stated purpose and objectives of the Zoonotic Disease Report. On a purely descriptive level, it appeared that both ESSENCE and DMSS would be well suited for inclusion in the ZDR. However, further evaluation through the use of system queries revealed several concerns. Because ESSENCE was unable to produce query results in a timely manner it was unable to be included in further analysis. This raised significant questions about the utility of the ESSENCE system for use in the ZDR at this time. Although DMSS was able to produce the requested query, several issues were revealed during the query process that eliminated it from being included in the next stages of analysis. The first issue arose when our request for social security numbers (SSNs) for use as unique identifiers was denied. Although with additional justification SSNs were provided, the time involved in the formalized query request process made the timeliness of DMSS data questionable. Next, during the matching procedures we discovered that the SSNs associated with each case represented the sponsoring service member, not the actual patient. Eventually we received the appropriate unique identifier (SSN for the patient) for each case, but the multiple data requests and justifications involved in obtaining the appropriate unique identifiers- reaffirmed the necessity of USAPHC being able to conduct system queries themselves. Additionally, the cumbersome data request process the overall utility of the DMSS system for use in the ZDR questionable. As outlined in chapter 3, the data in DMSS are very extensive and extremely valuable, especially when conducting longitudinal studies on service members. The exclusion of DMSS for use in the ZDR should not be viewed as a judgment on the data source. Rather it highlighted potential issues involved with relying on external sources to fulfil data requests.

Of the three remaining data systems, M2 appeared to be best suited for use in the ZDR. Although it was created to improve business intelligence reporting, the sources used and the completeness of the medical data feeds allows it to capture data equally for service member and beneficiary populations from both military and civilian medical facilities. Also, USAPHC personnel can obtain direct access to M2, which will enable them to not only conduct the queries themselves, but to also retain essential data characteristics (i.e., unique identifiers). Unfortunately, M2 does

not automatically receive laboratory data or reportable medical events data, which could result in some missed cases. HL7 provides laboratory data, but because not all diagnoses require laboratory confirmation, many cases may be missed if HL7 alone were used as the sole data source. However, as discussed in the background section of this document, HL7 is not designed to serve as a standalone disease surveillance system; rather the utility of HL7 is as a tool to augment existing surveillance systems through the confirmation of existing medical event data. DRSi captures reportable medical event data, but also has limitations by design. The system was created to capture Medical Event Reports filed at Army installations as mandated by the Armed Forces Reportable Medical Events Guidelines and Case Definitions. But because only 66 conditions are currently identified as reportable, and because medical event reports must be filed independently in addition to the regular electronic medical record, DRSi should not be viewed as a standalone surveillance system. Again, the utility of this system is as a way to augment data in a more robust medical data system.

Conclusions

This study recommends the use of M2 as the primary source of military human medical data for use in the Public Health Command Zoonotic Disease Report. The data system represents the most comprehensive and reliable source of both medical and population data; capturing not only the entire target population, but also having the potential to capture cases from both military and civilian healthcare facilities. In addition, M2 can be directly queried by USAPHC personnel which greatly improves not only the quality of the data but also the timeliness of the information received. Lastly, because the data in M2 retains unique identifiers, queries can be cross referenced and supplemented with information from DRSi and HL7.

Section II: An Evaluation of Data System Quality

Capture-Recapture Methodology

Capture-recapture (CR) techniques have long been used in wildlife biology to estimate population numbers. Because capturing an entire wildlife population is impractical, an estimation of the true population size is determined mathematically based on repeat samples drawn from the population (23). The basic approach involves a series of trapping events, each one attempting to capture as many animals as possible. The first trapping event

produces the first sample, the “capture.” These animals are counted, tagged, and released back into the population. The second trapping event produces the second sample. In addition to counting the total number captured in this second sample, the number of previously tagged animals is also counted. These previously tagged animals trapped in the second sample are the “recaptured” animals.

The pictorial below illustrates the application of capture-recapture techniques in wildlife biology (24). The four black birds in the top flock represent the “capture” sample. The one black bird in the lower flock represents the “recapture.” From the 1:4 dilution of marked birds, the total population size is determined to be four times that of the sample size.

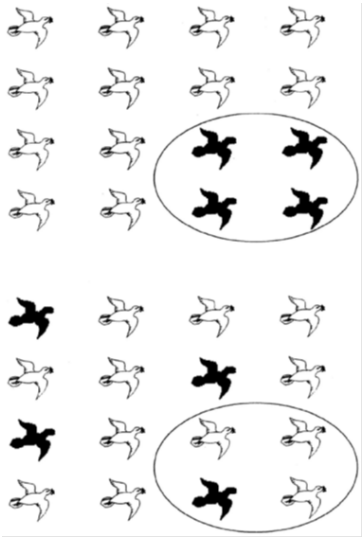

Figure 4.1: Diagram Illustrating a Basic Capture-Recapture (Source: GV Gill et al)

The table below depicts capture-recapture in a 2 x 2 format.

		First Sample "Capture"		
		Yes	No	Total
Second Sample "Recapture"	Yes	a	b	Z
	No	c	x	
	Total	Y		N

Figure 4.2: Capture-Recapture in Wildlife Populations

For a closed population, with equal catchability:

Y = number of animals marked in the original sample (a + c)

Z = the total number of animals captured in the second sample (a + b)

a = the number of animals from the first sample re-captured in the second sample

x = the number of animals missed by all trapping

N = the unknown population size (a + b + c + x)

Using the same logic as the bird example above, the ratio of recapture to full second capture is equal to the ratio of the first capture to the full (unknown) population. This formula is shown by:

$$a/Z = Y/N$$

The full population size can then be calculated by:

$$N = YZ/a$$

CR methods in epidemiology are attempts to estimate or adjust for incomplete case ascertainment using information from overlapping lists of cases from distinct sources (23). Potential uses include 1) estimation of population effected when the investigator has incomplete data available from two or more sources; 2) refinement of prevalence or incidence estimates derived from attempted exhaustive population surveys; 3) attempted evaluation of completeness of sources when data are received from many different sources; and 4) attempts to derive plausible upper or lower limits on estimates of the total population affected by an affliction (18, 23, 25).

There are several underlying assumptions when determining if CR methodology is appropriate for data source comparisons. Several of these are implicit, such as ensuring that the sources cover the same geographic area and time period, the case definitions used are consistent across data sources, and matching of cases from different sources is done appropriately (23). More explicit assumptions include: 1) equal "catchability" of cases, that for any single source, each case in the population has the same probability of ascertainment; 2) independence of sources, that for two sources comparisons, ascertainment of any case by the sources is independent; and 3) closed populations, that the population being sampled does not have any additions or losses during the study period (23).

		First Source (S1)		
		Yes	No	Total
Second Source (S2)	Yes	a	b	Z
	No	c	x	
	Total	Y		N

Figure 4.3 Capture-Recapture in Data Source Comparisons

Y = number of cases in the original sample (a + c)

Z = number of cases in the second sample (a + b)

a = the number of cases found in both sources

x = the number of cases missed by both sources

N = the unknown total number of cases (a + b + c + x)

Again, simple algebraic formulas can be done. The full number of cases can then be calculated by:

$$N = YZ/a$$

The number of cases missed by both sources can be estimated by:

$$x = bc/a$$

Lastly, the completeness of each data source can be calculated by:

$$S1_{completeness} = a/Z$$

$$S2_{completeness} = aY$$

or

$$S1_{completeness} = Y/N$$

$$S2_{completeness} = Z/N$$

As seen above, calculating CR estimates using only two sources is easily done by hand. The same concepts apply when comparing "k" sources, producing a table with 2^k cells; however the computations become too complicated to do by hand. Log-linear modeling permits the analysis of multiple sources at once (23, 26).

Log-linear Modeling

Basic Concept

General linearized models (GLMs) allow linear modeling to be applied to non-continuous (discrete) count or rate data, data where values are restricted to 0 and 1, and data where the variance is not constant or where variance may

be more than the mean (27, 26). Components of GLM include a linear component, a link function (how the expected value is related to the linear predictor), and a variance function which represents the relationship between the variance of responses as they relate to the mean (27). Log-linear models are a form of GLM that model the association and interaction patterns among categorical variables or the conditional relationship of the categorical variables (26). There is no distinction made between independent and dependent variables in log-linear modeling; all are treated as "response" variables (26). The outputs of such modeling are estimates of the cell frequencies for each response variable and their interaction variables. For clarity, one can consider the independent variable the estimated frequency.

Consider a contingency table with "i" and "j" categories, both being binomial categorical data and the frequency of the observation is the conditional data within the table.

The basic premise is that estimates of cell probabilities can be derived from the product of marginal probabilities of two independent variables (26). Where π is the probability of cell ij:

$$\pi_{ij} = \pi_{i\text{-}}. \pi_{.\text{-}j}$$

In terms of estimating the expected cell frequencies (μ_{ij}), given cell ij where n is the observed cell frequency:

$$\mu_{ij} = n\pi_{ij} = n(\pi_{i+} \pi_{+j})$$

Taking the natural logarithms gives the log-linear model of independence:

$$\log(\mu_{ij}) = \log(n) + \log(\pi_{i+}) + \log(\pi_{+j})$$

In a more standard notation, where A and B are two categorical variables:

$$\log(\mu_{ij}) = \lambda + \lambda_i^A + \lambda_j^B$$

λ = is the overall effect, or the grand mean of the logarithms of the expected counts; it ensures that the expected cell counts under the fitted model add up to the total sample size n

λ_i^A = represent the "main" effect variable A

λ_j^B = represent the "main" effect variable B

Including the combined effect (interaction) of two variables gives the following equation:

$$\log(\mu_{ij}) = \lambda + \lambda_i^A + \lambda_j^B + \lambda_{ij}^{AB}$$

λ_I^A = the main effect for variable A

λ_j^B = the main effect for variable B

λ_{ij}^{AB} = the interaction effect for variables A and B

The above model is considered a saturated model because of the inclusion of all possible one-way and two-way effects (26). Since the model has the same amount of variables as the number of cells in the contingency table, the expected cell frequencies will exactly match the observed frequencies since there are no degrees of freedom remaining to create uncertainty (26).

Capture-recapture and Log-linear Modeling to Evaluate Data Systems

Capture-recapture methods can be used in epidemiology as a way to estimate the total number of cases in a population when no complete list of cases exists. To do so, the data from multiple incomplete lists of cases (all representing the same source population) are used to estimate the total number of cases in the population that were missed by all lists. In other words, epidemiology uses capture-recapture to estimate the total number of cases missed

by multiple sources by adjusting for incomplete case ascertainment using lists of cases retrieved from sources that attempt to capture the same population (23). A major concern however, is the potential for dependence between sources, as this would violate one of the required assumptions for the use of CR methods. The application of log-linear models to data from multiple sources allows for the adjustment of these potential dependencies through the inclusion of their interaction terms.

The below is an example of the lay-out for a contingency table for three sources:

		Source 3			
		Yes		**No**	
		Source 2			
		Yes	**No**	**Yes**	**No**
Source 1	**Yes**	a	c	e	g
	No	b	d	f	x

Figure 4.4: Three-way Contingency Table for Log-linear Modeling

a = number of cases identified by all three sources

b = number of cases identified by sources 2 and 3

c = number of cases identified by source 1 and 3

d = number of cases identified by source 3 only

e = number of cases identified by sources 1 and 2

f = number of cases identified by source 2 only

g = number of cases identified by source 1 only

x = number of cases not identified by any sources (unknown)

By using “1” to represent sources that contributed to the number of cases and “0” to represent sources that did not contribute to the number of cases the data can be rearranged into the below table:

Source 1	Source 2	Source 3	Cell Letter
1	1	1	a
0	1	1	b
1	0	1	c
0	0	1	d
1	1	0	e
0	1	0	f
1	0	0	g
0	0	0	x

Figure 4.5: Dichotomized Identification of Cell Data

The strategy of log-linear modeling is to fit models to observed frequencies. Main effects models assume all sources are independent and that no interactions exist, whereas models containing interaction terms evaluate dependence between variables. Because the number of cases missed by all sources, denoted as x, is unknown, the three-way interaction is not included in the model. Instead the model is used to estimate that situation and x will equal μ, the expected frequency of missing cases. The number of observations is therefore 2^k-1, where k is the number of sources. Because 2^k-1 is equal to the number of variables in our model with all two-way interactions plus μ, and thus has zero degrees of freedom, the model may be called “saturated” but to avoid confusion it will be called “full”. The below table depicts the eight different log-linear models that are possible for a three source analysis:

Model Variables	Degrees of freedom
1, 2, 3*	3
1, 2, 3, 1-2	2
1, 2, 3, 1-3	2
1, 2, 3, 2-3	2
1, 2, 3, 1-2, 1-3	1
1, 2, 3, 1-2, 2-3	1
1, 2, 3, 1-3, 2-3	1
1, 2, 3, 1-2, 2-3, 1-3**	0

1 = Source 1, 2 = Source 2, 3 = Source 3
1-2 = interaction term for Source 1 and 2, 1-3 = interaction term for Source 1 and 3, etc.
* Main Effects Model
** Full Model
Figure 4.6: Three Source Models

Each model produces a set of expected cell frequencies that may or may not resemble the observed frequencies. These expected frequencies are used to aid in selecting the preferred model. Typical model selection techniques, such as forward or backward selection, can be employed to identify the most parsimonious model that fits the data according to pre-specified criteria. Other techniques used to identify the best model may include simply running all potential models and choosing the one that best fits the data. It is important to note, however, that a hierarchy of models exists and that the incorporation of any interaction term in the model necessitates the inclusion of each main effect variable involved. The use of non-hierarchical modeling is not recommended because it provides no statistical procedure for selecting the best model. For log-linear modeling the specified criteria for model fit generally includes formal goodness-of-fit statistics such as the likelihood-ratio test statistic (G^2), the Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), chi-square statistic (χ^2), or the change in deviance. In addition, the investigator may have knowledge of known dependence between sources, and therefore may select a model that includes a certain interaction even if of questionable statistical significance (23, 28).

The purpose of this part of the project was to use log-linear modeling with capture-recapture case count data to estimate the number of cases not captured within any of the data systems, and subsequently determine completeness of reporting for each data system.

Methods

Based on the analysis of the medical data system attributes conducted in Section I of this chapter, the remaining evaluation is limited to the following systems: M2, DRSi, and HL7.

Capture-Recapture

Case Count Data Collection

Case count data from Section I were used to select the most suitable disease for use in the subsequent capture-recapture analysis. Because of the low numbers of leishmaniosis and leptospirosis cases, the complete lack of HL7 reported hantavirus cases, and the large number of borreliosis cases, capture recapture was done using only the borreliosis data. The initial case count data showed a larger discrepancy between data sources than expected, with almost thirteen times the number of cases from M2. Further investigation revealed a large portion of the M2 cases came from civilian medical facilities through purchased care claims data. Because M2 is the only system that has complete access to purchased care records, the inclusion of these cases in capture recapture would violate the "equal catchability" assumption; therefore these cases were removed from the M2 case count data. A two week outpatient rule was applied to all cases found in each data system, so that two or more Lyme disease outpatient medical encounters occurring no more than 14 days apart would be counted as one case.

Case Matching

Case count data used social security numbers (SSN) as unique identifiers. Each dataset was assigned an indicator variable that represented presence/absence in the original data source. When a SSN appeared more than once within

a single data source, all of the data associated with the case were merged into one case and the first date of diagnosis was retained. Next, data from all sources were merged into a single database using the merge function in SPSS. Capture-recapture methods were applied to the combined dataset using SAS 9.3 (SAS), wherein a recapture event was defined as a case appearing in more than one data source. The SAS code used to conduct capture-recapture matching can be found in **Appendix B-4** along with the associated output. These steps were conducted by the U.S. Army Public Health Command, Epidemiology and Disease Surveillance Portfolio's Disease Epidemiology Program personnel.

Data Source Over and Underlap Analysis

The capture-recapture data were used to conduct an analysis of data source over- and underlaps. Overlap and underlap between data sources was calculated based on the following equations, where "a" through "g" represents observed (Figure 4.5):

$N_{obs} = a+b+c+d+e+f+g$

$N_{DRSi} = a+c+e+g$ $N_{HL7} = a+b+e+f$ $N_{M2} = a+b+c+d$

Overlap

% DRSi data simultaneously captured by HL7 = $(a+e)/ N_{DRSi}$

% DRSi data simultaneously captured by M2 = $(a+c)/ N_{DRSi}$

% HL7 data simultaneously captured by DRSi = $(a+d)/ N_{HL7}$

% HL7 data simultaneously captured by M2 = $(a+b)/ N_{HL7}$

% M2 data simultaneously captured by DRSi = $(a+c)/ N_{M2}$

% M2 data simultaneously captured by HL7 = $(a+b)/ N_{M2}$

Underlap

% DRSi data missed by HL7 = (c+g)/ N_{DRSi}

% DRSi data missed by M2 = (e+g)/ N_{DRSi}

% HL7data missed by DRSi = (b+f)/ N_{HL7}

% HL7data missed by M2 = (e+f)/ N_{HL7}

% M2data missed by DRSi = (b+d)/ N_{M2}

% M2data missed by HL7 = (c+d)/ N_{M2}

The step-by-step process for the over and underlap calculations can be found in **Appendix C-4**

.

Log-Linear Modeling

Capture-recapture data were converted into the format illustrated in **Appendix D-4** to allow for its use in log-linear modeling. The PROC GENMOD procedure in SAS version 9.3 was then used to perform log-linear modeling; this SAS code is described in **Appendix E-4.** In order to facilitate the comparison of the three medical data systems, all three were included in each model (main effects). Since the model with all three two-way interactions was significant the best model was chosen by comparing all models to it. With three potential data sources there were eight possible models (using all of the data) excluding the three-way interactions, seven of which data existed for. The model for absence from all data sources was the unknown being determined. Taking into account potential over-dispersion, modeling was done using both Poisson and negative binomial distributions. The negative binomial distribution was found not to be appropriate for this data because the variance differed between all of the models making them incomparable which was also demonstrated by the inability of the procedure to estimate the dispersion parameter by maximum likelihood; therefore the Poisson distribution was identified as the appropriate distribution. Model selection was based on several Goodness of Fit tests, to include χ^2, deviance, and AIC. In addition, the closeness between observed and fitted observations was considered. Lastly, knowledge of each data source (see "A

Review of Existing Data Sources", Chapter 3) allowed the incorporation of known dependence between data sources and further aided in the final model selection.

Reporting Completeness Calculations

Once the appropriate model was selected, completeness reporting to each data source was calculated based on the following equations:

$$N = a + b + c + d + e + f + g + x$$

N = total number of observed cases

x = the number of missing cases predicted by the model

Therefore, reporting completeness to each data system is as follows:

$$DRSi = (a + c + e + g)/ N$$

$$HL7 = (a + b + e + f)/ N$$

$$M2 = (a + b + c + d)/ N$$

The step-by-step process for the completeness calculations can be found in **Appendix F-4**.

Results

Capture-Recapture

The results of the capture-recapture are listed in Table 4.4 below. Only nine percent of the total cases were captured simultaneously by all three data sources, and approximately eighteen percent by any two data sources. Of the 73%

that were captured by any single source, 55% were captured by M2 alone, 15% by HL7 alone, and 2.5% by DRSi alone. Of the 763 total cases captured in one data source, 79% of the cases were captured by M2, 38% were captured by HL7, and 18% were captured by DRSi.

Table 4.4: Borreliosis Capture-Recapture Events (M2, DRSi, HL7)

	DRSi	HL7	M2	Cases
Cases Captured in all 3 Data Sources	x	x	x	95
		x	x	112
	x		x	42
Cases Captured in 2 Data Sources	x	x		28
			x	575
		x		161
Cases Captured in 1 Data Source	x			27
Subtotal	*192*	*396*	*824*	
Total Number of Borreliosis Cases Observed				**1040**

In addition, the capture-recapture data were used to conduct an analysis of data source over- and underlap. Table 4.5 shows the underlap between the sources. The table is read as the percent of cases from the column data source that were missed by the row data source. For example, of the cases captured by DRSi 36% were not found in HL7 and 29% were not found in M2. A table of the overlaps is available in **Appendix C-4**.

Table 4.5: Capture-Recapture Underlap

	Cases Captured by		
Cases Missed by	**HL7**	**DRSi**	**M2**
HL7	NA	36%	75%
DRSi	69%	NA	83%
M2	48%	29%	NA

Table is read as % of cases from column data source that were missed by row data source.

Log-Linear Modeling

The results of the log-linear modeling are shown in the tables below. Table 4.6 shows the number of observations associated with each data source or combination of data sources, both actual and predicted. The italicized model containing both DRSi-HL7 and M2-DRSi interaction terms appear to most closely fit the observed CR data.

Table 4.6: Comparison of Capture-Recapture Data to Predicted Data by Model

		Predicted							
Data Source(s)	**Actual CR Data**	**Full**	**Main Effects**	**DRSi-M2**	**M2-HL7**	**DRSi-HL7**	***DRSi-HL7, M2-DRSi***	**M2-DRSi, M2-HL7**	**DRSi-HL7, HL7-M2**
HL7	161	161	138	161	158	127	*161*	161	130
DRSi	27	27	56	40	27	32	*20*	27	27
M2	575	575	507	503	517	575	*575*	514	575
HL7, M2	112	112	203	184	173	146	*112*	173	143
DRSi, M2	42	42	82	100	101	37	*49*	103	42
DRSi, HL7	28	28	22	15	31	57	*35*	28	59
DRSi, HL7, M2	95	95	33	37	34	66	*88*	34	64

Each row specifies the source of data and each column in the "Predicted" section specifies the interaction terms included in the model.
All models include each data source (3 main effect variables).
The Main Effects model, aka independent model does not include any interactions.
The Full model includes all main effects and all two way interactions.
Each subsequent model includes the 3 main effects in addition to the interaction terms listed in the table.

Table 4.7 summarizes the three criteria used to assess model fit. Based on these criteria, the italicized model containing the both DRSi-HL7 and M2-DRSi interaction terms again appears to best fit the observed data.

Table 4.7: Comparison of Models by Goodness of Fit Criterion

Criteria	Full	Main Effects	DRSi-M2	M2-HL7	DRSi-HL7	*DRSi-HL7, M2-DRSi*	M2-DRSi, M2-HL7	DRSi-HL7, HL7-M2
Degrees of Freedom	0	3	2	2	2	*1*	1	1
Deviance	0	182.17	164.64	149.46	48.34	*5.67*	46.58	149.09
x^2	0	208.08	181.56	174.04	46.44	*5.79*	44.56	170.88
AIC	0	234.13	218.60	203.42	102.30	*61.63*	203.42	205.05

Deviance, where smaller deviance indicates better model fit, or smallest variance left unexplained.
AIC (Akaike Information Criterion) compares log-likelihoods of two models taking into account complexity of the model; where smaller AIC values indicate better model fit.

Table 4.8 shows the estimated number of missing cases predicted by each model. This estimation was then used to calculate the estimated total number of borreliosis cases by each model.

Table 4.8: Estimates of Borreliosis for Army Installations, April 1, 2011 to April 30, 2012

Estimate	Full	Main Effects	DRSi-M2	M2-HL7	DRSi-HL7	*DRSi-HL7, M2-DRSi*	M2-DRSi, M2-HL7	DRSi-HL7, HL7-M2
Missing Cases	1803	345	441	1	139	*827*	155	370
Total Cases	2843	1385	1481	149	1179	*1867*	1195	1410

Estimate of total number of Lyme borreliosis cases = N_{obs} + model estimate for missing cases; from Table 4 N_{obs} = 1040.

The parameter estimates for the selected model (**Appendix G-4**) is represented by the below equation where y is the estimated number of cases missed by the respective source or source combinations (29).

log(y) = 6.7173 -3.7333(DRSi) -1.6359(HL7) - 0.3629 (M2) + 2.2139 (DRSi-HL7) + 1.2756 (DRSi-M2)

In order to interpret these coefficients, one must take the natural log of the estimates to find the incidence risk associated with the use of the data source (29). The resulting transformation produced the following equation:

$$y = 826.58 - 0.024(DRSi) - 0.195(HL7) - 0.695(M2) - 9.150(DRSi\text{-}HL7) - 3.580(DRSi\text{-}M2)$$

Reporting Completeness Calculations

The selected model estimates that from the period of April 1, 2011 to April 30, 2012, 827 cases of Lyme borreliosis at Army medical facilities worldwide were missed by the combined data sources of M2, DRSi, and HL7. This estimate brings the total number of cases for this time period up to 1867. Based on this, the estimated reporting completeness of each data source alone and in combination were calculated.

Table 4.9: Summary of Selected Data Source Completeness

Data Source	Reporting Completeness (%)
DRSi	10
HL7	21
M2	44
DRSi and HL7	25
DRSi and M2	47
M2 and HL7	54
DRSi and HL7 and M2	56

Discussion

Again, the overall purpose of this study was to determine which military human medical data system to include in the USAPHC ZDR. The focus of this section was to evaluate each system's data completeness through the use of capture-recapture and log-linear modeling.

Capture-recapture methods in epidemiology are attempts to estimate the extent of incomplete case ascertainment from distinct sources, using information from overlapping lists (23). Although it always relates to estimations of the population, there are many different potential applications of the method in epidemiology, including: 1) estimation when the investigator has clearly incomplete data available from two or more sources; 2) refinement of prevalence or incidence estimates derived from attempted exhaustive population surveys; 3) attempted evaluation only of source completeness or of completeness of a registry that receives reports from many different sources; and 4) attempts to derive only plausible upper or lower limits on the total affected (23). We used capture-recapture in the third application of evaluating source completeness in order to assist USAPHC in determining the most complete human medical data (collection) system for use in the ZDR.

Implicit in capture-recapture analysis are three basic assumptions. The first assumption is the correct diagnosis of each case. Only once we are confident in this first assumption can we move onto the next two assumptions: appropriate matching of cases across data sources and assurance that all cases included are of the same time and space unit (23). In other words, the investigators must have a high level of confidence in the case count data used. As mentioned in the methods section, this level of confidence was not achieved with the DMSS case count data at the time of this analysis.

Table 4.10 provides a summary of the data source analysis done in this study.

Table 4.10: Summary of Selected Data Source Completeness

Data Source	Sole Contribution to Cases (%)	HL7 Cases Missed (%)	DRSi Cases Missed (%)	M2 Cases Missed (%)	Decrease in Missed Cases (%)	Reporting Completeness (%)
DRSi	26	69	NA	83	2	10
HL7	15	NA	36	75	20	21
M2	55	48	29	NA	70	44

Analysis of the capture-recapture data shows us that M2 is the most comprehensive single data source, with 55% of total captured cases coming from this data source alone. It also showed that the percent of cases missed by M2 was lower than the percent of cases missed by the other sources, again indicating that M2 is the most comprehensive data source. More specifically, M2 only missed 48% of the HL7 cases and 29% of the DRSi cases, whereas DRSi missed 69% of the HL7 cases and 83% of the M2 cases, and HL7 missed 36% of the DRSi cases and 75% of the M2 cases. This analysis showed that there is substantial underlap between these sources, and that reliance on a single data source will lead to underestimation of total cases.

Log-linear modeling was then done to get a better estimate for the total number of cases missed by all three data sources. Model selection was based on comparing each model individually to the full or 'saturated' model. Because the 'saturated' model includes all main effects and all two-way interactions, it has a separate parameter for each observation, and therefore should fit the data most completely. Model fit was done by assessing the deviance, where smaller deviance indicates better model fit, or smallest variance left unexplained. In addition, the AIC was also assessed. The AIC compares the log-likelihoods of the two models, but also takes into account the complexity of the model and/or sample size. Smaller AIC values indicate better model fit. The goodness of fit was not adequate for any of our models, but this is likely due to our small sample size (n = 7), meaning that any deviation seems significant as the test is not very robust. Model predicted case count values were compared to the observed values; models that more closely predicted the observed values were considered better fits. Model complexity (number of variables) was also considered, with the ultimate goal of finding the simplest possible model that explains the most variability in the data (the most parsimonious model).

Applying the above mentioned methods led to the selection of a model that included the main effects of M2, HL7, and DRSi and the interaction terms of DRSi-M2 and DRSi-HL7. According to this model, the number of cases missed by all data sources (when all parameters are held at zero) is 827. The parameter estimates for the model are interpreted (29) as the impact each variable has on determining the number of missed cases. Therefore, the model predicts that using the case count data from DRSi alone decreases the number of missed cases by 2.4%. Using HL7 alone decreases the number of missed cases by 19.5%. Using M2 decreases the number of missed cases by 69.6%.

When both DRSi and HL7 are used the number of missed cases decreases 9 fold, and when both DRSi and M2 are used the number of missed cases decreases by more than 3 fold.

Based on the estimated number of missing Lyme borreliosis cases, the reporting completeness of M2 was determined to be forty-four percent; this was twenty-three percent higher than that of HL7 and thirty-four percent higher than DRSi. Because all three systems theoretically encompass the same population, and because we limited the M2 data sources to only those at military treatment facilities, the differences in reporting completeness are likely due to differences in data types. Since DRSi data only include reportable medical events, the observed low reporting completeness of DRSi may reflect poor compliance with medical event reporting policy or it may be because not all physician entered ICD entries are associated with confirmed diagnoses that would require an associated reportable medical event report. The HL7 data only include laboratory confirmed diagnoses of Lyme borreliosis. In contrast to DRSi data, HL7 data come directly from the electronic hospital information system; therefore it does not require any additional reports to be filed. The lower reporting completeness of HL7 may be explained by the fact that Lyme borreliosis cases do not require laboratory confirmation in order to be considered a case. The M2 cases, in comparison, are neither limited to those diagnosed via laboratory methods nor to those with an accompanying RME record. For the purpose of this study, the cases reported to M2 included all inpatient and outpatient medical event encounters associated with a diagnosis of Lyme borreliosis at military treatment facilities.

Expanding completeness calculations to include more than one data source (Table 4.9), indicates the combined use of all three data sources produces the most completely captures the number of Lyme disease cases in this population. Specifically, the use of all three data sources simultaneously improves the reporting completeness by 12% as compared to relying on the M2 data source alone (from 44% to 56%).

Table 4.11: Capture-Recapture Overlap

	Cases Captured by		
Cases Captured by	**HL7**	**DRSi**	**M2**
HL7	NA	64%	25%
DRSi	30%	NA	17%
M2	52%	71%	NA

Table is read as % of cases from column data source that were also captured by row data source.

The inclusion of interaction terms in the selected model implies conditional independence of sources. Specifically, DRSi is not independent of HL7 and DRSi is not independent of M2, but HL7 and M2 are independent of each other given DRSi. Applying what we know about these data sources and looking at the overlap analysis done in Table 4.11, this relationship may be explained by the fact that DRSi has a case finding module with the capability to incorporate existing HL7 laboratory data (30% of the HL7 cases were also found in DRSi). Also, because DRSi tracks reportable medical events, it is possible that physicians want to be certain of their diagnosis before submitting their medical event report, making laboratory confirmation of these cases more likely (64% of the DRSi cases were also found in HL7). The DRSi to M2 interaction is interesting, as we know that M2 does not receive direct data feeds of the medical event reports contained in DRSi. In this case these two sources would show negative dependence if the cases in DRSi did not have a chance of entering the M2 data source. However, this is not the case, as DRSi reports exist separately from patient's regular electronic medical records. Therefore the 17% of the M2 cases that are also represented in DRSi and the 71% of the DRSi cases that are also in M2 probably represent chance overlap in cases.

Conclusions

Based on these findings this section of the study recommends the use all three data sources in combination when assessing the number of cases associated with a particular disease. This study finds M2 to be the most comprehensive military human medical data source, but because of substantial underlap between all three sources, and as indicated by log-linear modeling, it recommended that data from all three sources are incorporated in order to produce the most complete and reliable disease data.

Overall conclusion

The aim of this study was to conduct a systematic comparison of the human medical data systems most commonly used for public health surveillance in the U.S. Army in order to recommend the one(s) most suited for use in the USAPHC ZDR. This is the first comprehensive assessment of military human medical data systems for the purpose of zoonotic disease reporting. Although the study is geared towards meeting the specific goals of the U.S. Army Public Health Command Zoonotic Disease Report, the approach used to evaluate and compare the data systems can be adapted for use in evaluating data systems in general. One of the primary strengths of the study was the comprehensive approach used to conduct the assessment. The approach included the following three phases: 1) detailed system descriptions, 2) a comparison of specific data systems attributes, and 3) an evaluation of each system's data quality. As listed in chapter 3, the study started with detailed system descriptions based on a combination of literature reviews, interviews with system administrators and end users, and one-on-one demonstrations of system use. The knowledge acquired about each system during this phase of the study permitted the informed assessments of each data system's attributes conducted during the next two phases, as outlined in this chapter. In addition, the use of a standardized framework based on the specific goals and objectives of the ZDR, was critical to the successful comparison of the different data systems. Due to significant differences between the systems, a direct comparison could not have otherwise been conducted. A great deal was learned by conducting the case count queries. The queries not only permitted the subsequent capture-recapture and log-linear modeling analysis, but also a determination of data system "user-friendliness" such as simplicity, flexibility, timeliness, and stability. In fact, had the analysis ended with the data system attribute comparison, a completely different data system would have been recommended. The capture-recapture and log-linear modeling analysis permitted a quantitative assessment of the data quality of the data sources for the remaining data systems. The findings, in combination with the background knowledge of each system, facilitated the selection of final three data sources. Another strength was the selection of a disease with a case definition listed in the Armed Forces Reportable Medical Event Guidelines. This eliminated arbitrary variation in case definition from the list of reasons for differences in case counts, and allowed us to focus on the differences attributed to target populations, data sources, and data types. Lastly, because the overarching objective of the ZDR is to improve force health protection, a goal common to all of the systems analyzed, there was a large degree of cooperation among all participants in this study.

This study was conducted with the very specific goals of the ZDR in mind, focusing more on external validity than internal validity. In other words, the data systems that were found to best suit the goals of the ZDR are merely that, better sources of data for the specific purpose of the ZDR. This does not mean that the other systems are inadequate, only that the other systems did not meet the needs of the ZDR at this time. In addition, the results of the capture-recapture and subsequent log-linear modeling had several limitations. Due to the lack of reliable case count data from DMSS and ESSENCE during the study period, it was not possible to fully assess all five of the original data systems. In addition, because the credibility of any estimate is only as good as the accuracy of the information used, the issue of data accuracy must be discussed. Although all sources used case definitions based on the Armed Forces Reportable Medical Event Guidelines, each individual case was entered in the system by a person, therefore mistakes in data entry are always possible. Because we did not have access to individual records, we were unable to confirm the accuracy of the data contained in our queries. Also, although social security numbers were used as unique identifiers for case matching, mistakes in data entry related to these numbers could lead to inaccurate matching and therefore inaccurate capture-recapture estimates.

Recommendations

This study recommends the combined use of M2, HL7, and DRSi as the data sources of military human medical data for use in the Public Health Command Zoonotic Disease Report. Although the M2 data system represents the most comprehensive and reliable source of both medical and population data, substantial underlaps with the other two data sources and the results of log-linear modeling indicate the most reliable predictor of case count data to be the combined use of all three sources.

References

1. Permanent Order 275-002, 2 October 2009, Headquarters U.S. Army Medical Command Fort Sam Houston Texas, TX 782345-6000.

2. U.S. Army Public Health Command website accessed May 16, 2013 at http://phc.amedd.army.mil/

3. Department of Defense (DoD) Directive 6490.02E (8 Feb 2012)

4. Centers for Disease Control and Prevention: Framework for evaluating public health surveillance systems for early detection of outbreaks. MMWR 2004;53(No.RR-5)Atlanta, Georgia, Centers for Disease Control and Prevention.

5. The Military Health System website accessed May 20, 2013 at http://www.health.mil/About_MHS/Organizations/Index.aspx

6. The Military Health System: Framework for National Health Information Exchange Poster, Developed by the Office of the MHS Chief Information Officer, February, 2007

7. Ruberton MV and Brundage JF (2002). The Defense Medical Surveillance System and Department of Defense Serum Repository: Glimpses of the Future of Public Health Surveillance. Am. J. of Public Health; 92(12):1900-1904.

8. M2 Factsheet posted on the Military Health System website accessed January 13, 2013 at http://www.health.mil/MHSCIO/programs_products/DHSS/DHSS-Products/M2.aspx

9. M2 Information Briefing posted on the Military Health System website accessed January 13, 2013 at http://www.health.mil/MHSCIO/programs_products/DHSS/DHSS-Products/M2.aspx

10. Personal interviews with Nikki Jordan, December 2012, Epidemiologist, U.S. Army Public Health Command, Army Institute of Public Health, Epidemiology and Disease Surveillance Portfolio, email: nikki.jordan@us.army.mil.

11. ESSENCE Factsheet posted on the Military Health System website accessed December 10, 2013 at http://www.health.mil/MHSCIO/programs_products/DHSS/DHSS-Products/M2.aspx

12. ESSENCE Information Briefing by U.S. Army Public Health Command, Army Institute of Public Health, Epidemiology and Disease Surveillance Portfolio, updated September 12, 2013.

13. Personal interview with Paul Pietrusiak, January 2013, Epidemiologist, U.S. Army Public Health Command, Army Institute of Public Health, Epidemiology and Disease Surveillance Portfolio, email: paul.pietrusiak@us.army.mil.

14. Centers for Disease Control and Prevention. Updated guidelines for evaluating public health surveillance systems: recommendations from the guidelines working group. MMWR 2001;50(No.RR-13)Atlanta, Georgia, Centers for Disease Control and Prevention..

15. DRSi for HPOs Information Briefing by U.S. Army Public Health Command, Army Institute of Public Health, Epidemiology and Disease Surveillance Portfolio, updated April 2012.

16. Personal interview with Zachary McCormic, Epidemiologist, U.S. Army Public Health Command, Army Institute of Public Health, Epidemiology and Disease Surveillance Portfolio, email: zachary.mccormic@us.army.mil.

17. Institute of Medicine. 2007. "Global Infectious Disease Surveillance and Detection: Assessing the Challenges—finding solutions", Workshop Summary. Washington DC: The National Academies Press.

18. Navy and Marine Corps Public Health Center. Capture-Recapture Methods: Technical Report. September 19, 2007.

19. Meyer et al.: Antimicrobial resistance surveillance in the AFHSC-GEIS network. BMC Public Health 2011 11(Suppl 2):S8.

20. Personal interview with Gosia Nowak, Infectious Disease Epidemiologist, Navy and Marine Corps Public Health Center, Epi Data Center, e-mail: gosia.nowak@med.navy.mil

21. Personal interview with LCDR Rhonda Lizewski, Armed Forces Health Surveillance Center, March 2013, e-mail: rhonda.a.lizewski.mil@mail.mil

22. Armed Forces Reportable Medical Events Guidelines and Case Definitions, Armed Forces Health Surveillance Center (AFHSC), March 2012.

23. Hook, EB and Regal, RR (1995). "Capture-recapture methods in epidemiology: methods and limitations." Epidemiol Reb 17(2):243-64.

24. Gill, GV, Ismail AA, et al (2001). "The use of capture-recapture techniques in determining the prevalence of type 2 diabetes." Qjm 94(7):341-6.

25. Hook, EB and Regal, RR (2000). "Accuracy of Alternative Approaches to Capture-Recapture Estimates of Disease Frequency: Internal Validity Analysis of Data from Five Sources." Am J of Epi 152(8): 771-779.

26. PENN State Online Courses: STAT 504- Analysis of Discrete Data. Accessed March 2013 at https://online courses.science.psu.edu/stat504.

27. SAS OnlineDoc™: Version 8. Chapter 29. The GENMOD Procedure. Accessed April 2013.

28. Abeni, DD, Brancato, G, Perucci, CA (1994). "Capture-recapture to estimate the size of the population with Human Immunodeficiency Virus Type 1 Infection." Epidemiology 5:410-414.

29. Dohoo I, Martin W and Stryhn H. Chapter 27: Concepts of Infectious Disease Epidemiology. Veterinary Epidemiological Research. Prince Edward Island, Canada: VER Inc, 2009;715-725.

Appendices

Appendix A-4

Query Process: Specific Case Definitions

The USAPHC Epidemiology and Disease Surveillance Portfolio, Disease Epidemiology Program personnel has direct access to the DRSi, M2, and ESSENCE data systems. This permitted query requests to remain "informal" and internal to the study team. Obtaining the data contained within the DMSS and HL7 systems, on the other hand, required "formal" request processes.

The following parameters were applied to all data queries performed or requested:

1. All cases of hantavirus, leptospirosis, leishmaniosis, and borreliosis that met the case definition.
2. Population to include all Service members and other beneficiaries.
3. Limit cases counts to those associated with Army Installations.
4. Limit to cases reported or diagnosed and laboratory tests between 1 April 2011 to 30 April 2012.

In addition, the following system specific criteria were applied:

DMSS- Formal request to Armed Forces Health Surveillance Center for all cases of leptospirosis, leishmaniosis, borreliosis, and hantavirus with the following case definitions based on the Armed Forces Reportable Medical Events Guidelines (22):

Leptospirosis: Case definition includes all ICD-9 codes associated with leptospirosis (100.0, 100.8, 100.9) in any of the diagnostic positions; 1 inpatient encounter, 1 reportable event, or 2 outpatient visits within 2 weeks. Incident rule of 1 case per lifetime; inclusion in this study required the diagnosis to occur between 1 April 2011 to 30 April 2012.

Borreliosis: Case definition includes ICD-9 code 088.81 associated with borreliosis in any of the diagnostic positions; 1 inpatient encounter, 1 reportable event, or 2 outpatient visits within 2 weeks. Incident rule of 1 case per lifetime; inclusion in this study required the diagnosis to occur between 1 April 2011 to 30 April 2012.

Leishmaniosis: Case definition includes all ICD-9 codes associated with leishmaniosis (085.0-085.9) in any of the diagnostic positions; 1 inpatient encounter, 1 reportable event, or 2 outpatient visits within 2 weeks. Incident rule of 1 case per lifetime; inclusion in this study required the diagnosis to occur between 1 April 2011 to 30 April 2012.

Hantavirus: Case definition includes ICD-9 code 079.81 associated with hantavirus in any of the diagnostic positions; 1 inpatient encounter, 1 reportable event, or 2 outpatient visits within 2 weeks. Incident rule of 1 case per lifetime; inclusion in this study required the diagnosis to occur between 1 April 2011 to 30 April 2012.

HL7- Formal request to Navy and Marine Corps Public Health Center (EpiData Center) for all laboratory confirmed cases of leptospirosis, leishmaniosis, borreliosis, and hantavirus as defined in the Armed Forces Reportable Medical Events Guidelines (22):

Leptospirosis: Any of the following

• Fourfold or greater increase in *Leptospira* agglutination titer between acute and convalescent serum specimens obtained ≥ 2 weeks apart and studied at the same laboratory;

• Demonstration of *Leptospira* in a clinical specimen by immunofluorescence;

• Isolation and typing from blood or other clinical materials by culture of pathogenic leptospires;

• Positive serology, preferably by the Microscopic Agglutination Test (MAT). Ideally, the panel of *Leptospira* strains used for antigens should be representative of the locally occurring strains; or

• Detection of leptospiral DNA by PCR.

Leishmaniosis
1. Cutaneous and Mucosal/Mucocutaneous:
• Positive parasitology (stained smear or culture from the lesion); or
• PCR-positive.

2. Visceral:
• Positive parasitology (stained smears from bone marrow, spleen, liver, lymph node, blood or culture of the organism from a biopsy or aspirated material); or
• Positive serology (rK39 assay).

Borreliosis:
For the purposes of surveillance, the definition of a qualified laboratory assay is
• Positive Culture for *B. burgdorferi*; or
• Two-tier testing in accordance with CDC Notifiable Infectious Diseases criteria where:
-Positive IgM is sufficient only when ≤30 days from symptom onset
-Positive IgG is sufficient at any point during illness

• Single-tier IgG immunoblot seropositivity in accordance with CDC and DoD criteria; or
• CSF antibody positive for *B. burgdorferi* by Enzyme Immunoassay (EIA) or Immunofluorescence Assay (IFA), when the titer is higher than it was in serum

Hantavirus:
• Detection of hantavirus-specific IgM or rising titers of hantavirus-specific IgG; or
• Detection of hantavirus-specific RNA sequence by polymerase chain reaction (PCR) in clinical specimens; or
• Detection of hantavirus antigen by immunohistochemistry.

DRSi- Informal request to USAPHC Epidemiology and Disease Surveillance Portfolio personnel for all confirmed Medical Event Reports based on the Armed Forces Reportable Medical Event case definitions for leptospirosis (reference ICD-9s:100.0, 100.8, 100.9), borreliosis (reference ICD-9 088.81), leishmaniosis (reference ICD-9s: 085.0-085.9), and hantavirus (reference ICD-9 code 079.81) (22).

Leptospirosis: Case definition includes a clinically compatible case that is laboratory-confirmed.

Laboratory confirmation includes any of the following:
• Fourfold or greater increase in *Leptospira* agglutination titer between acute and convalescent serum specimens obtained ≥ 2 weeks apart and studied at the same laboratory;
• Demonstration of *Leptospira* in a clinical specimen by immunofluorescence;
• Isolation and typing from blood or other clinical materials by culture of pathogenic leptospires;
• Positive serology, preferably by the Microscopic Agglutination Test (MAT). Ideally, the panel of *Leptospira* strains used for antigens should be representative of the locally occurring strains; or
• Detection of leptospiral DNA by PCR.

Borreliosis: Case definitions include
• A case that meets the clinical criteria for diagnosis of EM with a known exposure;
• A case of EM with laboratory evidence of infection and without a known exposure; or
• A case with at least one late manifestation that has laboratory evidence of infection.

Laboratory evidence of infection includes any of the following:
• Positive Culture for *B. burgdorferi*; or
• Two-tier testing in accordance with CDC Notifiable Infectious Diseases criteria where:
-Positive IgM is sufficient only when ≤30 days from symptom onset

-Positive IgG is sufficient at any point during illness

• Single-tier IgG immunoblot seropositivity in accordance with CDC and DoD criteria; or
• CSF antibody positive for *B. burgdorferi* by Enzyme Immunoassay (EIA) or Immunofluorescence Assay (IFA), when the titer is higher than it was in serum.

Exposure is defined as having been (≤ 30 days before onset of EM) in wooded, brushy, or grassy areas (i.e., potential tick habitats) in a county in which Lyme disease is endemic. A history of tick bite is not required.

Leishmaniosis: Any of the following
1. Cutaneous and Mucosal/Mucocutaneous:
A case that has a clinically compatible lesion with parasitological confirmation of the diagnosis (positive smear or culture).
2. Visceral:
A case exhibiting clinical signs with serological and/or parasitological confirmation of leishmaniosis.

Hantavirus: Case definition includes a clinically compatible case that is laboratory-confirmed.

Laboratory confirmation includes any of the following:
• Detection of hantavirus-specific IgM or rising titers of hantavirus-specific IgG; or
• Detection of hantavirus-specific RNA sequence by polymerase chain reaction (PCR) in clinical specimens; or
• Detection of hantavirus antigen by immunohistochemistry.

M2- Informal request to USAPHC Epidemiology and Disease Surveillance Portfolio personnel using case definitions in accordance with the Armed Forces Reportable Medical Events Guidelines (22) for leptospirosis, borreliosis, leishmaniasis, and hantavirus as outlined in the DMSS section (see above).

ESSENCE- Informal request to USAPHC Epidemiology and Disease Surveillance Portfolio personnel.

The initial intent was to use groups of syndromes to perform a syndromic query. However, without access to unique identifiers it would not be possible to apply incidence rules or confirm diagnoses and therefore the data would not be comparable to the other sources. Instead the query was performed with user-defined syndrome groups based on the following ICD-9 codes:

Leptospirosis: Case definition includes all health encounter entries listing ICD-9 codes associated with leptospirosis (100.0, 100.8, 100.9). Inclusion in this study required the entry occurred between 1 April 2011 to 30 April 2012.

Borreliosis: Case definition includes all health encounter entries listing ICD-9 code 088.81 associated with borreliosis. Inclusion in this study required the entry occurred between 1 April 2011 to 30 April 2012.

Leishmaniosis: Case definition includes all health encounter entries listing ICD-9 codes associated with leishmaniosis (085.0-085.9). Inclusion in this study required the entry occurred between 1 April 2011 to 30 April 2012.

Hantavirus: Case definition includes all health encounter entries listing ICD-9 code 079.81 associated with hantavirus. Inclusion in this study required the entry occurred between 1 April 2011 to 30 April 2012.

Appendix B-4

SAS Code for Capture Recapture

```
proc freq;
table drsi*hl7*m2/list ;
run;
```

SAS Output from Capture Recapture (PROC FREQ)

DRSi	HL7	M2	Frequency	Percent	Cumulative Frequency	Cumulative Percent
1	1	1	95	9.13	95	9.13
0	1	1	112	10.77	207	19.90
1	0	1	42	4.04	249	23.94
0	0	1	575	55.29	824	79.23
1	1	0	28	2.69	852	81.92
0	1	0	161	15.48	1013	97.40
1	0	0	27	2.60	1040	100

Appendix C-4

Over and Underlap Analysis

Capture Recapture Data				
DRSi	**HL7**	**M2**	**Cell Letter**	**Frequency**
1	1	1	a	95
0	1	1	b	112
1	0	1	c	42
0	0	1	d	575
1	1	0	e	28
0	1	0	f	161
1	0	0	g	27

Cells a through g represents observed data.
Cell x represents the expected or estimated number of missing cases given the selected model.

N_{obs} = a+b+c+d+e+f+g = 1040
N_{DRSi} = a+c+e+g = 192 N_{HL7} = a+b+e+f = 396 N_{M2} = a +b+c+d = 824

Percent of data capture simultaneously by DRSi and HL7 = (a+e)/ N_{obs} = (95 + 28)/1040 = 12%
Percent of data capture simultaneously by DRSi and M2 = (a+c)/ N_{obs} = (95 + 42)/1040 = 13%
Percent of data capture simultaneously by HL7 and M2 = (a+b)/ N_{obs} = (95+112)/1040 = 20%

Overlap
Percent of DRSi data simultaneously captured by HL7 = (a+e)/ N_{DRSi} = (95 + 28)/192 = 64%
Percent of DRSi data simultaneously captured by M2 = (a+c)/ N_{DRSi} = (95 + 42)/192 = 71%

Percent of HL7 data simultaneously captured by DRSi = (a+d)/ N_{HL7} = (95 + 28)/396 = 30%
Percent of HL7 data simultaneously captured by M2 = (a+b)/ N_{HL7} = (95 + 112)/396 = 52%

Percent of M2 data simultaneously captured by DRSi = (a+c)/ N_{M2} = (95 + 42)/824 = 17%
Percent of M2 data simultaneously captured by HL7 = (a+b)/ N_{M2} = (95 + 112)/824 = 25%

Capture-Recapture Overlap			
	Cases Captured by		
Cases Captured by	**HL7**	**DRSi**	**M2**
HL7	NA	64%	25%
DRSi	30%	NA	17%
M2	52%	71%	NA

Table is read as % of cases from column data source that were also captured by row data source.

Underlap
Percent of DRSi data missed by HL7 = (c+g)/ N_{DRSi} = (42 + 27)/192 = 36%
Percent of DRSi data missed by M2 = (e+g)/ N_{DRSi} = (28 + 27)/192 = 29%

Percent of HL7data missed by DRSi = (b+f)/ N_{HL7} = (112 + 161)/396 = 69%
Percent of HL7data missed by M2 = (e+f)/ N_{HL7} = (28 + 161)/396 = 48%

Percent of M2data missed by DRSi = (b+d)/ N_{M2} = (112 + 575)/824 = 83%
Percent of M2data missed by HL7 = (c+d)/ N_{M2} = (42 + 575)/824 = 75%

Capture-Recapture Underlap			
	Cases Captured by		
Cases Missed by	**HL7**	**DRSi**	**M2**
HL7	NA	36%	75%
DRSi	69%	NA	83%
M2	48%	29%	NA

Table is read as % of cases from column data source that were missed by row data source.

Appendix D-4

Borreliosis Capture Recapture Data in SAS PROC GENMOD Format

DRSi	HL7	M2	w	r
1	1	1	1	95
0	1	1	1	112
1	0	1	1	42
0	0	1	1	575
1	1	0	1	28
0	1	0	1	161
1	0	0	1	27
0	0	0	0	

The above table is used to import the data into SAS. With k = 3 sources there are 2^k data points or observations or 8 observations, yet, in the case of capture-recapture there is only seven as we are estimating the final observation. Number of cases is designated by 'r' and 'w' designates the weight. The number of cases missed by all data sources will be '0', as this observation is unknown. Weights of '1' are assigned to all observed data points and '0' to the unknown number of missed cases. This weighting ensures that the unknown number of missed cases will not be included in the analysis.

Appendix E-4

Using SAS 9.3 PROC GENMOD for Three Source Analysis

It is necessary to specify the response (y) and weight variables; here they were respectively designated as 'r' and 'w'. Also, the error distribution must be specified; for our model selection this distribution was specified as Poisson. Type 3 analysis does not depend on the order in which the terms for the model are specified, which is appropriate. The class statement used in the saturated model (and subsequent models) specifies the classification variables to be used in the analysis. The 'ref= '0' param = ref' specifies one-way comparisons with '0' as the referent group.

The below coding was used to create the eight log-linear models for this three source analysis.

```
/* Main Effects Model*/
title 'Main Effects Only';
proc genmod;
class DRSI (REF = '0' param = ref) HL7 (REF = '0' param = ref) M2 (REF = '0' param = ref);
weight w;
model r= DRSi HL7 M2/p dist = poisson type3;
output out=predicted p=pred;
run;

/* Saturated Model*/
title 'Saturated Model-p2';
proc genmod;
class DRSI (REF = '0' param = ref) HL7 (REF = '0' param = ref) M2 (REF = '0' param = ref);
weight w;
model r= DRSi HL7 M2 DRSi*HL7 DRSi*M2 HL7*M2/p dist = poisson type3;
output out=p2 p=pred;
run;

/* Model with DRSi*HL7 Intx-p3*/
title 'Model with DRSi*HL7 Intx';
proc genmod;
class DRSI (REF = '0' param = ref) HL7 (REF = '0' param = ref) M2 (REF = '0' param = ref);
weight w;
model r= DRSi HL7 M2 DRSi*HL7/p dist = poisson type3;
output out=p3 p=pred;
run;

/* Model with DRSi*M2 Intx-p4*/
title 'Model with DRSi*M2 Intx';
proc genmod;
class DRSI (REF = '0' param = ref) HL7 (REF = '0' param = ref) M2 (REF = '0' param = ref);
weight w;
model r= DRSi HL7 M2 DRSi*M2/p dist = poisson type3;
output out=p4 p=pred;
run;

/* Model with HL7*M2 Intx-p5*/
title 'Model with HL7*M2 Intx';
proc genmod;
class DRSI (REF = '0' param = ref) HL7 (REF = '0' param = ref) M2 (REF = '0' param = ref);
weight w;
model r= DRSi HL7 M2 HL7*M2/p dist = poisson type3;
output out=p5 p=pred;
run;

/* Model with DRSi*HL7 and DRSi*M2 Intx-p6*/
title 'Model with DRSi*HL7 and DRSi*M2 Intx';
proc genmod;
class DRSI (REF = '0' param = ref) HL7 (REF = '0' param = ref) M2 (REF = '0' param = ref);
weight w;
model r= DRSi HL7 M2 DRSi*HL7 DRSi*M2/p dist = poisson type3;
output out=p6 p=pred;
run;

/* Model with DRSi*M2 and HL7*M2 Intx-p7*/
```

```
title 'Model with DRSi*HL7 and HL7*M2 Intx';
proc genmod;
class DRSI (REF = '0' param = ref) HL7 (REF = '0' param = ref) M2 (REF = '0' param = ref);
weight w;
model r= DRSi HL7 M2 DRSi*HL7 HL7*M2/p dist = poisson type3;
output out=p7 p=pred;
run;

/* Model with DRSi*HL7 and HL7*M2 Intx-p8*/
title 'Model with DRSi*M2 and HL7*M2 Intx';
proc genmod;
class DRSI (REF = '0' param = ref) HL7 (REF = '0' param = ref) M2 (REF = '0' param = ref);
weight w;
model r= DRSi HL7 M2 DRSi*M2 HL7*M2/p dist = poisson type3;
output out=p8 p=pred;
run;
```

Appendix F-4

Completeness Calculations

Capture Recapture Data				
DRSi	**HL7**	**M2**	**Cell Letter**	**Estimated Data**
1	1	1	a	95
0	1	1	b	112
1	0	1	c	42
0	0	1	d	575
1	1	0	e	28
0	1	0	f	161
1	0	0	g	27
0	0	0	x	827

Cells a through g represents observed data.
Cell x represents the expected or estimated number of missing cases given the selected model.

N_{obs} = a+b+c+d+e+f+g = 1040
N_{DRSi} = a+c+e+g = 192 N_{HL7} = a+b+e+f = 396 N_{M2} = a +b+c+d = 824

The selected model predicts the number of missed cases, x, is 827. Therefore, the actual total number of cases predicted by the model is:

N = a+b+c+d+e+f+g+x = 1040 + 827 = 1867

Completeness Reporting

Reporting Completeness to DRSi = (a+c+e+g)/N = 192/1867 = 0.103 or 10.3%
Reporting Completeness to HL7 = (a+b+e+f) /N = 396/ 1867 = 0.212or 21.2%
Reporting Completeness to M2 = (a+b+c+d) /N = 824/1867 = 0.441 or 44.1%

Summary for Data Source Reporting Completeness Analysis

Data Source	Reporting Completeness
DRSi	10.3%
HL7	21.2%
M2	44.1%

Appendix G-4

Selected Output from Log-linear Modeling

Predicted Capture-Recapture Data by Model
Each row specifies the source of data
Each column in the "Predicted" section specifies the interaction terms included in the model

Actual		Predicted							
Data Source	**CR Data**	**Full**	**Main Effects**	**DRSi-M2**	**M2-HL7**	**DRSi-HL7**	**DRSi-HL7, M2-DRSi**	**M2-DRSi, M2-HL7**	**DRSi-HL7, HL7-M2**
DRSi, HL7, M2	**95**	95.00	32.64	36.63	33.72	65.86	**87.77**	34.42	64.30
HL7, M2	**112**	112.00	202.91	183.67	173.28	146.19	**112.00**	172.58	142.70
DRSi, M2	**42**	42.00	81.56	100.37	100.50	36.95	**49.23**	102.58	42.00
M2	**575**	575.00	506.89	503.33	516.50	575.00	**575.00**	514.42	575.00
DRSi, HL7	**28**	28.00	22.24	14.70	30.78	57.14	**35.23**	28.00	58.70
HL7	**161**	161.00	138.21	161.00	158.22	126.81	**161.00**	161.00	130.30
DRSi	**27**	27.00	55.55	40.30	27.00	32.05	**19.77**	27.00	27.00
	x	1802.83	345.26	441.21	138.76	498.8	**826.56**	155.25	369.64

All models include each data source (3 main effect variables).
The Main Effects model, aka independent model does not include any interactions.
The Full model includes all main effects and all two way interactions.
Each subsequent model includes the 3 main effects in addition to the interaction terms listed in the table.
X represents the estimated number of missing cases based on the model.

Log-Linear Parameter Estimates
for Model with DRSi-HL7 and M2-DRSi

				Wald 95% CI			
Parameter	**DF**	**Estimate**	**Std Error**	**Lower Limit**	**Upper Limit**	**Wald x2**	**pr>x2**
Intercept	1	6.72	0.1299	6.46	6.97	2673.27	<0.0001
DRSi	1	-3.73	0.2106	-4.15	-3.32	314.3	<0.0001
HL7	1	-1.64	0.1033	-1.84	-1.43	250.86	<0.0001
M2	1	-0.36	0.123	-0.60	-0.12	8.7	0.0032
DRSi-HL7	1	2.21	0.1825	1.86	2.57	147.24	<0.0001
DRSi-M2	1	1.28	0.2015	0.88	1.67	40.05	<0.0001

Chapter 5: A Model Using Service Member Pet Dogs as Sentinels in Zoonotic Disease Surveillance

Introduction

American Veterinary Medical Association One Health Imperative

In 2008 the American Veterinary Medical Association (AVMA) published the *One Health: A New Professional Imperative (*1). The stated vision of the document is "to promote and improve the health of humans, animals and our environment, individually and collectively, by encouraging and ensuring the acceptance and adoption of One Health and its associated activities". It defines One Health as the collaborative effort of multiple disciplines-working locally, nationally, and globally – to attain optimal health for people, animals and our environment. The seventy-one page document discusses how the health of people, animals, and the environment are interconnected. It highlights that of the approximately 1,500 diseases recognized in humans; almost 60% are due to multi-host pathogens characterized by their movement across species lines and that over the last 3 decades, nearly 75% of the new emerging human infectious diseases have been zoonotic. The document also mentions how the degradation of the environment creates favorable settings for the expansion of existing infectious diseases for both human and animal health. The AVMA One Health Imperative is dedicated to finding a holistic, collaborative approach to contemporary health issues through the convergence of human, animal, and environmental domains (1).

One Health in the Military

The AVMA also discusses the numerous organizations that have already recognized the need to integrate human, animal, and environmental health and have taken steps to develop new programs and partnerships to support that integration. Some such organizations include the Centers for Disease Control and Prevention (CDC), Environmental Protection Agency (EPA), United States Agency for International Development (USAID), Wildlife Conservation Society (WCS), and National Science Foundation (NSF) (1). They also applaud the U.S. military for their history of

linking health operations in monitoring, surveillance, and laboratory systems. Briefly, in 1986, the Army established a data center to support its HIV-related screening, clinical care, and epidemiological research programs (2). By 1993, the system transitioned to the Army Medical Surveillance System, expanding its scope to include all illnesses and injuries of public health or military operational importance. In 1997, the Army Medical Surveillance System transitioned to the Defense Medical Surveillance System, and the Army Medical Surveillance Activity (AMSA) was assigned responsibility for its operation (2). In 2008, the Deputy Secretary of Defense established the Armed Forces Health Surveillance Center (AFHSC), by partnering AMSA, the DoD Global Emerging Infectious Disease Surveillance and Response System (DoD-GEIS), and the Global Health Surveillance Activity (3). In 2011, under the direction of the U.S. Army Surgeon General, the former U.S. Army Veterinary Command (VETCOM) joined with the former U.S. Army Center for Health Promotion and Preventive Medicine (USACHPPM) to create the new U.S. Army Public Health Command (USAPHC) (4). The mission statement for this new Command is: "Promote health and prevent disease, injury, and disability of Soldiers and military retirees, their Families, and Department of the Army civilian employees; and assure effective execution of full spectrum veterinary service for Army and Department of Defense Veterinary missions"(5). The command combines the technical skills of multiple disciplines, including veterinarians, physicians, entomologists, laboratory specialists, epidemiologists, and many others in order to achieve the USAPHC mission. Although not explicitly stated as such, the mission of the USAPHC fully embodies the One Health concept.

In 2011, the Epidemiology and Disease Surveillance Portfolio, Disease Epidemiology Program within the USAPHC initiated development of a Zoonotic Disease Report (ZDR) to provide U.S. Army public health personnel with critical health information regarding the presence and spread of zoonotic pathogens and to create opportunities for improved preventive medicine strategies. The audience of the ZDR is US Army Public Health Command Regional and District level Commanders, as well as veterinarians and public health and preventive medicine professionals throughout the Army. The report combines zoonotic disease information from diverse sources, including: extant Army/DoD human disease databases, data generated by the Laboratory Services Portfolio for arthropod borne disease surveillance as well as rabies specimen testing, and animal data from intergovernmental organizations' public access databases. The goal is that animal data will eventually include zoonotic diseases diagnosed in animals

seen at U.S. Army Veterinary Treatment Facilities once the web-based electronic medical record known as the Remote Online Veterinary Record (ROVR) is fully deployed.

The overall aim of this study was to demonstrate the utility of zoonotic disease data from pet dogs seen at U.S. Army Veterinary Treatment Facilities as sentinel surveillance for military human populations. Based on feedback from a stakeholder survey, where USAPHC commanders and public health officials were asked to list zoonotic diseases of concern, Lyme disease was selected as the zoonotic disease for this pilot study. Therefore, the specific aim of this study was to evaluate canine *Borrelia* seroprevalence data as sentinel surveillance for Lyme disease in military populations, serving as a model for the ZDR.

Background

Animals as Sentinels

The goal of using animals as sentinels for human disease surveillance is to estimate disease risk in order to make informed recommendations on preventive practices and therapeutic protocols. In order to serve as sentinels, the animal models must be susceptible to the pathogen and must produce a measurable symptom or response that indicates exposure or infection (6). It is also important that the diagnostic samples are easily obtained and that accurate diagnostic tests are available (6). Therefore, in order for dogs to serve as effective sentinels for Lyme disease in humans, they should be more likely to be exposed or infected than their human counterparts and/or should have a more rapidly detectable disease progression. These criteria ensure the risk assessment information from the sentinel animals is timely and allows for protection of the human counterparts.

Military Pets as Sentinels

Pets are very important to military populations. With frequent moves (every 2-3 years), the family pet is often the most reliable friend a military child can have. Also, the consistency of a family pet can greatly help when a parent, spouse, or significant other is deployed overseas for extended periods of time. Pets in the military are eligible for medical care at U.S. Army Veterinary Treatment Facilities (VTFs). These VTFs provide military pets with medical

care at reduced rates. Before each move these military pets receive free comprehensive health physicals in conjunction with mandatory health certificates from the departure site and if they move on to a military installation they may also be required to undergo in-processing physicals at the arrival site. These frequent physical examinations and associated laboratory work provides the U.S. Army VTFs the opportunity to build a robust animal disease database. Also, because they spend their time in the same geographic location and precise environmental areas as their owners, military pets have the potential to serve as appropriate indicators for disease than civilian (general population) pet dogs.

One application of the animal disease data from the Army VTFs is at the group level for area prevalence and incidence estimates, where information about the existence of a zoonotic pathogen in the pet population may alert human physicians to the risk of the disease in humans. Another application is at the individual animal level, where a pet may serve as a marker for environmental exposures in the owners; for example, a dog that is diagnosed with lead poisoning should trigger an investigation into where the dog was exposed, and may reveal lead based paints having been used in the house.

Lyme disease and Military Relevance

Lyme disease is the most commonly reported vector-borne illness in U.S. and the seventh most common Nationally Notifiable disease (7). Caused by the spirochetal bacterium *Borrelia burgdorferi*, it is transmitted to humans through the bite of infected *Ixodes* spp. ticks. Successful transfer of the pathogen from the invertebrate host to the vertebrate host requires the tick to remain attached during the 36 to 48 hour feeding time. Typical symptoms in humans include fever, headache, fatigue, and a characteristic skin rash called erythema migrans (EM). If left untreated, infection can spread to joints, the heart, and the nervous system. Sixty percent of those infected develop chronic, severe arthritis (66). Even with treatment, 10-20% may develop Post-Treatment Lyme Disease Syndrome (PTLDS) (7). In 1998, direct medical costs associated with Lyme disease were estimated to be $2,970 per case (8). Indirect costs, including associated non-medical costs and projected losses in productivity, were estimated to be $5,202 per case. For the same time period mean productivity loss per clinically defined late stage patient was $9,108 (8).

For the military, the cost associated with lost man-power productivity may also be felt in disrupted military training, deployment, and operations. Because of the significant impact of Lyme disease to military operations and its overall potential to threaten public health, a Department of Defense Lyme disease program was established. The program, overseen by the Armed Forces Health Surveillance Center (AFHSC), involved listing Lyme disease as one of the Department of Defense Tri-Service reportable events (9). In addition, AFHSC is required to compile an Annual Lyme Report, which provides information for cases of Lyme disease diagnosed during the last 10 years. The report presents the total cases of Lyme disease, by calendar year, and also provides detail by Service, for active component, Reserve/Guard and other beneficiaries. In 2011 the report showed that Lyme disease cases increased over 1.5 times from 2001 to 2007 (10). A detailed analysis of the communities affected showed that the majority of cases come from military sites in the Northeastern United States and Germany (11).

Dogs as Sentinels for human Lyme disease

Human Lyme disease is diagnosed based on symptoms, physical findings (e.g. EM lesion), and the possibility of exposure to infected ticks (7, 9). Laboratory testing is based on antibody testing but can be problematic as results merely indicate exposure and not current infection. Treatment is most successful if implemented in the early stages of disease, but many believe PTLDS exists, where even with treatment symptoms may persist for extended periods or may never resolve (7, 9). Due to concerns over side effects and lack of efficacy, human Lyme disease vaccines were withdrawn from the market in 2002 (7). Currently prevention strategies are limited to avoiding exposure such as using insect repellent, applying pesticides, reducing tick habitat, and if bitten, promptly removing the tick (7). Given the difficulty in diagnosing human Lyme disease, its potential to become a chronic condition, and its significant economic impact, it is clear that efforts must be focused preventing exposure. In order to do this one must be knowledgeable about the presence of the risk.

There have been many studies (12, 13, 14, 15) exploring the utility of dogs as sentinels for Lyme disease in humans. Because of the difficulties associated with detecting the *Borrelia* antigen, most of these studies have used sero-positivity as their indicator for risk. Arguments for their use include that they develop high circulating levels of antibodies to *Borrelia*, are easily sampled, and appear to be at greater risk (almost six times as likely as people to be

exposed to infected ticks than people) of infection than people (16). One reason for their greater risk of infection is their higher likelihood to be in direct contact with the ticks; more time outdoors and more time running through tall grasses, bushes and trees. Another reason is that attached ticks are less likely to be detected because of the animal's hair coat, which allows the tick the full 36-48 hours needed to transfer the pathogen to them. There are, however, some potential limitations too. First, 95% of infected dogs remain without clinical signs; which means that they may not produce outwardly visible signs to warn us of the risk of disease (17). This may not be much of an issue in regions where clinics practice routine annual tick titer screening, but may still be an issue in regions where *Borrelia* is less common or just now emerging. Second, unlike their human counter parts, many dogs are on routine flea and tick control. This may mean that some humans are actually at a higher risk than their canine counter parts. Third, as with humans, canine sero-positivity does not indicate incident cases but rather a history of exposure and therefore cannot be relied upon as a definitive indicator of Lyme disease risk for a precise location or time. Studies show that canine antibodies to *Borrelia* remain elevated for at least 69 weeks (18), which means the sero-positive dog may actually have been exposed more than a year earlier and potentially at an entirely different location. For this reason, it is important to acquire travel history information when trying to use canine sero-positivity as a predictor of Lyme disease risk in a given area.

The Borrelia genospecies of concern in the United States, *B. burgdorferi* sensu sticto (Bb) (17). Studies using dog serology in sentinel surveillance for human Lyme disease in the United States are most likely detecting Bb, not one of the other *Borreilae* genospecies. Unless otherwise indicated, for the remainder of this document, the Borrelia pathogen will be referred to as Bb, the most common disease causing genospecies in the United States.

Structure of this Chapter

Although using dogs as sentinels for Lyme disease risk in humans is not straight forward, careful consideration of both the advantages and limitations this approach can still make it a powerful adjunct to human Lyme disease surveillance efforts. Doing so involves three specific objectives, each of which will be addressed in a separate section. The first objective (Section I) is to investigate the utility of using military pet dogs as sentinel surveillance for military human Lyme disease by determining the association between Bb seroprevalence in military pet dogs and

military human Borreliosis data. The second objective is to compare military pet dog data to published civilian pet dog data in order to determine if there is a difference between Bb seroprevalence by location. (Section II). And finally, the third objective (Section III) evaluates the validity of the military pet dog Bb seroprevalence data by investigating potential confounders to the association between military pet dog Bb seroprevalence and test result location (installation).

This overall approach is outlined in the below diagram.

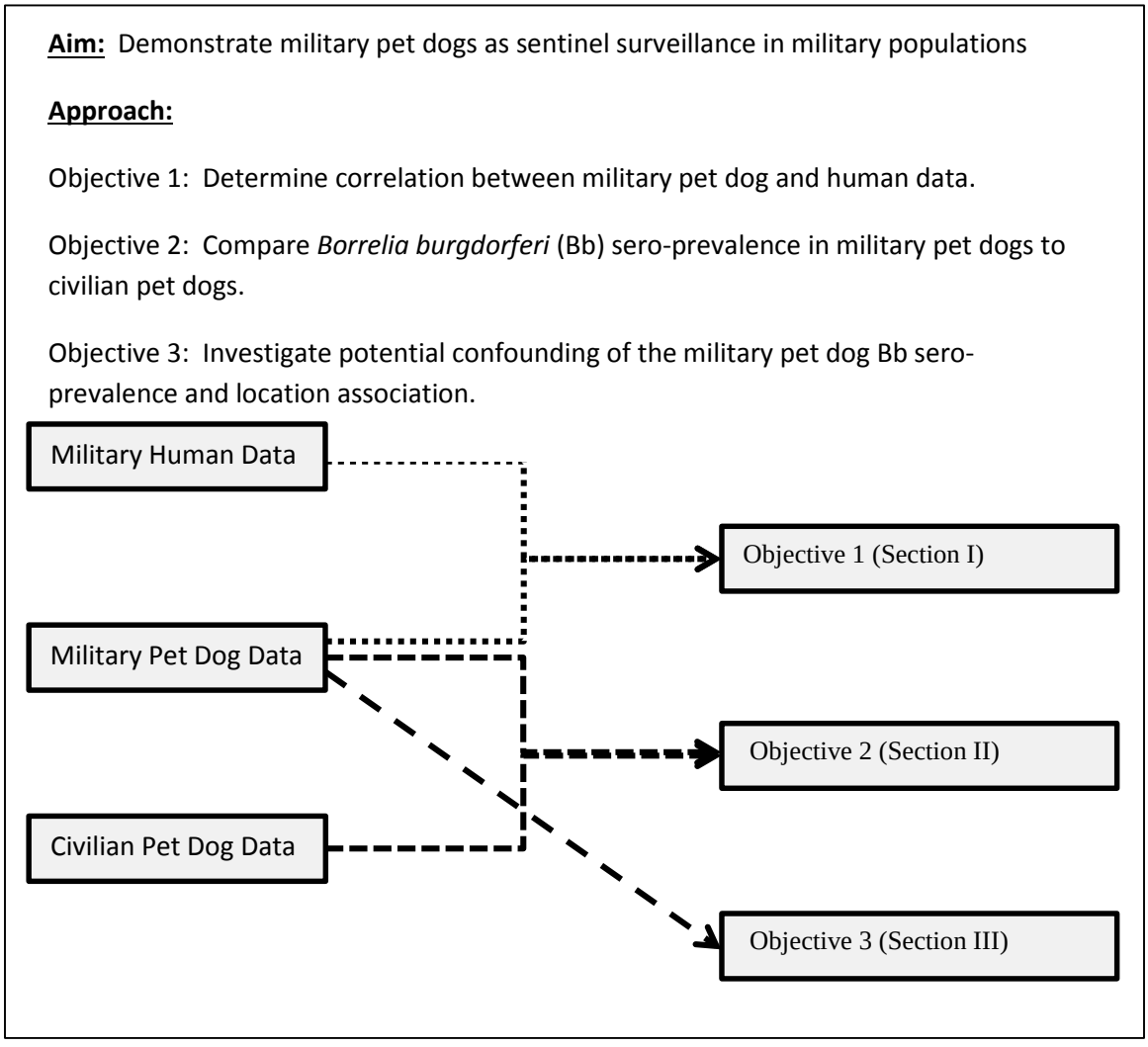

Figure 5.1: Overall Study Approach

Section I: Determining the Association between Military Pet Dog and Human Borreliosis Data

Studies Using Canine Bb Seroprevalence data in Sentinel Surveillance for Human Borreliosis in Civilian Populations

Many studies have demonstrated canine Bb prevalence as sentinel surveillance for human Lyme disease at the state or regional level. A 1991 study of pet dogs in Massachusetts showed estimates of the prevalence of Bb antibody in pet dog populations to be a sensitive, reliable, and convenient measure of the potential risk to people in the environment for the state of Massachusetts (13). Olson et al. used both coyote and pet dog serology, to confirm that Bb was not enzootic in the San Diego area, finding very low or insignificant rates of infection in these canine populations which corresponded to low rates observed in human populations (14). A Dutch study of hunting dogs and their owners showed the same order of seroprevalence of antibodies against Bb in hunters and hunting dogs, indicating that estimates of seroprevalence among hunting dogs are predictive of the risk of LB in humans (19). A 2004 study showed that both canine seroprevalence and human case count data showed fewer Lyme diseases cases from North Carolina as compared to Virginia, Maryland, and Pennsylvania, supporting the utility of dogs as a sentinel to characterize the risk of Bb transmission to humans in a defined geographical location (15).

In 2009 a national clinic-based serologic survey of pet dog prevalence and geographic distribution for four vector-borne diseases (*Dirofilaria immitis*, *Anaplasma phagocytophilum*, *Ehrlichia canis*, and *Borrelia burgdorferi*) was done (20). Results showed that over 5% of tested dogs across the United States were positive for antibodies to *B. burgdorferi.* The same study showed the highest sero-prevalence in dogs to be in the Northeastern U.S. (11.6%) which correlates with the etiologic foci of development for the disease in humans. In September of 2011 the Centers for Disease Control and Prevention published an article that further discussed this correlation, investigating whether or not canine serology could be used as an adjunct to human surveillance for Lyme disease (12). The study confirmed an overall correlation between canine sero-prevalence and existing reports on human incidence, and

suggested that surveillance using dogs as sentinels may be a sensitive marker for increased risk of Lyme disease for humans.

The Utility of Military Pet Dog Seroprevalence data in Sentinel Surveillance for Human Borreliosis in Military Populations

Equivalent studies had never been performed in military populations, but rising concerns over the continued increase in Lyme disease cases in military populations in 2011 provided the motivation to do so (21). In the military, a diagnosis of Lyme disease is a Tri-Service Reportable Event (9). This means that although all branches of service are responsible for implementing their own reporting system, data collection, and quality assurance, the collected data are all integrated into the Defense Medical Surveillance System (DMSS) database. Once in DMSS the data from all branches of services are available for further analysis. At the initiation of this study, according to DMSS, there appeared to be an increase in the number of reported human cases of Lyme disease at installations both in the northeastern part of the United States and in Germany. In addition, the 2011 creation of the USAPHC and their subsequent release of monthly ZDRs set the stage for exploring the use of military pet dogs as sentinel surveillance for zoonotic disease in military personnel, specifically in the case of Lyme disease.

In order to determine if military pet dogs can serve as sentinel surveillance for Lyme disease for military human populations, the correlation between canine Bb seroprevalence and human Lyme disease case counts must be investigated. Thus the aim of this phase of the study is to estimate the correlation between Bb seroprevalence in military pet dogs and military human Borreliosis data across selected military installations.

Methods

Site Selection

Military site selection was based on the rate of Lyme disease by military installation in the human active duty population for the ten years before the onset of this study, 1 January 2001 to 31 December 2011. Case count data derived from DMSS where cases of Lyme disease were defined as health encounters with an ICD-9 of 088.81 in any diagnostic position, including one inpatient encounter, one reportable event, or two outpatient encounters within 60 days of each other; limited to incident cases (the earliest encounter defining the case). Denominator data (person-time) was based on the number of active duty assigned to the installation on an annual basis. The resulting rate data, cases per 1,000 person-years, allowed for the adjustment of unequal population sizes across the different locations. Because population data per installation is only obtainable for active duty personnel in DMSS, this rate data was limited to the active military component.

Installation selection was based on calculated cumulative human Lyme disease rates and then weighted (2 x 2011 cases) rates for all of the installations. The purpose of the weighting was to place greater emphasis on installations with recent increases in Lyme disease; this was used to define emergent installations. Installations were then sorted on the weighted rates and a list of 20 installations with rates higher than 15/100,000 and 15 installations with rates less than 6/100,000 were then compiled in order to guide the selection of military veterinary treatment facilities. In order to have 95% confidence in detecting an average seroprevalence of 7% across all installations, it was determined that each location should be able to provide data on at least 280 dogs over the course of the one year study (19). The final list of participating military veterinary facilities was determined by working directly with USAPHC Veterinary Animal Medicine Program personnel, where only facilities with sufficient workload and manpower were selected to participate.

Data Collection

Human Military Lyme Disease Data

Sources

Military human case count data derived from two different sources. For consistency purposes, DMSS was again queried for Lyme disease case count data. In addition, based on data from a separate study that indicates M2 (Military Health System Management Analysis and Reporting Tool) is a more suitable source of human military data for zoonotic disease reporting, a query for Lyme disease cases from M2 was also submitted. There are two main difference between the systems; M2's ability to collect data from civilian medical facilities (non-military referral clinics) and M2's ability to pull data from all military health system beneficiaries; active duty Service members and retirees and the dependents of both populations. Both systems were queried by third party personnel; M2 by the U.S. Army Public Health Command and DMSS by the Armed Forces Health Surveillance Center (AFHSC). Queries were conducted based on written requests containing identical case definitions.

The data provided from DMSS were in the form of counts and rates. Case count data was determined by the number of cases of Lyme disease in the active duty population per military installation for the one year period from 1 November 2012 to 31 October 2013. Rate data was calculated using annual person-time by installation as determined by number of active duty records on file for that installation in the Defense Manpower Data Center (DMDC) for the one year period. The DMDC serves to collate personnel, manpower, training, financial, and other data for the Department of Defense, cataloguing it for purposes of healthcare, retirement funding and other administrative needs (22). In order to adjust for unequal population sizes across the different locations, the rate data were calculated as the number of cases per 100,000.

In order to be comparable, the data from M2 data were also in the form of counts and rates. Cases count data were again determined by the number of cases of Lyme disease per military installation for the one year period from 1 November 2012 to 31 October 2013. However, because the population data in M2 comes from the Defense Enrollment Eligibility Reporting System (DEERS), data were not limited to the active duty component. DEERS is a

computerized database that maintains personnel and benefits information for active, retired, and reserve service personnel as well as their family members and DoD civil service personnel and eligible contractors (23). The enormity of this potential population exceeded he capabilities of the M2 query module, therefore the data was instead retrieved directly from the Military Health System Repository (MDR). The MDR case count data included active duty, retirees, and their dependents. In order to adjust for unequal population sizes across the different locations, rates were calculated. In order to adjust for unequal population sizes across the different locations, the rate data were again calculated as the number of cases per 100,000.

Case Definition

In both medical data systems a confirmed case of Lyme disease was documented with an ICD-9 code of 088.81 and was clinically defined based on Armed Forces guidelines and case definition as 1) a diagnosis of erythema migrans (EM) by an experienced clinician with a known exposure (history of time spent in endemic area), 2) a diagnosis of EM with laboratory evidence (culture, IgG, IgM) of infection and without a known exposure, or 3) a case with at least one late manifestation (neuromuscular, neurologic, cardiovascular) that has laboratory evidence of infection. Case count data for both medical data systems were based on a surveillance definition of a health encounter with the ICD-9 code of 088.81 in any diagnostic position, including 1 inpatient encounter, 1 reportable event, or 2 outpatient encounters within 60 days of each other; each one applied a rule of one diagnosis per life time.

Canine Military Borellia burgdorferi Seroprevalence Data

Source

Military canine Bb seroprevalence data were collected from selected military veterinary facilities from 1 November 2012 through 31 October 2013. Although both government-owned Military Working Dogs (MWDs) and privately-owned pet dogs are seen at military veterinary facilities, this study was restricted to only the pet dog population. This restriction was done for several reasons; 1) MWDs are on a very high level of flea and tick control, therefore are at a reduced risk for tick bites, 2) MWDs receive daily grooming, therefore any tick that does bite is likely to be removed before the Bb pathogen is transferred, 3) MWDs generally train and live in well maintained environments and may be less likely to be in areas where ticks live. The pet dog population included dogs belonging to active

duty Service members, retirees and Reservists on orders. All participating veterinary facilities were provided the same Letter of Instruction (LOI) explaining the study details. Although owner participation was voluntary, participants were offered a discount if they agreed to fill out in a short survey before their pet's examination; owners were limited to only one pet per household. Details of the survey are provided in Section III of this chapter. Upon completion, the survey was handed to veterinary facility staff, who then annotated the IDEXX 4Dx Plus® test results on the last page of the survey along with the date and facility name.

Case Definition

Canine *Borrelia* sero-positivity was determined utilizing the in-house IDEXX 4Dx Plus® serological test already routinely used in all military veterinary facilities. The test is designed to detect the presences of serum antibody to C6, a synthetically produced peptide that is encoded by specific surface lipoproteins of *B. burgdorferi.* The genes of this surface lipoprotein (IR6) are only expressed during infection and replication of the spirochete in the mammalian host therefore; the presence of serum antibodies to C6 indicates host invasion and infection with *B. burgdorferi* allowing differentiation between vaccination and true infection (24, 17). In addition, the IR6 surface protein is genetically, structurally, and antigenically highly conserved among many *B. burgdorferi* strains. Experimentally, the C6 antibody response is detectable three to five weeks post infection and stays positive for at least 69 weeks. This is earlier than with conventional assays and even before the onset of clinical lameness. Most importantly perhaps, unlike previous ELISA tests, the C6 antibodies detected in the SNAP 4DX are not increased in dogs infected with *Dirofilaria* spp., *Babesia* spp., *Ehrlichia* spp., *Rickettsia* spp., or *Leptospira* spp. (25, 17). The IDEXX 4Dx Plus® is a qualitative test that claims a sensitivity of 98.8% and a specificity of 100%. Both diagnostic test results (with clinical signs) and routine health screening test results (without clinical signs) were included in this study. Seroprevalence status was associated with the date and location of the current test, there was no way to differentiate between incident versus prevalent cases.

Data Analysis

Since military human hospitals and military veterinary facilities have different catchment areas, analyses of the association between human and veterinary data were conducted in two different ways. Overall there are fewer

veterinary facilities in the military than there are human hospitals, therefore the catchment areas associated with veterinary facilities tend to be much larger, encompassing more than one military installation.

The following diagram depicts the differences in catchment areas between military human hospitals and veterinary facilities.

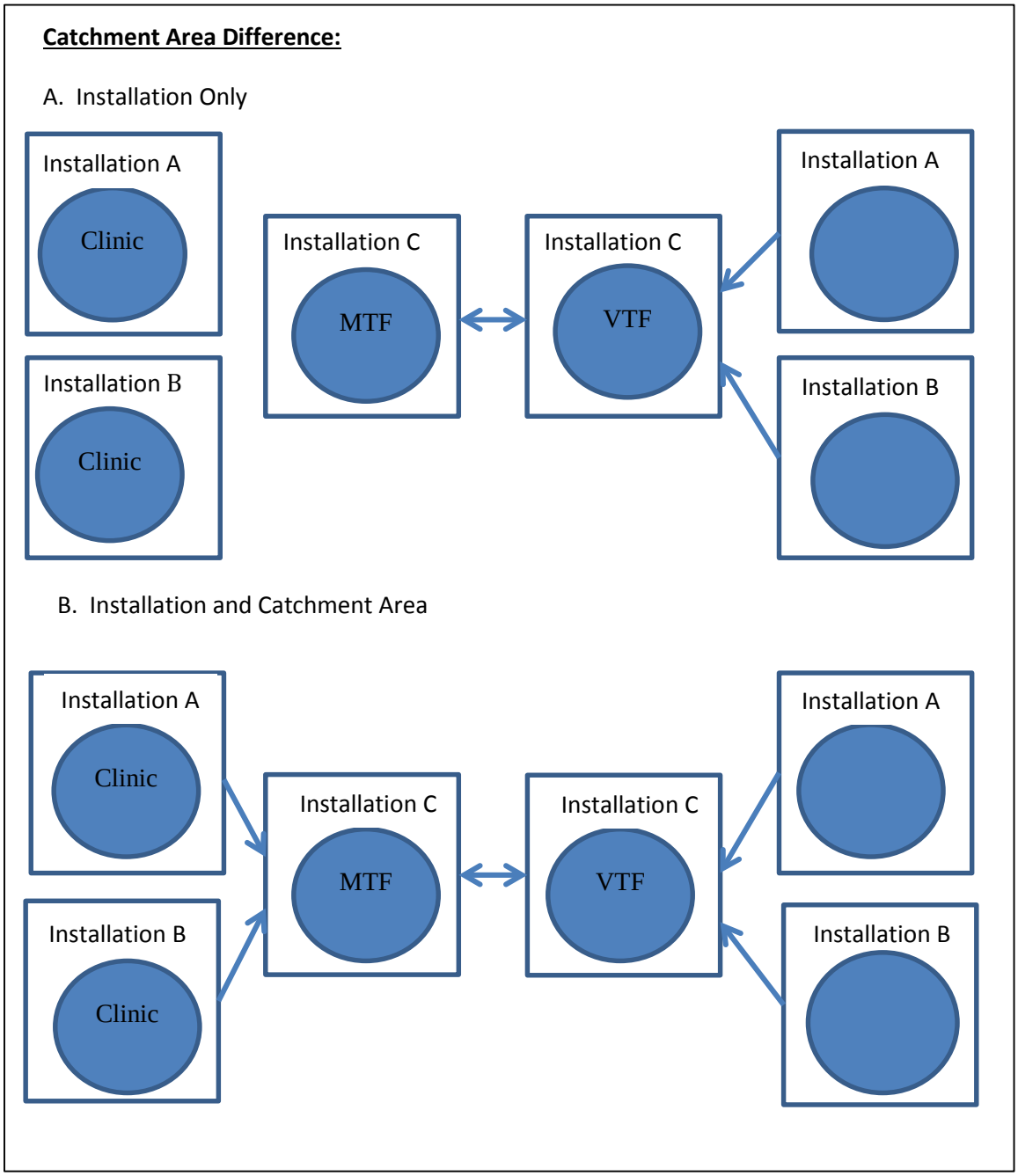

Figure 5.2: Explanation of Catchment Areas: MTF = Medical Treatment Facility providing human care, VTF = Veterinary Treatment Facility providing animal care

The installation only analysis means the human data were limited to that specified by the human medical facility location (Figure 5.2A). The installation and catchment analysis expanded the human data to include the installations reflected by the veterinary clinic catchment area (Figure 5.2B).

The above mentioned approach created two different data sets for the human rate data. Each data set was first checked from normality using MiniTab. The correlation of military pet dog and military human Borreliosis data was estimated using the Spearman Rank Correlation since the data were not normally distributed as part of the assumption to calculate bivariate correlation (26). First both the human and dog data sets were ranked from 1 to 15 by installation or catchment area, with 1 signifying the area with the highest incidence rate (human) or seroprevalence (dog). The Spearman Rank Correlation Coefficient was then calculated.

Results

Human Military Lyme Disease Data

DMSS Data (1 November 2012 through 31 October 2013)

There were only 20 confirmed cases of human Lyme disease resulting from the DMSS query. The highest rates of Lyme disease were reported from Germany (Stuttgart) and the Northeastern United States (Virginia, New York, and Virginia). Overall, eight of the installations (53%) reported having no cases over the one year period; two of which came from locations with a high expected prevalence based on previous years. The low rates of disease reported by this data system for this period of time made conclusion based on the data questionable; therefore it was decided to also perform the remaining analyses on the initial ten year DMSS query used to select study site locations.

MDR Data (1 November 2012 through 31 October 2013)

The MDR query resulted in 91 confirmed cases of human Lyme disease. Again the highest rates of Lyme disease came from installations in Germany and the Northeastern United States. It is unclear why no cases were reported from Walther Reed Medical center, but it is suspected that the site was somehow missed in the query process. Table 5.1 shows the case count and disease rate data for the two human data sources for the one year study period. For completeness purposes, the table includes data from the installation of interest (in bold) as well as data from

installations in the catchment area (indented but not bold), however, analysis by catchment area was not performed due to the lack of rate data for installations without cases.

Table 5.1: Comparison of Lyme Disease Data for two Military Human Medical Data Systems (2011 to 2012)

	Confirmed Lyme Disease Cases (#)		Lyme Disease Rate (per 100,000 person-year)	
Installation	**DMSS**	**MDR**	**DMSS**	**MDR**
Dog Center Europe				
Kaiserslautern	0	3	0.0	88.6
Landstuhl	0	2	0.0	26.0
Ramstein AB	0	2	0.0	41.2
Stuttgart	1	9	89.4	120.8
Spangdahlem AB	0	3	0.0	38.1
Fort Belvoir	2	4	50.1	14.8
Pentagon/Navy Annex	0	5	0.0	70.7
Aberdeen Proving Ground	0	3	0.0	50.6
NAWC Patuxent River	0	5	0.0	49.5
MCB Quantico	4	6	63.3	37.3
Walter Reed	0	0	0.0	0.0
Fort Drum	5	10	34.0	30.0
USMA West Point	0	6	0.0	56.5
MCB Camp Lejeune	3	6	10.6	20.4
Fort Lee	2	3	29.7	15.4
NAS Sigonella	0	0	0.0	0.0
NS Rota	0	0	0.0	0.0
Fort Bragg	2	10	4.9	9.2
Fort Carson	0	1	0.0	1.6
Fort Hood	1	4	2.8	4.1
MCB Camp Pendleton	0	1	0.0	3.0
McChord AFB	0	1	0.0	1.0
Fort Lewis	0	0	0.0	0.0
NTC Great Lakes	0	1	0.0	57.0
TOTAL	**20**	**91**		

* The lack of rate data for installations without cases, prevented catchment area data.

DMSS Data (1 January 2001 to 31 December 2011)

There were 1390 confirmed cases of human Lyme disease reported DoD wide from 1 January 2001 to 31 December 2011; 322 (23.17%) of these cases from installations included in this study. The highest rates of Lyme disease came from Germany and the Northeastern United States while the lowest rates came from Southern Europe (Spain and Italy) and Western, Southwestern, and Midwestern U.S. (Washington, California, Texas, Illinois).

Table 5.2: Human Military Lyme Disease Data from 2001 to 2011: DMSS Data Source

Installation	Confirmed Lyme Disease Cases (#)	Lyme Disease Rate (per 100,000 person-year) Installation	Catchment
Dog Center Europe			77.9
Kaiserslautern	13	114.8	
Landstuhl	8	63.1	
Ramstein AB	37	55.7	
Stuttgart	13	89.8	89.8
Spangdahlem AB	14	43.4	43.4
Fort Belvoir	9	32.9	41.3
Pentagon/Navy Annex	1	53.8	
Aberdeen Proving Ground	16	51.8	
NAWC Patuxent River	12	42.1	
MCB Quantico	24	33.8	
Walter Reed	10	33.2	
Fort Drum	29	21.2	100.6
USMA West Point	24	179.9	
MCB Camp Lejeune	46	18.0	18.0
Fort Lee	10	17.8	17.8
NAS Sigonella	0	0.0	0.0
NS Rota	0	0.0	0.0
Fort Bragg	25	5.3	5.3
Fort Carson	8	4.0	4.0
Fort Hood	12	2.5	2.5
MCB Camp Pendleton	2	0.8	0.8
McChord AFB	1	2.6	2.4
Fort Lewis	6	2.2	
NTC Great Lakes	2	1.1	1.1
TOTAL	**322**		

Canine Military Lyme Disease Data

Overall there were 3776 Bb seroprevalence test results collected from 3996 surveys during the one year study period. There were 109 sero-positive dogs contributing to an average sero-positivity of 2.89% study-wide; 3.30% seroprevalence (31/937) in Europe and 2.75% seroprevalence (78/2839) in the U.S. Approximately 5% of the submitted surveys were missing the requested test results, with nearly twice as many of the missing results coming from areas of expected high prevalence. The highest Bb seroprevalence came from Germany (Kaiserslautern, Stuttgart) and the Northeastern United States (Virginia, New York) while the lowest rates came from Southern Europe (Italy) and Western, Southwestern, and Midwestern U.S. (Washington, California, Texas, Illinois, Colorado).

Table 5.3: Canine *B. burgdorferi* Seroprevalence by Military Installation (Highest to Lowest)

	Sero Negative		Sero Positive		Total
Installation	**n**	**%**	**n**	**%**	
Fort Drum, NY	304	90.8%	31	9.3%	335
Fort Belvior, VA	204	95.3%	10	4. 7%	214
Stuttgart, DE	204	96.2%	8	3.8%	212
Kaiserslautern, DE	313	96.3%	12	3.7%	325
Greaty Lakes. IL	244	96.8%	8	3.2%	252
Rota, ES	65	97.0%	2	3.0%	67
Spangdahlem, DE	298	97.1%	9	2.9%	307
Camp Pendelton, CA	42	97.7%	1	2.3%	43
Fort Carson, CO	280	98.3%	5	1.8%	285
JBLM, WA	345	98.3%	6	1.7%	351
Camp Lejeune, NC	355	98.3%	6	1.7%	361
Fort Lee, VA	297	98.3%	5	1.7%	302
Fort Bragg, NC	520	99.1%	5	1.0%	525
Fort Hood, TX	170	99.4%	1	0.6%	171
Sigonella, IT	26	100.0%	0	0.0%	26

Comparison of Canine to Human Lyme Disease Data: Sensitivity of the Canine Model

Using sero-positivity alone, far more dogs tested positive for exposure to *Borrelia burgdorferi* than incident human cases were captured both by the MDR and DMSS data systems. Comparing canine sero-positivity data (sero-positivity for every 100,000 dogs tested) by installation with human rate data, canine estimates were between 42 to 272 and 42 to 1672 times greater than reported DMSS and MDR human data for the same time period, and 45 to 4660 times greater than the reported DMSS data for the ten year period.

Estimating the Association

Animal Data to Human Data from DMSS: 1 November 2012 through 31 October 2013

Because DMSS reported no cases of human Lyme disease for this one year period for over half of the sites (8/15 or 53%), the correlation between military pet dog and military human Borreliosis data using the Spearman Rank Correlation only pertains to 46.7 % (7/15) of the data and therefore does not give a complete picture. This estimation, is limited to these seven locations showed a very strong positive correlation between human Lyme disease rate and military pet dog Bb seroprevalence by military installation ($r_s = 0.821$). Because of the lack of rate data for the installations without cases, it was not possible to expand the human disease data to reflect the catchment area of the veterinary facilities.

Animal Data to Human Data from MDR: 1 November 2012 through 31 October 2013

The MDR data source reported two sites with no cases of human Lyme disease for this one year period, therefore the association between animal Bb seroprevalence and human Borreliosis data using the MDR data source only pertains to 87 % (13/15) of the data and therefore also does not give a complete picture. The resulting correlation was strong ($r_s = 0.538$). It was not possible to expand the human disease data to reflect the catchment area of the veterinary facilities because of the lack of rate data for the installations without cases.

Animal Data to Human Data from DMSS: 1 January 2001 to 31 December 2011

There was a strong positive correlation between human Lyme disease rate and military pet dog Bb seroprevalence by military installation ($r_s = 0.525$). When expanding the human disease data to reflect the catchment area of the veterinary facilities the correlation was slightly higher ($r_s = 0.664$). Both approaches indicate consistency in the rankings of Lyme disease in humans and Bb seroprevalence in dogs; the similarity in the resulting correlation coefficients indicates that either approach is suitable.

Table 5.4: Correlation of Human Lyme Disease Rate to Military Pet Dog *B. burgdorferi* Seroprevalence

Comparison	Applicability*	Spearman's rho (95% CI)
DMSS 1 year data	46.7 % (7/15)	0.821 (0.178, 0.972)
MDR 1 year data	87 % (13/15)	0.538 (-0.018, 0.839)
DMSS 10 year data (Installation Only)	100% (15/15)	0.525 (0.018, 0.817)
DMSS 10 year data (Catchment)	100% (15/15)	0.664 (0.230, 0.877)

*Because installations with no human Lyme disease cases were not included in this analysis, calculated estimates have varying applicability/relevance. Data are presented as the percent of installations (number of installation used over total) analyzed.

Discussion

There have been several studies exploring the potential to use dogs as sentinels for Lyme disease risk in humans. Dogs are well suited for this role for the following reasons. First, they are susceptible to the causative agent and when exposed produce an easily sampled antibody response, sometimes accompanied by clinical signs similar to those in people. Second, they have a greater likelihood of being in tick habitats and therefore a greater tendency for tick exposure. Third, because of their hair coat, an attached tick is more likely to be permitted to feed the full 36-48 hours required to transmit the *Borrelia* spirochete.

To our knowledge, this is the first study to investigate the utility of pet dogs belonging to service members as sentinels for Lyme disease risk in military populations. Estimates showed that there is a strong positive association between human incidence rates of Lyme disease and canine Bb seroprevalence in military human and canine populations. In addition, Bb sero-positivity per 100,000 dogs was consistently much higher than Lyme disease incidence per 100,000 regardless of human data source used, indicating that dogs are a more sensitive indicator of pathogen presence than their human counterparts. Because of a concurrent study investigating the differences

between existing military human disease data sources, this study utilized two different human data sources to evaluate the association between canine and human Lyme disease. One difference between the two is how each system receives its population (denominator) data. DMSS pulls its population data from DMDC, Defense Manpower Data Center, limiting the personnel data to that of active duty service members only. MDR on the other had pulls its population data from DEERS, Defense Enrollment Eligibility Reporting System, which captures anyone eligible to receive medical care in the military system to include but not limited to active duty, retirees, and their beneficiaries. Another difference is that DMSS is generally limited to cases seen at military treatment facilities (MTFs), whereas MDR is able to incorporate those cases seen at associated civilian medical facilities. These differences may explain why both the number of cases and the overall incidence rates from the MDR query are so much higher than those resulting from the DMSS query. In fact, the extremely low number of cases resulting from the one year DMSS query limited the usefulness of the data in this study, as the calculated correlation coefficient (r_s = 0.821) only applies to 7 of the 15 study sites. Analysis using the MDR data, on the other hand, didn't produce as high of an association (r_s = 0.538), but because it used all 15 sites it is more representative of the true relationship between military human Lyme disease and canine seroprevalence data. The differences in samples (specific installations compared) and sample sizes (number of installations compared) between the one year DMSS and MDR data may have led to either over and/or underestimation of these estimates, making direct comparison of the correlation coefficients inappropriate. The ten year DMSS data estimated a similar association (r_s = 0.524) as the MDR data. This may indicate that, especially for low prevalent diseases, MDR is a more sensitive data source. There was not a notable difference between the straight by installation association and one done incorporating the catchment areas (using the ten year DMSS data). This is likely due to the fact that VTFs are generally located on larger installations which also happen to have the larger MTFs. These larger MTFs are also more likely to see the more complicated cases and have more advanced diagnostic capabilities, making them more likely to be the location where regional Lyme disease cases are diagnosed.

This study compared two different measures of disease, human incidence and canine prevalence. This was unavoidable due to differences inherent in the data sources used. The human data came from queries of existing medical systems, where it was possible to limit the search to incident cases by location for the given study period. To adjust for differences in installation sizes incident rates per 100,000 were calculated based on person-years. No

such data system currently exists on the animal side; therefore, the canine data was actively collected during the study period. In order to have the smallest impact on current clinic operations, and maintain comparability to studies in the literature, this study utilized the extant in-house canine serological Bb antibody test, the IDEXX 4Dx Plus®. As discussed above, this antibody test is qualitative, where a positive test result merely indicates the dog has had a history of exposure to the Bb pathogen. For this reason it was not possible to determine whether or not a sero-positive dog was recently exposed (incident case) or if they were exposed several months ago; canine incidence was not possible. Canine seroprevalence was calculated as the number of sero-positive results per total number of tests run by installation. Using canine seroprevalence to indicate the presence of *Borrelia* risk has several limitations, however. Because exposed animals may have elevated antibodies for more than a year, sero-positivity at one location may actually result from exposure from another. Other animal and owner factors too have been associated with canine Bbseroprevalence; such as dog breed, preventive medicine practices, recreational activities, and canine living environment. All of these variables are potential confounders to the canine sero-prevalance to installation association and therefore must be considered when drawing conclusions on human Lyme disease risk based on canine seroprevalence alone. Section III of this chapter investigates the impact of these and several other potential confounders in much more detail.

Conclusions

This study found that seroprevalence estimates in military pet dogs not only reflect incidence rates in military human populations, but they also appear to be a more sensitive indicator of *B. burgdorferi* presence then their human counter parts. The findings recommend the use of service member pet dogs as a sensitive, reliable, and convenient measure of the risk of Lyme disease in sentinel surveillance in human military populations.

Section II: A Comparison of *Borrelia burgdorferi* Seroprevalence in Military and CivilianPet Dogs

Civilian Pet Dog *Borrelia burgdorferi* Seroprevalence

From 2001 to 2007 IDEXX Laboratories, in conjunction with marketing new in-house ELISA tests, conducted a rebate driven national clinic-based serological survey aimed at documenting canine infection with, or exposure to, four vector-borne disease agents, *Dirofilaria immitis, Borrelia burgdorferi, Ehrlichia canis,* and *Anaplasma phagocytophilu* (20). The data were compiled into a comprehensive database that was then analyzed to assess geographic trends in rates of positive tests for all four disease agents, finding that percent positive test results varied by agent in different regions of the United States. Looking exclusively at the *B. burgdorferi* (Bb) data, they found a significant difference in sero-positivity at the regional level where titers to Bb were most common in the Northeast (11.6%) as compared to the Midwest (4.0%), West (1.4%), and Southeast (1.0). The findings were published in Veterinary Parasitology in 2009, and included percent positive data at the state level (12). The data from this publication was used to represent the civilian pet dog Bb seroprevalence for the comparison study conducted in this chapter.

In 2011, the Centers for Disease Control and Prevention conducted a study investigating the use of this canine serology as adjunct to human Lyme disease surveillance (12). The study indicated canine Bb seroprevalence to be a sensitive but nonspecific marker of human risk.

Military Pet Dog *Borrelia burgdorferi* Seroprevalence

Currently, there is no equivalent comprehensive pet dog database in the military. The expectation is that observed geographical distributions of Bb sero-positivity in military pet dogs will be similar to those seen in the Bowman study; however, there are several things about the military pet dogs that have the potential to make this population significantly different from the civilian pet dogs in the Bowman study (20). These differences may affect the use of

these dogs in sentinel surveillance. Just like their human owners, military pet dogs are subject to frequent moves. These moves are generally every two to three years and involve massive relocations, generally to different states and often to completely different countries. Also, military owners often deploy to locations "unaccompanied", meaning without family members or pets. During these unaccompanied moves military pets are often fostered out to families or friends, potentially in other states or countries. Another significant difference is that military pets are eligible to receive inexpensive veterinary care at military operated veterinary facilities, making flea and tick prevention affordable and potentially contributing to a pet population with a higher than expected level of preventive medical care. The extensive travel history of military pet dogs along with the high potential for flea and tick prevention raises concerns over the potential to use these dogs as sentinels for human Lyme disease.

In order to determine the potential variation between military and civilian pet Bb seroprevalence, population seroprevalence data collected from military Veterinary Treatment Facilities (VTFs) were compared to the extant 2009 data.

Methods

Data Collection

Civilian Pet Dog Borrelia burgdorferi Data

Civilian pet dog Bb seroprevalence data came from the article published by Bowman et al. in 2009 (20). The data analyzed came from a comprehensive sero-survey of U.S. veterinary clinics conducted by IDEXX Laboratories, Westbrook, ME from 2001 to 2007. Participating clinics received rebates for using a commercial in-house ELISA assay when they submitted a log of all test results. The assay was designed to simultaneously detect qualitatively canine antibodies to multiple vector-borne diseases. From 2001-2006 all dogs were tested with the IDEXX 3DX SNAP test, a commercial in-house C6-based assay for simultaneous qualitative detection of canine antibodies to *E. canis* and *B. burgdorferi*, and to *D. immitis* antigen, in canine serum, plasma, or whole blood. In 2006 IDEXX added a fourth analyte for the detection of *A. phagocytophilum* antibody and changed the name of the test to the IDEXX 4DX SNAP Test. This test was used to collect the remainder of the data set, from 2006-2007. A total of 982,336

dogs were tested for *B. burgdorferi* antibody nation-wide. Percent positive test results were calculated by dividing the number of dogs reported positive for each agent by the total number of dogs tested. Data were collated by county of residence of each dog tested according to postal zip code provided with each record, and then assembled into state and regional groups. Data from the state level was used in this comparative study.

Military Pet Dog Borrelia Data

The military pet dog Bb seroprevalence data for this study were actively collected from participating military VTFs. Due to funding constraints it was not possible to survey all VTFs, instead 15 sites were selected based on human incident Lyme disease data (described in detail in Section I), and ensuring VTFs with both high and low expected prevalence were selected. The military pet dog data were collected from questionnaires designed specifically for this study from November 1, 2012 through October 31, 2013. Facilities were asked to provide data from at least 25 dogs a month, restricted to military-owned companion dogs and only one pet per household. Canine Bb seroprevalence was determined utilizing the IDEXX 4Dx SNAP test routinely used in VTFs. Test results from both wellness exams and diagnostic screening were utilized. *B. burgdorferi* seroprevalence was calculated as the total number of positive test results divided by the total number of dogs tested per site. The VTF level data represents the canine seroprevalence for that particular military installation or community; this installation level data was then used in the subsequent comparative study.

Analysis

Differences in Bb seroprevalence between military pet dogs and civilian pet dogs were examined through the use of the chi-squared test for proportions with a Bonferroni adjustment for multiple comparison (p-values less than 0.005 considered statistically significant).

Results

European locations were not represented in the Bowman study, and were therefore removed from this analysis. For eight of the ten locations there was no statistically significant difference between the Bb seroprevalence for civilian

owned vs military owned pet dogs. However, for two of the locations there is a significant difference. Specifically, the military pet dog prevalence at Fort Lee, VA was lower than expected for the state of Virginia and the military pet dog prevalence at Great Lakes, IL was higher than expected for the state of Illinois.

Table 5.5: Comparison of Civilian Pet Dog to Military Pet Dog *Borrelia burgdorferi* Seroprevalence by Location

Location	Civilian Prevalence by State	Military Prevalence by Installation	p-value
Fort Belvoir, VA	6.7	4.7	0.240
Fort Drum, NY	7.1	9.3	0.128
MCB Camp Lejeune, NC	1.3	1.7	0.505
Fort Lee, VA	6.7	1.7	0.000*
Fort Bragg, NC	1.3	1.0	0.525
Fort Carson, CO	0.4	1.8	0.010
Fort Hood, TX	0.2	0.6	0.159
MCB Camp Pendleton, CA	1.8	2.3	0.81
McChord AFB, WA	0.0	1.7	0.982
NTC Great Lakes, IL	1.0	3.2	0.000*

*significant p-values after adjustment

Discussion

The overall aim of this phase of the study is to demonstrate the utility of service member pet dogs in sentinel surveillance for military populations. There have been several studies (12, 13, 14, 15) in civilian populations showing the use of canine *B. burgdorferi* seroprevalence in sentinel surveillance for human Lyme disease risk, but to date none done looking at this relationship in military populations. In order to justify the investment in establishing a zoonotic disease surveillance program using service member pet dogs, one must demonstrate that there is a meaningful difference between existing civilian pet dog and military pet dog data.

For eight of the ten study locations there was no significant difference; military pet dog seroprevalence matched that of the civilian pet dogs. This finding suggests that service member pet dogs, like their civilian counterparts, are sensitive markers for Lyme disease risk and indicates their potential use in surveillance. Five of the military locations were overseas, and therefore did not have civilian data available for comparison. Because military populations are not limited to living within the United States, this identified a significant gap, that reliance on

civilian pet data available at the U.S. level is insufficient for risk projections for military populations at the global level. The study also indicated significant difference for two of the U.S. locations, specifically Fort Lee as compared to state of Virginia and the Great Lakes Naval Air Station as compared to the state of Illinois. Investigation into reasons for these discrepancies highlight why military pet dog data are a more appropriate indicator of Lyme disease risk in military populations than that collected from civilian dogs.

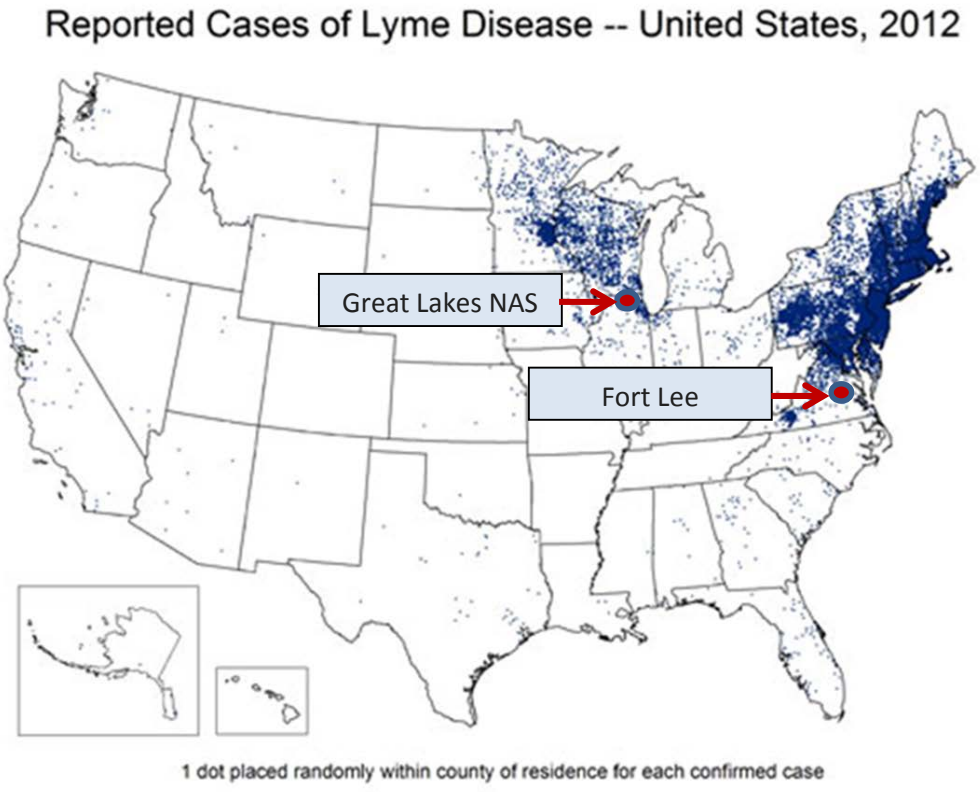

Figure 5.3: Map showing location of military installations where significant difference between military and civilian pet dog *B. burgdorferi* sero-prevalence exist.

Figure 5.3 overlays these locations on a map depicting the geographic distribution of Lyme disease cases reported to the CDC in 2012 (7). From this map one can see the Great Lakes Naval Air Station is located in the area of Illinois with the highest number of reported cases. The fact that the military pet dog data was restricted to samples taken from this exact location, not diluted by averaging Bb seroprevalence with data from the lower risk areas in the southern part of the state, explains why the prevalence is higher than expected. The number of Lyme disease cases in the Midwest has been increasing, with the southern movement of deer ticks from Michigan and Wisconsin (27).

The data obtained from the military pet dog population at this location is not only a more specific indicator of the risk for the military population, but it may also be predictive of the southern creep in this area, serving as an early warning for cases in human service members in the area. A similar, but opposite, phenomenon may be occurring at Fort Lee, VA where the military pet dog seroprevalence was lower than expected for the state of Virginia. Looking at Figure 5.7 one can see that the military installation is located in an area of the state with fewer reported cases of human Lyme disease

Conclusions

This study found reassuring similarities between military pet dog and civilian pet dog *B. burgdorferi* seroprevalence data for 8 of the 15 study locations. In addition, it showed that there is a meaningful difference between 2 of the locations. Furthermore, investigation into the difference appear to indicate that military pet dogs may be more appropriate indicators of Lyme disease risk for military populations than civilian pet dog data. It is recommended that the military utilize data from their pet animal populations as opposed to extant data from civilian pet populations, in sentinel surveillance for zoonotic diseases.

Section III: Evaluating the Military Pet Dog *Borellia burgdorferi* Seroprevalence and Location Association: Investigation of Confounders

Risk Factors for Borrelia burgdorferi

Lyme borreliosis is the most prevalent human vector-borne disease of temperate zones of the northern hemisphere, with notable hot spots in the eastern United States and Central Europe (28, 29). [85,000 cases reported annually in Europe in 2006; 15-20,000 annually in US in 2006 (WHO)]. The bacterial pathogen is maintained in a horizontal transmission cycle between is vector, hard bodied *Ixodes* spp. ticks, and vertebrate reservoir hosts. Although the specific *Borrelia* genospecies, *Ixodes* species, and vertebrate host vary between geographic locations, the overall life cycle is the same. The most common *Borreilia* genospecies in the United States is *B. burgdorferi*, therefore unless

otherwise indicated, for the remainder of this document, the *Borrelia* pathogen will be referred to as Bb. The slow-feeding *Ixodes* ticks have a life cycle that lasts two years and involves three separate blood feeding events, one per life stage (30). The first feeding is on birds and small mammals by the larvae in the summer, then the following spring/summer the nymph feed on rodents and larger mammals, and lastly the adults feed on large mammals again in the fall. Nymphs are the most likely source of infection in both dogs and humans (17, 31, 32). Favored habitats for questing (host-seeking) nymphs include meadows and fallow land, (31), dense wood-land, and areas with dense shrubs or bushes (32). Exposures can be either peridomestic (interface between neighborhoods and wilderness areas) or recreational (32). An example of a peridomestic exposure would be an animal or human that is bitten when spending time in their backyard which adjoins a wooded area of public land. Recreational exposures generally involve traveling to a natural area that harbors the ticks.

Previous studies investigating the use of dogs as sentinels in Lyme surveillance have discussed various risk factors for infection such as breed, recreational use/activity, living environment/residence lifestyle, sex, age, and knowledge/evidence of tick exposure (12, 13, 14, 15). Many of these factors make intuitive sense, for example, dog breeds with long, dark hair coats may be less likely to have an attached tick detected and removed. Animals that recreate outdoors or live in certain environments may have a higher likelihood tick exposure. At least one study has indicated that male dogs may be predisposed to exposure (33). The age of the dog at the time of sampling may be a proxy for cumulative time of exposure (13). In the diagnosis of human Lyme disease a history of tick exposure in a known endemic area can be sufficient to confirm a clinical diagnosis. Although this is not enough in canine diagnosis, the knowledge of tick exposure and/or the lack of tick prevention can be helpful.

The aim of this section is to investigate if the observed associations between seroprevalence and location (installation) are confounded by other risk factors for canine Bb seroprevalance. To do this the relationship was evaluated with regard to potential Lyme disease risk factors to include travel history as a proxy for the true origin of exposure.

Methods

Site Selection

The methods for selecting participating military veterinary treatment facilities have been described in detail in Section I of this Chapter. Briefly, VTFs were selected to ensure equal representation of both high and low expected Bb seroprevalence based on the rate of Lyme disease in the human active duty population. In order to have 95% confidence in detecting an average Bb seroprevalence of 7% across all selected installations, it was determined that each location should be able to provide data on at least 280 dogs over the course of the one year study (20). Only facilities with sufficient workload and manpower, as determined by USAPHC Animal Medicine Program personnel, were enrolled in the study.

Data Collection

Pet Health Questionnaire

Although U.S. Army VTFs are in the process of establishing a web-based electronic medical record system, at the time of this study, one was not yet in place. Without this capability, it was not possible to mine animal disease data from the headquarters level, therefore active surveillance of each participating VTF was undertaken in the form of a questionnaire instead. The use of the questionnaire served two purposes, one was to collect the military pet dog Bb seroprevalence data and the other was to collect the risk factor data from pet owners. The staff of all participating VTFs was provided the same Letter of Instruction (LOI) explaining the study protocol. Specifically, that the study was restricted to pet dogs (no government-owned working dogs), that eligible pets included both symptomatic and asymptomatic dogs requiring the IDEXX 4Dx Plus ® test, and that owners volunteering for the study were to receive a $5 discount, but were limited to one pet per survey and one survey per household. In addition, staff was instructed to offer study participation to all eligible owners upon arrival at the VTF on a first come first served basis at a rate of approximately 25 questionnaires per month. Once the questionnaire was completed the survey was handed to facility staff, who then annotated the IDEXX 4Dx Plus® test results on the last page along with the date and facility name. Data were collected from selected military veterinary facilities from 1 November 2012 through

31 October 2013. Questionnaires were collected on a monthly basis and physically mailed to the researchers for manual entry into an EXCEL spreadsheet designed for the study. All questions were multiple choice.

Military Pet Dog B. burgdorferi Seroprevalence Data

Canine Bb sero-positivity was determined utilizing the qualitative in-house IDEXX 4Dx Plus® serological test already routinely used in all military VTFs. No additional training or instruction on test use was provided. Both diagnostic test results (symptomatic) and routine health screening test results (asymptomatic) were included in this study. Seroprevalence data for all four vector-borne diseases included in the IDEXX 4Dx Plus® test was collected simultaneously. Test results (positive or negative) were annotated on the last page of the questionnaire by VTF staff along with date and location of the current test.

Military Pet Dog B. burgdorferi Risk Factor Data

Questions on risk factors came both from a review of current literature and a survey of both military and civilian veterinarians. All questions were limited to multiple choice answers. Owners were blinded to the specific goal of the survey, only being told that the information they were providing would lead to a better understanding of military pet dog populations and ultimately in better health care for their pet. They were also told that all questions pertained to the pet receiving the current IDEXX 4Dx Plus® test (Heartworm, Lyme, Ehrlichiosis, Anaplasmosis) and that participation was limited to only one dog per family. Specifically, the questionnaire included 23 multiple choice questions broken into 6 broad categories; Sponsor, Sponsor Residency History, General Pet Care Information, Pet Demographics (General Information About Your Pet), Dog Residency History, and General Pet Activity Information. The resulting questionnaire produced 84 total variables and one dichotomous outcome. The complete questionnaire is available in **Appendix A.-5**

1. Sponsor (service member associated with the pet) service component. Owners were asked the service component of the pet's sponsor; active duty, Reserve component, National Guard, retired, or other. It is expected that animals belonging to active duty service members move more frequently and may be less representative of the prevalence of the area they are living in.

2. Sponsor Residency History. Owners were asked to select places they have lived for more than a three month period from a list of first U.S. then European regions. U.S. regions were based on the regional breakdown

used in a study looking at civilian pet dog seroprevalence across the U.S. (20), whereas the European regional breakdown was based on a project funded by the European Commission (28). Three months was selected as the minimum cut off because it is not uncommon for service members to spend up to this time period in a location for the purpose of temporary duty or training; periods longer than this are more likely to involve true relocations. The question is geared towards capturing the owner's attitudes about flea and tick prevention use. Owners from Lyme endemic areas are suspected to be more likely to use such preventatives on their pets, and therefore we may expect lower prevalence from dogs that belong to owners that have lived in Lyme endemic areas.

3. General Pet Care Information. The first question asks about where the pet receives his/her primary preventive medicine care in order to determine how useful the information gathered from military VTFs will be in sentinel surveillance for zoonotic diseases in military populations. This section also asks question about the tested dog's history of heartworm and flea/tick prevention. The heartworm questions were partially so the owners wouldn't figure out the true focus of the survey, but it also investigates the assumption that dogs belonging to owner's that follow veterinary recommended preventive medicine practices will have lower seroprevalence. Questions on flea/tick preventative use also investigated this assumption, but the main reason for these questions was to determine the potential impact of their use on seroprevalence. Lastly, owners were asked whether a tick had ever been found on their dog.

4. Pet Demographics (General Information About Your Dog). Although sex was not suspected to impact sero-prevelence, in order to be consistent with current literature, the question was included in the survey. Because age may be a proxy for cumulative time of exposure, pet age was asked. Age categories included less than 1 year, 1-3 years, 3-6 years, 6-8 years, and more than 8 years. Because of hair coat factors, certain breeds may be more prone to prolonged tick attachment. Breed may also serve as a proxy for potential occupational and/or recreational exposure, therefore owners were asked to select the one breed category their dog best fit in. Breed groupings were based on AKC guidelines; each choice included an extensive list of examples. Lower seroprevalence may be expected in small, less athletic breeds or in breeds with certain hair coats. Dog weight was asked as a potential proxy for breed; weight categories were aimed at reflecting breed categories as closely as possible.

5. Dog Residency History. Owners were asked to select places that their dog had lived first in the last 5 years, then in the last 2 years. The regions listed matched the breakout described in the Sponsor Residency History

questions. The 5 year versus 2 year break down was based on the knowledge that canine antibodies to *B. burgdorferi* remain elevated for at least 69 weeks (18), therefore true origin of exposure may be better determined with this level of detail. In addition owners were asked how old the dog was when it joined their family, and if the dog had ever lived outside of the immediate family. These questions were geared towards assessing the reliability of the owner answers about the dog's potential for exposure.

6. General Pet Activity Information. Owners were asked questions geared towards determining the potential for their dogs to be in direct contact with ticks: the number of hours spent outdoors, because primarily outdoor dogs are expected to have higher seroprevalence; the type of environment, because dogs exposed to environments that support tick habitats (bushes, shrubs, forests) are expected to have higher seroprevalence; and recreational activity types and practices, because certain activities are more or less likely to put dogs at risk.

Data Analysis

SAS 9.3 was used to conduct the data analysis in this section. Potential confounders to the association between military pet dog seroprevalence and test result location were first investigated by using chi-squared tests for proportions to identify statistically significant associations (p-value < 0.05) between study variables and seroprevalence and then study variables and installation. In addition, univariate analysis identified variables as significant based on ORs, 95% CIs and p-values (< 0.25 as significant). A chi-squared test for proportions was used to evaluate how strongly potential confounders are related to each other in order to avoid redundancy in the model (p-value < 0.0001). Bivariable logistic regression, with the installation variable constant, was then performed to further investigate potential confounding by these risk factors. Each bivariate model (adjusted) was compared to the installation alone (crude) model. Changes in ORs of more than 10% ($[OR_{crude}-OR_{adjusted})]/OR_{adjusted}$) signified potential confounding by the variable necessitating its inclusion in the multivariable regression analysis. Multivariable logistic regression was then performed with Bb seroprevalence as the outcome of interest, installation as the exposure of interest, and the potential confounders identified through the descriptive statistics and bivariable logistic regression as the covariates. The most parsimonious model was determined by removing covariates from the model one at a time and evaluating the resulting impact on installation ORs. Once again, changes in more than three installation's ORs of more than 10% ($[OR_{full}-OR_{reduced})]/OR_{reduced}$) signified confounding by the variable

necessitating its inclusion in the final multivariable regression analysis. The covariates in the final model where determined to be identified confounders to the seroprevalence to installation association.

Results

Descriptive Statistics

Participation in the study was voluntary and clinics were not asked to record data for pet owners that declined participation, therefore response rates were not calculated. Overall, 3996 total surveys were collected, 94.5% (3776) of which included Bb sero-positivity results. All 15 selected installations participated, but nine fell short of the full one year study period. The calculated sample size for the full study period was 278/installation; seven of the locations fell short of this goal and eight exceeded it.

Military Pet Dog B. burgdorferi Seroprevelence Data

Table 5.6 shows the overall seroprevalence for the study population. There were 109 sero-positive dogs contributing to an overall sero-positivity of 2.89% study-wide.

Table5. 6: Military Pet Dog *B. burgdorferi* Seroprevalence Summary Statistics

Overall	2.89% (109/3776)
Average (standard deviation)	2.74% (+/- 2.21)
Median	2.33%
Range	0.00% - 9.25%

Analysis by installation shows 3.30% seroprevalence (31/937) in Europe and 2.75% seroprevalence (78/2839) in the U.S. The highest Bb seroprevalence came from the Northeastern and Midwestern United States (New York, Virginia, Illinois) and Germany (Kaiserslautern, Stuttgart) while the lowest rates came from Southern Europe (Italy) and Western, Southwestern, and Midwestern U.S. (Texas, North Caroline, and New Jersey). Table 5.7 shows seroprevalence by installation in addition to the number of observations for each category. It is important to note the

lack of seropositive results for the Signoella installation. Because the intent is to conduct a by installation analysis of sero-positivity, it would be inappropriate to combine this installation with any others and the addition of a "1" in order to include it in further modeling would be inaccurate, therefore the Sigonella installation was removed from further analysis.

Table 5.7: Military Pet Dog *B. burgdorferi* Seroprevalance by Installation

Installation	Sero Negative		Sero Positive		Total
	n	%	n	%	
Fort Drum	304	90.8%	31	9.3%	335
Fort Belvior	204	95.3%	10	4.7%	214
Stuttgart	204	96.2%	8	3.8%	212
Kaiserslautern	313	96.3%	12	3.7%	325
Great Lakes	244	96.8%	8	3.2%	252
Rota*	65	97.0%	2	3.0%	67
Spangdahlem	298	97.1%	9	2.9%	307
Camp Pendelton*	42	97. 7%	1	2.3%	43
Fort Carson	280	98.3%	5	1.8%	285
JBLM	345	98.3%	6	1.7%	351
Camp Lejeune	355	98.3%	6	1.7%	361
Fort Lee	297	98.3%	5	1.7%	302
Fort Bragg	520	99.1%	5	1.0%	525
Fort Hood*	170	99.4%	1	0.6%	171
Sigonella**	26	100.0%	0	0.0%	26

*Sparse cells for Sigonella, Fort Hood, Camp Pendelton, and Rota
**"0" cell for Sero+ from Sigonella forced to removal of this installation from further analysis
(collapsing with other installation is inappropriate and adding "1" would be inaccurate.

Military Pet Dog B. burgdorferi Seroprevalence Risk Factor Data (Characteristics of the Study Population)

The majority of those surveyed were Active Duty (73%). Overall 97% of owners reported living in the United States with the highest percent of owners (71%) claiming a history of living in the SE region. Forty-five percent reported living in Europe, with 37% claiming a history of living in Central Europe. Eighty-six percent of those participating in the survey claimed the surveying VTF as the location where the pet receives his/her primary preventive medicine care. More than eighty percent of pet owners claimed a history of flea, tick and heartworm preventative use and over sixty percent claimed their pets were current on these same preventives. Over fifty percent stated they had never seen a tick on their dog. The dog population was equally split between male neuter and female spayed dogs with an overall 60% falling between 1 and 6 years of age. Most dogs were reported to

come from one of four breeds; toy, sporting, herding, or working. The majority of owners reported their dogs lived in either wooded or brushy environments and spend between one to eight hours outside a day. The most commonly reported recreational activity was neighborhood walks. More than 80% of dogs were reported to have lived in the US within the last 5 years with the majority reporting a history of time in the SE US. Almost 75% reported a history of living in Europe within the last 5 years, with the most commonly reported region being Central Europe. Only thirty-two percent of owners reported their pets had prior ownership and 47% reported their dogs were more than three months old when joining the family. Sero-positivity for the three other vector borne diseases included in the commercial test were 8.26% for *Anaplasma* spp. (*A. platys* and/or *A. phagocytophilum*), 11.01% for *Ehrlichia* spp. (*E. ewingii* and/or *E. canis*), and 4.59% for *Dirofilaria immitus*. Analysis of each variable by seroprevalence resulted in cell sparcity (<5) for 31 of the variables, eight of which resulted in zero cells. A complete listing of study variable distribution by seroprevalence can be found in Tables A-M in Appendix B-5.

Univariate Analysis: Potential for Risk Factors to be Confounders to the Installation-Bb Seroprevalence Association

Based on this analysis 15 potential confounders were identified. A complete listing of Odds of Bb Sero-positivity by risk factor (ORs and 95% CIs) can be found in Tables I-XII in **Appendix C-5**.

Table 5.8: Significant Risk Factors Associated Military Pet Dog *B. burgdorferi* Seroprevalence (ORs, 95% CIs, and p-values)

Characteristic	OR*	95% CI	p-value
Owner Lived in Northeastern U.S. Region			
Yes*	1.000		
No	0.282	(0.192, 0.414)	<0.0001
Owner Lived in Central European Region			
Yes	1.000		
No	0.674	(0.460 , 0.987)	0.0427
History of Tick Exposure			<0.0001
Yes*	1.000		
No	0.236	(0.147, 0.379)	<0.0001
Not Known	1.619	(0.878, 2.987)	0.1227
Breed Group			<0.0001
Sporting*	1.000		
Herding	1.004	(0.587, 1.717)	0.9885

Working	0.721	(0.400, 1.299)	0.2758
Hound	0.424	(0.195, 0.922)	0.922
Terrier	0.122	(0.029, 0.509)	0.0039
Toy	0.210	(0.100, 0.438)	<0.0001
Non-Sporting	0.496	(0.206, 1.191)	0.1168
Other	0.538	(0.208, 1.390)	0.2004
Weight Group (pounds)			<0.0001
56 to 75*	1.000		
Less than 15	0.250	(0.125, 0.498)	<0.0001
16 to 35	0.409	(0.223, 0.750)	0.0035
36 to 55	0.807	(0.466, 1.400)	0.4460
More than 75	1.041	(0.611, 1.771)	0.8835
Dog Lived in Northeastern U.S. in past 5 years			
Yes	1.000		
No	0.230	(0.156, 0.341)	<0.0001
Dog Lived in Southeastern U.S. in past 5 years			
Yes	1.000		
No	1.538	(1.050, 2.2520)	<0.0001
Dog Lived in Central European Region in past 5 years			
No	1.000		
Yes	0.650	(0.431, 0.982)	0.0410
Lived in Western U.S. in past 2 years			
Yes	1.000		
No	2.379	(1.300, 4.356)	0.0050
Lived in Northeastern U.S. in past 2 years			
Yes	1.000		
No	0.214	(0.144, 0.320)	<0.0001
Lived in Southeastern U.S. in past 2 years			
Yes	1.000		
No	1.807	(1.217, 2.682)	0.0033
Prior Ownership			
Yes*	1.000		
No	0.675	(0.458, 0.996)	0.0477
Dog Lives in Wooded Environment			
Yes*	1.000		
No	0.530	(0.345, 0.813)	0.0037
Dog Recreates through Wilderness Walks			
Yes*	1.000		
No	0.384	(0.262, 0.563)	<0.0001
Anaplasma			
Yes*	1.000		
No	0.049	(0.021, 0.113)	<0.0001
Dirofilaria			

Yes*	1.000		
No	0.183	(0.070, 0.480)	0.0005
Ehrlichia			
Yes*	1.000		
No	0.2212	(0.118, 0.418)	<0.0001

* Referent Category
* Odds of Borrelia Sero-positivity by Characteristic

Refining Variable Selection

Quasi-complete separation ($p < 0.0001$) was found for breed (Bre) by weight (Wt), when the dog joined the family (JOIN) by previous owner status (POwn), and several different combinations of preventive medicine practices (current flea and tick preventive use (CFTP), history or flea and tick preventive use (HFTP), current heartworm preventive use (HHW), and current heartworm preventive use (CHW). Because of the sparsity of cells in the breed variable, weight was selected over breed. Previous owner status (POwn) was selected over JOIN because it was slightly more significant. None of the preventive medicine practice variables were significantly associated with sero-positivity.

The questions about dog residency history included overlapping time periods (within the last 5 years and within the last 2 years, indicating the existence of co-linearity between these two variables. In order to determine the more significant risk factor, odds ratios based on proportions were calculated (Table5. 9) and the variable with the higher magnitude of association with sero-positivity was considered the more significant risk factor.

Table 5.9: Odds Ratio Estimates and 95% CI s for Identified Dog Residency Risk within 5 vs 2 years of *B. burgdorferi* Sero-positivity status

Residency Location	Residency Time Period*	OR**	95% CI
Northeastern U.S.	5 years	4.3408	(2.9357, 6.4187)
	2 years	4.6613	(3.1225, 6.984)
Southeastern U.S.	5 years	0.6502	(0.4440, 0.9523)
	2 years	0.5534	(0.3729, 0.8214)
Central Europe	5 years	1.5375	(1.0178, 2.3229)
	2 years	1.4406	(0.9470, 2.1915)

* within last 5 years or within last 2 years
**referent category is no history of residing in location for time period

Dogs that lived in northeastern U.S. within the last 5 years were 4 times more likely to be Bb sero-positve than those with no history of living in the northeastern U.S. in last 5 years, and the magnitude of association is more with more recent residency history. Dogs that lived in southeastern U.S. within the last 5 years were 2/3 less likely to be Bb sero-positve than those with no history of living in the southeastern U.S. in last 5 years, and the magnitude of association is less with more recent residency history; recent history of southeastern U.S. residency is more protective. Dogs that lived in Central Europe within the last 5 years were 1.5 times more likely to be Bb sero-positve than those with no history of living in the Central Europe in last 5 yr, with the magnitude of association being more for the more distant exposure has stronger association. Based on these findings it was determined to use the more recent residency data for northeastern and southeastern US, but the more distant residency history for the data on Central European residency history.

In addition, sponsor and dog residency histories were found to be closely related, indicating dependence between these two variables. This dependency means that dogs share the same travel and residency history as their military sponsors (owners). In order to determine which risk factor to include in the modeling process, odds ratios based on proportions were calculated (Table 5.10) and the variable with the higher magnitude of association with Bb sero-positivity was included in the modeling process.

Table 5.10: Odds Ratio Estimates and 95% CI s for Identified Residential History Risk, Dog vs Owner, and *B. burgdorferi* Sero-positivity status

Residency Location	Resident*	OR**	95% CI
Northeastern U.S.	Dog (within last 2 years)	4.6613	(3.1225, 6.984)
	Owner	3.5483	(2.4139, 5.2159)
Southeastern U.S.	Dog (within last 2 years)	0.5534	(0.3729, 0.8214)
	Owner	0.6711	(0.6711, 0.4577)
Central Europe	Dog (within last 5 years)	1.5375	(1.0178, 2.3229)
	Owner	1.4846	(1.0131, 2.1757)

* residency history for dog vs. owner
**referent category is no history of residing in location for time period

The table shows that dog residential history data shows a slightly higher magnitude of association with Bb sero-positivity status than the human data does. Based on these findings it was determined to only use the dog residency data.

Regression Analysis

Bivariate Models: Installation Constant, Adjustment for Identified Risk Factors

Based on the combined findings of the descriptive analyses, investigation of the associations of each risk factor with both Bb seroprevalence and installation, determinations of inter-variable redundancies, and evaluation of odds ratios the following 12 variables were identified as potential risk factors for Bb sero-positivity in the study population: History of tick exposure (Tick), Dog's weight (Wt), history of previous owner (POwn), living in a wooded environment (Wood), recreating through wilderness walks (WWalk), history of dog living in the Northeastern US within the last 2 years (DgNERes2), history of dog living in the Southeastern US within the last 2 years (DgSERes2), history of dog living in the Western US within the last 2 years (DgWRes2), history of dog living in the Central Europe within the last 5 years (DgCRes5), co-infection with *D. immitus* (Diro), co-infection with *A.platys* and/or *A. phagocytophilum* (Ana), and co-infection with *E. canis* and/or E. ewingii (Ehr).

Because the overall goal of this study is to be able to use installation level canine Bb seroprevalence data as sentinel surveillance for human borreliosis, it is important to ensure that the seroprevalence reported for the installation is truly a reflection of the risk in the area, and not merely a reflection of the distribution of these identified risk factors. In order to better understand the potential impact of each variable on the installation to seroprevalence association, bivariate analyses were done. Only the Diro (co-infection with *D. immitus*) variable was determined to NOT be a potential confounder, indicating the other eleven variables could be potential confounders and therefore dictating their inclusion in the subsequent multivariable regression analysis. **Appendix D-5** has the results of these bivariant analyses.

Multivariate Models: Installation Constant, Adjustment for Identified Risk Factors

Eight risk factors remained potential confounders of the installation to Bb seroprevalence association. The resulting logistic regression model had the following form:

$$Y_{BOR} = \beta + \beta_{INST} + \beta_{TICK} + \beta_{WT} + \beta_{DGWRES2} + \beta_{DGNERES2} + \beta_{DGCEURES5} + \beta_{WWALK\ T} + \beta_{ANA} + \beta_{IEHR} + e$$

The $\chi 2$ goodness-of-fit statistic (log likelihood ratio) for the model was 185.014 with 25 degrees of freedom and a resulting p-value of <0.0001, suggesting that the model adequately explained the Bb seroprevalence outcome. Odds ratios for all factors and the full regression process can be found in **Appendix E-5**. Based on this final model the odds ratios for the risk factor of interest, installation, were determined (Table 5.11).

Table 5.11: Odds of Military Pet Dog *Borrelia burgdorferi* Sero-positivity by Installation as Compared to Fort Drum, NY

Installation (sero-positive/total tested)	**Crude OR**	**95% CI**	**Adjusted* OR**	**95% CI**
Kaiserslautern, DE (12/325)	0.376	(0.190, 0.745)	0.324	(0.094, 1.116)
Stuttgart. DE (8/212)	0.384	(0.173, 0.853)	0.234	(0.062, 0.883)
Rota, SP (2/67)	0.302	(0.070, 1.292)	0.570	(0.115, 2.828)
Spangdahlem, DE (9/298)	0.296	(0.139, 0.633)	0.262	(0.072, 0.948)
Fort Belvoir, VA (10/204)	0.481	(0.231, 1.002)	0.623	(0.258, 1.505)
Fort Bragg, NC (5/520)	0.094	(0.036, 0.245)	0.177	(0.058, 0.537)
Fort Carson, CO (5/280)	0.175	(0.067, 0.456)	0.605	(0.168, 2.179)
Fort Hood, TX (1/170)	0.058	(0.008, 0.426)	0.116	(0.012, 1.084)
Fort Lee, VA (5/297)	0.165	(0.063, 0.430)	0.263	(0.085, 0.81)
Camp Lejeune, NC (6/355)	0.166	(0.068, 0.402)	0.219	(0.071, 0.68)
Camp Pendleton, CA (1/42)	0.233	(0.031, 1.755)	0.883	(0.091, 8.588)
JBLM, WA (6/345)	0.170	(0.070, 0.414)	0.724	(0.201, 2.609)
Great Lakes, IL (8/244)	0.321	(0.145, 0.712)	0.759	(0.283, 2.034)

* Adjusted for the following Dog Characteristics: History of Tick Exposure, Weight, Recent (within 2 years) History of Living in Western, Recent (within 2 years) History of Living in Northeastern US, History (within 5 years) of Living in Central Europe, Recreation through Wilderness Walks, co-infection with *A. phagocytophilumum/A.platys* , and co-infection with *E.canis/ E, ewingii.*

As expected, all Odds Ratios are less than one, however the wide confidence intervals bring a level of uncertainty to these estimates.

Discussion

The overall purpose of this study was to demonstrate canine *B. burgdorferi* seroprevalence as an adjunct to sentinel surveillance for human Lyme borreliosis in military populations. The focus of this section was to determine whether the observed installation to seroprevalence association is influenced by the distribution of identified risk factors. Based on the findings of this section, eight of the 83 risk factors investigated were identified to be potential confounders, however, given the width of the confidence intervals, meaningful conclusions about confounding are difficult. This may be due to small sample sizes and overall low number of sero-positive observations. Several of the installations (seven of the fifteen) produced fewer than the requested number of surveys leading to sample sizes that were underpowered. This means that for these installations there may not have been enough samples taken to detect the true prevalence for the area. In addition, seven of the fifteen installations had five or fewer total sero-positive animals over the one year study period. The low sample sizes in conjunction with the low sero-positive cell counts may have contributed to instability of the model estimates, and the resulting inability to make conclusions on the potential for confounding. In addition, this approach only allows for adjustment by identified risk factors, there may still be unmeasured common features at the installation level that make these dogs more likely to have the same seroprevalence status. In other words, the potential for additional confounding of seroprevalence status at the installation level by other unmeasured risk factors may exist and therefore further contribute to the uncertainty of these estimates.

Although not the purpose of this study, analysis of the characteristics of the study population revealed several interesting trends. For example, the study did not find the sponsor's service component or preferred primary pet care facility to be significantly associated with seroprevalence. It also didn't matter whether or not the dog was previously owned or at which age the dog joined the family. However, the data showed that more than 84% of the dogs joined the family at less than a year of age and that over 65% had never had prior owners, meaning that the answers provided by these owners where likely to be accurate and reliable. Consistent with current literature, sex

was not found to be a risk factor. Inconsistent with the literature, however, age was not found to be associated with seroprevalence. This discrepancy could not be resolved even when categories were collapsed. With over 45% of study dogs fell in the high risk categories (three to six and six to eight years of age), it is not clear why age as a proxy for cumulative exposure was not a risk factor.

Overall 80% of the study dogs were reported to have a history of flea and tick product use and over 60% were reported to be currently on the preventatives. It was interesting, however, that the use of these flea and tick prevention products was not found to be associated with seroprevalence in this study population. This lack of association may be due to inaccurate owner reporting, either intentionally or unintentionally. It could also be due to owner compliance issues such as inappropriate application or use of the products. Of greatest concern would be if this lack of association between flea and tick product use and Bb seroprevalence indicates true product failure. In 2012 a study found no significant association between a history of acaricide use and the presence of ticks in their study population of pet dogs brought into veterinary clinics throughout the United Kingdom (34). These findings together indicate the need to re-evaluate the effectiveness of the currently recommended flea and tick preventive products.

The number of hours spent outdoors and whether or not most outdoor activities occurred on a leash were determined to not be a risk factor for Bb seroprevalence, however certain types of recreational activities (walks through the wilderness) and the type of outdoor living environment (wooded) did. Also consistent with the literature, breed was found to be a risk factor, with the highest prevalence occurring in the sporting, herding, and working breeds.

Sponsor residency history data was expected to be inversely associated with Bb seroprevalence based on preventive medicine practices (sponsor's from Lyme endemic areas such as the NE U.S. were suspected to have dogs with lower seroprevalence based on knowledge of risk), however was not. Furthermore, the data collected on sponsor residency history and dog residency history was determined to co-linear with the dog data having a higher magnitude of association.

In order to investigate the potential impact of travel history on seroprevalence, owners were asked to select where their dog lived both in the last 5 years, and in the last 2 years. The time windows were based on the knowledge that canine antibodies to *B. burgdorferi* remain elevated for at least 69 weeks (18). It is interesting to note that for the U.S. regions the 2 year window has a higher magnitude of association with seroprevalence when compared to the 5 year window, but the opposite is noticed in Europe. This difference could reflect that the exposure that occurred in Europe were longer ago (declining risk in the area; higher risk more than two years ago) or differences between the *Borreliae* strains found in Europe as compared to the United States. For example, perhaps the American strains are able to invade host tissues and evade serological detection more than the European strains, so that after two years antigen levels become low enough to produce false negative results. Because *B. burgdorferi* is transmitted by ticks, it is not surprising that a history of tick exposure was identified as being associated with Bb seroprevalence. Also, because the *Ixodes* spp. ticks that carry the Bb pathogen are also known to carry *Anaplasma* spp., it is not surprising that *Anaplasma* spp. sero-positivity is also associated with Bb sero-positivity. *Ehrlichia canis* (as tested for by the study test kits), on the other hand, is carried by a different tick, *Rhipicephalus* spp, but because these ticks have overlapping geographic distributions, and because factors that put dogs at risk for tick bites would put them at risk for all tick bites, it does make sense that *E. canis* sero-positivity would be associated with Bb sero-positivity.

The final model indicates that eight of the 83 potential risk factors are confounders of the observed Bb seroprevalence to installation association. It is intuitive that co-infection with *Anaplasma* spp. is confounding because the shared vector (and rodent reservoirs) actually leads to cases of co-exposure, making the pathogen associated with both the installation and Bb seroprevalence. As discussed above, the confounding by *E. canis* co-infection is likely due to overlapping geographic distributions and shared risk factors for tick exposure. It is also intuitive, then, that a history of tick exposure would only be a risk factor for Bb in regions where the pathogen exists, therefore making it associated with both installation and sero-positivity. The confounding by the variables representing recent dog residency history (Western and Northeastern U.S. within last 2 years) may be explained by the fact that most of these dogs are still at the indicated location (association with installation) and that these regions have the highest (NE) and lowest (W) Bb risk (association with Bb seroprevalence). A similar phenomenon may explain the confounding by the history of living in Central Europe, although for this variable the history or residency was more than 2 years ago. One reason for this difference (within 2 years vs. within 5 years) may be due to the fact

that overseas assignments tend to be for at least three years, meaning dogs with this variable are probably still in the indicated region (association with installation). However, if this is the case, the owners should also have indicated that the dogs are in this location at the 2 year mark. The difference is that the more recent time period does not appear to be associated with Bb sero-positivity to the same magnitude that the more distant time period does, possibly for the regional differences in *Borrelea* species as mentioned above. The last two confounders are less intuitive and more likely to be related to regional differences in owner attitudes and behaviors. For example, although one may expect dog weight, as a proxy for breed, to be associated with Bb seroprevalence, the association with installation is harder to understand. Perhaps certain dog breeds are more popular in certain regions of the country, for example short haired breeds in hot climates and long haired breeds in cold climates. In terms of the confounding by wilderness walks, this is likely more associated with regional variations in environment and climate. For example, *Ixodes* spp. ticks live in temperate climates. People that live in temperate climates tend to participate in outdoor activities such as hiking, which brings them in contact with *borrelia* infected ticks. High risk installations are found in these temperate climates. Therefore, wilderness walks as a proxy for climate, is associated with both Bb seroprevalence and installation.

Conclusions

Because of the limited power of this study, there was an inability to robustly evaluate whether canine *B. burgdorferi* seroprevalence is influenced by the distribution of identified risk factors. In order to better make this assessment, the study should be repeated for a longer period in order to meet the required sample size (many of the installations were underpowered). In addition, because of the low overall Bb seroprevalence in this population (likely due to the high level of flea and tick prevention), the target sample size should be increased, or a different zoonotic disease should be selected.

Overall Conclusions

The purpose of this chapter was to demonstrate the use of Service member pet dogs in sentinel surveillance for Lyme disease in military populations in order to recommend to the incorporation of these animals in zoonotic

disease surveillance programs in the military. Doing so involved three specific objectives: 1) determining the association between *B. burgdorferi* seroprevalence in military pet dogs and military human Borreliosis data, 2) comparing military pet dog Bb seroprevalence data to published civilian pet dog data in order to evaluate potential differences in the populations and determine the most appropriate source of data for military populations , and 3) investigating the validity of the military pet dog seroprevalence data by evaluating potential confounders to the association between military pet dog seroprevalence and test result location (installation).

This is the first study to investigate the utility of pet dogs belonging to service members as sentinels for Lyme disease risk in military populations. There are many reasons to examine this relationship. One is the fact that military populations are not limited to living within the United States, making reliance on pet data available at the U.S. level insufficient for risk projections for military populations at the global level. Also, because military personnel tend to live on or around the specific military installation they are assigned to, the use of existing canine prevalence data at the state level may not accurately represent the risk of Lyme disease at the installation level. Because of the shared recreational habits, environmental exposures, and travel history that exist between pet owners and their pets, it is logical to assume that service member pet dogs are better able to reflect shared exposures and therefore better represent the risk of Lyme disease as compared to civilian pet dogs in the same location. Lastly, the physical exams and routine laboratory work associated with the issuing of health certificates required for frequent movement of the owners contributes to a robust animal disease database; capturing evidence of exposures in asymptomatic dogs. On the other hand, seroprevalence estimates based on military pet dogs may also have several potential short comings. The high mobility of the population has the potential to lead to inaccurate geographic risk assessment if travel history is not controlled for. Also, the low cost of preventive medicine products at military run veterinary treatment facilities has the potential decrease the prevalence of disease in these animals and therefore decrease the sensitivity of the sentinel surveillance. The inclusion of stray animals that do not receive regular veterinary care and may not be as mobile is one approach that may increase the seroprevalence estimates and corresponding sensitivity, because these animals are more likely to roam free and be exposed to infected ticks.

This study found that Bb seroprevalence estimates in military pet dogs not only reflect incidence rates in military human populations, but they also appear to be a more sensitive indicator of *B. burgdorferi* presence then their human counter parts. In addition, it found that military pet dog seroprevalence data are more accurate markers of

Lyme disease risk for military populations than civilian pet dog data. Unfortunately, this study lacked the power to fully evaluate whether canine *B. burgdorferi* seroprevalence was influenced by the distribution of identified risk factors. In order to confidently make this determination, the study should be repeated for a longer period in order to meet the required sample size, or a different zoonotic disease should be selected.

Recommendations

Although using dogs as sentinels for Lyme disease risk in humans is not straight forward, careful consideration of both the advantages and limitations this approach can still make this a powerful adjunct to human Lyme disease surveillance efforts. This study has showed that service member pet dogs can serve as sensitive and convenient measures of the risk of Lyme disease in human military populations, and recommends the incorporation of service member pet animal data as a powerful adjunct to zoonotic disease surveillance in military populations.

References

1. American Veterinary Medical Association One Health: A New Professional Imperative, One Health Initial Task Force: Final Report, July 15, 2008.

2. Ruberton MV and Brundage JF (2002). The Defense Medical Surveillance System and Department of Defense Serum Repository: Glimpses of the Future of Public Health Surveillance. Am. J. of Public Health; 92(12):1900-1904.

3. DeFraites, Robert F. The Armed Forces Health Surveillance Center: Enhancing the Military Health System's Public Health Capabilities. BioMed Centrtal Public Health. 2011; 11(Suppl 2)S1 (http://www.biomedcentral.com/1471-2458/11/S2/S1)

4. Permanent Order 275-002, 2 October 2009, Headquarters U.S. Army Medical Command Fort Sam Houston Texas, TX 782345-6000.

5. U.S. Army Public Health Command website accessed May 16, 2013 at http://phc.amedd.army.mil/

6. McCluskey B. Use of sentinel herds in monitoring and surveillance systems. In: Animal disease surveillance and survey systems, methods and applications. 1st ed. Ames, Iowa: Blackwell Publishing, 2003; 119-123.

7. Center for Disease Control and Prevention. Lyme Disease Home Page. http://www.cdc.gov/lyme/index/html. Accessed 17Dec2013.

8. Zhang, Xinzhi, et al Economic Impact of Lyme Disease. Emerging Infectious Diseases. 2006; 12(4):653-660.

9. Tri-Service Reportable Events Guidelines & Case Definitions, June 2009.

10. Armed Forces Health Surveillance Center. Annual Lyme Disease Report. April 2011.

11. Medical Surveillance Monthly Report. Lyme Disease Among U.S. Military Members, Active and Reserve Component, 2001-2008. Vol 16. No 07. July 2009: 2-4. http://www.afhsc.mil.

12. Mead P, Goel R, Kugeler K. Canine Serology as Adjunct to Human Lyme Disease Surveillance. Emerging Infectious Diseases. 2011; 17:1710-1712.

13. Lindenmayer JM, Marshall D, Onderdonk AB. Dogs as Sentinels for Lyme Disease in Massachusetts. Am J of Public Health 1991; 81:1448-1455.

14. Olson PE, Kallen AJ, Bjorneby JM, Creek JG. Canines as Sentinels for Lyme Disease in San Diego County, California. J VET Diagn Invest 2000; 12:126-129.

15. Duncan AW, Correa MT, Levine JF, Breitschwerdt EB. The Dog as a Sentinel for Human Infection: Prevalence of *Borrelia burgdorferi* C6 Antibodies in Dogs from Southeastern and Mid-Atlantic States. Vector-Borne and Zoonotic Diseases 2004; 4:221-229.

16. Eng TR, Wilson ML, Spielman A, and Lastavica CC. Greater risk of *Borellia burgdorferi* infection in dogs than in people. J of Inf Disease 1988; 158:1410-1411.

17. Greene CE, Straubinger RK, and Levy SA. Borreliosis. In: Infectious Diseases of the Dog and Cat. 4th ed. St. Louis, Missouri: Elsevier Sauders Co, 2012; 447-465.

18. Philipp MT, Bowers LC, Fawcett PT, et al. Antibody response to IR6, a conserved immunodominant region of the VIsE lipoprotein, wanes rapidly after antibiotic treatment of *Borrelia burgdorferi* infection in experimental animals and humans. J Infect Dis. 2001; 184:870-878.

19. Goossens HAT, van den Bogaard AE, and Nohlmans MKE. Dogs as Sentinels for Human Lyme Borreliosis in the Netherlands. J Clin Microbiol 2001; 39:844-848.

20. Bowman D, Little SE, Lorentzen L, Shields J, Sullivan MP, Carlin EP. Prevalence and geographic distribution of *Dirofilaria immitis*, *Borrelia burgdorferi*, *Ehrlichia canis*, and *Anaplasma phagocytophilum* in dogs in the United States: Results of a national clinic-based serologic survey. Veterinary Parasitology 2009; 160:138-148.

21. Surveillance Snapshot: Lyme Disease among Beneficiaries of the Military Health System, 2001-2012. Medical Surveillance Monthly Report. 2013; 20(8):23.

22. https://www.dmdc.osd.mil/appj/dwp/index.jsp; accessed 12/16/14

23. http://www.tricare.mil/DEERS/; accessed 12/16/13

24. Liang FT, Jacobson RH, Straubinger RK, Grooters A, Philipp MT. Characterization of a *Borrelia burgdorferi* VlsE Invariable Region Useful in Canine Lyme Disease Sero-diagnosis by Enzyme-Linked Immunosorbent Assay. J of Clin Microbiology 2000; 38:4160-4166.

25. Levy SA, O'Connor TP, Hanscom JL, et al. Quantitative Measurement of C6 Antibody following Antibiotic Treatment of *Borrelia burgdorferi* Antibody-Positive Nonclinical Dogs. Clinical and Vaccine Immunology 2008;15:115-119.

26. Moyses, S and Nieto FJ. Chapter 8: Quality Assurance and Control. Epidemiology: Beyond the Basics 3rd Ed. Jones & Bartlett Learning, LLC, an Ascend Learning Company, USA: 2014;350-353.

27. Rydzewski J, Mateus-Pinilla N, Warner RE, Hamer S, Weng HY. *Ixodes scapularis* and *Borrelia burgdorferi* among diverse habitats within a natural area in east-central Illinois. Vector Borne Zoonotic Disease 2011; 10:1351-1358.

28. Lindgren E and Jaenson T. Lyme borreliosis in Europe: influences of climate and climate change, epidemiology, ecology and adaptation measures. World Health Organization Europe in conjunction with the "Climate Change and Adaptation Strategies for Human Health" (cCASHh) Project from 1 May 2001 to 31 July 2004. WHO Europe, 2006; document number EUR/04/5046250.

29. Diuk-Wasser, Maria A et al. Human Risk of Infection with *Borrelia burgdorferi*, the Lyme Disease Agent, in Eastern United States. Am. J. Trop. Med. Hyg. 2012; 86(2):320–327.

30. Gatewood, Anne G et al. Climate and Tick Seasonality Are Predictors of *Borrelia burgdorferi* Genotype Distribution. Applied and Environmental Microbiology. 2009; 75(8):2476–2483.

31. Richter D. and Matuschka FR. Differential Risk for Lyme Disease along Hiking Trail, Germany. Emerging Infectious Diseases. 2011; 17(9): 1704-1707

32. Peisman, J and Eisen, L. Prevention of Tick-borne Diseases. Annu. Rev. Entomol. 2008; 53:323-43.

33. Cohen ND, et al. Clinical and epizootiologic characteristics of dogs seropositive for *Borrelia burgdorferi* in Texas: 110 cases (1988). Journal of the American Veterinary Medical Association. 1990; 197(7):893-89834.

34. Smith FD, Ballantyne R., Morgan ER., et al. Estimating Lyme disease risk using pet dogs as sentinels. Comparative immunology, microbiology and infectious diseases. 2012; 35(2):163-167.

Appendices

Appendix A-5

Pet Dog
Health Care Questionnaire

Thank you for volunteering to participate in this survey.
The information you provide will lead to better understanding of our military pet dog population's unique life styles and will ultimately assist us in providing better health care to your pet.

All questions about your pet should be answered for the dog that is receiving the IDEXX 4Dx Plus® test (Heartworm, Lyme, Ehrlichiosis, Anaplasmosis) at today's visit.

Participation is limited to only one dog per family.

Sponsor

1. Please circle the sponsor's service component:
    a. Active Duty Army, Navy, Air Force, or Marine
    b. Reserve Component
    c. National Guard
    d. Retired Military
    e. Other

Sponsor Residency History

2. Including where you currently live, which of the US regions have you **ever** lived in for more than 3 months?
Circle all that apply.
    a. I have never lived in the US
    b. Mid-Western United States (IA, IL, IN, KS, MI, MN, MO, ND, NE, OH, SD, WI)
    c. South Eastern United States (AL, AR, FL, GA, KY, LA, MS, NC, OK, SC, TN, TX, VA, WV)
    d. Western United States (AZ, CA, CO, ID, MT, NM, NV, OR, UT, WA, WY)
    e. North Eastern United States (CT, DE, MA, MD, ME, NH, NJ, NY, PA, RI, VT, DC)
    f. Other United States regions (AK, HI, Puerto Rico)

3. Including where you currently live, which of the European regions/countries have you **ever** lived in for more than 3 months?
Circle all that apply.
    a. I have never lived in Europe.
    b. Sweden, Netherland, and/or the Baltic States (Lithuania, Estonia, Latvia)
    c. Northern Europe (Denmark, Finland, Ireland, Norway, the United Kingdom)
    d. Southern Europe and/or the Balkans (Spain, Italy, Bulgaria, Croatia, Greece, Turkey, Bosnia)
    e. Eastern Europe (Armenia, Azerbaijan, Russia, Ukraine)
    f. Central Europe (Austria, Czechoslovakia, Germany, Hungary, Poland, Switzerland, Belgium, France)
    g. Other European countries not listed

General Pet Care Information

All questions about your pet should be answered for the dog that is receiving the IDEXX 4Dx Plus® test (Heartworm, Lyme, Ehrlichiosis, Anaplasmosis) at today's visit.

4. Where does your dog receive his/her **primary** preventive medical care? (Routine exams, tests, vaccinations)
    a. This Military Veterinary Treatment Facility
    b. Another Military Veterinary Treatment Facility
    c. Civilian Veterinary Facility

5. Has your dog **ever** been on Heartworm Prevention (for example: Interceptor®, Sentinel®, Heartgard®, Heartgard Plus®)?
    a. Yes
    b. No
    c. I do not know

6. Is your dog **currently** on Heartworm Prevention (for example: Interceptor®, Sentinel®, Heartgard®, Heartgard Plus®)?
    a. Yes
    b. No
    c. I do not know

7. Has your dog **ever** been on Flea and Tick Prevention (for example: Program®, Sentinel®, Frontline®, Flea/Tick Collars, Flea/Tick Shampoos or Dips)?
    a. Yes
    b. No
    c. I do not know

8. Is your dog **currently** on Flea and Tick Prevention (for example: Program®, Sentinel®, Frontline®, Flea/Tick Collars, Flea/Tick Shampoos or Dips)?
    a. Yes
    b. No
    c. I do not know

9. Has a tick ever been found on your dog?
    a. Yes
    b. No
    c. I do not know

General Information About Your Dog

10. What is the gender of you dog?
    a. Male Intact
    b. Male Neutered
    c. Female Intact
    d. Female Spayed

11. Approximately how old is your dog?
    a. Less than 1 year
    b. 1-3 years
    c. 3-6 years
    d. 6-8 years
    e. more than 8 years

12. In what breed category would you place your dog?
Please note, these are not complete listings; select the category that best suits your dog.
For mixed breeds select the one that you feel applies most.
    a. Sporting (Spaniels, Retrievers, Pointers, Setters, Brittany, Vizsla, Weimaraner)
    b. Herding (Collies, Cattle Dogs, Australian Shepherds, Malinois, Sheep Dogs, Corgi, German Shepherd)
    c. Working (Akita, Malamute, Mountain Dogs, Boxers, Mastiffs, Pinschers, Schnauzers, Danes, Pyrenees, Newfoundland, Rottweiler, St. Bernard, Samoyed, Husky)
    d. Hound (Beagles, Hounds, Dachshunds, Whippets, Ridgebacks)
    e. Terrier (Airedale, Staffordshire, Carin, Bull, Russell, Scottish, Fox)
    f. Toy (Chihuahua, King Charles Spaniel, Italian Greyhound, Maltese, Pekingese, Pomeranian, Pug, Yorkie, Shih Tzu, Poodle, Papilon)
    g. Non-Sporting (Bichon, Boston Terrier, Bulldog, Shar Pei, Chow, Dalmatian, Keeshund, Standard Poodle)
    h. Other

13. How much does your dog weigh?
    a. Less than 15 pounds
    b. 16-35 pounds
    c. 36-55 pounds
    d. 56-75 pounds
    e. More than 76 pounds

Dog Residency History

14. Including where your dog currently lives, which of the US regions has your dog lived in **within the past 5 years**?

Circle all that apply.

a. He/she has never lived in the US
b. Mid-West United States (IA, IL, IN, KS, MI, MN, MO, ND, NE, OH, SD, WI)
c. South Eastern United States (AL, AR, FL, GA, KY, LA, MS, NC, OK, SC, TN, TX, VA, WV)
d. Western United States (AZ, CA, CO, ID, MT, NM, NV, OR, UT, WA, WY)
e. North Eastern United States (CT, DE, MA, MD, ME, NH, NJ, NY, PA, RI, VT, DC)
f. Other United States regions (AK, HI, Puerto Rico)

15. Including where your dog currently lives, which of the US regions has your dog lived in **within the past 2 years**?

Circle all that apply.

a. He/she has never lived in the US
b. Mid-West United States (IA, IL, IN, KS, MI, MN, MO, ND, NE, OH, SD, WI)
c. South Eastern United States (AL, AR, FL, GA, KY, LA, MS, NC, OK, SC, TN, TX, VA, WV)
d. Western United States (AZ, CA, CO, ID, MT, NM, NV, OR, UT, WA, WY)
e. North Eastern United States (CT, DE, MA, MD, ME, NH, NJ, NY, PA, RI, VT, DC)
f. Other United States regions (AK, HI, Puerto Rico)

16. Including where your dog currently lives, which of the European regions/countries has your dog lived in **within the past 5 years**?

Circle all that apply.

a. My dog has never lived in Europe.
b. Sweden, Netherland, and/or the Baltic States (Lithuania, Estonia, Latvia)
c. Northern Europe (Denmark, Finland, Ireland, Norway, the United Kingdom)
d. Southern Europe and/or the Balkans (Spain, Italy, Bulgaria, Croatia, Greece, Turkey, Bosnia)
e. Eastern Europe (Armenia, Azerbaijan, Russia, Ukraine)
f. Central Europe (Austria, Czechoslovakia, Germany, Hungary, Poland, Switzerland, Belgium, France)
g. Other European countries not listed

17. Including where your dog currently lives, which of the European regions/countries has your dog lived in **within the past 2 years**?

Circle all that apply.

a. My dog has never lived in Europe.
b. Sweden, Netherland, and/or the Baltic States (Lithuania, Estonia, Latvia)
c. Northern Europe (Denmark, Finland, Ireland, Norway, the United Kingdom)
d. Southern Europe and/or the Balkans (Spain, Italy, Bulgaria, Croatia, Greece, Turkey, Bosnia)
e. Eastern Europe (Armenia, Azerbaijan, Russia, Ukraine)
f. Central Europe (Austria, Czechoslovakia, Germany, Hungary, Poland, Switzerland, Belgium, France)
g. Other European countries not listed

18.a. How old was your dog when he/she joined your family?)

a. less than 3 months old
b. 3 to 6 months old
c. 6 months to 1 year old
d. more than a year old

18.b. Other than when he/she was a newborn puppy, has your dog ever belonged to someone other than you?

a. Yes
b. No

General Pet Activity Information

19. How many hours a day is your dog outside?
    a. Less than an hour
    b. 1 to 8 hours
    c. More than 8 hours but not all day
    d. All day but not all night
    e. All day and night

20. When your dog is outside, what best describes his/her surroundings?
Circle all that apply.
    a. Concrete, dirt, gravel
    b. Grass, field, meadow
    c. Wooded or forested
    d. Shrubs and bushes
    e. Other

21. What type of recreational activities does your dog do?
Circle all that apply.
    a. Neighborhood walks
    b. Wilderness walks (for example hiking, fishing, hunting, camping)
    c. Dog Parks
    d. None of the above

22. Have you ever used your dog for sport (for example hunting, retrieving, field trials)?
    a. Yes
    b. No

23. When recreating outdoors, is your dog primarily on leash?
    a. Yes
    b. No

****

END OWNER SURVEY
THANK YOU FOR YOUR PARTICIPATION
PLEASE HAND COMPLETED SURVEY AND THIS SHEET TO THE RECEPTIONIST OR TECHNICIAN

****

THE FOLLOWING SECTION
IS TO BE COMPLETED BY
THE VETERINARY TREATMENT FACILITY STAFF

****

IDEXX 4Dx Plus® Test Results

24. *Borrelia burgdorferi* (Lyme Disease)
    a. Positive
    b. Negative

25. *Anaplasma phagocytophilum* (Anaplasmosis)
    a. Positive
    b. Negative

26. *Ehrlichia canis* (Ehrlichiosis)
    a. Positive
    b. Negative

27. *Dirofilaria immitis* (Heartworm Disease)
    a. Positive
    b. Negative

VTF Information

28. VTF Name and Location (Installation):_____

29. Survey collection date (DDMMYY): _____

Appendix B-5

Military Pet Dog *B. burgdorferi* Risk Factor Data (Characteristics of the Study Population)

Table A: *B. burgdorferi* Seroprevalence by Owner Service Component

Characteristic	Sero Negative (N=3667)		Sero Positive (N=109)		Total (N=3776)	
Sponsor's Service Component						
Active Duty	2667	72.7%	80	73.4%	2747	72.7%
Reserve Corps	15	0.4%	1	0.9%	16	0.4%
National Guard	6	0.2%	1	0.9%	7	0.2%
Retired Military	832	22.7%	23	21.1%	855	22.6%
Other	110	3.0%	2	1.8%	112	3.0%
Missing	37	1.0%	2	1.8%	39	1.0%

Table B: *B. burgdorferi* Seroprevalence by Ownership Characteristics

Characteristic	Sero Negative (N=3667)		Sero Positive (N=109)		Total (N=3776)	
Age of Dog When Joined Family						
< 3 months	1813	49.44%	41	37.61%	1854	49.10%
3 to 6 months	800	21.82%	28	25.69%	828	21.93%
6 months to a year	370	10.09%	12	11.01%	382	10.12%
> 1 year	603	16.44%	24	22.02%	627	16.60%
Missing	81	2.21%	4	1.00%	85	2.25%
Prior Ownership						
No	2422	66.05%	63	57.80%	2485	65.81%
Yes	1168	31.85%	45	41.28%	1213	32.12%
Missing	77	2.10%	1	0.92%	78	2.07%

Table C: *B. burgdorferi* Seroprevalence by Owner U.S. Residency History

Characteristic	Sero Negative (N=3667)		Sero Positive (N=109)		Total (N=3776)	
Owner Lived in U.S.						
No	107	2.92%	3	2.75%	110	2.91%
Yes	3560	97.08%	106	97.25%	3666	97.09%
Owner Lived in Midwest Region						
No	2591	70.66%	76	69.72%	2667	70.63%
Yes	1076	29.34%	33	30.28%	1109	29.37%
Owner Lived in West Region						
No	2101	57.29%	65	59.63%	2166	57.36%
Yes	1566	42.71%	44	40.37%	1610	42.64%
Owner Lived in Northeast Region						

No	2700	73.63%	48	44.04%	2748	72.78%
Yes	967	26.37%	61	55.96%	1028	27.22%
Owner Lived in Southeast Region						
No	1063	28.99%	41	37.61%	1104	29.24%
Yes	2604	71.01%	68	62.39%	2672	70.76%
Owner Lived in Other U.S. Regions						
No	3229	88.06%	101	92.66%	3330	88.19%
Yes	438	11.94%	8	7.34%	446	11.81%
Missing U.S. Residecy Data						
No	3623	98.80%	107	98.17%	3730	98.78%
Yes	44	1.20%	2	1.83%	46	1.22%

Table D: *B. burgdorferi* Seroprevalence by Owner European Residency History

Characteristic	**Sero Negative (N=3667)**		**Sero Positive (N=109)**		**Total (N=3776)**	
Owner Lived in Europe						
No	2025	55.22%	53	48.62%	2078	55.03%
Yes	1642	44.78%	56	51.38%	1698	44.97%
Owner Lived in Baltic Region						
No	3634	99.10%	106	97.25%	3740	99.05%
Yes	33	0.90%	3	2.75%	36	0.95%
Owner Lived in Northern Region						
No	3545	96.67%	107	98.17%	3652	96.72%
Yes	122	3.33%	2	1.83%	124	3.28%
Owner Lived in Southern Region						
No	3378	92.12%	100	91.74%	3478	92.11%
Yes	289	7.88%	9	8.26%	298	7.89%
Owner Lived in Eastern Region						
No	3649	99.51%	109	100.00%	3758	99.52%
Yes	18	0.49%	0	0.00%	18	0.48%
Owner Lived in Central Region						
No	2303	62.80%	58	53.21%	2361	62.53%
Yes	1364	37.20%	51	46.79%	1415	37.47%
Owner Lived in Other Region						
No	3601	98.20%	108	99.08%	3709	98.23%
Yes	66	1.80%	1	0.92%	67	1.77%
Missing Region European Residency Data						
No	3567	97.27%	106	97.25%	3293	87.21%
Yes	100	2.73%	3	2.75%	93	2.46%

Table E: *B. burgdorferi* Seroprevalence by Preventive Medicine Practices

Characteristic	Sero Negative (N=3667)		Sero Positive (N=109)		Total (N=3776)	
Primary Pet Care Facility						
Military Facility where survey completed	3148	85.85%	102	93.58%	3250	86.07%
Different Military Facility than survey completed	50	1.36%	1	0.92%	51	1.35%
Civilian Facility	309	8.43%	3	2.75%	312	8.26%
Missing	160	4.36%	3	2.75%	163	4.32%
History of Heartworm Preventative Use						
No History of Use	451	12.30%	10	9.17%	461	12.21%
Yes History of Use	2928	79.85%	88	80.73%	3016	79.87%
Owner not know	247	6.74%	10	9.17%	257	6.81%
Missing	41	1.12%	1	0.92%	42	1.11%
Current Heartworm Preventative Use						
No History of Use	1303	35.53%	45	41.28%	1348	35.70%
Yes History of Use	2235	60.95%	60	55.05%	2295	60.78%
Owner not know	90	2.45%	2	1.83%	92	2.44%
Missing	39	1.06%	2	1.83%	41	1.09%
History of Flea and Tick Prevatative Use						
No History of Use	395	10.77%	8	7.34%	403	10.67%
Yes History of Use	3100	84.54%	96	88.07%	3196	84.64%
Owner not know	132	3.60%	4	3.67%	136	3.60%
Missing	40	1.09%	1	0.92%	41	1.09%
Current Flea and Tick Prevatative Use						
No History of Use	1325	36.13%	41	37.61%	1366	36.18%
Yes History of Use	2258	61.58%	65	59.63%	2323	61.52%
Owner not know	48	1.31%	1	0.92%	49	1.30%
Missing	36	0.98%	2	1.83%	28	0.74%
History of Tick Exposure						
No	1990	54.27%	23	21.10%	2013	53.31%
Yes	1471	40.11%	72	66.06%	1543	40.86%
Owner not know	168	4.58%	13	11.93%	181	4.79%
Missing	38	1.04%	1	0.92%	39	1.03%

Table F: *B. burgdorferi* Seroprevalence by Dog Characteristics

Characteristic	Sero Negative (N=3667)		Sero Positive (N=109)		Total (N=3776)	
Gender						
Male Intact	549	14.97%	16	14.68%	565	14.96%
Male Neuter	1345	36.68%	40	36.70%	1385	36.68%

Female Intact	336	9.16%	7	6.42%	343	9.08%
Female Spayed	1318	35.94%	44	40.37%	1362	36.07%
Missing	119	3.25%	2	1.83%	121	3.20%
Age(years)						
Less than 1	172	4.69%	2	1.83%	174	4.61%
1 to 3	1134	30.92%	27	24.77%	1161	30.75%
3 to 6	1119	30.52%	33	30.28%	1152	30.51%
6 to 8	572	15.60%	18	16.51%	590	15.63%
more than 8	579	15.79%	26	23.85%	605	16.02%
Missing	91	2.48%	3	2.75%	94	2.49%
Breed Group						
Sporting	693	18.90%	36	33.03%	729	19.31%
Herding	441	12.03%	23	21.10%	464	12.29%
Working	454	12.38%	17	15.60%	471	12.47%
Hound	363	9.90%	8	7.34%	371	9.83%
Terrier	316	8.62%	2	1.83%	318	8.42%
Toy	826	22.53%	9	8.26%	835	22.11%
Non-Sporting	233	6.35%	6	5.50%	239	6.33%
Other	179	4.88%	5	4.59%	184	4.87%
Missing	162	4.42%	3	2.75%	165	4.37%
Weight Group (pounds)						
Less than 15	942	25.69%	11	10.09%	953	25.24%
16 to 35	836	22.80%	16	14.68%	852	22.56%
36 to 55	583	15.90%	22	20.18%	605	16.02%
56 to 75	706	19.25%	33	30.28%	739	19.57%
More than 75	514	14.02%	25	22.94%	539	14.27%
Missing	86	2.35%	2	1.83%	88	2.33%

Table G: *B. burgdorferi* Seroprevalence Environment Characteristics

Characteristic	**Sero Negative (N=3667)**		**Sero Positive (N=109)**		**Total (N=3776)**	
Dirt						
No	2505	68.31%	82	75.23%	82	2.17%
Yes	1162	31.69%	27	24.77%	1189	31.49%
Missing	0	0.00%	0	0.00%	0	0.00%
Grass						
No	501	13.66%	18	16.51%	519	13.74%
Yes	3166	86.34%	91	83.49%	3257	86.26%
Missing	0	0.00%	0		0	0.00%
Wood						

No	3053	83.26%	79	72.48%	3132	82.94%
Yes	614	16.74%	30	27.52%	644	17.06%
Missing	0	0.00%	0	0.00%	0	0.00%
Bush						
No	2889	78.78%	82	75.23%	2971	78.68%
Yes	778	21.22%	27	24.77%	805	21.32%
Missing	0	0.00%	0	0.00%	0	0.00%
Other						
No	3598	98.12%	106	97.25%	3704	98.09%
Yes	69	1.88%	3	2.75%	72	1.91%
Missing	0	0.00%	0	0.00%	0	0.00%
Missing						
No	3628	98.94%	108	99.08%	3736	98.94%
Yes	39	1.06%	1	0.92%	40	1.06%
Missing	0	0.00%	0	0.00%	0	0.00%

Table H: *B. burgdorferi* Seroprevalence Recreational Activities

Characteristic	**Sero Negative (N=3667)**		**Sero Positive (N=109)**		**Total (N=3776)**	
Neighborhood Walks						
No	513	13.99%	18	16.51%	531	14.06%
Yes	3154	86.01%	91	83.49%	3245	85.94%
Missing	0	0.00%	0	0.00%	0	0.00%
Wilderness Walks						
No	2581	70.38%	52	47.71%	2633	69.73%
Yes	1086	29.62%	57	52.29%	1143	30.27%
Missing	0	0.00%	0		0	0.00%
Dog Parks						
No	2620	71.45%	82	75.23%	2702	71.56%
Yes	1047	28.55%	27	24.77%	1074	28.44%
Missing	0	0.00%	0	0.00%	0	0.00%
No Recreational Activities						
No	3386	92.34%	102	93.58%	3488	92.37%
Yes	281	7.66%	7	6.42%	288	7.63%
Missing	0	0.00%	0	0.00%	0	0.00%
Missing						
No	3629	98.96%	108	99.08%	3737	98.97%
Yes	38	1.04%	1	0.92%	39	1.03%
Missing	0	0.00%	0	0.00%	0	0.00%
Sporting						

No	3522	96.05%	102	93.58%	3624	95.97%
Yes	112	3.05%	6	5.50%	118	3.13%
Missing	33	0.90%	1	0.92%	34	0.90%
Leash						
No	758	20.67%	26	23.85%	784	20.76%
Yes	2843	77.53%	81	74.31%	2924	77.44%
Missing	66	1.80%	2	1.83%	68	1.80%
Daily Time Spend Outdoors (hours)						
< 1	1059	28.88%	27	24.77%	1086	28.76%
1 to 8	2417	65.91%	77	70.64%	2494	66.05%
> 8 (not all day)	35	0.95%	2	1.83%	37	0.98%
All day and night	41	1.12%	1	0.92%	42	1.11%
Missing	56	1.53%	1	0.92%	57	1.51%

Table I: *B. burgdorferi* Seroprevalence by Dog U.S. Residency History in last 5 years

Characteristic	**Sero Negative (N=3667)**		**Sero Positive (N=109)**		**Total (N=3776)**	
Lived in U.S. in past 5 years						
No	406	11.07%	12	11.01%	418	11.07%
Yes	3261	88.93%	97	88.99%	3358	88.93%
Missing/Unclear	0	0.00%	0	0.00%	0	0.00%
Lived in MW in past 5 years						
No	3101	84.57%	90	82.57%	3191	84.51%
Yes	566	15.43%	19	17.43%	585	15.49%
Lived in W in past 5 years						
No	2680	73.08%	88	80.73%	2768	73.31%
Yes	987	26.92%	21	19.27%	1008	26.69%
Lived in NE in past 5 years						
No	3139	85.60%	63	57.80%	3202	84.80%
Yes	528	14.40%	46	42.20%	574	15.20%
Lived in SE in past 5 years						
No	1526	41.61%	57	52.29%	1583	41.92%
Yes	2141	58.39%	52	47.71%	2193	58.08%
Lived in Other U.S. Regions in past 5 years						
No	3510	95.72%	108	99.08%	3618	95.82%
Yes	157	4.28%	1	0.92%	158	4.18%
Missing Dog Residecy Data (for last 5 years)						
No	3592	97.95%	106	97.25%	3698	97.93%
Yes	75	2.05%	3	2.75%	78	2.07%

Table J: *B. burgdorferi* Seroprevalence by Dog U.S. Residency History in last 2 years						
Characteristic	**Sero Negative (N=3667)**		**Sero Positive (N=109)**		**Total (N=3776)**	
Lived in U.S. in past 2 years						
No	640	17.45%	19	17.43%	659	17.45%
Yes	3027	82.55%	90	82.57%	3117	82.55%
Lived in MW in past 2 years						
No	3248	88.57%	93	85.32%	3341	88.48%
Yes	419	11.43%	16	14.68%	435	11.52%
Lived in W in past 2 years						
No	2833	77.26%	97	88.99%	2930	77.60%
Yes	834	22.74%	12	11.01%	846	22.40%
Lived in NE in past 2 years						
No	3247	88.55%	68	62.39%	3315	87.79%
Yes	420	11.45%	41	37.61%	461	12.21%
Lived in SE in past 2 years						
No	1791	48.84%	69	63.30%	1860	49.26%
Yes	1876	51.16%	40	36.70%	1916	50.74%
Lived in Other U.S. Regions in past 2 years						
No	3574	97.46%	109	100.00%	3683	97.54%
Yes	93	2.54%	0	0.00%	93	2.46%
Missing Dog Residecy Data (for last 2 years)						
No	3501	95.47%	102	93.58%	3603	95.42%
Yes	166	4.53%	7	6.42%	173	4.58%

Table K: *B. burgdorferi* Seroprevalence by Dog U.S. Residency History in last 5 years						
Characteristic	**Sero Negative (N=3667)**		**Sero Positive (N=109)**		**Total (N=3776)**	
Lived in Europe in past 5 years						
No	2708	73.85%	74	67.89%	2782	73.68%
Yes	959	26.15%	35	32.11%	994	26.32%
Lived in Baltic Region in past 5 years						
No	3653	99.62%	108	99.08%	3761	99.60%
Yes	14	0.38%	1	0.92%	15	0.40%
Lived in Northern Region in past 5 years						
No	3649	99.51%	109	100.00%	3758	99.52%
Yes	18	0.49%	0	0.00%	18	0.48%
Lived in Southern Region in past 5 years						
No	3547	96.73%	108	99.08%	3655	96.80%
Yes	120	3.27%	1	0.92%	121	3.20%
Lived in Eastern Region in past 5 years						

No	3662	99.86%	109	100.00%	3771	99.87%
Yes	5	0.14%	0	0.00%	5	0.13%
Lived in Central Region in past 5 years						
No	2832	77.23%	75	68.81%	2907	76.99%
Yes	835	22.77%	34	31.19%	869	23.01%
Lived in Other Region in past 5 years						
No	3652	99.59%	109	100.00%	3761	99.60%
Yes	15	0.41%	0	0.00%	15	0.40%
Missing Region Data for past 5 years						
No	3572	97.41%	107	98.17%	3679	97.43%
Yes	95	2.59%	2	1.83%	97	2.57%

Table L: *B. burgdorferi* Seroprevalence by Dog European Residency in Last 2 Years

Characteristic	**Sero Negative (N=3667)**		**Sero Positive (N=109)**		**Total (N=3776)**	
Lived in Europe in past 2 years						
No	2743	74.80%	76	69.72%	2819	74.66%
Yes	924	25.20%	33	30.28%	957	25.34%
Lived in Baltic Region in past 2 years						
No	3657	99.73%	108	99.08%	3765	99.71%
Yes	10	0.27%	1	0.92%	11	0.29%
Lived in Northern Region in past 2 years						
No	3656	99.70%	109	100.00%	3765	99.71%
Yes	11	0.30%	0	0.00%	11	0.29%
Lived in Southern Region in past 2 years						
No	3564	97.19%	107	98.17%	3671	97.22%
Yes	103	2.81%	2	1.83%	105	2.78%
Lived in Eastern Region in past 2 years						
No	3662	99.86%	109	100.00%	3771	99.87%
Yes	5	0.14%	0	0.00%	5	0.13%
Lived in Central Region in past 2 years						
No	2846	77.61%	77	70.64%	2923	77.41%
Yes	821	22.39%	32	29.36%	853	22.59%
Lived in Other Region in past 2 years						
No	3654	99.65%	109	100.00%	3763	99.66%
Yes	13	0.35%	0	0.00%	13	0.34%
Missing Region Data for past 2 years						
No	3553	96.89%	105	96.33%	3658	96.88%
Yes	114	3.11%	4	3.67%	118	3.13%

Table M: *B. burgdorferi* Seroprevalence by Co-infection

Infectious Agent	Sero Negative (N=3667)		Sero Positive (N=109)		Total (N=3776)	
Anaplasma						
No	3649	99.51%	100	91.74%	3749	99.28%
Yes	16	0.44%	9	8.26%	25	0.66%
Missing	2	0.05%	0	0.00%	2	0.05%
Dirofilaria						
No	3631	99.02%	104	95.41%	3735	98.91%
Yes	32	0.87%	5	4.59%	37	0.98%
Missing	4	0.11%	0	0.00%	4	0.11%
Ehrlichia						
No	3568	97.30%	97	88.99%	3665	97.06%
Yes	98	2.67%	12	11.01%	110	2.91%
Missing	1	0.03%	0	0.00%	1	0.03%

Appendix C-5

Odds of *B. burgdorferi* Sero-positivity by Risk Factor (ORs, 95% CIs)

Table I: Odds of *B. burgdorferi* Sero-positivity by Owner Characteristics

Characteristic	OR	95% CI	p-value
Sponsor's Service Component			**0.4393**
Retired Military*	1.000		
Active Duty	1.085	(0.678, 1.736)	0.7336
Reserve Corps	2.412	(0.305, 19.039)	0.4037
National Guard	6.030	(0.697, 52.132)	0.1026
Other	0.658	(0.153, 2.828)	0.5734

*Referent category

Table II: Odds of *B. burgdorferi* Sero-positivity by Ownership Characteristics

Characteristic	OR	95% CI	p-value
Age of Dog When Joined Family			**0.1244**
> 1 year*	1.000		
< 3 months	0.568	(0.340, 0.948)	0.0305
3 to 6 months	0.879	(0.505, 1.532)	0.6501
6 months to a year	0.815	(0.403, 1.649)	0.5692
Prior Ownership			
Yes*	1.000		
No	0.675	(0.458, 0.996)	0.0477

*Referent category

Table III: Odds of *B. burgdorferi* Sero-positivity by Owner U.S. Residency History

Characteristic	OR	95% CI	p-value
Owner Lived in U.S.			
Yes*	1.000		
No	0.942	(0.294, 3.014)	0.9193
Owner Lived in MW Region			
Yes	1.000		
No	0.956	(0.632, 1.448)	0.83332
Owner Lived in W Region			
Yes	1.000		
No	1.010	(0.747, 1.623)	0.6268
Owner Lived in NE Region			
Yes	1.000		
No	0.282	(0.192, 0.414)	<0.0001
Owner Lived in SE Region			

Yes	1.000		
No	1.477	(0.996, 2.190)	0.0524
Owner Lived in Other U.S. Regions			
Yes	1.000		
No	1.713	(0.828, 3.542)	0.1468
Missing U.S. Residecy Data			
Yes	1.000		
No	0.650	(0.155, 2.715)	0.5545

*"Yes" as referent category for each category

Table IV: Odds of *B. burgdorferi* Sero-positivity by Owner European Residency History

Characteristic	OR	95% CI	p-value
Owner Lived in Europe			
Yes*	1.000		
No	0.767	(0.524, 1.124)	0.1735
Owner Lived in Baltic Region			
Yes	1.000		
No	0.321	(0.097, 1.063)	0.0628
Owner Lived in Northern Region			
Yes	1.000		
No	1.841	(0.449, 7.544)	0.3963
Owner Lived in Southern Region			
Yes	1.000		
No	0.951	(0.476, 1.900)	0.8860
Owner Lived in Eastern Region			
Yes	1.000		
No	>999.999	(<0.001, >999.999)	0.9830
Owner Lived in Central Region			
Yes	1.000		
No	0.674	(0.460 , 0.987)	0.0427
Owner Lived in Other Region			
Yes	1.000		
No	1.979	(0.272, 14.392)	0.5000
Missing Region European Residency Data			
Yes	1.000		
No	0.991	(0.309, 3.174)	0.9872

*"Yes" as referent category for each category

Table V: Odds of *B. burgdorferi* Sero-positivity by Dog U.S. Residency History in last 5 years

Characteristic	OR	95% CI	p-value
Primary Pet Care Facility			**0.1117**
Military Facility where survey completed*	1.000		
Different Military Facility	0.617	(0.084, 4.512)	0.6345
Civilian Facility	0.300	(0.094, 0.950)	0.0407
History of Heartworm Preventative Use			**0.4152**
No History of Use*	1.000		
Yes History of Use	1.355	(0.699, 2.627)	0.3675
Owner not know	1.826	(0.750, 4.447)	0.1849
Current Heartworm Preventative Use			**0.4166**
No History of Use*	1.000		
Yes History of Use	0.777	(0.525, 1.151)	0.2083
Owner not know	0.644	(0.154, 2.696)	0.5467
History of Flea and Tick Preventative Use			**0.5209**
No History of Use*	1.000		
Yes History of Use	1.529	(0.738, 3.169)	0.2535
Owner not know	1.496	(0.443, 5.049)	0.5161
Current Flea and Tick Preventative Use			**0.8841**
No History of Use*	1.000		
Yes History of Use	0.930	(0.626, 1.383)	0.7211
Owner not know	0.674	(0.091, 4.998)	0.6996
History of Tick Exposure			**<0.0001**
Yes*	1.000		
No	0.236	(0.147, 0.379)	<.0001
Owner Not Know	1.619	(0.878, 2.987)	0.1227

* Referent category

Table VI: Odds of *B. burgdorferi* Sero-positivity by by Dog Characteristics

Characteristic	OR	95% CI	p-value
Gender			**0.7058**
Male Intact*	1.000		
Male Neuter	1.020	(0.567, 1.838)	0.9462
Female Intact	0.715	(0.291, 1.756)	0.4640
Female Spayed	1.145	(0.641, 2.047)	0.6467
Age(years)			**0.1244**
more than 8*			
less than 1	0.259	(0.061, 1.102)	1.102
1 to 3	0.530	(0.307, 0.917)	0.917

3 to 6	0.657	(0.389, 1.109)	1.109
6 to 8	0.701	(0.380, 1.292)	1.292
Breed Group			**<0.0001**
Sporting*	1.000		
Herding	1.004	(0.587, 1.717)	0.9885
Working	0.721	(0.400, 1.299)	0.2758
Hound	0.424	(0.195, 0.922)	0.922
Terrier	0.122	(0.029, 0.509)	0.0039
Toy	0.210	(0.100, 0.438)	<0.0001
Non-Sporting	0.496	(0.206, 1.191)	0.1168
Other	0.538	(0.208, 1.390)	0.2004
Weight Group (pounds)			**<0.0001**
56 to 75*	1.000		
Less than 15	0.250	(0.125, 0.498)	<0.0001
16 to 35	0.409	(0.223, 0.750)	0.0035
36 to 55	0.807	(0.466, 1.400)	0.4460
More than 75	1.041	(0.611, 1.771)	0.8835

*Referent category

Table VII: Odds of *B. burgdorferi* Sero-positivity by Environment Characteristics

Characteristic	OR	95% CI	p-value
Dirt			
Yes*	1.000		
No	0.3424	0.2247 2.3226	0.1275
Grass			
Yes*	1.000		
No	1.250	0.748 2.091	0.3948
Wood			
Yes*	1.000		
No	0.530	(0.345, 0.813)	0.0037
Bush			
Yes*	1.000		
No	0.818	(0.526, 1.272)	0.3725
Other			
Yes*	1.000		
No	0.678	(0.210, 2.187)	0.5151
Missing			
Yes*	1.000		
No	1.161	(0.158, 8.528)	0.8834

*Referent category

Table VIII: Odds of *B. burgdorferi* Sero-positivity by Recreational Activity Characteristics

Characteristic	OR	95% CI		p-value
Neighborhood Walks				
Yes*	1.000			
No	1.216	0.727	2.034	0.4555
Wilderness Walks				
Yes*	1.000			
No	0.384	0.262	0.563	<0.0001
Dog Parks				
Yes*	1.000			
No	1.214	(0.781	1.886)	0.3893
None of the Listed Recreational Activites				
Yes*	1.000			
No	1.209	(0.557	2.626)	0.6312
Missing				
Yes*	1.000			
No	1.131	(0.154	8.312)	0.9038
Dog is used for Sporting				
Yes*	1.000			
No	0.541	(0.232	1.258)	0.1534
Dog primarily recreates on Leash				
Yes*	1.000			
No	1.204	0.768	1.886	0.4178
Daily Time Spend Outdoors (hours)				**0.7154**
All day and night*	1.000			
< 1	1.045	0.139	7.881	0.9657
1 to 8	1.306	0.177	9.619	0.7932
> 8 (not all day)	0.695	0.042	11.431	0.7989

*Referent category

Table IX: Odds of *B. burgdorferi* Sero-positivity by Dog U.S. Residency History in last 5 years

Characteristic	OR	95% CI		p-value
Lived in U.S. in past 5 years				
Yes*	1.000			
No	0.994	0.541	1.826	0.9837
Lived in MW in past 5 years				
Yes	1.000			
No	0.865	0.523	1.430	0.5706

Lived in W in past 5 years			
Yes	1.000		
No	1.543	0.953, 2.498	0.0774
Lived in NE in past 5 years			
Yes	1.000		
No	0.230	0.156 0.341	<0.0001
Lived in SE in past 5 years			
Yes	1.000		
No	1.538	1.050 2.2520	0.0270
Lived in Other U.S. Region in past 5 years			
Yes	1.000		
No	4.830	(0.670, 34.817)	0.1189
Missing U.S. Residency Data			
Yes	1.000		
No	0.737	(0.229, 2.374)	0.6092

*Referent category

Table X: Odds of *B. burgdorferi* Sero-positivity by Dog U.S. Residency History in last 2 years

Characteristic	**OR**	**95% CI**	**p-value**
Lived in U.S. in past 2 years			
Yes*	1.000		
No	0.998	(0.604, 1.650)	0.9953
Lived in MW in past 2 years			
Yes	1.000		
No	0.749	(0.437, 1.286)	0.2953
Lived in W in past 2 years			
Yes	1.000		
No	2.379	(1.300, 4.356)	0.0050
Lived in NE in past 2 years			
Yes	1.000		
No	0.214	(0.144, 0.320)	<0.0001
Lived in SE in past 2 years			
Yes	1.000		
No	1.807	(1.217, 2.682)	0.0033
Lived in Other U.S. Region in past 2 years			
Yes	1.000		
No	>999.999	(<0.001, >999.999)	0.9834
Missing U.S. Residency Data			
Yes	1.000		
No	0.691	(0.316, 1.509)	0.3537

*Referent category

Table XI: Odds of *B. burgdorferi* Sero-positivity by by Dog European Residency History in past 5 years

Characteristic	OR	95% CI	p-value
Lived in Europe in past 5 years			
Yes	1.000		
No	0.749	(0.497, 1.127)	0.1650
Lived in Baltic Region in past 5 years			
Yes	1.000		
No	0.414	(0.054, 3.176)	0.3962
Lived in Northern Region in past 5 years			
Yes	1.000		
No	>999.999	(<0.001, >999.999)	0.9834
Lived in Southern Region in past 5 years			
Yes	1.000		
No	3.654	(0.506, 26.394)	0.1990
Lived in Eastern Region in past 5 years			
No	1.000		
Yes	>999.999	(<0.001, >999.999)	0.9834
Lived in Central Region in past 5 years			
No	1.000		
Yes	0.650	(0.431, 0.982)	0.0410
Lived in Other Region in past 5 years			
No	1.000		
Yes	>999.999	(<0.001, >999.999)	0.9834
Missing Region European Residency Data in past 5 years			
No	1.000		
Yes	1.421	(0.346, 5.837)	0.6258

*Referent category

Table XII: Odds of *B. burgdorferi* Sero-positivity by by Dog European Residency History in past 2 years

Characteristic	OR	95% CI	p-value
Lived in Europe in past 2 years			
Yes	1.000		
No	0.776	(0.512, 1.175)	0.2307
Lived in Baltic Region in past 2 years			
Yes	1.000		
No	0.295	(0.037, 2.327)	0.2468
Lived in Northern Region in past 2 years			
Yes	1.000		
No	>999.999	(<0.001, >999.999)	0.9834

Lived in Southern Region in past 2 years			
Yes	1.000		
No	1.549	(0.377, 6.348)	0.5454
Lived in Eastern Region in past 2 years			
No	1.000		
Yes	>999.999	(<0.001, >999.999)	0.9834
Lived in Central Region in past 2 years			
No	1.000		
Yes	0.694	(0.456, 1.056)	0.0880
Lived in Other Region in past 2 years			
No	1.000		
Yes	>999.999	(<0.001, >999.999)	0.9834
Missing Region European Residency Data in past 2 years			
No	1.000		
Yes	0.842	(0.305, 2.325)	0.7403

*Referent category

Table XIII: Odds of *B. burgdorferi* Sero-positivity by Co-infection

Co-Infection	OR	95% CI	p-value
Anaplasma			
Yes*	1.000		
No	0.049	(0.021, 0.113)	<0.0001
Dirofilaria			
Yes*	1.000		
No	0.183	(0.070, 0.480)	0.0005
Ehrlichia			
Yes*	1.000		
No	0.2212	(0.118, 0.418)	<0.0001

*Referent category

Appendix D-5

Bivariate Analyses: Odds of *B. burgdorferi* Sero-positivity by Installation Adjusting for Identified Risk Factors

Note: Analyses do not include Signonella due to "0" sero-positivity.
ORs are as compared to Fort Drum NY, the installation with the highest prevalence

CRUDE

Table a: Installation Only Model (CRUDE)					
Parameter	**DF**	**Estimate**	**Standard Error**	**Pr > ChiSq**	**OR_{CRUDE}**
Intercept	1	-2.2827	0.1885	<.0001	
Kaiserslautern, GE	1	-0.9786	0.3494	0.0051	0.376
Stuttgart. GE	1	-0.956	0.4067	0.0188	0.384
Rota, SP	1	-1.1985	0.7422	0.1064	0.302
Spangdahlem, GE	1	-1.2171	0.3873	0.0017	0.296
Fort Belvoir, VA	1	-0.7328	0.3748	0.0505	0.481
Fort Bragg, NC	1	-2.3617	0.4873	<.0001	0.094
Fort Carson, CO	1	-1.7426	0.489	0.0004	0.175
Fort Hood, TX	1	-2.8527	1.0203	0.0052	0.058
Fort Lee, VA	1	-1.8016	0.4888	0.0002	0.165
Camp Lejeune, NC	1	-1.7976	0.4528	<.0001	0.166
Camp Pendleton, CA	1	-1.4549	1.0292	0.1575	0.233
JBLM, WA	1	-1.7691	0.4529	<.0001	0.17
Great Lakes, IL	1	-1.135	0.4058	0.0052	0.321

Coefficient, SE, and p-values look reasonable. Rota, Spain and Camp Pendleton, CA may be questionable (n= 67 and n = 43, respectively)

ADJUSTED (Tick)

Table b: Bivariate Model with Tick (Adjusted)						
Parameter	**DF**	**Estimate**	**Standard Error**	**Pr > ChiSq**	**$OR_{ADJUSTED}$**	**Change in OR (Cr-Ad)/Ad**
Intercept	1	-1.6903	0.2149	<.0001		
Kaiserslautern, GE	1	-1.2501	0.3583	0.0005	0.286	0.31
Stuttgart. GE	1	-1.2953	0.4155	0.0018	0.274	0.40
Rota, SP	1	-1.2197	0.7521	0.1049	0.295	0.02
Spangdahlem, GE	1	-1.4788	0.3935	0.0002	0.228	0.30
Fort Belvoir, VA	1	-1.0692	0.3824	0.0052	0.343	0.40
Fort Bragg, NC	1	-2.3813	0.4904	<.0001	0.092	0.02
Fort Carson, CO	1	-1.5204	0.4944	0.0021	0.219	-0.20
Fort Hood, TX	1	-2.5642	1.0245	0.0123	0.077	-0.25

Fort Lee, VA	1	-1.9852	0.4933	<.0001	0.137	0.20
Camp Lejeune, NC	1	-2.0934	0.4922	<.0001	0.123	0.35
Camp Pendleton, CA	1	-1.5546	1.0365	0.1336	0.211	0.10
JBLM, WA	1	-1.668	0.4591	0.0003	0.189	-0.10
Great Lakes, IL	1	-1.0968	0.4111	0.0076	0.334	-0.04
Tick	1	-1.4373	0.2524	<.0001	0.238	

Most change in OR > 10% → confounding by Tick (Type III Analysis p-value < 0.0001)

ADJUSTED (Wt in lbs)

Table c: Bivariate Model with Wt in lbs (Adjusted)						
Parameter	**DF**	**Estimate**	**Standard Error**	**Pr > ChiSq**	**$OR_{ADJUSTED}$**	**Change in OR (Cr-Ad)/Ad**
Intercept	1	-1.8266	0.2444	<.0001		
Kaiserslautern, GE	1	-0.9097	0.3538	0.0101	0.403	-0.07
Stuttgart. GE	1	-0.9943	0.411	0.0156	0.37	0.04
Rota, SP	1	-1.1693	0.7468	0.1174	0.311	-0.03
Spangdahlem, GE	1	-1.1915	0.3916	0.0023	0.304	-0.03
Fort Belvoir, VA	1	-0.6794	0.3797	0.0736	0.507	-0.05
Fort Bragg, NC	1	-2.3955	0.4902	<.0001	0.091	0.03
Fort Carson, CO	1	-1.7507	0.4923	0.0004	0.174	0.01
Fort Hood, TX	1	-2.7136	1.0227	0.008	0.066	-0.12
Fort Lee, VA	1	-1.7147	0.492	0.0005	0.18	-0.08
Camp Lejeune, NC	1	-1.991	0.4914	<.0001	0.137	0.21
Camp Pendleton, CA	1	-1.542	1.0338	0.1358	0.214	0.09
JBLM, WA	1	-1.7885	0.4565	<.0001	0.167	0.02
Great Lakes, IL	1	-1.0333	0.4102	0.0118	0.356	-0.10
Wt: 16-35 vs. < 15	1	-1.4207	0.3546	<.0001	0.242	
Wt: 36-55 vs. < 15		-0.9467	0.3121	0.0024	0.388	
Wt: 56-75 vs. < 15		-0.1974	0.285	0.4884	0.821	
Wt: >76 vs. < 15		0.0484	0.2759	0.8606	1.05	

Two change in OR >10%; Type III Analysis p-value < 0.0001 →Confounding

ADJUSTED (DgSERes2)

Table d: Bivariate Model with DgSERes2 (Adjusted)						
Parameter	**DF**	**Estimate**	**Standard Error**	**Pr > ChiSq**	**$OR_{ADJUSTED}$**	**Change in OR (Cr-Ad)/Ad**
Intercept	1	-2.4441	0.2646	<.0001		
Kaiserslautern, GE	1	-0.988	0.3496	0.0047	0.372	0.01
Stuttgart. GE	1	-0.9632	0.4069	0.0179	0.382	0.01

Rota, SP	1	-1.2434	0.7439	0.0946	0.288	0.05
Spangdahlem, GE	1	-1.2502	0.389	0.0013	0.286	0.03
Fort Belvoir, VA	1	-0.617	0.3968	0.12	0.54	-0.11
Fort Bragg, NC	1	-2.2178	0.5141	<.0001	0.109	-0.14
Fort Carson, CO	1	-1.749	0.4891	0.0003	0.174	0.01
Fort Hood, TX	1	-2.7063	1.0342	0.0089	0.067	-0.13
Fort Lee, VA	1	-1.6672	0.512	0.0011	0.189	-0.13
Camp Lejeune, NC	1	-1.6589	0.4795	0.0005	0.19	-0.13
Camp Pendleton, CA	1	-1.4798	1.0297	0.1507	0.228	0.02
JBLM, WA	1	-1.7803	0.4531	<.0001	0.169	0.01
Great Lakes, IL	1	-1.1762	0.4082	0.004	0.308	0.04
DgSERes2	1	0.2245	0.2521	0.3732	1.252	

Four changes in OR > 10% → confounding by DgSERes2 (Type III Analysis p-value = 0.3732)

ADJUSTED (DgWRes2)

Table e: Bivariate Model with DgWRes2 (Adjusted)						
Parameter	**DF**	**Estimate**	**Standard Error**	**Pr > ChiSq**	**$OR_{ADJUSTED}$**	**Change in OR (Cr-Ad)/Ad**
Intercept	1	-3.2733	0.4535	<.0001		
Kaiserslautern, DE	1	-0.988	0.3498	0.0047	0.372	0.01
Stuttgart. GE	1	-0.9649	0.4072	0.0178	0.381	0.01
Rota, SP	1	-1.2039	0.7428	0.1051	0.3	0.01
Spangdahlem, DE	1	-1.2251	0.3877	0.0016	0.294	0.01
Fort Belvoir, VA	1	-0.772	0.3753	0.0397	0.462	0.04
Fort Bragg, NC	1	-2.3945	0.4876	<.0001	0.091	0.03
Fort Carson, CO	1	-0.953	0.5719	0.0957	0.386	-0.55
Fort Hood, TX	1	-2.8872	1.0205	0.0047	0.056	0.04
Fort Lee, VA	1	-1.8437	0.4892	0.0002	0.158	0.04
Camp Lejeune, NC	1	-1.8285	0.4531	<.0001	0.161	0.03
Camp Pendleton, CA	1	-0.6593	1.0744	0.5395	0.517	-0.55
JBLM, WA	1	-0.9713	0.5436	0.074	0.379	-0.55
Great Lakes, IL	1	-1.1694	0.4062	0.004	0.311	0.03
DgWRes2	1	1.0685	0.4318	<.0001	2.911	

Changes in OR >> 10% for three → confounding by DgWRes2 (Type III Analysis p-value = 0.0133)

ADJUSTED (DgNERes2)

Table f: Bivariate Model with DgNERes2 (Adjusted)						
Parameter	**DF**	**Estimate**	**Standard Error**	**Pr > ChiSq**	**$OR_{ADJUSTED}$**	**Change in OR (Cr-Ad)/Ad**

Intercept	1	-2.1445	0.1909	<.0001		
Kaiserslautern, DE	1	-0.1306	0.4233	0.7577	0.878	-0.57
Stuttgart. DE	1	-0.1461	0.4649	0.7533	0.864	-0.56
Rota, SP	1	-0.364	0.7793	0.6404	0.695	-0.57
Spangdahlem, DE	1	-0.3532	0.4582	0.4408	0.702	-0.58
Fort Belvoir, VA	1	-0.0951	0.4101	0.8165	0.909	-0.47
Fort Bragg, NC	1	-1.4721	0.5496	0.0074	0.229	-0.59
Fort Carson, CO	1	-0.8919	0.5443	0.1013	0.41	-0.57
Fort Hood, TX	1	-1.9791	1.0503	0.0595	0.138	-0.58
Fort Lee, VA	1	-0.969	0.541	0.0733	0.379	-0.56
Camp Lejeune, NC	1	-0.972	0.5074	0.0554	0.378	-0.56
Camp Pendleton, CA	1	-0.5582	1.0617	0.5991	0.572	-0.59
JBLM, WA	1	-0.8973	0.516	0.082	0.408	-0.58
Great Lakes, IL	1	-0.2456	0.4793	0.6084	0.782	-0.59
DgNERes2	1	-1.0773	0.2994	0.0003	0.341	

Changes in OR > 10% → confounding by DgNERes2 (Type III Analysis p-value = 0.0003)

ADJUSTED (DgCRes5)

Table g: Bivariate Model with DgCRes5 (Adjusted)						
Parameter	**DF**	**Estimate**	**Standard Error**	**Pr > ChiSq**	**$OR_{ADJUSTED}$**	**Change in OR (Cr-Ad)/Ad**
Intercept	1	-1.4919	0.4214	0.0004		
Kaiserslautern, DE	1	-1.7235	0.5033	0.0006	0.178	1.11
Stuttgart. DE	1	-1.696	0.5436	0.0018	0.183	1.10
Rota, SP	1	-1.1618	0.7429	0.1178	0.313	-0.04
Spangdahlem, DE	1	-1.9615	0.5302	0.0002	0.141	1.10
Fort Belvoir, VA	1	-0.7288	0.3754	0.0522	0.482	0
Fort Bragg, NC	1	-2.3381	0.4877	<.0001	0.097	-0.03
Fort Carson, CO	1	-1.7534	0.4895	0.0003	0.173	0.01
Fort Hood, TX	1	-2.8423	1.0205	0.0054	0.058	0
Fort Lee, NJ	1	-1.8209	0.4895	0.0002	0.162	0.02
Camp Lejeune, NC	1	-1.7456	0.4542	0.0001	0.175	-0.05
Camp Pendleton, CA	1	-1.3993	1.0299	0.1743	0.247	-0.06
JBLM, WA	1	-1.7499	0.4533	0.0001	0.174	-0.02
Great Lakes, IL	1	-1.0945	0.4068	0.0071	0.335	-0.04
DgCRe5	1	-0.8465	0.412	0.0399	0.429	

Changes in OR > 10% → confounding by DgCRes5 (Type III Analysis p-value = 0.0399)

ADJUSTED (POwn)

Table h: Bivariate Model with POwn (Adjusted)						
Parameter	**DF**	**Estimate**	**Standard Error**	**Pr > ChiSq**	**$OR_{ADJUSTED}$**	**Change in OR (Cr-Ad)/Ad**
Intercept	1	-2.0151	0.2205	<.0001		
Kaiserslautern, DE	1	-0.9297	0.3503	0.008	0.395	-0.05
Stuttgart. DE	1	-0.8959	0.408	0.0281	0.408	-0.06
Rota, SP	1	-1.1588	0.7433	0.119	0.314	-0.04
Spangdahlem, DE	1	-1.1902	0.3879	0.0022	0.304	-0.03
Fort Belvoir, VA	1	-0.6825	0.3761	0.0696	0.505	-0.05
Fort Bragg, NC	1	-2.3538	0.4876	<.0001	0.095	-0.01
Fort Carson, CO	1	-1.7122	0.4895	0.0005	0.18	-0.03
Fort Hood, TX	1	-2.8522	1.0206	0.0052	0.058	0.00
Fort Lee, VA	1	-1.8017	0.4892	0.0002	0.165	0.00
Camp Lejeune, NC	1	-1.959	0.4886	<.0001	0.141	0.18
Camp Pendleton, CA	1	-1.4658	1.0302	0.1548	0.231	0.01
JBLM, WA	1	-1.7652	0.4533	<.0001	0.171	-0.01
Great Lakes, IL	1	-1.1078	0.4066	0.0064	0.33	-0.03
POwn	1	-0.4267	0.2018	0.0345	0.653	

Only one change in OR > 10%--> questionable confounding by POwn (Type III Analysis p-value = 0.0345)

ADJUSTED (Wood)

Table i: Bivariate Model with Wood (Adjusted)						
Parameter	**DF**	**Estimate**	**Standard Error**	**Pr > ChiSq**	**$OR_{ADJUSTED}$**	**Change in OR (Cr-Ad)/Ad**
Intercept	1	-1.712	0.2698	<.0001		
Kaiserslautern, DE	1	-1.1066	0.3542	0.0018	0.331	0.14
Stuttgart. DE	1	-1.1953	0.4184	0.0043	0.303	0.27
Rota, SP	1	-1.1195	0.7432	0.132	0.326	-0.07
Spangdahlem, DE	1	-1.2486	0.3882	0.0013	0.287	0.03
Fort Belvoir, VA	1	-0.8553	0.3791	0.0241	0.425	0.13
Fort Bragg, NC	1	-2.4208	0.4884	<.0001	0.089	0.06
Fort Carson, CO	1	-1.6852	0.4899	0.0006	0.185	-0.05
Fort Hood, TX	1	-2.7924	1.0208	0.0062	0.061	-0.05
Fort Lee, NJ	1	-1.9147	0.4917	<.0001	0.147	0.12
Camp Lejeune, NC	1	-1.8398	0.4537	<.0001	0.159	0.04
Camp Pendleton, CA	1	-1.3835	1.03	0.1792	0.251	-0.07
JBLM, WA	1	-1.8332	0.4543	<.0001	0.16	0.06
Great Lakes, IL	1	-1.1296	0.4063	0.0054	0.323	-0.01
Wood	1	-0.6633	0.2323	0.0043	0.515	

Four changes in OR > 10%--> confounding by Wood (Type III Analysis p-value = 0.0043)

ADJUSTED (WWalk)

Table j: Bivariate Model with WWalk (Adjusted)						
Parameter	**DF**	**Estimate**	**Standard Error**	**Pr > ChiSq**	**$OR_{ADJUSTED}$**	**Change in OR (Cr-Ad)/Ad**
Intercept	1	-1.6735	0.2209	<.0001		
Kaiserslautern, DE	1	-1.1521	0.3537	0.0011	0.316	0.19
Stuttgart. DE	1	-1.2322	0.4131	0.0029	0.292	0.32
Rota, SP	1	-1.1172	0.7452	0.1338	0.327	-0.08
Spangdahlem, DE	1	-1.4163	0.3919	0.0003	0.243	0.22
Fort Belvoir, VA	1	-0.705	0.3772	0.0616	0.494	-0.03
Fort Bragg, NC	1	-2.3458	0.4885	<.0001	0.096	-0.02
Fort Carson, CO	1	-1.881	0.4918	0.0001	0.152	0.15
Fort Hood, TX	1	-2.7463	1.0213	0.0072	0.064	-0.09
Fort Lee, VA	1	-1.7087	0.4906	0.0005	0.181	-0.09
Camp Lejeune, NC	1	-1.7424	0.4544	0.0001	0.175	-0.05
Camp Pendleton, CA	1	-1.3934	1.0323	0.1771	0.248	-0.06
JBLM, WA	1	-1.8501	0.455	<.0001	0.157	0.08
Great Lakes, IL	1	-1.0867	0.4079	0.0077	0.337	-0.05
WWalk	1	-0.9533	0.2032	<.0001	0.385	

Four changes in OR > 10%--> confounding by DgWRes2 (Type III Analysis p-value< 0.0001)

ADJUSTED (Ana)

Table k: Bivariate Model with Ana (Adjusted)						
Parameter	**DF**	**Estimate**	**Standard Error**	**Pr > ChiSq**	**$OR_{ADJUSTED}$**	**Change in OR (Cr-Ad)/Ad**
Intercept	1	0.6289	0.4829	0.1928		
Kaiserslautern, DE	1	-0.9633	0.355	0.0067	0.382	-0.02
Stuttgart. DE	1	-1.0727	0.4223	0.0111	0.342	0.12
Rota, SP	1	-1.1149	0.7435	0.1338	0.328	-0.08
Spangdahlem, DE	1	-1.2507	0.3954	0.0016	0.286	0.03
Fort Belvoir, VA	1	-0.6962	0.3799	0.0668	0.498	-0.03
Fort Bragg, NC	1	-2.338	0.4908	<.0001	0.097	-0.03
Fort Carson, CO	1	-1.7565	0.4956	0.0004	0.173	0.01
Fort Hood, TX	1	-2.8647	1.0253	0.0052	0.057	0.02
Fort Lee, NJ	1	-1.7179	0.4907	0.0005	0.179	-0.08
Camp Lejeune, NC	1	-1.7924	0.458	<.0001	0.167	-0.01
Camp Pendleton, CA	1	-1.3713	1.0302	0.1831	0.254	-0.08

JBLM, WA	1	-1.8016	0.4603	<.0001	0.165	0.03
Great Lakes, IL	1	-1.0981	0.4104	0.0075	0.333	-0.04
Ana	1	-2.9952	0.4581	<.0001	0.05	

Only one change in OR > 10%--> questionable confounding by Ana (Type III Analysis p-value <0.0001)

ADJUSTED (Diro)

Table l: Bivariate Model with Diro (Adjusted)

Parameter	DF	Estimate	Standard Error	Pr > ChiSq	$OR_{ADJUSTED}$	Change in OR (Cr-Ad)/Ad
Intercept	1	-0.4501	0.536	0.401		
Kaiserslautern, DE	1	-0.9487	0.3509	0.0069	0.387	-0.03
Stuttgart. DE	1	-0.9226	0.4082	0.0238	0.398	-0.04
Rota, SP	1	-1.2134	0.7458	0.1038	0.297	0.02
Spangdahlem, DE	1	-1.1621	0.3887	0.0028	0.313	-0.05
Fort Belvoir, VA	1	-0.6984	0.3763	0.0635	0.497	-0.03
Fort Bragg, NC	1	-2.4036	0.4898	<.0001	0.09	0.04
Fort Carson, CO	1	-1.7401	0.4905	0.0004	0.175	0.00
Fort Hood, TX	1	-2.8541	1.0218	0.0052	0.058	0.00
Fort Lee, VA	1	-1.8286	0.4912	0.0002	0.161	0.02
Camp Lejeune, NC	1	-1.8116	0.4547	<.0001	0.163	0.02
Camp Pendleton, CA	1	-1.3999	1.0298	0.174	0.247	-0.06
JBLM, WA	1	-1.74	0.454	0.0001	0.176	-0.03
Great Lakes, IL	1	-1.0758	0.4071	0.0082	0.341	-0.06
Diro	1	-1.8877	0.5237	0.0003	0.151	

No changes in OR > 10% → no confounding by Diro (Type III Analysis p-value = 0.0003)

ADJUSTED (Ehr)

Table m: Bivariate Model with Ehr (Adjusted)

Parameter	DF	Estimate	Standard Error	Pr > ChiSq	$OR_{ADJUSTED}$	Change in OR (Cr-Ad)/Ad
Intercept	1	-0.5039	0.3844	0.1898		
Kaiserslautern, DE	1	-1.0591	0.3525	0.0027	0.347	0.08
Stuttgart. DE	1	-1.0321	0.41	0.0118	0.356	0.08
Rota, SP	1	-1.3615	0.7507	0.0697	0.256	0.18
Spangdahlem, DE	1	-1.2106	0.3878	0.0018	0.298	-0.01
Fort Belvoir, VA	1	-0.8757	0.3813	0.0216	0.417	0.15
Fort Bragg, NC	1	-2.4542	0.4898	<.0001	0.086	0.09
Fort Carson, CO	1	-1.9028	0.4949	0.0001	0.149	0.17
Fort Hood, TX	1	-2.8548	1.021	0.0052	0.058	0.00

Fort Lee, VA	1	-2.0671	0.5003	<.0001	0.127	0.30
Camp Lejeune, NC	1	-1.9782	0.4597	<.0001	0.138	0.20
Camp Pendleton, CA	1	-1.7157	1.0424	0.0998	0.18	0.29
JBLM, WA	1	-1.8252	0.4547	<.0001	0.161	0.06
Great Lakes, IL	1	-1.2617	0.411	0.0021	0.283	0.13
Ehr	1	-1.7996	0.3403	<.0001	0.165	

Changes in OR > 10% → confounding by Ehr (Type III Analysis p-value <0.0001)

Appendix E-5

Multivariate Analyses: Odds of *B. burgdorferi* Sero-positivity by Installation Adjusting for Identified Risk Factors

Note: Analyses do not include Signonella due to "0" seropositivity.
ORs are as compared to Fort Drum NY, the installation with the highest prevalence

CRUDE (Full Model)

Outcome = *B. burgdorferi* Sero-positivity
Exposure of Interest = Installation (w/o Sigonella) as compared to Fort Drum NY (INST "9")
Potential confounders = tick wt dgseres2 dgwres2 dgneres2 dgcres5 pown wood wwalk ana ehr

Table A1: Full Model
Type 3 Analysis of Effects

Effect	DF	Wald Chi-Square	Pr > ChiSq
INST	13	16.277	0.2345
Tick	2	25.3572	<.0001
Wt	4	15.6877	0.0035
DgSERes2	1	0.5294	0.4668
DgWRes2	1	5.3494	0.0207
DgNERes2	1	5.8201	0.0158
DgCRes5	1	0.9059	0.3412
POwn	1	1.4229	0.2329
Wood	1	0.5342	0.4648
WWalk	1	5.5423	0.0186
Ana	1	30.4136	<.0001
Ehr	1	7.291	0.0069

Table A2: Full Model (CRUDE)				
Parameter	**DF**	**Maximum Likelihood Estimate**	**Standard Error**	**OR_{CRUDE}**
Intercept	1	2.7017	0.8716	
Kaiserslautern, DE	1	-1.0837	0.6509	0.338
Stuttgart. DE	1	-1.48	0.6975	0.228
Rota, SP	1	-0.6117	0.8291	0.542
Spangdahlem, DE	1	-1.3014	0.6761	0.272
Fort Belvoir, VA	1	-0.4477	0.4571	0.639
Fort Bragg, NC	1	-1.6854	0.5814	0.185
Fort Carson, CO	1	-0.4583	0.6615	0.632
Fort Hood, TX	1	-2.0933	1.1622	0.123
Fort Lee, VA	1	-1.3101	0.5831	0.27
Camp Lejeune, NC	1	-1.4616	0.5926	0.232
Camp Pendleton, CA	1	-0.1186	1.1619	0.888
JBLM, WA	1	-0.3334	0.6598	0.716
Great Lakes, IL	1	-0.3618	0.5132	0.696
History of Tick (Yes vs. No)	1	-1.0858	0.2673	0.338
History of Tick (Don't Know vs. No)	1	0.6557	0.3522	1.926
Dog Weight in pounds (<15 vs. 56 to 75)	1	-1.2224	0.3717	0.295
Dog Weight in pounds (16 to 35 vs. 56 to 75)	1	-0.9193	0.3271	0.399
Dog Weight in pounds (36 to 55 vs. 56 to 75)	1	-0.245	0.2971	0.783
Dog Weight in pounds (> 75 vs. 56 to 75)	1	-0.2031	0.2957	0.816
Dog lived in SE U.S. within last 2 yrs (No vs. Yes)	1	0.1935	0.2659	1.213
Dog lived in W U.S. within last 2yrs (No vs. Yes)	1	-1.1238	0.4859	0.325
Dog lived in NE U.S. within last 2yrs (No vs. Yes)	1	0.7893	0.3272	2.202
Dog lived in Central Europe within last 5 yrs (No vs. Yes)	1	-0.4719	0.4958	0.624
Prior Ownership (No vs. Yes)	1	-0.2617	0.2194	0.77
Wooded living environment (No vs. Yes)	1	-0.1887	0.2582	0.828
Recreates through Wilderness Walks (No vs. Yes)	1	-0.5463	0.2321	0.579

Co-infected with *Anaplasma* spp. (No vs. Yes)	1	-2.7521	0.499	0.064
Co-infected with *Ehrlichia* spp. (No vs. Yes)	1	-1.0414	0.3857	0.353

Reasonable coefficients and SEs
Rota, Spain and Camp Pendleton, CA may be questionable (n= 67 and n = 43, respectively)

REDUCED (without DgSERes2)

Outcome = *B. burgdorferi* Sero-positivity
Exposure of Interest = Installation (w/o Sigonella) as compared to Fort Drum NY (INST "9")
Potential confounders = tick wt dgwres2 dgneres2 dgcres5 pown wood wwalk ana her

Table B1: Reduced Model (without DgSERes2) Type 3 Analysis of Effects

Effect	DF	Wald Chi-Square	Pr > ChiSq
INST	13	19.7432	0.1018
Tick	2	25.1889	<.0001
Wt	4	15.7045	0.0034
DgWRes2	1	5.1036	0.0239
DgNERes2	1	6.7454	0.0094
DgCRes5	1	0.9728	0.324
POwn	1	1.4969	0.2212
Wood	1	0.5243	0.469
WWalk	1	5.4892	0.0191
Ana	1	30.6385	<.0001
Ehr	1	7.1242	0.0076

Table B2: Reduced Model (without DgSERes2)					
Parameter	**DF**	**Maximum Likelihood Estimate**	**Standard Error**	**OR_{ADJ}**	**Change in OR (Cr-Adj)/Adj**
Intercept	1	2.8073	0.8596		
Kaiserslautern, DE	1	-1.0427	0.6466	0.353	-0.04
Stuttgart. DE	1	-1.443	0.6953	0.236	-0.03
Rota, SP	1	-0.5456	0.8243	0.579	-0.06
Spangdahlem, DE	1	-1.2322	0.6682	0.292	-0.07
Fort Belvoir, VA	1	-0.4886	0.4544	0.614	0.04
Fort Bragg, NC	1	-1.7703	0.5688	0.17	0.09
Fort Carson, CO	1	-0.4447	0.6602	0.641	-0.01
Fort Hood, TX	1	-2.2047	1.1591	0.11	0.12
Fort Lee, VA	1	-1.3724	0.5768	0.253	0.07
Camp Lejeune, NC	1	-1.5378	0.5827	0.215	0.08
Camp Pendleton, CA	1	-0.0845	1.1586	0.919	-0.03
JBLM, WA	1	-0.3192	0.659	0.727	-0.02
Great Lakes, IL	1	-0.2902	0.5036	0.748	-0.07
History of Tick (Yes vs. No)	1	-1.0706	0.2659	0.343	-0.01
History of Tick (Don't Know vs. No)	1	0.6719	0.3512	1.958	-0.02
Dog Weight in pounds (<15 vs. 56 to 75)	1	-1.2297	0.3715	0.292	0.01
Dog Weight in pounds (16 to 35 vs. 56 to 75)	1	-0.9218	0.327	0.398	0.00
Dog Weight in pounds (36 to 55 vs. 56 to 75)	1	-0.2586	0.2964	0.772	0.01
Dog Weight in pounds (> 75 vs. 56 to 75)	1	-0.2147	0.2953	0.807	0.01
Dog lived in W U.S. within last 2yrs (No vs. Yes)	1	-1.0913	0.4831	0.336	2.61
Dog lived in NE U.S. within last 2yrs (No vs. Yes)	1	0.8331	0.3208	2.3	-0.86
Dog lived in Central Europe within last 5 yrs (No vs. Yes)	1	-0.4875	0.4943	0.614	2.59
Prior Ownership (No vs. Yes)	1	-0.2681	0.2192	0.765	-0.18
Wooded living environment (No vs. Yes)	1	-0.1869	0.2582	0.83	-0.07
Recreates through Wilderness Walks (No vs. Yes)	1	-0.5437	0.2321	0.581	0.43
Co-infected with *Anaplasma* spp. (No vs. Yes)	1	-2.7642	0.4994	0.063	8.19

Co-infected with *Ehrlichia* spp. (No vs. Yes)	1	2.8073	0.8596	0.358	-0.82

Changes in Institution ORs barely > 10% → NO confounding by DgSERes2; NOT return to model

REDUCED (without DgSERes2 Wood)

Outcome = *B. burgdorferi* Sero-positivity
Exposure of Interest = Installation (w/o Sigonella) as compared to Fort Drum NY (INST "9")
Potential confounders = tick wt dgwres2 dgneres2 dgcres5 pown wwalk ana ehr

Table C1: Reduced Model (without DgSERes2 and Wood) Type 3 Analysis of Effects

Effect	DF	Wald Chi-Square	Pr > ChiSq
INST	13	19.3367	0.113
Tick	2	25.9649	<.0001
Wt	4	15.6329	0.0036
DgWRes2	1	5.0496	0.0246
DgNERes2	1	6.8424	0.0089
DgCRes5	1	0.9801	0.3222
POwn	1	1.5588	0.2118
WWalk	1	6.8189	0.009
Ana	1	30.925	<.0001
Ehr	1	7.3502	0.0067

Table C2: Reduced Model (without DgSERes2 and Wood)					
Parameter	**DF**	**Maximum Likelihood Estimate**	**Standard Error**	**OR_{adj}**	**Change in OR (Cr-Adj)/Adj**
Intercept	1	2.6931	0.8427		
Kaiserslautern, DE	1	-1.015	0.6446	0.362	-0.02
Stuttgart. DE	1	-1.3768	0.6871	0.252	-0.06
Rota, SP	1	-0.5601	0.8231	0.571	0.01
Spangdahlem, DE	1	-1.2328	0.667	0.291	0.00
Fort Belvoir, VA	1	-0.4582	0.4516	0.632	-0.03
Fort Bragg, NC	1	-1.7432	0.5672	0.175	-0.03
Fort Carson, CO	1	-0.4749	0.6572	0.622	0.03
Fort Hood, TX	1	-2.2204	1.1599	0.109	0.01
Fort Lee, VA	1	-1.3368	0.5731	0.263	-0.04
Camp Lejeune, NC	1	-1.5169	0.581	0.219	-0.02
Camp Pendleton, CA	1	-0.1125	1.1586	0.894	0.03
JBLM, WA	1	-0.3111	0.6581	0.733	-0.01
Great Lakes, IL	1	-0.2962	0.5033	0.744	0.01
History of Tick (Yes vs. No)	1	-1.0843	0.2649	0.338	0.01
History of Tick (Don't Know vs. No)	1	0.6763	0.3512	1.967	0.00
Dog Weight in pounds (<15 vs. 56 to 75)	1	-1.2215	0.371	0.295	-0.01
Dog Weight in pounds (16 to 35 vs. 56 to 75)	1	-0.9184	0.3272	0.399	0.00
Dog Weight in pounds (36 to 55 vs. 56 to 75)	1	-0.2574	0.2964	0.773	0.00
Dog Weight in pounds (> 75 vs. 56 to 75)	1	-0.2	0.2943	0.819	-0.01
Dog lived in W U.S. within last 2yrs (No vs. Yes)	1	-1.0831	0.482	0.339	-0.01
Dog lived in NE U.S. within last 2yrs (No vs. Yes)	1	0.8377	0.3202	2.311	0.00
Dog lived in Central Europe within last 5 yrs (No vs. Yes)	1	-0.4878	0.4927	0.614	0.00
Prior Ownership (No vs. Yes)	1	-0.2734	0.2189	0.761	0.01
Recreates through Wilderness Walks (No vs. Yes)	1	-0.5855	0.2242	0.557	0.49
Co-infected with *Anaplasma* spp. (No vs. Yes)	1	-2.772	0.4985	0.063	8.22
Co-infected with *Ehrlichia* spp.	1	-1.0405	0.3838	0.353	-0.82

(No vs. Yes)					

No changes in OR > 10% → NO confounding by Wood; NOT return Wood to model

REDUCED (without DgSERes2, Wood, and DgCRes5)

Outcome = *B. burgdorferi* Sero-positivity

Exposure of Interest = Installation (w/o Sigonella) as compared to Fort Drum NY (INST "9")

Potential confounders = tick wt dgwres2 dgneres2 pown wwalk ana her

Table D1: Reduced Model (without DgSERes2 , Wood and DgCRes2) Type 3 Analysis of Effects

Effect	DF	Wald Chi-Square	Pr > ChiSq
INST	13	18.5773	0.1368
Tick	2	26.1419	<.0001
Wt	4	15.464	0.0038
DgWRes2	1	5.0239	0.025
DgNERes2	1	7.1922	0.0073
POwn	1	1.5905	0.2073
WWalk	1	6.7741	0.0092
Ana	1	31.6165	<.0001
Ehr	1	7.3162	0.0068

Table D2: Reduced Model (without DgSERes2, Wood, and DgCRes5)					
Parameter	**DF**	**Maximum Likelihood Estimate**	**Standard Error**	**OR_{ADJ}**	**Change in OR (Cr-Adj)/Ad**
Intercept	1	2.2525	0.7184		
Kaiserslautern, DE	1	-0.5641	0.4562	0.569	-0.36
Stuttgart. DE	1	-0.9226	0.5127	0.397	-0.37
Rota, SP	1	-0.5416	0.8213	0.582	-0.02
Spangdahlem, DE	1	-0.7757	0.4828	0.46	-0.37
Fort Belvoir, VA	1	-0.4625	0.4515	0.63	0.00
Fort Bragg, NC	1	-1.734	0.566	0.177	-0.01
Fort Carson, CO	1	-0.4498	0.6538	0.638	-0.03
Fort Hood, TX	1	-2.098	1.1194	0.123	-0.11
Fort Lee, VA	1	-1.3095	0.5716	0.27	-0.03
Camp Lejeune, NC	1	-1.5329	0.5806	0.216	0.01
Camp Pendleton, CA	1	-0.1293	1.1578	0.879	0.02
JBLM, WA	1	-0.3188	0.6581	0.727	0.01
Great Lakes, IL	1	-0.3027	0.503	0.739	0.01
History of Tick (Yes vs. No)	1	-1.0969	0.2647	0.334	0.01
History of Tick (Don't Know vs. No)	1	0.6608	0.3508	1.936	0.02
Dog Weight in pounds (<15 vs. 56 to 75)	1	-1.2156	0.3713	0.297	-0.01
Dog Weight in pounds (16 to 35 vs. 56 to 75)	1	-0.9107	0.327	0.402	-0.01
Dog Weight in pounds (36 to 55 vs. 56 to 75)	1	-0.2564	0.2961	0.774	0.00
Dog Weight in pounds (> 75 vs. 56 to 75)	1	-0.1933	0.2944	0.824	-0.01
Dog lived in W U.S. within last 2yrs (No vs. Yes)	1	-1.0769	0.4804	0.341	-0.01
Dog lived in NE U.S. within last 2yrs (No vs. Yes)	1	0.8558	0.3191	2.353	-0.02
Prior Ownership (No vs. Yes)	1	-0.2758	0.2187	0.759	-0.19
Recreates through Wilderness Walks (No vs. Yes)	1	-0.5835	0.2242	0.558	0.36
Co-infected with *Anaplasma* spp. (No vs. Yes)	1	-2.8078	0.4994	0.06	8.28
Co-infected with *Ehrlichia* spp. (No vs. Yes)	1	-1.0387	0.384	0.354	-0.82

Changes in OR > 10% → confounding by DgCRes2; keep in model

FINAL MODEL

REDUCED (without DgSERes2, Wood, and POwn)

Outcome = *B. burgdorferi* Sero-positivity

Exposure of Interest = Installation (w/o Sigonella) as compared to Fort Drum NY (INST "9")

Potential confounders = tick wt dgwres2 dgneres2 dgcres5 wwalk ana ehr

Table E1: Reduced Model (without DgSERes2, Wood, and POwn) Type 3 Analysis of Effects

Effect	DF	Wald Chi-Square	Pr > ChiSq
INST	13	19.5047	0.1083
Tick	2	29.0135	<.0001
Wt	4	16.2616	0.0027
DgWRes2	1	4.6877	0.0304
DgNERes2	1	7.0413	0.008
DgCRes5	1	1.3929	0.2379
WWalk	1	6.6774	0.0098
Ana	1	30.4467	<.0001
Ehr	1	8.4122	0.0037

Table E2: Reduced Model (without DgSERes2, Wood, and POwn)					
Parameter	**DF**	**Maximum Likelihood Estimate**	**Standard Error**	**OR_{ADJ}**	**Change in OR (Cr-Adj)/Ad**
Intercept	1	2.5443	0.8324		
Kaiserslautern, DE	1	-1.126	0.6306	0.324	0.12
Stuttgart. DE	1	-1.4511	0.6767	0.234	0.08
Rota, SP	1	-0.5627	0.8176	0.57	0.00
Spangdahlem, DE	1	-1.3397	0.6564	0.262	0.11
Fort Belvoir, VA	1	-0.4737	0.4504	0.623	0.01
Fort Bragg, NC	1	-1.7332	0.5668	0.177	-0.01
Fort Carson, CO	1	-0.5022	0.6535	0.605	0.03
Fort Hood, TX	1	-2.1577	1.142	0.116	-0.06
Fort Lee, VA	1	-1.3348	0.5738	0.263	0.00
Camp Lejeune, NC	1	-1.5183	0.5778	0.219	0.00
Camp Pendleton, CA	1	-0.1242	1.1605	0.883	0.01
JBLM, WA	1	-0.3223	0.6537	0.724	0.01
Great Lakes, IL	1	-0.2757	0.5028	0.759	-0.02
History of Tick (Yes vs. No)	1	-1.1017	0.2647	0.332	0.02
History of Tick (Don't Know vs. No)	1	0.7708	0.3451	2.161	-0.09
Dog Weight in pounds (<15 vs. 56 to 75)	1	-1.2311	0.3705	0.292	0.01
Dog Weight in pounds (16 to 35 vs. 56 to 75)	1	-0.946	0.3279	0.388	0.03
Dog Weight in pounds (36 to 55 vs. 56 to 75)	1	-0.2415	0.2956	0.785	-0.02
Dog Weight in pounds (> 75 vs. 56 to 75)	1	-0.201	0.2927	0.818	0.00
Dog lived in W U.S. within last 2yrs (No vs. Yes)	1	-1.0343	0.4777	0.355	-0.05
Dog lived in NE U.S. within last 2yrs (No vs. Yes)	1	0.8494	0.3201	2.338	-0.01
Dog lived in Central Europe within last 5 yrs (No vs. Yes)	1	-0.5645	0.4783	0.569	0.08
Recreates through Wilderness Walks (No vs. Yes)	1	-0.5779	0.2236	0.561	0.36
Co-infected with *Anaplasma* spp. (No vs. Yes)	1	-2.6813	0.4859	0.068	7.19
Co-infected with *Ehrlichia* spp. (No vs. Yes)	1	-1.1042	0.3807	0.331	-0.81

Changes in OR > barely 10%; NO confounding byPOwn; NOT return to model

FINAL MODEL

Table F1: Odds of *B. burgdorferi* Sero-positivity Controlling for Identified Risk Factors		
Parameter	**OR**	**95% CI**
Kaiserslautern, DE vs. Fort Drum NY	0.324	(0.094,1.116)
Stuttgart. DE vs. Fort Drum NY	0.234	(0.062, 0.883)
Rota, SP vs. Fort Drum NY	0.57	(0.115, 2.828)
Spangdahlem, DE vs. Fort Drum NY	0.262	(0.072, 0.948)
Fort Belvoir, VA vs. Fort Drum NY	0.623	(0.258, 1.505)
Fort Bragg, NC vs. Fort Drum NY	0.177	(0.058, 0.537)
Fort Carson, CO vs. Fort Drum NY	0.605	(0.168, 2.179)
Fort Hood, TX vs. Fort Drum NY	0.116	(0.012, 1.084)
Fort Lee, VA vs. Fort Drum NY	0.263	(0.085, 0.81)
Camp Lejeune, NC vs. Fort Drum NY	0.219	(0.071, 0.68)
Camp Pendleton, CA vs. Fort Drum NY	0.883	(0.091, 8.588)
JBLM, WA vs. Fort Drum NY	0.724	(0.201, 2.609)
Great Lakes, IL vs. Fort Drum NY	0.759	(0.283, 2.034)
History of Tick (Yes vs. No)	0.332	(0.198, 0.558)
History of Tick (Don't Know vs. No)	2.161	(1.099, 4.251)
Dog Weight in pounds (<15 vs. 56 to 75)	0.292	(0.141, 0.604)
Dog Weight in pounds (16 to 35 vs. 56 to 75)	0.388	(0.204, 0.738)
Dog Weight in pounds (36 to 55 vs. 56 to 75)	0.785	(0.44, 1.402)
Dog Weight in pounds (> 75 vs. 56 to 75)	0.818	(0.461, 1.452)
Dog lived in W U.S. within last 2yrs (No vs. Yes)	0.355	(0.139, 0.907)
Dog lived in NE U.S. within last 2yrs (No vs. Yes)	2.338	(1.249, 4.379)
Dog lived in Central Europe within last 5 yrs (No vs. Yes)	0.569	(0.223, 1.452)
Recreates through Wilderness Walks (No vs. Yes)	0.561	(0.362, 0.87)
Co-infected with *Anaplasma* spp.(No vs. Yes)	0.068	(0.026, 0.177)
Co-infected with *Ehrlichia* spp. (No vs. Yes)	0.331	(0.157, 0.699)

Chapter 6: Recommendations and Future Directions

Research Recommendations

Military Human Disease Data System Review

The military has several human disease data sources; therefore the first step in creating this combined report was to determine the one best suited for use in the USAPHC ZDR. In order to do so, a systematic review of five of the most commonly used human military data systems was conducted. The data systems reviewed include M2, DMSS, ESSENCE, DRSi, and HL7. The approach involved the following three phases: 1) detailed system descriptions, 2) a comparison of specific data systems attributes, and 3) an evaluation of each system's data quality.

A standardized framework based on CDC guidelines was used to compile data from each military data system (1). The resulting document provides a comprehensive comparison of the key attributes and capabilities of each system. This document can be used as a reference to assist military public health personnel in selecting the most appropriate data system to use for specific monitoring or surveillance activities. The knowledge acquired during this phase permitted the assessments of each data system's attributes during the next two phases.

Based on the specific goals and objectives of the ZDR, it is recommended that the USAPHC ZDR uses M2 as the primary source of military human medical data. This data system represents the most comprehensive and reliable source of both medical and population data; capturing not only the entire target population, but also having the ability to capture cases from military and civilian healthcare facilities. In addition, M2 can be directly queried by USAPHC personnel, which greatly improves the quality of the data and the timeliness of the information. The M2 does not automatically incorporate Reportable Medical Event or laboratory data, but because it retains unique identifiers, queries of M2 can be cross referenced and supplemented with information from sources that do (DRSi and HL7, respectively). These recommendations were presented to the USAPHC Disease Epidemiology Program in January of 2013; subsequently these sources were used to provide human disease data for the first publication of the ZDR in March of 2013.

An evaluation of data quality, through calculations of data source completeness, established M2 as the most complete source of military human medical data. Specifically, a comparison of Lyme disease case count data showed 55% of total captured cases came from the M2 system alone; in addition, the percent of cases missed by M2 was lower than the percent of cases missed by any other source. The investigation also revealed a substantial underlap between sources, indicating that reliance on a single data source could lead to significant underestimation of total cases. These findings again support the recommendation that USAPHC use M2 as the primary source of human disease data, supplemented with data from DRSi and HL7 as appropriate. This phase of the analysis involved the use of case count data resulting from data system queries; due to technical issues associated with producing the requested data, DMSS and ESSENCE were not included.

Military Pet Dogs as Sentinels in Lyme Disease Surveillance

Currently the animal disease information utilized in the ZDR is limited to that taken from the public access databases of intergovernmental organizations. Although there is value in this information, the current pilot study using canine *Borrelia burgdorferi* (Bb) seroprevalence showed the advantages of using animal disease data from military populations as sentinel surveillance for zoonotic disease military populations. The study involved three parts: 1) an investigation of the correlation between Bb seroprevalence in military pet dogs and military human Lyme disease data, 2) a comparison of military pet dog Bb seroprevalence data to published civilian pet dog data in order to determine if meaningful differences exist indicating the more appropriate population for use in sentinel surveillance, and 3) a determination of the validity of the military pet dog seroprevalence data by location.

Based on the correlation found between the animal Bb seroprevalence and human Lyme disease data, military pet dogs are recommended in sentinel surveillance in human military populations. The pilot study found that seroprevalence estimates in military pet dogs consistently reflect incidence rates in military human populations, both at areas of high and low expected prevalence. In addition, the prevalence estimates in the canine populations were consistently higher than the incidence rates seen in the human populations, demonstrating dogs as a more sensitive indicator of *B. burgdorferi* presence then their human counter parts. Finally, the pet dog data used in this study came

from routine in-house diagnostic testing, showing the use of animal data from military veterinary facilities as a convenient measure of the risk of Lyme disease in sentinel surveillance in human military populations.

The comparison of military pet dog seroprevalence data to that published on civilian pet dogs undoubtedly supports the recommendation of military animal data as a more appropriate indicator of human disease risk in military populations. Using disease data collected from military pet dog populations offers a more reliable way to acquire data from locations where civilian pet dog data may be limited or where barriers to accessing such data may exist. Also, the use of military pet dog data proved to be a more accurate marker of disease risk in military populations. In particular, the higher seroprevalence of military pet dogs in Illinois more accurately reflected the military incidence of Lyme disease in the state (at the Great Lakes Naval Air Station) than the civilian pet dog data did.

The military pet dog Bb seroprevalence estimates were unable to be confidently evaluated for confounding by the distribution of identified canine risk factors; therefore the use of these animals valid indicators of Lyme disease risk for military populations could not be definitively determined. In order to make this determination the study should be repeated for a longer period in order to meet the required sample size (many of the installations were underpowered). In addition, because of the low overall *B. burgdorferi* seroprevalence in this population (likely due to the high overall level of flea and tick prevention), the target sample size should be increased, or a different zoonotic disease should be selected.

Future Directions

The creation of the ZDR truly embodies the concept of One Health, combining health data from human, animal, and entomological sources in order to provide military public health officials and commanders with a powerful health risk assessment tool. The use of military human disease data systems permits sensitive and specific reflection of the zoonotic disease status of this population world-wide. Unfortunately, the same cannot be confidently said about the animal disease data contained in the report. Because military veterinary facilities have historically lacked a standardized electronic record keeping system capable of capturing relevant disease data from their own animal

populations, the animal disease data contained in the ZDR is currently limited to that available from public access databases.

The pilot study conducted in this dissertation demonstrates pet dogs belonging to Service members as a more accurate indicator of zoonotic disease than civilian pets living in the same locations. The current deployment of a web-based electronic medical record known as the Remote Online Veterinary Record (ROVR) has the potential to provide animal disease data from military veterinary facilities. It is strongly recommended that this animal disease data source becomes incorporated into the ZDR.

Incorporating ROVR in the ZDR

The combined review of existing public health surveillance systems and literature using animals as sentinels in zoonotic disease surveillance revealed several factors to consider before making risk assessments based on animal disease data. The following discussion is not all inclusive, rather it is intended to highlight a few of the factors that should be considered before incorporating military animal data into the ZDR. When possible, specific examples discussing ROVR's capabilities will be used.

In order to accurately characterize the risk of disease in humans it is important to understand the animal population represented in the data system. The animal population captured in ROVR may include pets, government owned animals, and sometimes strays; all of which have varying degrees of zoonotic diseases exposure and therefore may play different roles in risk assessment. For this reason it is important that information about the presence of zoonotic diseases in military animals includes details on from which animal population the data came.

Making appropriate risk comparisons between locations with unequal population sizes require data to be presented as ratios or proportions rather than counts. In order to do this requires knowledge of denominator data such as total population at risk or total population tested. Ideally, the data from ROVR would include not only the number of positive animals for the zoonotic disease but also the total number of animals at risk for the disease or the total number of animals tested for that particular disease.

To most accurately assess the risk of disease at a certain locality, the animal data should include details of prior travel and residency locations as well as preventive medicine practices (i.e. vaccination status, flea and tick preventative use). Some of this information is already captured in ROVR; specifically, the travel history data is captured under the animal transfer screen and vaccination histories and preventive medicine purchases are captured under the prevention screens. In addition, ROVR contains animal demographic data such as species, breed, sex, and age, all of which may play a role in disease risk in animal populations.

The review also showed the importance of standardized case definitions in ensuring consistency in data quality and accuracy in risk assessments across locations. Unfortunately, although the DoD has reportable medical event guidelines and case definitions for human disease, there is currently no veterinary equivalent (2). In the absence of these standards, in addition to the disease name, the animal disease data from ROVR should include details such as the method of diagnosis (laboratory test, physical exam findings, and clinical assessment) and the diagnostic category assigned (suspect, probable, confirmed).

A Zoonotic Disease Surveillance System Combining Human and Animal Disease Data

To this point, these recommendations have been focused on ways to refine disease data collection and analysis methods before their incorporation in disease reports, specifically the USAPHC ZDR. Although reports like these are a valuable public health tool, reliance on them or other forms of published data for disease risk assessments has inherent limitations. The following outlines some of these limitations, using the current ZDR as an example. The intent of this discussion is not to point out flaws of the report, but rather to provide concrete justification for the continued investment in methods for improved zoonotic disease surveillance practices in the military.

Due to the time involved in collecting, compiling, reviewing, and publishing disease reports, the data presented are generally somewhat dated and therefore have practical limitations in terms of implementing preventive strategies. Currently the ZDR is published on a quarterly basis, with additional data published on a supplemental basis as needed, issues with data timeliness limit its use to situational awareness only.

It is not possible to include all diseases in a publication; therefore reports are generally limited to conditions considered to be important by those publishing the report. The diseases reported in the ZDR are based on a stakeholder survey that was conducted in 2011. In order to ensure the report continues to contain the most reverent diseases for the target population, however, would require periodic stakeholder surveys to be conducted.

Published reports are limited in the amount of details they can contain; information such as patient demographics, clinical presentation, or pathogen subtype is generally not included. Although the ZDR attempts to discuss pertinent details on disease trends, the majority of the data are limited to summary case count data by location.

Based on these limitations, the best approach to conducting health risk assessments is a health surveillance system that permits users to access the health data themselves. The deployment of the web-based electronic veterinary record system provides the military with the opportunity to fully embrace the next phase One Health; merging data from both human and animal populations into one centrally accessible database. This would permit users to query the database to their own (or their commander's) specifications. Details of the query fields and filters would still need to be determined, but at a minimum it is recommended they include zoonotic disease name, location of diagnosis, species affected (to include human), and data of diagnosis. In addition, they system could incorporate features found to be useful in existing systems, such as the user-defined alter capability of ESSENCE.

It is recommended that USAPHC employ the technical expertise available within its command to compile a task force focused on designing a comprehensive zoonotic disease data system utilizing data from animal and human populations simultaneously. Ultimately, the goal should be also incorporate entomological, laboratory, and environmental data as well. In doing so the USAPHC has the unique opportunity to be the first organization employing this comprehensive approach to zoonotic disease surveillance; leading the way in fully embracing the concept of One Health.

References

1. Centers for Disease Control and Prevention. Updated guidelines for evaluating public health surveillance systems: recommendations from the guidelines working group. MMWR 2001;50 (No.RR-13) Atlanta, Georgia, Centers for Disease Control and Prevention.

2. Tri-Service Reportable Events Guidelines & Case Definitions, June 2009.