SYSTEMS ENGINEERING MODELS FOR SIGNATURE INJURIES OF MODERN MILITARY CONFLICTS

A Dissertation Presented

by

Hande Muşdal

to

The Department of Mechanical and Industrial Engineering

in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in the field of

Industrial Engineering

Northeastern University
Boston, Massachusetts

August 2013
Abstract

Modern military conflicts are producing dramatic increases in a new class of “silent” injuries, in part due to the changing manners by which war is waged and better protective equipment. Foremost among these are traumatic brain injury (TBI), posttraumatic stress disorder (PTSD), depression, and various mental health issues. The estimated prevalence of TBI is significantly higher in the military when compared to the general population, with the vast majority of soldiers being diagnosed with mild TBI which occurs with no outward signs of trauma. PTSD, on the other hand, is an increasingly important problem among U.S. service members and one of the signature injuries in the Iraq and Afghanistan wars. Since associated problems with these disorders often are cognitive, emotional, and behavioral, many cases go undetected and untreated indefinitely, linked with significant psychological disorders, long-term disabilities, and economic burdens.

This dissertation presents several systems engineering models to optimize the overall design, effectiveness, and capacity of healthcare systems for detecting and treating silent injuries, such as TBI and PTSD, as well as a general health problem that is common among veterans, sleep apnea, by addressing the following needs: (1) sequential screening processes, (2) categorical diagnostic methods, and (3) care services location-allocation (network optimization) models.
The first focus of this dissertation is analyzing and optimizing the design of disease screening processes. Several probability and Monte Carlo simulation models are developed to investigate the current and proposed PTSD screening processes within the Veterans Health Administration (VHA). Results indicate that a more systematically designed system, which consists of a series of annual screenings along with a standardized confirmatory testing, results in lower false diagnosis rates, predictable performance, and reduced costs. Additionally, a sequential screening process for mild TBI is proposed and illustrated in order to give an insight into how the general approach is applied to other disorders.

The second area of focus is developing multi-state categorical diagnostic models, which combine different assessment tools and consist of multiple screenings over time, in order to improve the diagnosis reliability and accuracy. Three types of predictive models – fuzzy logic, logistic regression, and neural networks – are described and used to determine whether or not an individual has TBI and to categorize him into the most likely severity state. A numerical example is given for illustration purposes and results indicate that these models can help reduce the number of unrecognized and misdiagnosed cases.

Finally, the third part of this dissertation illustrates the use of location-allocation models in order to improve access to care and patient satisfaction while minimizing overall system cost. Several single and multi-period integer programming models are developed and used across a range of specialty care services, namely, PTSD treatment and sleep apnea testing services, within the VHA. Results indicate significant opportunity to simultaneously reduce total cost, reduce total travel distances, and increase within-network access.
Acknowledgments

To my mom and dad for their unconditional love and endless belief in me...

The pages of this dissertation hold far more than the conclusion of years of study; they also reflect the relationships with many great people to whom I owe my deepest gratitude for being a part of this journey.

First and foremost, I would like to express my sincere appreciation to my adviser, Dr. James C. Benneyan, for the tremendous support and opportunities he provided throughout my doctoral studies. Without his continuous guidance, encouragement, patience, and knowledge, this dissertation would not have been possible. His vision, wisdom, and expertise have made him a constant oasis of ideas, which has exceptionally inspired and enriched my growth as a student and researcher. I consider myself amazingly fortunate to have had such a great mentor.

I would like to thank Dr. Brian Shiner and Dr. Thomas P. Cullinane for accepting to be a part of my dissertation committee. I truly appreciate their constructive comments, time, and support. Their invaluable contribution made this dissertation a complete work. I also extend my appreciation to Dr. Bradley V. Watts for his time and feedback.
A special appreciation goes to my project teammate and good friend, Seda Sinangil, whose contribution to a part of this dissertation is indisputable. Her hard work and colorful personality have been a great source of motivation and joy for me even during the most challenging times.

I am truly thankful to my ex-office mates but forever friends, Serpil Mutlu and Zeynep Karakuş, for their invaluable friendship and support; it was their company that made this journey enjoyable. I would also like to thank all my other ex-office mates for the joy they brought to my life. Each and every one of them has great places in my heart.

I would like to thank Zeynep Damla and Çağdaş Önal for being the perfect company during our long Friday night dinners and weekend trips, making Boston a home away from home, and being a family away from family. My sincere thanks also go to my dear friends Aysun Taşeli, Aysun Sünnetçi, and Mehmet Erkan Ceyhan for enriching my life in so many ways.

I am deeply grateful to my mother-in-law, Sefanur Öndemir, and my father-in-law, Mehmet Öndemir, for loving and supporting me as their daughter. I would like to thank Aykut Öndemir, Mine Öndemir, Volkan Öndemir, and Dilek Öndemir for welcoming me warmly into their family. I am also thankful to Gözde Öndemir, Nazlı Öndemir, and Duru Öndemir for making my life more delightful.

My greatest gratitude goes to my family for shaping me into who I am today, loving me unconditionally, and believing in me no matter what. I thank my mother, İnci Muşdal, for taking care of me even across the other side of the world. I thank my father, Yılmaz Muşdal, for supporting every single decision that I made. I thank my brother, Uygar Muşdal, for always
being there for me and cheering me up even in the darkest times. He is the other part of my soul, the one I have fought the most, and the one I have loved the most. I am also thankful to my aunt, Günseli Gürbulak, and her husband, Talat Gürbulak, for caring for me no less than their own children and to my lovely cousins, Sezin Gürbulak and Ulaş Gürbulak, for making my life more beautiful every single day.

Last but not least, I am wholeheartedly grateful to my husband, Önder Öndemir, for his endless patience, dedicated support, and unwavering love. Since our first day together, he has been my constant source of comfort, joy, and inspiration, the best friend I could ever ask for, and the reason to look forward to each new day. I cannot thank him enough for standing by me always, during my best and even my worst.

A part of this dissertation was funded by the Center for Health Organization Transformation (National Science Foundation grant IIP-10341990) and the VISN 1 Veterans Engineering Resource Center (grant VA 241-P-1772).
Table of Contents

Abstract .......................................................................................................................................... iii
Acknowledgments ........................................................................................................................... v
Table of Contents ......................................................................................................................... viii
List of Figures ............................................................................................................................... xii
List of Tables .................................................................................................................................. xvii
List of Abbreviations ................................................................................................................... xix
Chapter 1 - Introduction ................................................................................................................. 1
  1.1 Silent Injuries ...................................................................................................................... 1
    1.1.1 Traumatic Brain Injury .............................................................................................. 2
    1.1.2 Posttraumatic Stress Disorder .................................................................................. 6
  1.2 Sleep Apnea ....................................................................................................................... 10
  1.3 Motivation for Work ......................................................................................................... 14
Chapter 2 - Sequential Screening Models .................................................................................. 17
  2.1 Background ....................................................................................................................... 17
  2.2 PTSD Screening Process within the VHA ....................................................................... 19
  2.3 Methodology .................................................................................................................... 23
    2.3.1 Notation and Assumptions ....................................................................................... 23
2.3.2 Overall System Sensitivity ........................................................................................ 28
  2.3.2.1 Probabilities of Detecting Truly Positive Patients and of a False-Negative .............. 28
  2.3.2.2 Expected Number and Distribution of False-Negatives ........................................... 30

2.3.3 Overall System Specificity ........................................................................................ 32
  2.3.3.1 Probabilities of Detecting Truly Negative Patients and of a False-Positive ............ 32
  2.3.3.2 Expected Number and Distribution of False-Positives ........................................... 34

2.3.4 Expected Cost ............................................................................................................ 35

2.4 Scenario Analysis .............................................................................................................. 38
  2.4.1 Effect of Scoring Methods on Overall System Accuracy, Cost, and Workload ........ 38
  2.4.2 Effect of False Diagnosis Costs on the Total Cost of Screening Protocols .......... 49
  2.4.3 Effect of Compliance Rates on Overall System Accuracy and Cost ......................... 49
  2.4.4 Effect of PTSD Prevalence on the Predictive Value of Screening Protocols ............ 52

2.5 Other Possible Applications: A Sequential Screening Model for Mild TBI ............... 53

2.6 Discussion ......................................................................................................................... 59

Chapter 3 - Multi-State Categorical Diagnostic Models ....................................................... 61

3.1 Background ....................................................................................................................... 61

3.2 Fuzzy Categorical Model .................................................................................................. 64
  3.2.1 Methodology .............................................................................................................. 64
  3.2.2 Numerical Example ................................................................................................... 68
  3.2.3 Model Accuracy ......................................................................................................... 71

3.3 Multinomial Logistic Regression Categorical Model ....................................................... 73
  3.3.1 Methodology .............................................................................................................. 73
  3.3.2 Numerical Example ................................................................................................... 75
  3.3.3 Model Accuracy ......................................................................................................... 76
3.4 Artificial Neural Network Model...................................................................................... 78

3.4.1 Methodology.............................................................................................................. 78
3.4.2 Numerical Example ................................................................................................. 81
3.4.3 Model Accuracy ....................................................................................................... 81

3.5 Model Comparison and Optimization........................................................................... 82
3.6 Discussion ......................................................................................................................... 84

Chapter 4 - Network Optimization Models ........................................................................ 86

4.1 Background....................................................................................................................... 86

4.2 Optimal Location and Use of In-Person and Video-Based PTSD Treatment Services
within the VHA.................................................................................................................. 90

4.2.1 VA New England PTSD Treatment Services ............................................................ 91
4.2.2 Methodology ............................................................................................................. 95

4.2.2.1 Model Overview and Assumptions .............................................................................. 95
4.2.2.2 Integer Programming Model ...................................................................................... 97
4.2.3 Application and Results .......................................................................................... 99

4.2.3.1 Application Details ................................................................................................... 99
4.2.3.2 Results .................................................................................................................... 99
4.2.4 Discussion ................................................................................................................. 104

4.3 Single and Multi-Period Location-Allocation Models for Sleep Apnea Testing Services
within the VHA.................................................................................................................. 106

4.3.1 VA New England Sleep Apnea Services ................................................................. 106
4.3.2 Methodology ............................................................................................................ 112

4.3.2.1 Model Overview and Assumptions ............................................................................ 112
4.3.2.2 Single-Period Model ............................................................................................... 113
4.3.2.3 Multi-Period Model ............................................................................................... 116
List of Figures

Figure 2-1 Current PTSD screening process in the VA................................................................. 21

Figure 2-2 Proposed PTSD screening process in the VA........................................................... 22

Figure 2-3 Probability distribution of cost of (a) a false-positive and (b) a false-negative result 26

Figure 2-4 Probability distributions of number of false diagnoses per 10,000 patients in the first year ........................................................................................................................................ 31

Figure 2-5 Probability distribution of total cost per 10,000 patients (obtained via simulation) 37

Figure 2-6 Comparison of past, current, and proposed screening protocols for PTSD identification in the VA ................................................................................................................................. 40

Figure 2-7 Effect of scoring method on overall (a) sensitivity and (b) specificity over a 20-year screening horizon .................................................................................................................. 41

Figure 2-8 Effect of scoring method on expected number of (a) false-negatives and (b) false-positives over a 20-year screening horizon ...................................................................................... 42

Figure 2-9 Effect of scoring method on expected (a) screening cost and (b) total cost over a 20-year screening horizon ........................................................................................................... 43
Figure 2-10 Decomposition of expected (a) total cost and (b) amount of screening tests assuming protocol 1 ........................................................... 44

Figure 2-11 Tradeoff between stopping criteria and expected (a) number of false diagnoses, and (b) cost at the end of screening horizon \((j = 20)\) assuming protocol 1 ........................................ 45

Figure 2-12 Effect of discontinued screening after three negative results on expected (a) number of false-negatives and (b) number of false-positives assuming protocol 1 ......................... 46

Figure 2-13 Effect of discontinued screening after three negative results on expected (a) screening cost and (b) total cost assuming protocol 1 ............................................................. 47

Figure 2-14 Effect of false-negative/false-positive cost ratio on total cost assuming protocol 1 with various stopping criteria ........................................................................................... 49

Figure 2-15 Effect of compliance rates on expected (a) number of false-negatives and (b) number of false-positives assuming protocol 1 with discontinued screening after three negative results ........................................................................................................... 51

Figure 2-16 Effect of compliance rates on expected (a) screening cost and (b) total cost assuming protocol 1 with discontinued screening after three negative results ......................... 52

Figure 2-17 Effect of PTSD prevalence on predictive value of protocol 1 with discontinued screening after three negative results ........................................................................................................... 53

Figure 2-18 A sequential screening process for mild TBI ................................................................. 55

Figure 2-19 Competing effect of number of screening tests on overall sensitivity versus specificity, and the role of convergence to bounds, where (a) both system bounds = 0.95 and (b) both system bounds = 0.80 .................................................................................................. 57
Figure 4-7 Sleep apnea care demand by state across VISN 1 (January 2009 - May 2010)........ 107

Figure 4-8 Estimated sleep apnea care demand by 3-digit zip code and all VA facilities across
VISN 1 (January 2009 - May 2010).................................................................................... 108

Figure 4-9 Fee visits for West Haven, (a) number of fee visits versus monthly utilization, (b)
weekly fee visits versus in-house visits............................................................................ 110

Figure 4-10 Utilization of sleep apnea capacity by medical center over time......................... 111

Figure 4-11 Travel distance distribution, VISN 1 (January 2009 - May 2010)....................... 111

Figure 4-12 Total projected sleep apnea demand by year for entire VISN 1 (New England).... 118

Figure 4-13 Optimized tradeoff between acceptable distance and (a) total cost, and (b) coverage
percentage for single-period model assuming current observed demands and case 3 ....... 121

Figure 4-14 Optimized tradeoff between acceptable distance and (a) total cost, and (b) coverage
percentage for single-period model assuming estimated true demand and candidate facilities
under case 3 ......................................................................................................................... 123

Figure 4-15 Demand nodes covered within-network and outside network and VA facilities that
provide service for single-period model assuming a 60-mile acceptable maximum distance
and candidate facilities under case 3 ................................................................................... 126

Figure 4-16 Optimized tradeoff between acceptable distance and (a) average annual cost, and (b)
coverage percentage for multi-period model assuming case 3........................................ 128

Figure 4-17 Optimal (a) annual cost and (b) coverage percentages considering different planning
horizons (both for case 3 and \( P = 3 \))......................................................... 131

Figure 4-18 Pseudo code for the simulation models with demand variability ...................... 132
Figure 4-19 Changes in average (a) total cost, (b) number of unanticipated fee patients, (c) house cost, and (d) fee cost assuming variable annual demands and single period model results for case 3 and $P = 3$ ........................................................................................................................................ 134

Figure 4-20 Changes in average (a) total cost, (b) number of unanticipated fee patients, (c) house cost, and (d) fee cost assuming variable monthly demands and single period model results for case 3 and $P = 3$ ........................................................................................................................................ 135
List of Tables

Table 2-1 Main characteristics of the PC-PTSD, PCL, and CAPS screening tests ...................... 22

Table 2-2 Minimum, average, and maximum cost estimates for a false-positive and false-
negative result under the assumption of different cases ........................................................ 26

Table 2-3 Summary of analytic distributions of cost, workload, and accuracy with comparison to
simulation results ................................................................................................................... 38

Table 2-4 Overall sensitivity, specificity, cost, and workload of alternate screening protocols at
important intervals (with discontinued screening after three negative results) ..................... 48

Table 2-5 Scenarios for patient compliance rates ....................................................................... 50

Table 2-6 Mathematical expressions of overall sensitivity, specificity, and cost for mild TBI
screening model ..................................................................................................................... 56

Table 3-1 Example of a rule-base for \( N = 2 \) screening test case .............................................. 68

Table 3-2 Rule-base for \( N = 4 \) screening test example ............................................................. 70

Table 3-3 Changes in diagnosis over time and according to defuzzification method ............. 71

Table 3-4 Illustrative portion of evaluation dataset ................................................................. 72

Table 3-5 Category probabilities for MLR approach ............................................................... 77
Table 3-6 Changes in final diagnosis over time based on the method used ........................................ 82
Table 3-7 Overall comparison of model performances ....................................................................... 83
Table 4-1 Representation of basic discrete facility location models .................................................. 89
Table 4-2 Optimal allocation of patient zip codes to VISN 1 facilities for a 30-mile maximum acceptable distance assuming estimated care demand in 2010 ........................................ 102
Table 4-3 Number of sleep beds, within-network and outside-network sleep apnea patients in VISN 1 (January 2009 - May 2010) ..................................................................................... 109
Table 4-4 Cost parameters used in mathematical models .................................................................. 119
Table 4-5 Multi-period model solution times based on different scenarios for case 1, 2, and 3 120
Table 4-6 Results for single-period model with 60-mile maximum acceptable travel distance ........................................ 122
Table 4-7 Optimal facility locations for single-period model based on estimated true demand 124
Table 4-8 Allocation of demand nodes to facilities for single-period model ................................. 125
Table 4-9 Results for multi-period model with 60-mile maximum acceptable travel distance . 127
Table 4-10 Optimal facility locations for multi-period model .......................................................... 129
Table 4-11 Allocation of demand nodes to facilities for multi-period model ................................. 130
# List of Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>ANN</td>
<td>Artificial Neural Networks</td>
</tr>
<tr>
<td>CAPS</td>
<td>Clinician-Administered PTSD scale</td>
</tr>
<tr>
<td>DoD</td>
<td>Department of Defense</td>
</tr>
<tr>
<td>FL</td>
<td>Fuzzy Logic</td>
</tr>
<tr>
<td>MLR</td>
<td>Multinomial Logistic Regression</td>
</tr>
<tr>
<td>OEF/OIF</td>
<td>Operations Enduring Freedom and Iraqi Freedom</td>
</tr>
<tr>
<td>PCL</td>
<td>PTSD Clinician Checklist</td>
</tr>
<tr>
<td>PC-PTSD</td>
<td>Primary Care PTSD Screen</td>
</tr>
<tr>
<td>PTSD</td>
<td>Posttraumatic Stress Disorder</td>
</tr>
<tr>
<td>TBI</td>
<td>Traumatic Brain Injury</td>
</tr>
<tr>
<td>VA</td>
<td>Veterans Affairs</td>
</tr>
<tr>
<td>VHA</td>
<td>Veterans Health Administration</td>
</tr>
<tr>
<td>VISN</td>
<td>Veterans Integrated Service Network</td>
</tr>
</tbody>
</table>
Chapter 1 - Introduction

This dissertation focuses on developing systems engineering models to optimize the overall design, effectiveness, and capacity of large integrated healthcare systems for detecting and treating two of most common military “silent” injuries, traumatic brain injury (TBI) and posttraumatic stress disorder (PTSD), as well as a general health problem that is common among veterans, sleep apnea. In this chapter, general information about the extended area of interest, motivation behind this research, and outline of the dissertation are provided.

1.1 Silent Injuries

More than 2 million U.S. service members have been deployed since October 2001 for Operations Enduring and Iraqi Freedom (OEF/OIF) in Afghanistan and Iraq, with many exposed to prolonged periods of deployment-related stress and traumatic events (Tanielian and Jaycox, 2008; Congressional Budget Office, 2012). Fortunately, advances in protective equipment and battlefield care have reduced the mortality rates that are historically lower than in earlier prolonged conflicts, such as Vietnam and Korea (Regan, 2004; Warden, 2006). However, there have been significant increases in a different kind of casualties – “silent” injuries such as mental health conditions and cognitive impairments resulting from deployment experiences. Two of
these conditions that affect military service members during deployment to a combat theater and afterward are traumatic brain injury (TBI) and posttraumatic stress disorder (PTSD).

Both TBI and PTSD are referred as the signature wounds of Iraq and Afghanistan conflicts, with disproportionately more casualties when compared to the physical injuries of combat. TBI is caused by sudden trauma to the head and is commonly sustained by soldiers exposed to explosions. PTSD, on the other hand, is an anxiety disorder induced by exposure to a traumatic event, such as witnessing an injury or death. Unlike physical wounds of war, these conditions often are invisible to the eye, remaining invisible to other service members, family members, and society in general. These conditions affect mood, thoughts, and behavior and may go unrecognized and unacknowledged indefinitely, resulting in serious medical and social consequences (Tanielian and Jaycox, 2008; Congressional Budget Office, 2012).

1.1.1 Traumatic Brain Injury

A traumatic brain injury (TBI) is any non-penetrating or penetrating structural injury or physiological disruption to the brain due to an external force (Langlois et al., 2006; Butler et al., 2008). Falls (35.2%), motor vehicle and traffic accidents (17.3%), impacts (16.5%), and assaults (10%) are the most common causes of civilian TBI. Among civilians, an estimated 1.7 million people per year sustain a TBI, of which approximately 52,000 die, 275,000 are hospitalized, and 1.365 million, which is nearly 80%, are treated by an emergency department (Faul et al., 2010). Military personnel and others working in combat zones also are at particular risk of TBI due to encounters with explosions, rocket-propelled grenades, improvised explosive devices, and land mines (Champion et al., 2009; Ling et al., 2009; Defense and Veterans Brain Injury Center, 2013). Since associated problems, such as related to thinking and memory, often are not visible
and awareness among the general public is limited, TBI frequently is referred to as the “silent
epidemic” (Martin et al., 2008; McCrea, 2008; Vaishnavi et al., 2009; Faul et al., 2010).

TBI is identified as one of the signature injuries among U.S. troops serving in the Iraq and
Afghanistan wars, accounting for a larger proportion of casualties than in previous wars.
Improved protective equipment has reduced the incidence of penetrating head injuries and death,
but there has been a significant increase in the incidence of closed-head TBI. Some estimates of
military TBI are as high as 22%, with an estimated 50% of injured combat soldiers suffering
some form of TBI from Operations Enduring and Iraqi Freedom (OEF/OIF) (Okie, 2005;
Warden, 2006; Jackson et al., 2008; Tanielian and Jaycox, 2008; Wojcik et al., 2010; Smith et
al., 2013). As of May 2013, a total of 273,859 soldiers suffering from TBI were reported by the
Defense and Veterans Brain Injury Center (i.e., medical diagnoses of TBI among U.S. forces
since 2000), with 58% sustaining their injury while in the Army, 15% while in the Marines, 14% 
while in the Navy, and 14% while in the Air Force (Defense and Veterans Brain Injury Center,
2013).

The severity of a TBI can be mild, moderate, or severe depending on the extent of the damage to
the brain. A person with a mild TBI may remain conscious or may experience a loss of
consciousness for a few seconds or minutes. Other symptoms of mild TBI include headache,
mental confusion, lightheadedness, dizziness, double and blurred vision or tired eyes, ringing in
the ears, bad or metallic taste in the mouth, fatigue or lethargy, a change in sleep patterns,
behavioral or mood changes, decreased coordination, and trouble with memory, concentration,
calculation, attention, or thinking. Worsening symptoms indicate a more severe injury, including
a more severe and permanent headache, repeated vomiting or nausea, personality change,
convulsions or seizures, an inability to awaken from sleep, dilation of one or both pupils of the eyes, slurred speech, weakness or numbness in the extremities, loss of coordination, and increased confusion, restlessness, or agitation (National Institute of Neurological Disorders and Stroke, 2013).

Most diagnosed TBIs (82%) are categorized as mild, while moderate (8%), severe (1%), or penetrating wounds (2%) are less common (Defense and Veterans Brain Injury Center, 2013). The diagnosis of TBI is relatively straightforward when symptoms are more visible, but mild TBIs are difficult to detect, both by those afflicted and by healthcare professionals. Moreover, a caregiver or a screening mechanism may attribute TBI symptoms to other medical or cognitive conditions (Heegaard and Biros, 2007; Anderson-Barnes et al., 2010). In combat zones, TBI often occurs in combination with more obviously life-threatening injuries; therefore, cases may go unrecognized, potentially putting the individuals or their units at risk (Butler et al., 2008; Tanielian and Jaycox, 2008). Additionally, when TBI occurs with no obvious symptoms, service members might not seek medical treatment (Martin et al., 2008; McCrea, 2008). Many TBI sufferers, especially if untreated, may endure medical, behavioral, and social consequences for many years, even a lifetime. Any delays in treatment may compromise recovery or result in significant cognitive, physical, and/or psychological impairment (Rao and Lyketsos, 2000; Maas et al., 2008; Anderson-Barnes et al., 2010; Faul et al., 2010). Consequently, soldiers who are exposed to blasts should be screened for TBI immediately following the event in order to minimize medical complications.

The existence and severity of TBI are evaluated in several ways: subjective assessments, e.g., questions about dizziness; external signs, e.g., lacerations, fractures, fluids; physical experiences,
e.g., duration of amnesia or unconsciousness; neurologic assessments, e.g., the Glasgow Coma Scores (GCS), Military Acute Concussion Evaluation (MACE), the Automated Neuropsychological Assessment Metrics (ANAM); and finally, clinical assessments, e.g., computed tomography (CT) scans, magnetic resonance imaging (MRI), and electroencephalogram (EEG) (U.S. Department of Health and Human Services, 2006; Elder et al., 2010; Maruta et al., 2010). In the military, testing typically follows a process of screening, identification, evaluation, confirmation, and stratification by severity (Butler et al., 2008).

Proper treatment of mild TBI requires early identification, monitoring and management of symptoms, and adequate rest. A person with signs of moderate or severe TBI, on the other hand, should receive medical attention as soon as possible. Since little can be done to reverse the initial brain damage caused by trauma, clinical care is first directed toward controlling secondary effects of TBI, such as brain swelling, bleeding in the head, and poor cerebral blood flow, that can lead to further neuronal damage. Primary concerns include insuring proper oxygen supply to the brain and the rest of the body, maintaining adequate blood flow, and controlling blood pressure. Approximately half of severely head-injured patients need surgery to remove or repair ruptured blood vessels or bruised brain tissue. Following acute critical care, these patients receive rehabilitation that involves individually tailored treatment programs in the areas of physical therapy, occupational therapy, speech/language therapy, physical medicine, psychology/psychiatry, and social support (Butler et al., 2008; National Institute of Neurological Disorders and Stroke, 2013).

Disabilities resulting from a TBI depend upon the severity of the injury, the location of the injury, and the age and general health of the patient. Some common disabilities include problems...
with cognition (thinking, memory, and reasoning), sensory processing (sight, hearing, touch,
taste, and smell), communication (expression and understanding), and behavior or mental health
(depression, anxiety, personality changes, aggression, acting out, and social inappropriateness).
More serious head injuries may result in stupor, an unconscious, unresponsive, unaware, and/or
unarousable state, and a vegetative or persistent vegetative state (National Institute of
Neurological Disorders and Stroke, 2013).

TBI incidence, screening processes, risk factors, treatment and care services, and consequences
have been studied extensively in the literature, including these by Nathanson et al. (2003),
Javouhey et al. (2006), Kaufman et al. (2006), Langlois et al. (2006), Côté et al. (2007), Guler et
al. (2008), Hoge et al. (2008), Jackson et al. (2008), Wu et al. (2008), Guler et al. (2009),
Elgmark Andersson et al. (2010), Faul et al. (2010), Rosenfeld and Ford (2010), Syam and Côté
(2010), Liao et al. (2012), and Asemota et al. (2013).

1.1.2 Posttraumatic Stress Disorder

Posttraumatic stress disorder (PTSD) is a severe and often disabling mental health disorder
resulting from traumatic events such as assaults, wars, disasters, terrorist attacks, and accidents
(American Psychiatric Association, 2000; Hoge et al., 2006; Milliken et al., 2007; Shalev, 2009).
PTSD is common among Americans with a lifetime prevalence of 3.7% in the general
population, yet only 7% of those affected seek treatment in their first year, with a lifetime
treatment-seeking rate of just over 50% (Kessler et al., 2005; Wang et al., 2005). In the military,
the prevalence is especially higher – up to 20% among Iraq and Afghanistan veterans – given
that exposure to combat increases the risk of PTSD (Prigerson et al., 2001; Hoge et al., 2004;
Ramchand et al., 2010).
PTSD can occur in anyone who has experienced a traumatic event, including war veterans and survivors of physical and sexual assaults, abuse, accidents, disasters, and many other serious events. However, not everyone who has experienced a traumatic event eventually develop PTSD. In fact, people who have not been through such traumas also can suffer from PTSD. Although it is difficult to predict whether or not a person will develop PTSD, a number of factors that increase the risk have been identified. These include having experienced dangerous events and traumas, having a history of mental illness and/or a family history of mental illness, the severity of the trauma experienced, and having little or no social support after the trauma (Kessler et al., 1995; Brewin et al., 2000; U.S. Department of Health and Human Services, 2013). Additionally, women are at increased risk of PTSD, with a lifetime prevalence of more than twice as high than men, mostly because they are more likely to experience the kinds of trauma that can trigger the condition such as sexual assaults (Halligan and Yehuda, 2000; Freedman et al., 2002; National Comorbidity Survey-Replication, 2007; Shalev, 2009).

PTSD symptoms usually develop shortly after the traumatic event, but may not appear until months or years have passed. In many survivors the symptoms subside, while they persist in others in the form of chronic PTSD (Shalev, 2009). These symptoms can be grouped into three categories, namely, re-experiencing, avoidance, and hyperarousal. Re-experiencing symptoms include recurrent images, thoughts, flashbacks, or dreams about the traumatic event and may cause problems in a person’s everyday routine. Avoidance is defined as making an effort to avoid thoughts, feelings, conversations, places, or people that are reminders of the traumatic event, feeling emotionally numb, losing interest in once-enjoyable activities, and having trouble remembering the important details of the traumatic event. Finally, hyperarousal includes symptoms of anxiety or heightened awareness of danger such as sleeplessness, irritability, being

PTSD is diagnosed based on signs and symptoms and a thorough psychological evaluation. Over the past several decades, considerable progress has been made in the development and empirical evaluation of assessment instruments, such as questionnaires, structured interviews, and psychophysiological procedures, for measuring PTSD (Foa and Yadin, 2011). The most commonly used PTSD assessment tools are Primary Care PTSD Screen (PC-PTSD), PTSD Clinician Checklist (PCL), and Clinician Administered PTSD Scale (CAPS). PC-PTSD is a short and simple screening questionnaire designed to detect the PTSD diagnosis in busy primary care clinics. It includes four yes/no questions which focus on capturing PTSD symptom clusters. The result of the PC-PTSD screening test is determined according to the number of “yes” answers (i.e., cut-off score). To date, PC-PTSD and its diagnostic qualities have been exclusively studied, indicating high sensitivity and specificity values for cut-off scores of 2 and 3 (Prins et al., 2003; Seal et al., 2008; van Dam et al., 2010). PCL, on the other hand, is a 17-item self-report measure of PTSD symptomatology, where respondents are asked to rate how much they have been bothered by each symptom in the past month on a scale ranging from 1 (not at all) to 5 (extremely). Different scoring procedures are used to yield either a continuous measure of PTSD symptom severity or a dichotomous (present/absent) indicator of diagnostic status. Dichotomous scoring methods include either an overall cut-off score or a symptom cluster scoring approach (American Psychiatric Association, 2000; Forbes et al., 2001; Norris and Hamblen, 2004; Keen et al., 2008). A commonly used cut-off score is 50, which is suggested as the optimally efficient cut-off score by Weathers et al. (1993). Finally, CAPS is widely considered the gold standard in PTSD assessment and was designed to overcome the limitations of other available PTSD
interviews. It is a 30-item structured interview, which requires clinicians to rate each PTSD symptom’s frequency and intensity using a scale ranging from 0 to 4. The most frequently used scoring rule is to record a symptom as present if it has a frequency of 1 or more and an intensity of 2 or more (Blake et al., 1990; Blake et al., 1995; Weathers et al., 1999; Weathers et al., 2001; Gray et al., 2004).

There are several treatments that were found to be effective with PTSD, including psychotherapy (talk therapy), medications, or both. Psychotherapy involves talking with a mental health professional either one-on-one or in a group. Talk therapy treatment for PTSD usually lasts 6 to 12 weeks, but can take more time. Many types of psychotherapy can help PTSD patients, some of which target PTSD symptoms directly while others focus on social, family, or job-related problems. Different therapies may be combined by therapists depending on each patient’s needs. One of the helpful therapies is called cognitive behavioral therapy (CBT) which consists of several parts, including exposure therapy, cognitive restructuring, and stress inoculation training. Medications, on the other hand, help control PTSD symptoms such as sadness, worry, anger, and feeling numb inside and also may make it easier to go through psychotherapy. The most commonly prescribed medications for treating PTSD patients are sertraline and paroxetine, both of which are antidepressants and also used to treat depression (U.S. Department of Health and Human Services, 2013; VA National Center for PTSD, 2013).

Most of the problems associated with PTSD can be addressed by effective treatments (Foa et al., 2008; U.S. Department of Veterans Affairs and Department of Defense, 2010). However, many individuals with PTSD do not seek mental health care for various reasons – the stigma associated with having a mental health problem, for example, or the inconvenience of undergoing additional
evaluation and treatment (Congressional Budget Office, 2012). Instead, many of them use primary care services, where PTSD may be underdetected or undertreated (Schnurr et al., 2013). If remain untreated or undertreated, PTSD results in significant morbidity and long-term disability (Schnurr et al., 2006; Hermann et al., 2012). Moreover, people with PTSD are at very high risk to develop a number of other mental health disorders (Orsillo et al., 1996; Keane and Kaloupek, 1997; Shalev et al., 1998), the most common ones being anxiety disorders (73.3%), depression (68.6%), and substance use disorders (10.5%) (Magruder et al., 2005). The co-occurrence of PTSD and another disorder may have a negative impact on the treatment of PTSD and compromise recovery.

There are numerous studies in the literature that focus on PTSD prevalence (Kessler et al., 1995; Hoge et al., 2004; Magruder et al., 2005; Ramchand et al., 2010; Shiner et al., 2012), risk factors (Brewin et al., 2000; Halligan and Yehuda, 2000; Freedman et al., 2002), symptoms (Bisson, 2007), consequences (Murdoch et al., 2005; Smith et al., 2005; Schnurr and Lunney, 2011), screening processes (Forbes et al., 2001; Prins et al., 2003; Gray et al., 2004), and treatment services (Rosenheck and Fontana, 2007; Hermes et al., 2012; Shiner et al., 2012; Schnurr et al., 2013).

1.2 Sleep Apnea

Sleep apnea is a type of sleep disorder characterized by pauses in breathing (called “apneas”) or instances of shallow or infrequent breathing (called “hypopneas”) during sleep. Each pause in breathing may last from at least ten seconds to minutes and occur 5 to 30 times or more an hour (Ancoli-Isreal and Avalon, 2006; Erman, 2006; U.S. Department of Health and Human Services,
There are two types of sleep apnea, namely, obstructive and central sleep apnea. In obstructive sleep apnea, breathing is abnormal because of narrowing or closure of the throat, while in central sleep apnea, it is because of a change in the breathing control and rhythm (Schmidt-Nowara, 2012). Obstructive sleep apnea is more common among adults with the estimates of prevalence being in the range of 3% to 7% in the general population (Morgenthaler et al., 2006; Punjabi, 2008). Military veterans are at particularly high risk of sleep disorders due to hazardous working conditions, inconsistent work hours, harsh environments, routine exposure to loud noises, and crowded sleeping spaces, with the most frequent diagnosis being obstructive sleep apnea (Seelig et al., 2010; Mysliwiec et al., 2013). Also, recent studies suggest that the increased incidence of sleep disorders among military personnel is potentially related to PTSD, depression, anxiety, or mild TBI (Sharafkhaneh et al., 2005; Lewis et al., 2009; Amin et al., 2010; McLay et al., 2010).

Sleep apnea is a serious condition that affects a person’s ability to safely perform normal daily activities as well as his/her long term health and has been linked with high blood pressure, heart problems, metabolic syndrome, diabetes, stroke, fatigue, headaches, snoring, daytime drowsiness, and memory problems (Ancoli-Isreal and Avalon, 2006; Erman, 2006). Although anyone can develop sleep apnea, there are certain factors that increase the risk, including increasing age, male sex, obesity, family history, menopause, craniofacial abnormalities, and certain health behaviors such as cigarette smoking or alcohol use (Young et al., 2004; Punjabi, 2008; Schmidt-Nowara, 2012; U.S. Department of Health and Human Services, 2012).

The main symptoms of sleep apnea are loud snoring, fatigue, and daytime sleepiness. However, some people have no symptoms or may not be aware of the symptoms such as snoring. Fatigue
and sleepiness, on the other hand, often are attributed to other reasons such as overwork and increasing age. As a result, some people may not recognize that they have a problem and therefore do not seek medical care. Other symptoms may include restless sleep, awakening with choking, gasping, or smothering, morning headaches, dry mouth or sore throat, awakening unrested, memory impairment, difficulty concentrating, and low energy, which again may go unrecognized or be attributed to other reasons (Schmidt-Nowara, 2012).

The diagnosis of sleep apnea is based on the conjoint evaluation of clinical symptoms (e.g., excessive daytime sleepiness and fatigue) and of the results of a formal sleep study such as polysomnography or reduced channels home based test. The overnight polysomnogram, the standard diagnostic test for obstructive sleep apnea, involves simultaneous recordings of multiple physiologic signals during sleep, including the electroencephalogram, electrooculogram, electromyogram, oronasal airflow, and oxyhemoglobin saturation. These recordings allow identification and classification of sleep-related apneas and hypopneas. Sleep apnea severity is typically assessed with the apnea–hypopnea index (AHI), which is the number of apneas and hypopneas per hour of sleep. Several additional measures of disease severity that characterize the degree of nocturnal hypoxemia (e.g., average oxyhemoglobin desaturation) and extent of sleep fragmentation (i.e., arousal frequency) also are used in the clinical and research areas (American Academy of Sleep Medicine, 1999; Punjabi, 2008; Schmidt-Nowara, 2012). The polysomnogram is attended in a laboratory by a sleep technician. However, home portable monitoring devices also are available to diagnose sleep apnea in individuals with a high probability of at least moderate or severe sleep apnea who do not have other illnesses or sleep disorders that may interfere with the results (Lee et al., 2008; Schmidt-Nowara, 2012).
The goals of treating sleep apnea are to restore regular breathing during sleep and to relieve symptoms such as loud snoring and daytime sleepiness. For milder cases, home remedies and lifestyle changes such as losing weight, quitting smoking, adjusting sleep position, avoiding alcohol and other sedatives, and using nasal sprays or allergy medicines to keep nasal passages open at night may be enough to treat or reduce symptoms. If these changes do not improve the symptoms or sleep apnea is moderate to severe, other treatments such as breathing devices or surgery are recommended. The most effective treatment for moderate to severe sleep apnea is to use a mechanical device – continuous positive airway pressure (CPAP) – to keep the upper airway open during sleep. A CPAP device uses an air-tight attachment to the nose, typically a mask, connected to a tube and a blower which generates the pressure (Gay et al., 2006; Schmidt-Nowara, 2012; U.S. Department of Health and Human Services, 2012). Although CPAP is the most common and reliable method of treating sleep apnea, some people find it uncomfortable and experience problems while using it. In these cases, other devices such as bi-level positive airway pressure (BPAP), expiratory positive airway pressure (EPAP), or oral devices may be recommended. Surgery, on the other hand, is usually only an option for patients who cannot tolerate or do not improve with non-surgical treatments such as CPAP or oral devices (Mayo Clinic, 2012; American Sleep Apnea Association, 2012a).

Although there are relatively simple, inexpensive, and effective modalities for diagnosing and treating sleep apnea, the vast majority of patients remain undiagnosed and therefore untreated because of the lack of awareness by the public and healthcare professionals. If left untreated, sleep apnea can have serious and life-shortening consequences, including high blood pressure, heart disease, stroke, diabetes, depression, memory problems, weight gain, impotence, and headaches. Moreover, untreated sleep apnea may be responsible for job impairment and motor
vehicle accidents, as studies have shown that people with severe sleep apnea are more than twice as likely to be involved in a motor vehicle accident as people without this condition (Ellen et al., 2006; George, 2007; Lee et al., 2008; Schmidt-Nowara, 2012; American Sleep Apnea Association, 2012b).

There are many studies in the literature that focus on sleep apnea prevalence, risk factors, symptoms, consequences, diagnosis, and treatment, including these by Sin et al. (1999), Beninati and Sanders (2001), Sharafkhaneh et al. (2004), El-Ad and Lavie (2005), Kjelsberg et al. (2005), Ellen et al. (2006), Ferguson et al. (2006), Gay et al. (2006), Freire et al. (2007), George (2007), Lee et al. (2008), Punjabi (2008), Matthews and Aloia (2009), Yaggi and Strohl (2010), Lloberes et al. (2011), Ryan et al. (2011), Mannarino et al. (2012), Vozoris (2012), Madani et al. (2013), and Mysliwiec et al. (2013).

1.3 Motivation for Work

Although “silent” injuries have generated widespread concern among policymakers, recent studies suggest that fundamental gaps remain about the mental health and cognitive needs of U.S. service members returning from Iraq and Afghanistan and the sufficiency of the care systems available to meet these needs. There is a strong belief that the Veterans Health Administration (VHA), the healthcare system within the Department of Veterans Affairs (VA), and the healthcare system for active-duty personnel within the Department of Defense (DoD) cannot adequately screen, diagnose, and treat the service members and veterans affected by TBI and PTSD, and yet little of the existing research is useful in guiding policy with regard to how
best to improve access to and the delivery of mental health care to those in need (Hoge et al., 2004; Tanielian and Jaycox, 2008; Congressional Budget Office, 2012).

The main part of this dissertation focuses on illustrating the use of systems engineering approaches to improve the effectiveness and cost of screening and diagnostic services for TBI and PTSD as well as patients’ access to these services within large integrated healthcare systems, such as VHA. This research is of great importance to the VHA – and potentially to the DoD as well – as it focuses on ensuring U.S. servicemen receive proper care for a large and growing healthcare concern. Substantial improvements in other healthcare sectors have resulted from systems engineering, suggesting significant opportunities for mental health care process improvements and cost reductions.

As a large integrated healthcare system, VHA provides comprehensive medical services to the veterans with not only combat-related conditions but also a broad range of other general health problems. Given that most veterans are exposed to situations and circumstances that have long-lasting effects on their health and thus more likely to develop chronic conditions (Yu et al., 2003), there are significant opportunities for further improvements both in quality and cost of care provided with the use of systems engineering methods to optimize the care processes for long-term health/general care problems along with those for mental health conditions. These opportunities also are illustrated in this dissertation, using sleep apnea as a case study.

The remainder of this dissertation is organized as follows. Chapter 2 presents several probability and Monte Carlo simulation models to determine the overall accuracy and cost of current and proposed screening processes for PTSD within the VHA, along with the illustration of a mild
TBI sequential screening process. In Chapter 3, three types of predictive diagnostic models – fuzzy logic, logistic regression, and neural networks – are developed to properly categorize TBI patients into the most likely severity state and a numerical example is provided for illustration purposes. Chapter 4 presents several single and multi-period integer programming models to determine the optimal locations and capacities for a range of specialty care services within the VHA, namely, PTSD treatment and sleep apnea testing services, which minimize overall system cost and total travel distances while providing appropriate within-network access. Finally, a summary of conclusions, limitations, and recommendations for future work are addressed in Chapter 5.
Chapter 2 - Sequential Screening Models

Among U.S. military members, mental disorders account for significant morbidity, disability, healthcare utilization, and attrition from military service. Veterans with untreated mental health problems may face severe consequences to their overall health and ability to fully reintegrate into the community, exacerbating the potential for chronic mental health conditions. Posttraumatic stress disorder (PTSD), for instance, causes significant physical and social problems along with tremendous medical and social costs if untreated or undertreated. This chapter presents several probability and Monte Carlo simulation models to develop a better understanding of current PTSD screening processes within the Veterans Health Administration (VHA), to determine the overall system accuracy, cost, and the total amount of resources for the current and proposed configurations, and to conduct a scenario analysis in order to thoroughly explore the potential improvements in the early diagnosis and care of veterans with PTSD. Also illustrated is a sequential screening process for traumatic brain injury (TBI) in order to give an insight into how the general approach is applied to other disorders.

2.1 Background

Mental health problems are among the most important contributors to the burden of disease and disability worldwide (Brundtland, 2000). Mental disorders are one of the leading causes of
disability in the U.S., with as many as 30% of the adult population suffering from a diagnosable mental disorder in a given year (Kessler et al., 1994; Kessler et al., 2005). The excess disability that is attributed to mental disorders is a result of high prevalence and chronicity, early age of onset, and the fact that they result in serious impairments (WHO International Consortium in Psychiatric Epidemiology, 2000). The magnitude of this burden also is attributed to the fact that fewer than half of individuals with these disorders receive treatment or initial treatment is generally delayed for many years, sometimes decades (Kessler et al., 1999; Kessler et al., 2001; Kohn et al., 2004; Wang et al., 2005). Undiagnosed and untreated mental illnesses have serious medical and social consequences, including but not limited to an illness that is more severe and difficult to treat, the development of co-occurring mental illnesses, worsening disability, unemployment, substance abuse, homelessness, and suicide (Norquist and Regier, 1996; Kessler et al., 2001). Individuals with untreated mental illnesses face an increased risk of having chronic medical conditions, yet they are less likely to follow-up with treatment for their health problems and ultimately die earlier than others without these conditions (Ciechanowski et al., 2000; Gehi et al., 2005; Colton and Manderscheid, 2006; Manderscheid et al., 2008). Moreover, the consequences extend beyond the personal cost and into the societal cost, as people with untreated mental illnesses earn less money and are prone to social isolation (Crisp et al., 2000; Goetzel et al., 2002). If identified early and treated effectively, most people with serious mental illnesses can significantly reduce the impact of their illness, improve their earning potential, and integrate into their community (Katon et al., 2005; Schennach et al., 2012; Schnurr and Lunney, 2012).

The effective treatment of mental health problems in primary care patients has been plagued for decades by the under-diagnosis and the under-treatment of mental and substance related disorders. Although there are marked gains in the identification and treatment of some mental
health disorders in many primary care settings, progress is slower for the anxiety disorders, particularly posttraumatic stress disorder (PTSD) (Magruder et al., 2005). A well-designed and effective screening program can help properly identify individuals with PTSD and give them access to effective treatments in a timely fashion. Mathematical modeling approaches, such as probability and simulation models, can be very useful at helping policy makers, clinicians, and researchers establish a better understanding of current PTSD screening processes, develop more effective alternatives both in terms of quality and cost, and further perform a scenario analysis to explore the possible improvements. These types of mathematical models have been extensively discussed by Benneyan et al. in the case of cervical cancer (Benneyan and Kaminsky, 1996; Kaminsky et al., 1997). Other applications include different health issues such as breast cancer (Ozekici and Pliska, 1991; Baker, 1998; Lee and Zelen, 1998; Warwick and Duffy, 2005; Maillart et al., 2008; Ayer et al., 2012), colorectal cancer (Knudsen et al., 2012), bladder cancer (Kent et al., 1989), hypertension (Donner et al., 1979), Down’s syndrome (Malone et al., 2005), and depression and dementia (Thibault and Steiner, 2004).

2.2 PTSD Screening Process within the VHA

PTSD is a serious mental health disorder which debilitates physical and social well-being and produces tremendous medical and social costs if untreated or undertreated (Schnurr et al., 2006; Hermann et al., 2012). The estimated prevalence of lifetime PTSD is increasingly higher in the military when compared to the general population, as it is one of the signature injuries in the Iraq and Afghanistan wars (Kessler et al., 2005; Magruder et al., 2005; Tanielian and Jaycox, 2008). In the Veterans Affairs (VA) population, PTSD is consistently associated with poorer health, decreased earnings, and social isolation (Schnurr and Green, 2004; Murdoch et al., 2005).
Fortunately, effective treatments, many of which were tested in the VHA, can address these problems if the veterans with PTSD are identified and treated in a timely fashion (Smith et al., 2005; Davis et al., 2012; Schnurr and Lunney, 2012). Therefore, the VA has invested significant resources in the development of a comprehensive PTSD screening system and instituted a yearly screening program for all primary care patients beginning in 2010 (Shiner et al., 2012). Since then, over 98% of primary care patients received screening in compliance with the protocol (U.S. Department of Veterans Affairs, 2012). Currently, each veteran is supposed to be screened annually by the Primary Care PTSD screen (PC-PTSD) for the first five years after his/her most recent date of separation from active duty, and then every five years afterwards. However, the date of separation is not known for all VA users and thus PC-PTSD screening continues annually beyond the first five years for those with an unknown date of separation. A clinical exam, which has a sensitivity (i.e., probability of diagnosing a PTSD patient) of 0.465 and a specificity (i.e., probability of testing negative in the absence of PTSD) of 0.966 (Magruder et al., 2005), confirms or disconfirms a positive result on the PC-PTSD. A schematic presentation of the current PTSD screening process in the VA is given in Figure 2-1.
While the PC-PTSD is a useful screening measure and applied appropriately in the VA, it represents only a first step in properly identifying PTSD patients. Confirmatory tests, such as the self-administered PTSD Checklist (PCL) and Clinician-Administered PTSD scale (CAPS) also play a role in PTSD identification after initial screening with the PC-PTSD. Therefore, this study proposes a more systematic approach where PCL and CAPS are used as confirmatory tests, as shown in Figure 2-2. The main characteristics of PC-PTSD, PCL, and CAPS are summarized in Table 2-1, with more detailed information presented earlier in Section 1.1.2.
Figure 2-2 Proposed PTSD screening process in the VA

Table 2-1 Main characteristics of the PC-PTSD, PCL, and CAPS screening tests

<table>
<thead>
<tr>
<th>Test</th>
<th>Cut-off score</th>
<th>Sensitivity*</th>
<th>Specificity*</th>
<th>Patient compliance rate</th>
<th>Avg cost</th>
<th>Time (minutes)</th>
<th>By whom</th>
</tr>
</thead>
<tbody>
<tr>
<td>PC-PTSD</td>
<td>2</td>
<td>0.91</td>
<td>0.72</td>
<td>0.90</td>
<td>$1</td>
<td>0.5</td>
<td>0.5 Administrative staff</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>0.84</td>
<td>0.90</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PCL</td>
<td>50</td>
<td>0.82</td>
<td>0.83</td>
<td>0.30</td>
<td>$25</td>
<td>8</td>
<td>8   Self-administered</td>
</tr>
<tr>
<td></td>
<td>DSM</td>
<td>1.00</td>
<td>0.92</td>
<td></td>
<td></td>
<td></td>
<td>(interpreted by a clinician)</td>
</tr>
<tr>
<td></td>
<td>Both</td>
<td>0.60</td>
<td>0.99</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CAPS</td>
<td>DSM</td>
<td>1.00</td>
<td>1.00</td>
<td>0.05</td>
<td>$200</td>
<td>70</td>
<td>50  Trained clinician</td>
</tr>
</tbody>
</table>

*Shiner et al. (2012)
Chapter 2 - Sequential Screening Models

2.3 Methodology

2.3.1 Notation and Assumptions

A mathematical model was developed to determine the overall system sensitivity, specificity, and total screening and false diagnosis costs under current and proposed process configurations for PTSD screening. The following notation is used in the probabilistic model, with regards to

\textit{Test characteristics:}

\begin{align*}
\alpha_s &= \text{probability that a PC-PTSD will determine a truly positive patient as negative}, \\
\alpha_p &= \text{probability that a PCL will determine a truly positive patient as negative}, \\
\alpha_c &= \text{probability that a CAPS will determine a truly positive patient as negative}, \\
\beta_s &= \text{probability that a PC-PTSD will determine a truly negative patient as positive}, \\
\beta_p &= \text{probability that a PCL will determine a truly negative patient as positive}, \\
\beta_c &= \text{probability that a CAPS will determine a truly negative patient as positive}, \\
u &= \text{probability that a PCL cannot make a definite diagnosis}, \\
r_s &= \text{patient compliance rate for PC-PTSD}, \\
r_p &= \text{patient compliance rate for PCL, and} \\
r_c &= \text{patient compliance rate for CAPS};
\end{align*}

\textit{Cost parameters:}

\begin{align*}
k_s &= \text{average cost of a PC-PTSD}, \\
k_p &= \text{average cost of a PCL}, \\
k_c &= \text{average cost of a CAPS}, \\
k_{FP} &= \text{average cost of a false-positive final result, and} \\
k_{FN} &= \text{average cost of a false-negative final result per year};
\end{align*}
**Policy performance measures:**

\[ j = \text{analysis horizon,} \]
\[ p_j = \text{overall sensitivity after } j \text{ years,} \]
\[ q_j = \text{overall specificity after } j \text{ years,} \]
\[ p_{FN|\text{Pos.}, j} = \text{probability of not detecting a truly positive patient after } j \text{ years,} \]
\[ p_{FP|\text{Neg.}, j} = \text{probability of identifying a truly negative patient as positive after } j \text{ years,} \]
\[ p_{TP, j} = \text{probability that any patient yields a true-positive after } j \text{ years,} \]
\[ p_{FN, j} = \text{probability that any patient yields a false-negative after } j \text{ years,} \]
\[ p_{TN, j} = \text{probability that any patient yields a true-negative after } j \text{ years,} \]
\[ p_{FP, j} = \text{probability that any patient yields a false-positive after } j \text{ years,} \]
\[ I_s = \text{number of PC-PTSD tests per any given patient,} \]
\[ I_p = \text{number of PCL tests per any given patient,} \]
\[ I_c = \text{number of CAPS tests per any given patient,} \]
\[ p_{\text{con}}^+ = \text{probability that diagnosis is confirmed by either PCL or CAPS given that the patient is positive,} \]
\[ p_{\text{con}}^- = \text{probability that diagnosis is confirmed by either PCL or CAPS given that the patient is negative,} \]
\[ p_{s_{\text{nc}}}^+ = \text{probability that patient takes PC-PTSD but result remains unconfirmed by neither PCL nor CAPS given that he/she is positive,} \]
\[ p_{s_{\text{nc}}}^- = \text{probability that patient takes PC-PTSD but result remains unconfirmed by neither PCL nor CAPS given that he/she is negative,} \]
\[ p_{p_{\text{nc}}}^+ = \text{probability that patient takes PCL but result remains unconfirmed given that he/she is positive, and} \]
\[ p_{p_{\text{nc}}}^- = \text{probability that patient takes PCL but result remains unconfirmed given that he/she is negative.} \]

The overall system sensitivity is defined as the probability that a truly positive patient will be determined correctly to be positive within the analysis horizon. Similarly, the overall system specificity is defined as the probability that a truly negative patient will be classified correctly as
negative at the end of the analysis horizon. In the proposed configuration, each patient is supposed to be screened by the PC-PTSD during $j \geq 1$ years. After a negative determination by the PC-PTSD in year $i < j$, no further action is taken and the patient is scheduled for another screen the next year. In any year $i < j$, if the PC-PTSD determines the patient to be positive, he/she is sent to the PCL for confirmation. If PCL cannot make a definite diagnosis (with probability $u$), the patient is sent to the CAPS.

The cost estimates for $k_s$, $k_p$, $k_c$, $k_{FP}$, and $k_{FN}$ are based on discussions with several psychiatrists in the VA. Specifically, the estimate for the cost of each test ($k_s$, $k_p$, and $k_c$) is based on the personnel needed for administration. False-positive and false-negative costs ($k_{FP}$ and $k_{FN}$), on the other hand, are more difficult to estimate and consist of treatment and disability costs as well as cost of anxiety and inconvenience caused to the patient. These costs vary from patient to patient and significantly different results might be obtained using different estimates. In order to make more accurate estimates, several cases for both false-positive and false-negative results were identified and for each case, minimum, average, and maximum cost values were determined, as given in Table 2-2. The probability distributions of these costs also are illustrated in Figure 2-3.
Table 2-2 Minimum, average, and maximum cost estimates for a false-positive and false-negative result under the assumption of different cases

<table>
<thead>
<tr>
<th>Case</th>
<th>Definition</th>
<th>Percentage</th>
<th>Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Min</td>
</tr>
<tr>
<td>1</td>
<td>Initial evaluation</td>
<td>40%</td>
<td>$300</td>
</tr>
<tr>
<td>2</td>
<td>Minimum treatment</td>
<td>20%</td>
<td>$1,000</td>
</tr>
<tr>
<td>3</td>
<td>Ongoing treatment</td>
<td>20%</td>
<td>$6,000</td>
</tr>
<tr>
<td>4</td>
<td>Partial disability and extensive treatment</td>
<td>10%</td>
<td>$30,000</td>
</tr>
<tr>
<td>5</td>
<td>Full disability and extensive treatment</td>
<td>10%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>False-negative</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>Minimal symptoms</td>
<td>25%</td>
<td>$1,000</td>
</tr>
<tr>
<td>2</td>
<td>Partial disability</td>
<td>50%</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Full disability (dysfunctional)</td>
<td>25%</td>
<td></td>
</tr>
</tbody>
</table>

(a) Probability distribution of cost of a false-positive result

(b) Probability distribution of cost of a false-negative result

Figure 2-3 Probability distribution of cost of (a) a false-positive and (b) a false-negative result
Other model assumptions include

(1) Each PC-PTSD, PCL, and CAPS test has a patient compliance rate (i.e., not all patients take these tests even though it is required) and these rates are assumed to remain constant over time.

(2) Each PC-PTSD, PCL, and CAPS test has a sensitivity of \(1 - \alpha_s\), \(1 - \alpha_p\), and \(1 - \alpha_c\), respectively. Thus, the probability that a PC-PTSD, PCL, and CAPS test will correctly determine a truly positive patient to be positive is \(1 - \alpha_s\), \(1 - \alpha_p\), and \(1 - \alpha_c\), respectively. These sensitivities are assumed to remain constant over time.

(3) Each PC-PTSD, PCL, and CAPS test has a specificity of \(1 - \beta_s\), \(1 - \beta_p\), and \(1 - \beta_c\), respectively. Thus, the probability that a PC-PTSD, PCL, and CAPS test will correctly determine a truly negative patient to be negative is \(1 - \beta_s\), \(1 - \beta_p\), and \(1 - \beta_c\), respectively. These specificities are assumed to remain constant over time.

(4) All tests are assumed to be independent of each other.

(5) The expected number of PC-PTSD tests per positive patient and per negative patient are denoted, respectively, as \(E[I_s | \text{Positive}]\) and \(E[I_s | \text{Negative}]\). Similarly, the expected number of PCL and CAPS tests per positive patient and per negative patient are denoted, respectively, as \(E[I_p | \text{Positive}], E[I_p | \text{Negative}], E[I_c | \text{Positive}], E[I_c | \text{Negative}], E[I_c | \text{Negative}]\).

(6) Screening and false diagnosis costs are assumed to remain constant over time.
2.3.2 Overall System Sensitivity

2.3.2.1 Probabilities of Detecting Truly Positive Patients and of a False-Negative

The overall sensitivity is a combined function of the individual PC-PTSD, PCL, and CAPS sensitivities, the patient compliance rates, the length of the analysis horizon, and the probability that a PCL cannot make a definite diagnosis. Note that a truly positive patient can be given a final correct positive determination in several manners. For example, assuming that the patient complies with the overall screening process, if the PC-PTSD in any year correctly determines the patient to be positive, the PCL subsequently confirms this determination. Alternatively, after a positive determination from the PC-PTSD, if the PCL cannot make a definite diagnosis, the CAPS confirms the positive determination.

By considering all possible probabilistic manners in which a truly positive patient is correctly identified as positive, a mathematical model of the overall probability of detecting a truly positive patient was developed based on the assumptions and notation given in Section 2.3.1. Thus, the probability of detecting a truly positive patient in \( j \) years is shown in Appendix A to be

\[
\text{Overall sensitivity} = \text{Probability of detecting a truly positive patient} = P\{\text{Positive determination}|\text{Patient is positive}\}
\]

\[
= \frac{\left(1-(1-a_p(1-a_c))r_p(1-u(1-r_c))\right)^j(1-a_p-u(1-a_p-r_c(1-a_c)))}{1-u(1-r_c)}.
\]

The probability of a false-negative (FN), traditionally defined as the conditional probability of incorrectly identifying a truly positive patient as negative in \( j \) years, simply is
Probability of a FN = Probability of not detecting a truly positive patient
\[ p_{\text{FN}} | \text{Pos}, j = P\{\text{Negative determination} | \text{Patient is positive}\} \]
\[ = 1 - \text{Overall sensitivity} = 1 - p_j \]
\[ = \frac{(1-r_s(1-\alpha_s)r_p(1-u(1-r_c)))(1-\alpha_p-u(1-\alpha_p-r_c(1-\alpha_c)))+(1-u)\alpha_p+ur_c\alpha_c}{1-u(1-r_c)}. \] (2-2)

The unconditional probabilities of a true-positive (TP) and a false-negative (FN) result after \( j \) years screening of any patient whose true state is not known \textit{a priori} are the above rates multiplied by the prevalence, \( p \), of PTSD in the population under consideration:

\[ P\{\text{Any patient yields a TP}\} = p_{\text{TP}, j} = p \times p_j \]
\[ = p \frac{(1-r_s(1-\alpha_s)r_p(1-u(1-r_c)))(1-\alpha_p-u(1-\alpha_p-r_c(1-\alpha_c)))+(1-u)\alpha_p+ur_c\alpha_c}{1-u(1-r_c)}, \] (2-3)

\[ P\{\text{Any patient yields a FN}\} = p_{\text{FN}, j} = p \times p_{\text{FN}|\text{Pos}, j} \]
\[ = p \frac{(1-r_s(1-\alpha_s)r_p(1-u(1-r_c)))(1-\alpha_p-u(1-\alpha_p-r_c(1-\alpha_c)))+(1-u)\alpha_p+ur_c\alpha_c}{1-u(1-r_c)}. \] (2-4)

As an example, given a prevalence of \( p = 0.115 \) (Magruder et al., 2005), PC-PTSD sensitivity of \( 1-\alpha_s = 0.84 \) (with a cut-off score of 3), PCL sensitivity of \( 1-\alpha_p = 1 \) (with DSM criteria), CAPS sensitivity of \( 1-\alpha_c = 1 \) (with DSM criteria), compliance rate of \( r_s = 0.90, r_p = 0.30, \) and \( r_c = 0.05 \) for PC-PTSD, PCL, and CAPS, respectively, and assuming that PCL cannot make a definite diagnosis with a probability of \( u = 0.10 \), then the probabilities of incorrectly determining a truly positive patient to be negative and of incorrectly determining any patient to be negative in the first year (\( j = 1 \)) are
\[ P\{\text{Negative determination|Patient is positive}\} = p_{FN|Pos, 1} \]
\[ = \frac{(1-r_s(1-a_s)r_p(1-u(1-r_c)))^j(1-a_p-u(1-a_p-r_s(1-a_s)))+(1-u)a_p+ur_s a_c}{1-u(1-r_c)} \]
\[ = \left(1-(0.90)(0.84)(0.30)(1-(0.10)(1-0.05))\right)^j(1-(0.10)(1-(0.05)(1)))+(1-0.10)(0)+(0.10)(0.05)(0) \]
\[ = 1 - \left(1 - ((0.90)(0.16) + 1 - 0.90)^j\right)0.30\left(1 - 0.10(1 - 0.05(1))\right) \]
\[ = 0.7947 \]

and

\[ P\{\text{Any patient yields a FN}\} = p_{FN, 1} = p \times p_{FN|Pos, 1} = (0.115)(0.7947) = 0.0914. \]

### 2.3.2.2 Expected Number and Distribution of False-Negatives

If \( n_p \) positive patients are screened independently, then for any given set of conditions the number of correct positive identifications is modeled with a binomial distribution with an expected value and standard deviation of \( n_pp_j \) and \( \sqrt{n_pp_j(1-p_j)} \), respectively. Similarly, the number of undetected positive patients out of \( n \) patients of unknown medical status also has a binomial distribution, here with an expected value and standard deviation of \( n_p^{FN, j} \) and \( \sqrt{n_p^{FN, j}(1-p^{FN, j})} \), respectively. (For completeness, also note that the companions to these two random variables – the number of false-negative diagnoses per \( n_p \) positive patients and the number of correct positive diagnoses per \( n \) patients of unknown status – are binomial with the same variances as above and with expectations \( n_pl^{FN|P, j} \) and \( n_TP, j \), respectively.)
These expectations, variances, and distributions can be used to predict the relative frequencies of the number of events of interest which will occur over time, such as the number of false-negatives per $n$ patients per year. For example, Figure 2-4 shows the probability distribution of the number of false-negative diagnoses per $n = 10,000$ patients for the previous example. Note that the expected number of false-negative diagnoses in the first year is

$$\text{Expected number of FN diagnoses per 10,000 patients} = np_{FN, 1} = 10,000 \times 0.0914 = 913.9579,$$

and the standard deviation is

$$\sqrt{np_{FN, 1} \left( 1 - p_{FN, 1} \right)} = \sqrt{10,000 \times 0.0914 \times (1 - 0.0914)} = \sqrt{830.43} \approx 28.8171,$$

which agrees with the location and spread of the corresponding probability mass in Figure 2-4.

[Figure 2-4 Probability distributions of number of false diagnoses per 10,000 patients in the first year ($p = 0.115, j = 1, u = 0.10, \alpha_s = 0.16, \alpha_p = 0, \alpha_c = 0, \beta_s = 0.10, \beta_p = 0.08, \beta_c = 0, r_s = 0.90, r_p = 0.30, r_c = 0.05$)]
2.3.3 Overall System Specificity

2.3.3.1 Probabilities of Detecting Truly Negative Patients and of a False-Positive

Similar to the above, a mathematical model of the overall probability of detecting a truly negative patient was developed based on the assumptions and notation given earlier in Section 2.3.1. Thus, the probability of identifying a truly negative patient as negative in \( j \) years is shown in Appendix B to be

\[
q_j = P\{\text{Negative determination|Patient is negative}\} = \left(1 - r_d \beta_p (1-u(1-r_c))\right) \frac{\left((1-u)\beta_p + ur_c \beta_c \right) + 1 - \beta_p - u \left(1-\beta_p - r_c(1-\beta_c)\right)}{1-u(1-r_c)}, \tag{2-5}
\]

Also as above, the probability of a false-positive (FP), traditionally defined as the conditional probability of incorrectly identifying a truly negative patient as positive, is

\[
P_{FP|Neg, j} = 1 - q_j = \frac{\left(1 - (1-r_d \beta_p (1-u(1-r_c)))\right) \left((1-u)\beta_p + ur_c \beta_c \right)}{1-u(1-r_c)}, \tag{2-6}
\]

and the corresponding unconditional true-negative (TN) and false-positive (FP) probabilities are

\[
P\{\text{Any patient yields a TN}\} = p_{TN, j} = (1-p) \times q_j
\]

\[
= (1-p) \left(1 - r_d \beta_p (1-u(1-r_c))\right) \frac{\left((1-u)\beta_p + ur_c \beta_c \right) + 1 - \beta_p - u \left(1-\beta_p - r_c(1-\beta_c)\right)}{1-u(1-r_c)}, \tag{2-7}
\]
and

\[ P\{\text{Any patient yields a FP}\} = p_{FP, j} = (1 - p) \times p_{FP|\text{Neg.}, j} \]

\[ = (1 - p) \left( \frac{\left(1 - \left(1 - r_s \beta_s r_p (1 - u(1 - r_c))\right)\right) \left(1 - u\beta_p + u r_c \beta_c\right)}{1 - u(1 - r_c)} \right) \]  

(2-8)

As an example, given a prevalence of \(p = 0.115\), PC-PTSD specificity of \(1 - \beta_s = 0.90\) (with a cut-off score of 3), PCL specificity of \(1 - \beta_p = 0.92\) (with DSM criteria), CAPS specificity of \(1 - \beta_c = 1\) (with DSM criteria), compliance rate of \(r_s = 0.90\), \(r_p = 0.30\), and \(r_c = 0.05\) for PC-PTSD, PCL, and CAPS, respectively, and assuming that PCL cannot make a definite diagnosis with a probability of \(u = 0.10\), then the probabilities of incorrectly determining a truly negative patient to be positive and of incorrectly determining any patient to be positive in the first year \((j = 1)\) are

\[ P\{\text{Positive determination|Patient is negative}\} = p_{FP|\text{Neg.}, 1} \]

\[ = \frac{\left(1 - \left(1 - r_s \beta_s r_p (1 - u(1 - r_c))\right)\right) \left(1 - u\beta_p + u r_c \beta_c\right)}{1 - u(1 - r_c)} \]

\[ = \frac{\left(1 - \left(1 - (0.90)(0.10)(0.30)(1 - (0.10)(1 - 0.05))\right)\right) \left(1 - (0.10)(0.08) + (0.10)(0.05)(0)\right)}{1 - (0.10)(1 - 0.05)} \]

\[ = 1 - \left( ((0.90)(0.90) + 1 - 0.90)^1 + \left(1 - ((0.90)(0.90) + 1 - 0.90)^1\right) \right) \times 0.30 \left(0.92 - 0.10(0.92 - (0.05)(1))\right) \]

\[ = 0.0019 \]

and

\[ P\{\text{Any patient yields a FP}\} = p_{FP, 1} = (1 - p) \times p_{FP|\text{Neg.}, 1} = (1 - 0.115)(0.0019) = 0.0017. \]
2.3.3.2 Expected Number and Distribution of False-Positives

Again, if \( n_n \) negative patients are screened independently, then the number of correct negative identifications is modeled with a binomial distribution with parameters \( n_n \) and \( q_j \), and the expected number and standard deviation of correct identifications per \( n_n \) truly negative patients, \( n_n q_j \) and \( \sqrt{n_n q_j (1 - q_j)} \), respectively, can be examined. Similarly, the number of undetected negative patients out of \( n \) patients of unknown medical status also has a binomial distribution, here with an expected value and standard deviation of \( np_{FP, j} \) and \( \sqrt{np_{FP, j} (1 - p_{FP, j})} \), respectively. (Again, note that the companions to these two random variables – the number of false-positive diagnoses per \( n_n \) negative patients and the number of correct negative diagnoses per \( n \) patients of unknown status – are binomial with the same variances as above and with expectations \( n_n p_{FP | N, j} \) and \( np_{TN, j} \), respectively.)

Figure 2-4 also shows the probability distribution of the number of false-positive diagnoses again per \( n = 10,000 \) patients for the previous example. Note that the expected number of false-positive diagnoses in this situation is

\[
\text{Expected number of FP diagnoses per 10,000 patients} = np_{FP, 1} = 10,000 \times 0.0017 = 17.2044,
\]

and the standard deviation is

\[
\sqrt{np_{FP, 1} (1 - p_{FP, 1})} = \sqrt{10,000 \times 0.0597 \times (1 - 0.0597)} = \sqrt{17.1748} \approx 4.1442,
\]
which again agrees with the corresponding probability mass in Figure 2-4.

### 2.3.4 Expected Cost

Based on the previous assumptions and notation, the expected cost per \( n \) patients screened after \( j \) years is shown in Appendix C to be

\[
EC_j = \text{Expected cost of all PC-PTSD tests per } n \text{ patients} + \\
\text{Expected cost of all PCL tests per } n \text{ patients} + \\
\text{Expected cost of all CAPS tests per } n \text{ patients} + \\
\text{Expected cost of all false-positives per } n \text{ patients} + \\
\text{Expected cost of all false-negatives per } n \text{ patients}
\]

\[
= (k_s \times \text{Expected number of PC-PTSD tests per } n \text{ patients}) + \\
(k_p \times \text{Expected number of PCL tests per } n \text{ patients}) + \\
(k_c \times \text{Expected number of CAPS tests per } n \text{ patients}) + \\
(k_{FP} \times \text{Expected number of false-positives per } n \text{ patients}) + \\
(k_{FN} \times \text{Expected number of false-negatives per } n \text{ patients})
\]

\[
= k_s E[I_{s,n}] + k_p E[I_{p,n}] + k_c E[I_{c,n}] + k_{FP} E[FP] + k_{FN} E[FN]
\]

\[
= k_s n \left[ p \left( \sum_{i=1}^{j} \left( \sum_{a=0}^{(a+i-1)!} \frac{(a+i-1)!}{a!(i-1)!} \right) (1 - r_s)^a + j p_{s,nc}^+ (1 - p_{con}^+)^{i-1} \right) + (1 - p) \left( \sum_{i=1}^{j} p_{s,nc}^- \sum_{a=0}^{(a+i-1)!} \frac{(a+i-1)!}{a!(i-1)!} (1 - r_s)^a + j p_{s,nc}^- (1 - p_{con}^-)^{i-1} \right) \right]
\]

\[
+ k_p n \left[ p \left( \sum_{i=1}^{j} \left( \sum_{a=0}^{(a+i-1)!} \frac{(a+i-1)!}{a!(i-1)!} \right) (1 - r_s (1 - a_s) r_p^a + j p_{p,nc}^+ (1 - p_{con}^+)^{i-1}) \right) + (1 - p) \left( \sum_{i=1}^{j} p_{p,nc}^- \sum_{a=0}^{(a+i-1)!} \frac{(a+i-1)!}{a!(i-1)!} (1 - r_s \beta_s r_p^a + j p_{p,nc}^- (1 - p_{con}^-)^{i-1}) \right) \right]
\]

\[
+ k_c n \left[ \frac{\beta_c}{1-u(1-r_c)} \left( p \left( 1 - (1 - p_{con}^+) \right) + (1 - p) \left( 1 - (1 - p_{con}^-) \right) \right) \right]
\]

\[
+ k_{FP} n (1 - p) \left[ \frac{(1-u(1-r_c))}{1-a(1-r_c)} \left( (1-u) \beta_p + \beta_{FP} \right) \right]
\]

\[
+ k_{FN} n p \left[ \frac{(1-r_s (1-a_s) r_p (1-u(1-r_c)))}{1-u(1-r_c)} \left( (1-u) a_p + \beta_{FP} \right) \right].
\]

\[(2.9)\]
As an example, the expected cost per \( n = 10,000 \) screened patients for the previous situation (with \( p = 0.115, j = 1, 1-\alpha_s = 0.84, 1-\alpha_p = 1, 1-\alpha_c = 1, 1-\beta_s = 0.90, 1-\beta_p = 0.92, 1-\beta_c = 1, r_s = 0.90, r_p = 0.30, r_c = 0.05, \) and \( u = 0.10 \)) and with cost estimates \( k_s = \$1, k_p = \$25, k_c = \$200, k_{FP} = \$13,720, \) and \( k_{FN} = \$15,250 \) is (note that \( p_{con}^+, p_{con}^-, p_{s\_nc}^+, p_{s\_nc}^-, p_{p\_nc}^+, \) and \( p_{p\_nc}^- \) were calculated as given in Appendix C)

\[
EC_j = k_s n \left( p (p_{con}^+ + p_{s\_nc}^+) + (1 - p) (p_{con}^- + p_{s\_nc}^-) \right)
+ k_p n \left( p (p_{con}^+ + p_{p\_nc}^+) + (1 - p) (p_{con}^- + p_{p\_nc}^-) \right)
+ k_c \left( \frac{ur_c}{1-u(1-r_c)} \right) \left( p (p_{con}^+) + (1 - p) (p_{con}^-) \right)
+ k_{FP} n \left( \frac{(r_{p}s_{p}(1-u(1-r_c))\beta_s)}{1-u(1-r_c)} \right)
+ k_{FN} n \left( \frac{u(1-u)(1-r_c)}{1-u(1-r_c)} \right)
\]

\[
= k_s (10,000) \left( (0.115)(0.2053 + 0.6947) + (1 - 0.115)(0.0244 + 0.8756) \right)
+ k_p (10,000) \left( (0.115)(0.2053 + 0.0215) + (1 - 0.115)(0.0244 + 0.0026) \right)
+ k_c (10,000) \left( (0.10)(0.05) \right) \left( (0.115)(0.2053) + (1 - 0.115)(0.0244) \right)
+ k_{FP} (10,000) \left( 1 - 0.115 \right) \left( 0.90 \right) \left( 0.10 \right) \left( 0.30 \right) \left( 1 - 0.10 \right) \left( 0.08 \right)
+ k_{FN} (10,000) \left( 0.115 \right) \left( 1 - 0.90 \right) \left( 0.84 \right) \left( 0.30 \right) \left( 1 - 0.10 \right) \left( 1 - 0.05 \right)
\]

\[
= (\$1)(10,000)(0.90) + (\$25)(10,000)(0.0499) + (\$200)(10,000)(0.00025) + (\$13,720)(10,000)(0.0017) + (\$15,250)(10,000)(0.00914)
\]

\[
= \$9,000 + \$12,494.25 + \$499.77 + \$236,044.37 + \$13,937,857.98 = \$14,195,896.36.
\]

Although it is not pursued in this study, the variance of total cost per \( n \) screened patients also can be obtained analytically, and the distribution has been examined via computer simulation. Figure 2-5 illustrates a distribution of 5,000 simulation replications of the cost per 10,000 screened patients for the above situation, which fits the normal distribution best at a significance level of
0.05 (with a $p$-value of 0.388). Further simulation results are summarized later in Table 2-3, indicating close agreement with the analytic results derived in this section.

Figure 2-5 Probability distribution of total cost per 10,000 patients (obtained via simulation)

($p = 0.115, j = 1, u = 0.10, k_s = $1, k_p = $25, k_c = $200, k_{FP} = $13,720, k_{FN} = $15,250, \alpha_s = 0.16, \alpha_p = 0, \alpha_c = 0, \beta_s = 0.10, \beta_p = 0.08, \beta_c = 0, r_s = 0.90, r_p = 0.30, r_c = 0.05, 5000 replications)$
Table 2-3 Summary of analytic distributions of cost, workload, and accuracy with comparison to simulation results – 5,000 replications ($p = 0.115$, $u = 0.10$, $k_i = $1, $k_p = $25, $k_c = $200, $k_{FP} = $13,720, $k_{FN} = $15,250, $\alpha_i = 0.16$, $\alpha_p = 0$, $\alpha_c = 0$, $\beta_i = 0.10$, $\beta_p = 0.08$, $\beta_c = 0$, $r_i = 0.90$, $r_p = 0.30$, $r_c = 0.05$)

<table>
<thead>
<tr>
<th>Year ($j$)</th>
<th>Per 10,000 patients ($n = 10,000$)</th>
<th>Analytic results</th>
<th>Simulation program results</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Expected value</td>
<td>Stdev</td>
</tr>
<tr>
<td>Total cost</td>
<td></td>
<td>$14,195,896$</td>
<td>$443,124.47$</td>
</tr>
<tr>
<td>$j = 1$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Workload (# screenings):</td>
<td>by admins (PC-PTSD)</td>
<td>9,000.00</td>
<td>30.0000</td>
</tr>
<tr>
<td></td>
<td>by clinicians (PCL)</td>
<td>499,7700</td>
<td>21.7897</td>
</tr>
<tr>
<td></td>
<td>by trained clinicians (CAPS)</td>
<td>2.4989</td>
<td>1.5806</td>
</tr>
<tr>
<td>Accuracy:</td>
<td># false-negatives</td>
<td>913.9579</td>
<td>28.8171</td>
</tr>
<tr>
<td></td>
<td># false-positives</td>
<td>17.2044</td>
<td>4.1442</td>
</tr>
<tr>
<td>Total cost</td>
<td>$34,569,800$</td>
<td>$1,194,883$</td>
<td>$34,551,996$</td>
</tr>
<tr>
<td>$j = 3$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Workload (# screenings):</td>
<td>by admins (PC-PTSD)</td>
<td>25,827.17</td>
<td>63.3129</td>
</tr>
<tr>
<td></td>
<td>by clinicians (PCL)</td>
<td>1332.3215</td>
<td>34.4606</td>
</tr>
<tr>
<td></td>
<td>by trained clinicians (CAPS)</td>
<td>6.6616</td>
<td>2.5801</td>
</tr>
<tr>
<td>Accuracy:</td>
<td># false-negatives</td>
<td>577.2752</td>
<td>23.3227</td>
</tr>
<tr>
<td></td>
<td># false-positives</td>
<td>50.3623</td>
<td>7.0787</td>
</tr>
<tr>
<td>Total cost</td>
<td>$47,592,761$</td>
<td>$1,800,087$</td>
<td>$47,580,935$</td>
</tr>
<tr>
<td>$j = 5$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Workload (# screenings):</td>
<td>by admins (PC-PTSD)</td>
<td>41,369.48</td>
<td>111.1222</td>
</tr>
<tr>
<td></td>
<td>by clinicians (PCL)</td>
<td>2005.5960</td>
<td>41.0852</td>
</tr>
<tr>
<td></td>
<td>by trained clinicians (CAPS)</td>
<td>10.0280</td>
<td>3.1651</td>
</tr>
<tr>
<td>Accuracy:</td>
<td># false-negatives</td>
<td>364.6193</td>
<td>18.7437</td>
</tr>
<tr>
<td></td>
<td># false-positives</td>
<td>81.9196</td>
<td>9.0138</td>
</tr>
</tbody>
</table>

2.4 Scenario Analysis

2.4.1 Effect of Scoring Methods on Overall System Accuracy, Cost, and Workload

The expected value expressions above also can be used to more precisely examine the joint effect of various scoring methods (i.e., test performances) and the length of screening horizon on
the overall system sensitivity, specificity, cost, and clinician workload. For this purpose, four different screening protocols (i.e., alternate scenarios for the test performances) were developed and their overall performance in terms of accuracy and cost was compared to the current process as well as the baseline which is the performance of clinician identification of PTSD before yearly PC-PTSD was implemented, given a prevalence of \( p = 0.115 \), compliance rate of \( r_s = 0.90 \), \( r_p = 0.30 \), and \( r_c = 0.05 \) for PC-PTSD, PCL, and CAPS, respectively, and assuming that PCL cannot make a definite diagnosis with a probability of \( u = 0.10 \). In the first protocol, the scoring method with the highest specificity is assumed to be used for PC-PTSD while the one with the highest sensitivity is chosen for PCL. In other words, the first protocol aims to eliminate the truly negative patients in the first step of the screening process (i.e., PC-PTSD) by using the scoring method with the highest specificity, and then to successfully confirm the truly positive patients in the second step (i.e., PCL or CAPS) by using the scoring method with the highest sensitivity. The second protocol is the opposite of the first one while in the third protocol, the scoring method with the highest sensitivity is used both for PC-PTSD and PCL. Finally, the last protocol illustrates the annual screening with PCL instead of PC-PTSD. These screening protocols are illustrated in Figure 2-6 with their overall sensitivity, specificity, and screening cost per patient after 5 years.
## Chapter 2 - Sequential Screening Models

### Figure 2-6 Comparison of old, current, and potential processes for PTSD case identification in the VA

<table>
<thead>
<tr>
<th>Stage 1</th>
<th>Stage 2</th>
<th>Stage 3</th>
<th>Year 5 results</th>
</tr>
</thead>
</table>
| Baseline | Clinician diagnosis of PTSD | | Sensitivity = 0.47  
Specificity = 0.97  
Cost’ = $1,500.00 |
| Current | Screening with PC-PTSD (High specificity: Cutoff of 3)  
Sensitivity = 0.84  
Specificity = 0.90 | Clinician diagnosis of PTSD  
Sensitivity = 0.47  
Specificity = 0.97 | | Sensitivity = 0.55  
Specificity = 0.98  
Cost’ = $528.93 |
| Protocol 1 | Screening with PC-PTSD (High specificity: Cutoff of 3)  
Sensitivity = 0.84  
Specificity = 0.90 | Reflex testing with PCL (High sensitivity: DSM)  
Sensitivity = 1.00  
Specificity = 0.92 | Confirm with CAPS (Gold Standard)  
Sensitivity = 1.00  
Specificity = 1.00 | | Sensitivity = 0.68  
Specificity = 0.99  
Cost’ = $9.35 |
| Protocol 2 | Screening with PC-PTSD (High sensitivity: Cutoff of 3)  
Sensitivity = 0.91  
Specificity = 0.72 | Reflex testing with PCL (High specificity: DSM-50)  
Sensitivity = 0.60  
Specificity = 0.99 | | | Sensitivity = 0.43  
Specificity = 1.00  
Cost’ = $13.76 |
| Protocol 3 | Screening with PC-PTSD (High sensitivity: Cutoff of 3)  
Sensitivity = 0.91  
Specificity = 0.72 | Reflex testing with PCL (High sensitivity: DSM)  
Sensitivity = 1.00  
Specificity = 0.92 | | | Sensitivity = 0.72  
Specificity = 0.98  
Cost’ = $13.76 |
| Protocol 4 | Screening with PCL (High sensitivity: DSM)  
Sensitivity = 1.00  
Specificity = 0.92 | Confirm with CAPS (Gold Standard)  
Sensitivity = 1.00  
Specificity = 1.00 | | | Sensitivity = 0.79  
Specificity = 1.00  
Cost’ = $139.49 |

*Expected screening cost per patient*

Figure 2-6 Comparison of past, current, and proposed screening protocols for PTSD identification in the VA

Figure 2-7 illustrates the effect of the test performances on the overall system sensitivity and specificity over a 20-year screening horizon, showing that the addition of the PC-PTSD increased the sensitivity of PTSD patient identification in the VA by almost 15% while maintaining a specificity of over 97%, when compared to annual clinical exams alone (i.e., baseline). Including triggered testing with the PCL and CAPS in the case of positive results adds further improvements in sensitivity as yearly screening continues (i.e., after 3 years for protocols 1, 3, and 4, after 10 years for protocol 2). However, these proposed screening protocols might result in losses of specificity if repeated over a 20-year horizon.
Figure 2-7 Effect of scoring method on overall (a) sensitivity and (b) specificity over a 20-year screening horizon.

Figure 2-8 plots the expected number of false-negatives and false-positives per 10,000 patients against years. As expected, protocol 4, which proposes yearly PCL screening, performs best in terms of sensitivity and specificity (i.e., generates less false-negative and false-positive results). Protocol 3 has the second best sensitivity but the worst specificity, while protocol 2 has the second best specificity but the worst sensitivity. Protocol 1, on the other hand, gives relatively good results in terms of both the number of false-negatives and false-positives.
Figure 2-8 Effect of scoring method on expected number of (a) false-negatives and (b) false-positives over a 20-year screening horizon

Figure 2-9 compares the expected screening and total costs per 10,000 patients over the years, indicating that protocol 4, which achieves the best sensitivity and specificity, also has the lowest total cost (i.e., total of screening, false-negative, and false-positive costs). Its screening cost, however, is significantly higher (up to $300,000 more per year), as well as the number of CAPS tests needed (up to 500 more CAPS per year) when compared to the other proposed protocols; this might generate an excessive workload considering that there may not be enough clinicians
with a working knowledge of PTSD in the VA to perform that many CAPS tests. Protocol 1, on the other hand, achieves a 29.94% and 56.65% decrease in the number of false-negatives and false-positives in the first 5 years when compared to the current process, and generates a reasonable screening cost and clinician workload, as illustrated in Figure 2-10.

Figure 2-9 Effect of scoring method on expected (a) screening cost and (b) total cost over a 20-year screening horizon

![Figure 2-9 Effect of scoring method on expected (a) screening cost and (b) total cost over a 20-year screening horizon](image-url)
Chapter 2 - Sequential Screening Models

Figure 2-10 Decomposition of expected (a) total cost and (b) amount of screening tests assuming protocol 1

The above examples illustrate the inherent tradeoff that decision makers face – that yearly screening significantly increases the probability of detecting true-positives, but has the opposite effect on the probability of identifying true-negatives. Figure 2-7, however, shows that the magnitude of this effect on reducing specificity is lower than on increasing sensitivity. On the other hand, Figure 2-8 illustrates that, when compared to the current process, protocol 1 decreases the number of false-negatives significantly, but generates more false-positives as yearly screening continues. Therefore, to consider “stopping yearly screening after a number of consecutive negative results” controls the increase in the number of false-positives while
maintaining the improvement achieved in terms of false-negatives. Figure 2-11 illustrates the tradeoff between the stopping criteria (i.e., number of negatives required to stop screening) and the expected number of false diagnoses and cost for protocol 1, indicating that discontinued screening after three or four negative results achieves a significant decrease in the number of false-positives and the screening cost, with a slight increase in the number of false-negatives and total cost. Note that stopping screening earlier decreases the number of false-positives further, but causes an enormous increase in the number of false-negatives and ultimately, the total cost.

Figure 2-11 Tradeoff between stopping criteria and expected (a) number of false diagnoses, and (b) cost at the end of screening horizon ($j = 20$) assuming protocol 1.
Figure 2-12 and 2-13 illustrate that under the assumption of protocol 1, a significant decrease in the number of false-positives is achieved by stopping yearly screening after three negative results, resulting in a reasonable increase in the number of false-negatives and the total cost. Furthermore, overall sensitivity, specificity, cost, and clinician workload of all screening protocols with discontinued screening after three negative results are summarized later in Table 2-4.
Figure 2-13 Effect of discontinued screening after three negative results on expected (a) screening cost and (b) total cost assuming protocol 1.
Table 2-4 Overall sensitivity, specificity, cost, and workload of alternate screening protocols at important intervals (with discontinued screening after three negative results)

<table>
<thead>
<tr>
<th>Performance measure</th>
<th>Year 1</th>
<th>Year 3</th>
<th>Year 5</th>
<th>Year 10</th>
<th>Year 20</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sensitivity</td>
<td>0.3796</td>
<td>0.5260</td>
<td>0.5470</td>
<td>0.5249</td>
<td>0.5252</td>
</tr>
<tr>
<td>Specificity</td>
<td>0.9892</td>
<td>0.9787</td>
<td>0.9786</td>
<td>0.9802</td>
<td>0.9802</td>
</tr>
<tr>
<td><strong>Expected cost:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Screening</td>
<td>$1,131,700</td>
<td>$3,202,506</td>
<td>$5,289,318</td>
<td>$9,898,422</td>
<td>$19,118,444</td>
</tr>
<tr>
<td>Total</td>
<td>$13,321,237</td>
<td>$33,049,644</td>
<td>$53,381,554</td>
<td>$103,562,981</td>
<td>$207,489,577</td>
</tr>
</tbody>
</table>

| Workload (# screenings): |        |        |        |         |         |
| by admins (PC-PTSD)      | 6,749.38 | 19,642.72 | 32,579.59 | 61,100.88 | 118,143.39 |
| by clinicians (clinical exam) | 3,749.83 | 10,609.54 | 17,522.46 | 32,791.07 | 63,334.34 |
| Sensitivity             | 0.2054 | 0.4981 | 0.6803 | 0.8545  | 0.9126  |
| Specificity             | 0.9981 | 0.9943 | 0.9933 | 0.9932  | 0.9932  |
| **Expected cost:**      |        |        |        |         |         |
| Screening               | $21,994 | $60,469 | $25,790 | $82,013 | $132,520 |
| Total                   | $14,183,610 | $34,542,493 | $47,285,676 | $64,961,683 | $83,421,454 |

| Workload (# screenings): |        |        |        |         |         |
| by admins (PC-PTSD)      | 8,999.98 | 25,828.22 | 30,893.70 | 32,103.35 | 32,410.12 |
| by clinicians (PCL)      | 499.57 | 1,331.66 | 1,687.45 | 1,918.78 | 1,993.00 |
| by trained clinicians (CAPS) | 2.52 | 6.75 | 8.55 | 9.69 | 10.05 |
| Sensitivity             | 0.1340 | 0.3188 | 0.4306 | 0.5379  | 0.5782  |
| Specificity             | 0.9993 | 0.9981 | 0.9976 | 0.9974  | 0.9974  |
| **Expected cost:**      |        |        |        |         |         |
| Screening               | $33,733 | $90,947 | $118,357 | $129,712 | $132,520 |
| Total                   | $15,301,406 | $40,813,516 | $61,731,975 | $105,332,802 | $181,515,773 |

| Workload (# screenings): |        |        |        |         |         |
| by admins (PC-PTSD)      | 9,000.64 | 24,764.01 | 32,037.39 | 34,327.71 | 34,729.19 |
| by clinicians (PCL)      | 951.19 | 2,545.40 | 3,319.82 | 3,668.55 | 3,761.20 |
| by trained clinicians (CAPS) | 4.76 | 12.74 | 16.62 | 18.35 | 18.81 |
| Sensitivity             | 0.2225 | 0.5298 | 0.7153 | 0.8931  | 0.9602  |
| Specificity             | 0.9946 | 0.9848 | 0.9804 | 0.9794  | 0.9793  |
| **Expected cost:**      |        |        |        |         |         |
| Screening               | $33,733 | $90,956 | $118,348 | $129,718 | $132,533 |
| Total                   | $14,334,987 | $34,437,863 | $46,407,750 | $60,972,252 | $71,671,584 |

| Workload (# screenings): |        |        |        |         |         |
| by admins (PC-PTSD)      | 8,999.60 | 24,763.53 | 32,034.03 | 34,325.83 | 34,727.65 |
| by clinicians (PCL)      | 951.01 | 2,545.40 | 3,319.43 | 3,668.59 | 3,761.40 |
| by trained clinicians (CAPS) | 4.79 | 12.76 | 16.64 | 18.39 | 18.85 |
| Sensitivity             | 0.2700 | 0.6111 | 0.7927 | 0.9407  | 0.9863  |
| Specificity             | 1.0000 | 1.0000 | 1.0000 | 1.0000  | 1.0000  |
| **Expected cost:**      |        |        |        |         |         |
| Screening               | $325,244 | $895,878 | $1,067,509 | $1,123,298 | $1,138,391 |
| Total                   | $13,135,053 | $29,882,365 | $38,675,902 | $47,948,290 | $52,893,837 |

| Workload (# screenings): |        |        |        |         |         |
| by clinicians (PCL)      | 8,998.83 | 25,724.37 | 30,194.43 | 31,014.51 | 31,197.53 |
| by trained clinicians (CAPS) | 501.37 | 1,263.84 | 1,563.24 | 1,739.68 | 1,792.26 |

*Per 10,000 patients
2.4.2 Effect of False Diagnosis Costs on the Total Cost of Screening Protocols

The total cost includes false-negative and false-positive costs, which are subject to considerable uncertainty and subjective discussion. Even though a range of possible cost estimates, given in Table 2-2, are considered to be as realistic as possible, this uncertainty still is a limitation and the exploration of how the overall system is affected by the changes in these costs is especially important. To this end, Figure 2-14 illustrates the effect of the ratio between false-negative and false-positive costs on the total cost under the assumption of protocol 1 with various stopping criteria, indicating that the false-negative cost has more influence on the total cost than the false-positive cost. The total cost significantly increases as the false-negative cost increases, but the same increase in the false-positive cost does not make any significant change.

![Figure 2-14 Effect of false-negative/false-positive cost ratio on total cost assuming protocol 1 with various stopping criteria](image)

2.4.3 Effect of Compliance Rates on Overall System Accuracy and Cost

The patient compliance rates in the above examples (Table 2-1) were determined based on discussions with several psychiatrists in the VA and intended to be as realistic as possible.
However, it is important to emphasize that these rates are not known with certainty and that it is necessary to explore how the system accuracy and cost are affected by various compliance values. For this purpose, several compliance rate scenarios, given in Table 2-5, were developed based on currently available data. The pessimistic and optimistic scenarios represent the lower and upper bounds, respectively, while realistic scenario reflects the current situation (i.e., the one used in the previous examples). Also shown for comparison is the “perfect” situation which assumes that all patients fully comply with each test (i.e., patients take the tests when it is required).

<table>
<thead>
<tr>
<th>Scenario</th>
<th>Compliance rates</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PC-PTSD ($r_5$)</td>
</tr>
<tr>
<td>Pessimistic (intolerable)</td>
<td>0.75</td>
</tr>
<tr>
<td>Realistic (current)</td>
<td>0.90</td>
</tr>
<tr>
<td>Optimistic (ideal)</td>
<td>1.00</td>
</tr>
<tr>
<td>Perfect (unrealistic)</td>
<td>1.00</td>
</tr>
</tbody>
</table>

Figure 2-15 and 2-16 illustrate the effect of patient compliance on the overall system sensitivity, specificity, and screening and total cost under the assumption of protocol 1 with discontinued screening after three negative results, indicating that, as expected, the number of false-negatives and thus total cost decrease as each test’s compliance rate increases. The number of false-positives and the cost of screening, however, increase significantly as patients comply more with each test, primarily because of conducting more PCL tests (note that only 10% of patients who take PCL cannot be determined as positive or negative and are tested with CAPS which is considered the gold standard for PTSD diagnosis). More specifically, improving patient compliance results in up to 99.18% and 91.48% decrease in the number of false-negatives and
the total cost, while generating an increase up to 268.62% and 324.53% in the number of false-positives and the cost of screening, respectively. Of note, even though these increases seem quite high, the number of false-positives and the cost of screening per year are still 7.41% and 97.33% less on average, respectively, when compared to the current process, even assuming the “perfect” patient compliance which generates the biggest increases.

Figure 2-15 Effect of compliance rates on expected (a) number of false-negatives and (b) number of false-positives assuming protocol 1 with discontinued screening after three negative results
2.4.4 Effect of PTSD Prevalence on the Predictive Value of Screening Protocols

The utility of a screening test and the rationale behind the screening programs are strongly determined by the prevalence of the disease in the population (Harris et al., 2001). Therefore, the performance of the screening protocols for various PTSD prevalence rates also is analyzed. For instance, Figure 2-17 illustrates the changes in the positive and negative predictive values of protocol 1 with discontinued screening after three negatives for a range of PTSD population
prevalence (i.e., from low to very high). As expected, using the same protocol in a population with higher prevalence increases the positive predictive value significantly while resulting in a slightly decreased negative predictive value. In other words, even though the corresponding protocol has good performance characteristics (i.e., sensitivity of 0.91 and specificity of 0.99), its positive predictive value is fairly low when applied to a population with lower prevalence, suggesting that the VA is an ideal place to evaluate the PTSD screening protocols given the higher prevalence when compared to the general population (Kessler et al., 2005; Magruder et al., 2005; Seal et al., 2007; Shiner et al., 2012).

![Figure 2-17 Effect of PTSD prevalence on predictive value of protocol 1 with discontinued screening after three negative results](image)

2.5 Other Possible Applications: A Sequential Screening Model for Mild TBI

While the main focus of this chapter is PTSD, similar sequential screening models can be developed for a range of other mental health conditions. In order to give an insight into how the general approach is applied to other disorders, this section illustrates a sequential screening probability model developed for mild traumatic brain injury (TBI). As one of the common
military injuries which occur with no visible symptoms and may go unrecognized and untreated, mild TBI would be a good example of how sequential screening can help reduce the number of unrecognized cases and improve the effectiveness and cost.

Similar to the above PTSD screening processes, the proposed mild TBI screening process consists of a number of screening tests which are conducted at predetermined time intervals (may vary depending on the extent of the injury) and a confirmation test which is used to follow-up on a positive determination of a screening test. As usual, the confirmation test tends to be more thorough and expensive than the screening tests and aims to confirm any positive diagnosis made by the screening tests. Confirmation of a negative diagnosis also might be necessary, especially in some settings such as combat zones given that the complicated nature and strict restrictions of war conditions. To this end, a decision stage is included in the screening process, which decides whether or not a patient should take the confirmation test even though he/she is determined as negative by all screening tests. Given that the soldiers work as teams (troops) and members of the same team are subject to the same or similar external effects that may cause TBI (i.e., mines, grenade blasts, etc.), the existence of a soldier with TBI in a small team is assumed to increase the likelihood of finding TBI on his teammates. Therefore, the screening process is extended such that the patients who are determined as negative by all screening tests are rescreened by the confirmation test if at least one of their teammates has diagnosed with TBI recently. The schematic presentation of the proposed screening process is given in Figure 2-18.
Mathematical expressions for overall accuracy and cost of mild TBI screening process can be built using joint, conditional, and expected value probability laws, similar to those for the PTSD screening process. Again, by considering all possible probabilistic manners in which a truly positive [negative] patient is correctly identified as positive [negative] and the expected number of screenings and false diagnoses, a mathematical model of the overall sensitivity, specificity, and expected total cost was developed under the assumption of both non-homogeneous and homogeneous screening tests, as summarized in Table 2-6. Derivation of these mathematical expressions is given in Appendix D along with the model assumptions and notation.
Table 2-6 Mathematical expressions of overall sensitivity, specificity, and cost for mild TBI screening model

<table>
<thead>
<tr>
<th>Overall sensitivity</th>
<th>Overall specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Expected total cost</td>
<td></td>
</tr>
</tbody>
</table>

Non-homogeneous screening tests

\[ p_j = \left(1 - \prod_{i=1}^{j} a_i (1 - r_i)\right) (1 - a_c) \]
\[ q_j = 1 - \beta_c \left(1 - \prod_{i=1}^{j} (1 - \beta_i) (1 - r_j)\right) \]
\[ EC_j = k_p \left[p \left(\frac{1 - a_j}{1 - a_c}\right) + (1 - p) \left(\frac{1 - (1 - \beta_j)}{\beta_c}\right)\right] + k_n \left(1 - \left(p \prod_{i=1}^{j} a_i (1 - r_i) + (1 - p) \prod_{i=1}^{j-1} (1 - \beta_i) (1 - r_j)\right)\right] + k_{FPn} \left(1 - \beta_c \right) \left(1 - \prod_{i=1}^{j} (1 - \beta_i) (1 - r_j)\right) + k_{FN} np \left(\alpha_c + (1 - \alpha_c) \prod_{i=1}^{j} a_i (1 - r_j)\right) \]

Homogeneous screening tests

\[ p_j = \left(1 - \alpha_j (1 - r_j)\right) (1 - a_c) \]
\[ q_j = 1 - \beta_c \left(1 - \left(1 - \beta_s\right) (1 - r_j)\right) \]
\[ EC_j = k_p \left[p \left(\frac{1 - a_j}{1 - a_c}\right) + (1 - p) \left(\frac{1 - (1 - \beta_j)}{\beta_c}\right)\right] + k_n \left(1 - \left(p \prod_{i=1}^{j} a_i (1 - r_i) + (1 - p) \prod_{i=1}^{j-1} (1 - \beta_i) (1 - r_j)\right)\right] + k_{FPn} \left(1 - \beta_c \right) \left(1 - \prod_{i=1}^{j} (1 - \beta_i) (1 - r_j)\right) + k_{FN} np \left(\alpha_c + (1 - \alpha_c) a_j / (1 - r_j)\right) \]

where \((1 - a_i)\) and \((1 - a_c)\) are the sensitivities and \((1 - \beta_i)\) and \((1 - \beta_s)\) are the specificities for each screening test \(i\) and the confirmation test, \(p\) is the population incidence rate, \(n\) is the number of patients to be screened, \(r_i\) and \(r_j\) are the rescreening rates, and \(k_n, k_c, k_{FN}, k_{FPn}\) is the cost of a screening test, the confirmation test, a false-positive result, and a false-negative result.

Similar to the case of PTSD screening, the above expressions can be used to more precisely examine the joint effect of various test accuracies and number of screening tests on the overall system sensitivity, specificity, and total cost. Figure 2-19, for example, illustrates the competing effect of additional screenings on overall sensitivity and specificity, and the rate at which they approach their respective bounds (confirmation test sensitivity and specificity which need not be the same), given an incidence of \(p = 0.15\) and rescreening rates of \(r_1 = 0.40\) and \(r_2 = 0.10\) and assuming that, for simplicity, all screening tests have identical sensitivities and specificities.

Note that the confirmation test sensitivity is an upper bound on the overall system sensitivity as the number of screening tests and/or their sensitivities increase. Similarly, the overall system specificity is bounded below by the confirmation test specificity. Thus, no screening policy of
this type, even with many near-perfect [very poor] screening tests, can have higher [lower] system sensitivity [specificity] than the final confirmation test (Benneyan and Kaminsky, 1996). As these examples illustrate, significant improvements in the probability of detecting truly positive patients clearly are possible in many cases by adding more screening tests. It, however, has the opposite effect of decreasing the probability of correctly classifying true-negatives. This effect is significantly greater as the inaccuracy of screening and confirmation tests increase.

Figure 2-19 Competing effect of number of screening tests on overall sensitivity versus specificity, and the role of convergence to bounds, where (a) both system bounds = 0.95 and (b) both system bounds = 0.80 ($p = 0.15$, $r_1 = 0.40$, $r_2 = 0.10$; both system bounds = $1 - \alpha_c = 1 - \beta_c$)
The expected cost model also can be used to evaluate a variety of screening policies for any set of values for test accuracies ($\alpha_s$, $\alpha_c$, $\beta_s$, $\beta_c$), test and false-diagnosis costs ($k_s$, $k_c$, $k_{FP}$, $k_{FN}$), rescreening rates ($r_1$, $r_2$), incidence rate ($p$), and number of patients to be screened ($n$). Moreover, the optimal number of screening tests ($j^*$) which minimizes the expected total cost can be found, such as by a simple search algorithm or more advanced optimization methods. In order to illustrate this use and the significant sensitivity of the optimal policy to input assumptions, Figure 2-20 compares the expected total costs for three different scenarios based on test accuracies where the following values were assumed: $k_s = $5, $k_c = $25, $k_{FP} = $750, $k_{FN} = $5000, $r_1 = 0.40, r_2 = 0.10, p = 0.15, and $n = 100.

![Figure 2-20 Comparison of expected total cost per 100 patients](image)

The above examples help illustrate the general approach and explore how the overall system is affected by the changes in various input assumptions. The figures used in these examples, however, are only for illustrative purposes and do not reflect the actual practice. Additional research is needed, for example, to better identify the screening and confirmation tests that can be used appropriately in the sequential screening process and then to estimate their accuracies.
and costs. Furthermore, a comprehensive sensitivity analysis, such as presented in Section 2.4, should be conducted especially for the inputs that are subject to considerable uncertainty and subjective discussion (e.g., false diagnosis costs).

2.6 Discussion

The design of any screening protocol should depend on the ability to detect truly positive and truly negative patients and all costs incurred as a result of this protocol. The mathematical models presented in this chapter can help policy makers, clinicians, and researchers understand these characteristics for various possible protocols in order to design more effective and minimum cost screening procedures for any particular situation. While the main focus of this study is PTSD, a focus on optimally sequencing screening tests is likely warranted in a range of other mental health disorders, such as illustrated in Section 2.5.

Given that untreated and undertreated PTSD result in significant physical and social problems as well as tremendous medical and social costs, a well-designed and effective screening program is vital in properly identifying and treating PTSD patients. This study investigated the current screening processes and some alternatives for screening PTSD within the VHA and further explored the effects of annual screenings, patient compliance, and prevalence on the expected values, variances, and probability distributions of false-negatives, false-positives, and incurring costs. Results indicate that a more intentionally designed system, which consists of a series of annual screening tests along with a standardized confirmatory testing, results in lower false diagnosis rates, predictable performance, and reduced costs. More specifically, an average of 55.07% and 18.10% decrease in the number of false-negative and false-positive diagnoses along
with a savings of 32.96% in total cost per year were shown to be possible by performing a standardized confirmatory testing (with the use of PCL and CAPS) to follow-up the positive results of the annual PC-PTSD screenings. These results, however, are very dependent on the model inputs. A scenario analysis, therefore, was performed in order to illustrate how the overall results are affected by various values of several model parameters, such as screening test sensitivities and specificities, false-diagnosis costs, patient compliance rates, and prevalence. This analysis also is useful to help decision makers explore the tradeoffs between annual screenings and false-negative rates, false-positive rates, and costs.

One of the possible limitations of this study is that the most of the model inputs are not known with certainty. This is especially true with respect to the costs of false-negative and false-positive results and patient compliance rates. Although all input estimations were based on discussions with several clinicians who have a thorough knowledge of PTSD and a sensitivity analysis was performed for a range of possible values, additional research could be useful to confirm these estimates or to better estimate the aforementioned inputs. Another limitation is the assumption that the VA has the capacity to perform the confirmatory screening and to provide the treatment that would be recommended after the screening.
Chapter 3 - Multi-State Categorical Diagnostic Models

Traumatic brain injury (TBI), one of the most common military injuries, is an enormous problem facing most U.S. soldiers today. Given that the vast majority of TBI diagnoses among military personnel are categorized as mild, many cases go undetected and untreated, resulting in significant psychological disorders, long-term disabilities, and economic burdens. Assessment tools for TBI, which are found to be efficient when the extent of the injury is more severe, often are insufficient when there are no obvious signs of trauma. This chapter presents several predictive diagnostic models to determine whether or not an individual has TBI and to categorize him into the most likely severity state. These models aim to improve the reliability and accuracy of TBI diagnoses by combining the results of different screening tests, which are conducted consecutively over time, with the use of several mathematical techniques, namely, fuzzy logic, logistic regression, and neural networks.

3.1 Background

Military personnel and others working in combat zones are at particular risk of traumatic brain injury (TBI) due to the fact that they are likely to be exposed to events that may cause TBI, such as concussive blasts or explosions (Champion et al., 2009; Ling et al., 2009; Defense and Veterans Brain Injury Center, 2013). In combat zones, TBI often occurs in tandem with
ostensibly life-threatening injuries; therefore, cases may go unrecognized, potentially putting the individuals or their units at risk (Butler et al., 2008; Tanielian and Jaycox, 2008). Additionally, when TBI occurs with no outward signs of trauma (i.e., mild cases), service members might not seek medical treatment (Martin et al., 2008; McCrea, 2008). Many TBI sufferers, especially if untreated, may experience medical, behavioral, and social problems for a long time, perhaps a lifetime. Any delays in treatment may compromise recovery or result in serious cognitive, physical, and/or psychological problems and even permanent damage (Rao and Lyketsos, 2000; Maas et al., 2008; Anderson-Barnes et al., 2010; Faul et al., 2010).

The existence and severity of TBI are evaluated in several ways, including subjective assessments, external signs, physical experiences, neurologic assessments (e.g., the Glasgow Coma Scores (GCS)), and clinical assessments (e.g., computed tomography (CT) scans, magnetic resonance imaging (MRI), and electroencephalogram (EEG)) (U.S. Department of Health and Human Services, 2006; Elder et al., 2010; Maruta et al., 2010). In the military, testing typically follows a process of screening, identification, evaluation, confirmation, and stratification by severity (Butler et al., 2008). While ideally all TBI would be detected soon after the time of injury, initial tests tend to be less accurate than more thorough and expensive post-deployment tests. GCS, developed by Teasdale and Jennett (1974), for instance, is a valuable first assessment of TBI severity and prognosis immediately after injury. However, it has several practical limitations (Wijdicks et al., 2005; Matis and Birbilis, 2008; Zuercher et al., 2009; Middleton, 2012), generally due to the fact that often it is not measured in emergency rooms or in hospitals where TBI patients are first transported (Thatcher et al., 2001). Conventional CT and MRI, the most frequent imaging modalities used in acute diagnosis and management, often cannot detect mild TBIs since generally the brain appears quite normal (Dhawan et al., 2006;
Shenton et al., 2012). EEG, on the other hand, is recommended to detect mild TBI even though it is not considered sensitive enough to distinguish mild and moderate TBIs (Thatcher, 2000; Haneef et al., 2013). In practice more than one method can be used for diagnosis to assure reliability and accuracy. Guler et al. (2008), for instance, developed a diagnostic system for determining the severity of civilian TBI by using fuzzy logic to combine information from GCS and EEG results. They found a significant relationship between the findings of neurologists and the systems output for normal, mild, and severe electroencephalography tracing data. In the study, fuzzy inferences were developed from GCS and EEG trauma score data, which were then used to produce a trauma severity degree and final diagnosis for each patient. However, neither an analysis on the accuracy or optimization of their approach nor a discussion on the longitudinal nature of increasingly accurate data was presented.

This study proposes a predictive modeling approach using fuzzy logic, multinomial logistic regression, and neural networks, which provides a comprehensive screening procedure as illustrated in Figure 3-1. The main concern of this procedure is to better identify mild cases which occur without any visible symptoms and to prevent any delays in treatment. Also, in the proposed procedure, the determination can change continuously as each screening test result becomes available, presumably getting more accurate over time, given that the symptoms of mild injuries may not be immediately recognized.
3.2 Fuzzy Categorical Model

3.2.1 Methodology

Fuzzy logic (FL), which was first introduced by Lotfi A. Zadeh (1965), is a computational paradigm for processing information in a way that resembles human reasoning in the presence of uncertainty (Ross, 1995). It has been successfully used in healthcare applications, including Veryha and Adamczyk (2005) in primary dental care services, Dalalah and Magableh (2008) for remote health service delivery, Guler et al. (2008) for detecting civilian TBI severity, Khan et al. (2008) for cancer prognosis, Medjahed et al. (2012) for remote healthcare monitoring, Torshabi et al. (2012) for tumor motion prediction, Czabanski et al. (2013) for classification of fetal heart rate tracings, and Kwiatkowska and Kielan (2013) for assessment of clinical depression.

Fuzzy logic emulates a person’s ability to categorize things by assessing the degree of membership rather than evaluating whether they satisfy some unambiguous definition (Lootsma, 1997). A fuzzy variable (here TBI severity categories, i.e., normal, mild, moderate, and severe) is characterized by its name tag, a set of fuzzy values (also known as linguistic values or labels),
and the membership function of these labels. These membership functions assign a membership value, \( \mu_{\text{Label}}(x) \), to a given real value \( x \) within some predefined range, \( R \) (known as the universe of discourse). Two basic fuzzy operations, “\( \text{and} \)” and “\( \text{or} \)”, are defined as (Guler et al., 2002):

\[
\text{Definition 1: } \mu_{A \text{ and } B}(x) = \mu_A(x) \land \mu_B(x) = \min\{\mu_A(x), \mu_B(x)\}
\]

and

\[
\text{Definition 2: } \mu_{A \text{ or } B}(x) = \mu_A(x) \lor \mu_B(x) = \max\{\mu_A(x), \mu_B(x)\},
\]

where \( A \) and \( B \) are fuzzy variables. Using these operators, fuzzy variables can be combined to form fuzzy-logic expressions.

The general fuzzy logic algorithm consists of three activities: fuzzification, fuzzy inference, and defuzzification. Fuzzification is the process of making a crisp numeric quantity fuzzy (Ross, 1995) by mapping it onto fuzzy membership functions. Typically, conventional trapezoid shapes, such as those shown in Figure 3-2, are used to calculate fuzzy membership functions or degrees of membership for each category as

\[
\mu(x_i) = \begin{cases} 
\frac{x_i - a}{b - a}, & a \leq x_i \leq b \\
1, & b < x_i < c \\
\frac{d - x_i}{d - c}, & c \leq x_i \leq d, \\
0, & \text{otherwise}
\end{cases}
\]  

(3-1)

where \( a, b, c, \) and \( d \) are the defining parameters of each membership function.
The second step, fuzzy inference, is the process of obtaining a fuzzy output using a rule-base that combines several inputs (here several screening tests), i.e., the results of the above fuzzification linguistic inputs. The rules in the rule-base define the connection between input and output fuzzy variables (Guler et al., 2002). The relationships and outputs obtained from the rule-base are interpreted using “and” and “or” operators (Guler et al., 2008).

Finally, defuzzification is the conversion of these fuzzy quantities to a crisp value $z^*$. Several methods exist for combining and defuzzifying fuzzy membership values, including the max-membership, center of gravity (COG), weighted average (WA), mean-of-maxima membership (MOM), center of sums (COS), center of largest (COL) area, and smallest or largest of maxima (SOM or LOM) methods (Ross, 1995). By example, maxima methods choose the set with the maximum membership, whereas in the COG method, a common and useful approach, the center of gravity of the area under the membership functions is calculated as
\[ z^* = \frac{\int \mu_C(x) \cdot x \, dx}{\int \mu_C(x) \, dx} \] \hspace{1cm} (3-2)

where \( C \) is the union of the membership functions. First, the results of the rules are added together and the top parts of the graphs are chopped off to form trapezoids. All of these trapezoids are then superimposed one over another, forming a single geometric shape. The centroid of this shape, called the fuzzy centroid, is then calculated and its \( x \) coordinate gives the defuzzified value.

As an extension to Guler et al. (2008), a more comprehensive procedure was developed by sequentially applying fuzzy inference longitudinally to a series of screening results as they become available. Applying a fuzzy diagnostic system sequentially \( N \) times (i.e., to \( N \) screening test results) eventually produces \( N \) trauma severity degree values \((X_1, X_2, \ldots, X_N)\) for each patient, which become the inputs to the fuzzification logic. Without the loss of generality, these values were assumed to be between 0 and 100 based on the expert-based scale from Guler et al. (2008). Four fuzzy sets – normal (no TBI), mild, moderate, and severe – were formed for these values using the four trapezoid membership functions shown in Figure 3-2 (which are different for each screening test, resulting in \( 4N \) membership functions). Note that the parameters of each membership function \((a_i, b_i, c_i, \text{ and } d_i)\) for fuzzy set \( i \), where \( i = \text{No(rmal), Mi(ld), Mo(derate),} \) and \( \text{Se(vere)} \)) typically might be determined by experts or with an optimization tuning routine, such as demonstrated later. Membership degrees for each category (i.e., for each of the four trauma severities) were then computed using Eq. (3-1).
For the fuzzy inference step, a rule-base which consists of four linguistic outputs – normal, mild, moderate, and severe – the same as the membership functions was used. Table 3-1 illustrates the rule-base for the case with two screening results \((N = 2)\). As seen, after the second screening there are \(4^2 = 16\) rules. To illustrate interpretation of this rule-base, if the output of screening test 1 is “mild” (column 2) and the output of screening test 2 is “moderate” (row 3), then the output of the fuzzy inference process is “moderate” (shown in italic font). According to the rule-base and using min-max operators, the fuzzy outputs of the inference process were obtained and represented by \(\mu^i_s(X_s)\), which means that according to the result \(X_s\) of screening test \(s\), a patient belongs to fuzzy set \(i\) with a membership degree of \(\mu^i_s\) evaluated at the value \(X_s\), where \(s = 1, 2, \ldots, N\) and \(i = \text{No, Mi, Mo, Se}\). These fuzzy values were then converted into crisp values using the defuzzification methods mentioned above.

### Table 3-1 Example of a rule-base for \(N = 2\) screening test case

<table>
<thead>
<tr>
<th>The output of the inference</th>
<th>Output of screening test 1</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Normal</td>
</tr>
<tr>
<td>Output of screening test 2</td>
<td>Normal</td>
</tr>
<tr>
<td>Mild</td>
<td>Mild</td>
</tr>
<tr>
<td>Moderate</td>
<td>Moderate</td>
</tr>
<tr>
<td>Severe</td>
<td>Moderate</td>
</tr>
</tbody>
</table>

### 3.2.2 Numerical Example

Consider the \(N = 4\) case with four screening tests, where the parameters of each membership function are given as \(a_{\text{No}} = b_{\text{No}} = 0, c_{\text{No}} = 15, d_{\text{No}} = 40, a_{\text{Mi}} = 15, b_{\text{Mi}} = 30, c_{\text{Mi}} = 40, d_{\text{Mi}} = 50, a_{\text{Mo}} = 30, b_{\text{Mo}} = c_{\text{Mo}} = 50, d_{\text{Mo}} = 75, a_{\text{Se}} = 50, b_{\text{Se}} = 70, \) and \(c_{\text{Se}} = 80\). For simplicity, the membership functions were assumed to have the same parameters for all four tests, although in
reality they likely would be different. The fuzzy categorical model was applied to categorize a patient with the trauma severity degrees of $X_1 = 45$, $X_2 = 56$, $X_3 = 65$, and $X_4 = 20$ obtained from the screening tests. According to the fuzzy membership functions and using Eq. (3-1), the membership degrees for each TBI category according to the screening test 1 were calculated as

$$
\mu_{Mi}^1(45) = \frac{d_{Mi} - X_1}{d_{Mi} - c_{Mi}} = \frac{50 - 45}{50 - 40} = 0.50
$$

and

$$
\mu_{Mo}^1(45) = \frac{X_1 - a_{Mo}}{b_{Mo} - a_{Mo}} = \frac{45 - 30}{50 - 30} = 0.75.
$$

Similarly, the membership degrees for each category according to the screening tests 2, 3, and 4 were calculated as $\mu_{Mo}^2(56) = 0.76$ and $\mu_{Se}^2(56) = 0.30$, $\mu_{Mo}^3(65) = 0.48$ and $\mu_{Se}^3(65) = 0.65$, and $\mu_{No}^4(20) = 0.80$ and $\mu_{Mi}^4(20) = 0.33$.

With four tests, now there are $4^4 = 256$ rules, although only $2^4 = 16$ of these rules are needed in this example (listed in Table 3-2) since each trauma severity degree has a membership degree different than zero for only two of the categories. Using this rule-base and min-max operators, this patient was determined to belong to the “mild” and “moderate” fuzzy sets with membership degrees from rules 1-2 and rules 3-16, respectively, of

$$
\mu_{Mi} = \max \left\{ \min \left( \mu_{Mi}^1(X_1), \ldots, \mu_{No}^4(X_4) \right), \min \left( \mu_{Mi}^1(X_1), \ldots, \mu_{Mi}^4(X_4) \right) \right\}
$$

$$
= \max \{ \min(0.5, 0.76, 0.48, 0.80), \min(0.5, 0.76, 0.48, 0.33) \} = 0.48
$$
and

\[ \mu_{Mo} = \max \left\{ \min \left( \mu_{M_1}^1(X_1), \ldots, \mu_{M_4}^4(X_4) \right), \min \left( \mu_{N_1}^1(X_1), \ldots, \mu_{N_4}^4(X_4) \right) \right\} \]
\[ = \max \{ \min(0.5, 0.76, 0.65, 0.80), \min(0.75, 0.30, 0.65, 0.33) \} = 0.65. \]

<table>
<thead>
<tr>
<th>Rule</th>
<th>Output of scr. test 1</th>
<th>Output of scr. test 2</th>
<th>Output of scr. test 3</th>
<th>Output of scr. test 4</th>
<th>Output of inference</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Mild</td>
<td>Moderate</td>
<td>Moderate</td>
<td>Normal</td>
<td>Mild</td>
</tr>
<tr>
<td>2</td>
<td>Mild</td>
<td>Moderate</td>
<td>Moderate</td>
<td>Mild</td>
<td>Moderate</td>
</tr>
<tr>
<td>3</td>
<td>Mild</td>
<td>Moderate</td>
<td>Severe</td>
<td>Normal</td>
<td>Moderate</td>
</tr>
<tr>
<td>4</td>
<td>Mild</td>
<td>Moderate</td>
<td>Severe</td>
<td>Mild</td>
<td>Moderate</td>
</tr>
<tr>
<td>5</td>
<td>Mild</td>
<td>Severe</td>
<td>Moderate</td>
<td>Normal</td>
<td>Moderate</td>
</tr>
<tr>
<td>6</td>
<td>Mild</td>
<td>Severe</td>
<td>Moderate</td>
<td>Mild</td>
<td>Moderate</td>
</tr>
<tr>
<td>7</td>
<td>Mild</td>
<td>Severe</td>
<td>Severe</td>
<td>Normal</td>
<td>Moderate</td>
</tr>
<tr>
<td>8</td>
<td>Mild</td>
<td>Severe</td>
<td>Severe</td>
<td>Mild</td>
<td>Moderate</td>
</tr>
<tr>
<td>9</td>
<td>Moderate</td>
<td>Moderate</td>
<td>Moderate</td>
<td>Normal</td>
<td>Moderate</td>
</tr>
<tr>
<td>10</td>
<td>Moderate</td>
<td>Moderate</td>
<td>Moderate</td>
<td>Mild</td>
<td>Moderate</td>
</tr>
<tr>
<td>11</td>
<td>Moderate</td>
<td>Moderate</td>
<td>Severe</td>
<td>Normal</td>
<td>Moderate</td>
</tr>
<tr>
<td>12</td>
<td>Moderate</td>
<td>Moderate</td>
<td>Severe</td>
<td>Mild</td>
<td>Moderate</td>
</tr>
<tr>
<td>13</td>
<td>Moderate</td>
<td>Severe</td>
<td>Moderate</td>
<td>Normal</td>
<td>Moderate</td>
</tr>
<tr>
<td>14</td>
<td>Moderate</td>
<td>Severe</td>
<td>Moderate</td>
<td>Mild</td>
<td>Moderate</td>
</tr>
<tr>
<td>15</td>
<td>Moderate</td>
<td>Severe</td>
<td>Severe</td>
<td>Normal</td>
<td>Moderate</td>
</tr>
<tr>
<td>16</td>
<td>Moderate</td>
<td>Severe</td>
<td>Severe</td>
<td>Mild</td>
<td>Moderate</td>
</tr>
</tbody>
</table>

Next, these fuzzy values were converted into a crisp value using any of the defuzzification methods mentioned above. The present study compared the results of four methods – COG, MOM, SOM, and LOM. Using the COG (see Eq. (3-2)) and MOM methods, the final trauma severity degrees were calculated to be 44.75 (shown below) and 50.88, respectively, both of which indicate a final diagnosis of “moderate trauma” using the expert-based scale from Guler et al. (2008). Using the SOM and LOM methods, however, the final trauma severity degrees were 43.00 and 58.75, respectively, corresponding to “moderate” and “severe” trauma on this scale.
Chapter 3 - Multi-State Categorical Diagnostic Models

\[ Q_{\text{COG}}^{\text{final}} = \frac{\int_{15}^{22.2} \left( x - 1 \right) dx + \int_{39.6}^{43} \left( x - \frac{3}{2} \right) dx + \int_{58.75}^{75} \left( x - \frac{5}{2} + 3 \right) dx}{\int_{15}^{22.2} \left( x \right) dx + \int_{39.6}^{43} \left( x \right) dx + \int_{58.75}^{75} \left( x \right) dx} = 44.75. \]

Table 3-3 illustrates how the final diagnosis changes over time as the result of each screening test is incorporated and as each defuzzification method is used. To automate this process, the predicted health states for each patient can be determined using MATLAB’s fuzzy logic tool.

<table>
<thead>
<tr>
<th>Time</th>
<th>Screening test 1</th>
<th>Screening test 2</th>
<th>Screening test 3</th>
<th>Screening test 4</th>
<th>Final diagnosis</th>
</tr>
</thead>
<tbody>
<tr>
<td>( t_1 )</td>
<td>45</td>
<td>(not yet available)</td>
<td>(not yet available)</td>
<td>(not yet available)</td>
<td>Moderate</td>
</tr>
<tr>
<td>( t_2 )</td>
<td>45</td>
<td>56</td>
<td>(not yet available)</td>
<td>(not yet available)</td>
<td>Severe</td>
</tr>
<tr>
<td>( t_3 )</td>
<td>45</td>
<td>56</td>
<td>63</td>
<td>(not yet available)</td>
<td>Severe</td>
</tr>
<tr>
<td>( t_4 )</td>
<td>45</td>
<td>56</td>
<td>63</td>
<td>20</td>
<td>Moderate</td>
</tr>
</tbody>
</table>

3.2.3 Model Accuracy

The predicted diagnoses for a large cohort of patients were compared to their ultimate known states which were determined via expert consensus. Table 3-4 illustrates a portion of a hypothetical dataset with 100 patients. Using the \( N = 4 \) case with four screening tests and membership functions from the previous section, Figure 3-3 summarizes the correlation between the actual and predicted TBI states. 77% of all results are correctly classified using all 4 tests, with a correlation coefficient of \( R = 0.837047 \), indicating a fairly strong agreement.
Table 3-4 Illustrative portion of evaluation dataset

<table>
<thead>
<tr>
<th>Patient #</th>
<th>Actual health state</th>
<th>Result of scr. test 1 ($X_1$)</th>
<th>Result of scr. test 2 ($X_2$)</th>
<th>Result of scr. test 3 ($X_3$)</th>
<th>Result of scr. test 4 ($X_4$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>4</td>
<td>54</td>
<td>84</td>
<td>78</td>
<td>51</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>70</td>
<td>84</td>
<td>44</td>
<td>21</td>
</tr>
<tr>
<td>3</td>
<td>4</td>
<td>52</td>
<td>68</td>
<td>45</td>
<td>79</td>
</tr>
<tr>
<td>4</td>
<td>3</td>
<td>50</td>
<td>79</td>
<td>2</td>
<td>46</td>
</tr>
<tr>
<td>5</td>
<td>2</td>
<td>10</td>
<td>18</td>
<td>33</td>
<td>54</td>
</tr>
<tr>
<td>6</td>
<td>4</td>
<td>80</td>
<td>61</td>
<td>64</td>
<td>47</td>
</tr>
<tr>
<td>7</td>
<td>3</td>
<td>77</td>
<td>36</td>
<td>44</td>
<td>54</td>
</tr>
<tr>
<td>8</td>
<td>2</td>
<td>39</td>
<td>41</td>
<td>35</td>
<td>11</td>
</tr>
<tr>
<td>9</td>
<td>2</td>
<td>10</td>
<td>24</td>
<td>12</td>
<td>10</td>
</tr>
<tr>
<td>10</td>
<td>3</td>
<td>69</td>
<td>7</td>
<td>21</td>
<td>40</td>
</tr>
<tr>
<td>11</td>
<td>2</td>
<td>20</td>
<td>42</td>
<td>34</td>
<td>9</td>
</tr>
<tr>
<td>12</td>
<td>4</td>
<td>50</td>
<td>84</td>
<td>65</td>
<td>41</td>
</tr>
<tr>
<td>13</td>
<td>3</td>
<td>49</td>
<td>17</td>
<td>20</td>
<td>25</td>
</tr>
<tr>
<td>14</td>
<td>4</td>
<td>81</td>
<td>75</td>
<td>86</td>
<td>88</td>
</tr>
<tr>
<td>15</td>
<td>4</td>
<td>54</td>
<td>64</td>
<td>74</td>
<td>53</td>
</tr>
</tbody>
</table>

Four categories for the health state of a patient: 1, 2, 3, and 4, representing “Normal”, “Mild”, “Moderate”, and “Severe” states, respectively.

Figure 3-3 Accuracy of fuzzy categorical model results over time
(The larger the bubbles, the more accurate the estimations)
Chapter 3 - Multi-State Categorical Diagnostic Models

3.3 Multinomial Logistic Regression Categorical Model

3.3.1 Methodology

Multinomial logistic regression (MLR) generalizes logistic regression to allow more than two discrete outcomes (i.e., categories) for the dependent variable (Wang, 2005). Its performance has been extensively discussed in the healthcare literature, including these by Haas-Wilson and Savoca (1990), Hossain et al. (2002), Peng and Nichols (2003), Zhu et al. (2010), and Upadhyaya et al. (2013).

MLR method builds a separate equation for each category and each equation provides a predicted probability of being in the corresponding or any lower category (Sentas et al., 2005). Here, it was used to categorize patients into the most probable state of severity by predicting cumulative probabilities for each severity state. The prediction equations are of the general form

\[ l(\pi_j) = b_0^{(j)} + \sum_{i=1}^{N} b_i^{(j)} x_i, \]  

(3-3)

where \( \pi_j \) is the cumulative probability for the \( j \)th category, the \( b_0^{(j)} \) and \( b_i^{(j)} \) terms are the regression coefficients for the \( j \)th category, \( x_i \) is the \( i \)th predictor variable (test result), and \( N \) is the number of tests. After estimating the model coefficients via maximum likelihood, the cumulative probability for each category \( j \) is calculated by solving Eq. (3-3) for \( \pi_j \). Non-cumulative probabilities are calculated via subtracting, i.e., \( p_j = \pi_j - \pi_{j-1} \), and the category with maximum probability is selected as the final diagnosis (Norris et al., 2006; Sentas and Angelis, 2006; Asgary et al., 2007). As with any logistic regression, the model predicts a transformation...
$l(\pi_j)$ of the desired probabilities, rather than the probabilities themselves. The function $l(.)$ is called *link function* (Sentas et al., 2005), with the most commonly used link functions being

- a *logit* function, $\log\left(\frac{\pi_j}{1-\pi_j}\right)$, used for evenly distributed categories,
- a *complementary log-log* function, $\log\left(-\log(1-\pi_j)\right)$, used when higher categories are more probable,
- a *log-log* function, $-\log\left(-\log(\pi_j)\right)$, used when lower categories are more probable, and
- a *probit* function, $\phi^{-1}(\pi_j)$ where $\phi$ is the Gaussian cumulative probability density, used when latent variable is normally distributed (Chan, 2005).

As shown in Figure 3-4, these link functions differ in their degree of agreement depending on the value of $\pi$.

![Figure 3-4 A graphical comparison of link functions](image)

Since all categories are evenly distributed, i.e., none of the categories are more probable, here the logit function was used as the link function,
\[
\log\left(\frac{\pi_j}{1 - \pi_j}\right) = b_0^{(j)} + \sum_{i=1}^{N} b_i^{(j)} x_i, \quad (3-4)
\]

where \( j = 1, 2, \ldots, q-1 \). Thus the exact probabilities for each category are in the form of

\[
p_j = \frac{e^{b_0^{(j)} + \sum_{i=1}^{n} b_i^{(j)} x_i}}{1 + e^{b_0^{(j)} + \sum_{i=1}^{n} b_i^{(j)} x_i}} - \frac{e^{b_0^{(j-1)} + \sum_{i=1}^{n} b_i^{(j-1)} x_i}}{1 + e^{b_0^{(j-1)} + \sum_{i=1}^{n} b_i^{(j-1)} x_i}}, \quad (3-5)
\]

where \( j = 1, 2, \ldots, q-1 \), and

\[
p_q = 1 - \sum_{j=1}^{q-1} p_j. \quad (3-6)
\]

### 3.3.2 Numerical Example

Again using the \( N = 4 \) screening tests case and the same dataset from the previous section, the MLR coefficients were calculated via MINITAB to produce

\[
\log\left(\frac{\pi_1}{1 - \pi_1}\right) = 1.117 - 0.062Q_1 - 0.036Q_2 - 0.046Q_3 - 0.035Q_4,
\]

\[
\log\left(\frac{\pi_2}{1 - \pi_2}\right) = 5.757 - 0.062Q_1 - 0.036Q_2 - 0.046Q_3 - 0.035Q_4,
\]

and

\[
\log\left(\frac{\pi_3}{1 - \pi_3}\right) = 9.791 - 0.062Q_1 - 0.036Q_2 - 0.046Q_3 - 0.035Q_4.
\]
By solving these equations, the cumulative probabilities were calculated and then with subtraction, the probabilities \( p_1, p_2, p_3, \) and \( p_4 \) were obtained. By example, for the same patient in Section 3.2.2 the probability of the category 1 (Normal) was calculated as

\[
p_1 = \pi_1 = \frac{e^{1.117-0.062(45)-0.036(56)-0.046(63)-0.035(20)}}{1 + e^{1.117-0.062(45)-0.036(56)-0.046(63)-0.035(20)}} = 0.0006.
\]

Similarly, the probabilities of the category 2, 3, and 4 were calculated to be \( p_2 = \pi_2 - \pi_1 = 0.0636 \), \( p_3 = \pi_3 - \pi_2 = 0.7307 \), and \( p_4 = 1 - \sum_{j=1}^{3} p_j = 0.2050 \), respectively. Since the probability of category 3 is the greatest, this patient was given a final diagnosis of “Moderate” (i.e., category 3).

### 3.3.3 Model Accuracy

Using the same evaluation data set as in Section 3.2.3, Table 3-5 summarizes the category probabilities for the first 15 patients. Figure 3-5 compares the actual and predicted states over time, with 73% of patients being correctly classified after all four tests and a correlation coefficient of \( R = 0.799164 \) nearly as strong as the fuzzy categorical model.
Table 3-5 Category probabilities for MLR approach

<table>
<thead>
<tr>
<th>Patient #</th>
<th>Prob. of category 1 ($p_1$)</th>
<th>Prob. of category 2 ($p_2$)</th>
<th>Prob. of category 3 ($p_3$)</th>
<th>Prob. of category 4 ($p_4$)</th>
<th>Final classification</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.000023</td>
<td>0.002383</td>
<td>0.117400</td>
<td>0.880194</td>
<td>4</td>
</tr>
<tr>
<td>2</td>
<td>0.000118</td>
<td>0.011911</td>
<td>0.395217</td>
<td>0.592755</td>
<td>4</td>
</tr>
<tr>
<td>3</td>
<td>0.000081</td>
<td>0.008261</td>
<td>0.313562</td>
<td>0.678095</td>
<td>4</td>
</tr>
<tr>
<td>4</td>
<td>0.001416</td>
<td>0.126670</td>
<td>0.764269</td>
<td>0.107645</td>
<td>3</td>
</tr>
<tr>
<td>5</td>
<td>0.027531</td>
<td>0.718158</td>
<td>0.248304</td>
<td>0.006007</td>
<td>2</td>
</tr>
<tr>
<td>6</td>
<td>0.000023</td>
<td>0.002385</td>
<td>0.117503</td>
<td>0.880089</td>
<td>4</td>
</tr>
<tr>
<td>7</td>
<td>0.000137</td>
<td>0.013859</td>
<td>0.430769</td>
<td>0.555234</td>
<td>4</td>
</tr>
<tr>
<td>8</td>
<td>0.008200</td>
<td>0.453112</td>
<td>0.518414</td>
<td>0.020274</td>
<td>3</td>
</tr>
<tr>
<td>9</td>
<td>0.217870</td>
<td>0.748631</td>
<td>0.032886</td>
<td>0.000614</td>
<td>2</td>
</tr>
<tr>
<td>10</td>
<td>0.003030</td>
<td>0.236395</td>
<td>0.707282</td>
<td>0.053293</td>
<td>3</td>
</tr>
<tr>
<td>11</td>
<td>0.028387</td>
<td>0.723232</td>
<td>0.242559</td>
<td>0.005822</td>
<td>2</td>
</tr>
<tr>
<td>12</td>
<td>0.000077</td>
<td>0.007847</td>
<td>0.302782</td>
<td>0.689293</td>
<td>4</td>
</tr>
<tr>
<td>13</td>
<td>0.012815</td>
<td>0.560656</td>
<td>0.413521</td>
<td>0.013009</td>
<td>2</td>
</tr>
<tr>
<td>14</td>
<td>0.000001</td>
<td>0.000117</td>
<td>0.065171</td>
<td>0.993365</td>
<td>4</td>
</tr>
<tr>
<td>15</td>
<td>0.000054</td>
<td>0.005503</td>
<td>0.234165</td>
<td>0.760278</td>
<td>4</td>
</tr>
</tbody>
</table>

Four categories for the health state of a patient: 1, 2, 3, and 4, representing “Normal”, “Mild”, “Moderate”, and “Severe” states, respectively.

Figure 3-5 Accuracy of MLR categorical model results over time
(The larger the bubbles, the more accurate the estimations)
3.4 Artificial Neural Network Model

3.4.1 Methodology

As a final approach, an artificial neural network (ANN) can be used to categorize patients into their most probable state of severity. ANN is a non-linear mathematical model that mimics the structure or functional aspects of biological neural networks. Its structure development is based on information that flows through it during a learning phase (Guler et al., 2009). ANNs have been successively used as decision support tools for the early detection, diagnosis, and prognosis of a wide range of diseases (Vaughn et al., 2000; Mangiameli et al., 2004; Lisboa and Taktak, 2006), such as blunt trauma (Marble and Healy, 1999), breast cancer (West and West, 2000; West et al., 2005), prostate cancer (Anagnostou et al., 2003; Remzi and Djavan, 2004; Cinar et al., 2009), skin diseases (Chang and Chen, 2009), traumatic brain injury (Guler et al., 2009), and liver disease (Lin and Chuang, 2010).

An ANN consists of an input layer, an output layer, and a number of hidden layers, with each layer containing a predetermined number of artificial neurons (see Figure 3-6). Each neuron receives an input from an upstream neuron and sends an output signal to a downstream neuron, with connections between neurons optimally weighted by a learning algorithm to create a desired final output.
The strength of the connection between an input $k$ and a neuron $j$ is noted by the value of its weight $w_{kj}$. To model the actual activity inside a neuron (see Figure 3-7), an adder sums up all the inputs modified by their respective weights as a linear combination. An activation function $\phi(.)$ then controls the amplitude of the neuron’s output, with an acceptable output range usually being between 0 and 1 or –1 and 1. The activity of a neuron is written as

$$o_j = \phi(n_{etj}) = \phi\left(\sum_{k=1}^{n} w_{kj}x_k\right),$$  \hspace{1cm} (3-7)$$

where $o_j$ is neuron $j$’s output, and $net_j$ is the summation of the weighted inputs received by neuron $j$. 

Figure 3-6 Generic structure of an ANN
The common activation functions include two soft non-linearity activation functions,

- a *sigmoid* function, \( o_j = \frac{1}{1+e^{-net_j}} \), and
- a *hyperbolic tangent* function, \( o_j = \tanh\left(\frac{net_j}{2}\right) \);

along with two hard non-linearity activation functions,

- a *signum* function, \( o_j = \begin{cases} 1, & net_j > \theta_j \\ \text{undefined}, & net_j = \theta_j \\ -1, & net_j < \theta_j \end{cases} \), and
- a *step* function, \( o_j = \begin{cases} 1, & net_j > \theta_j \\ 0, & net_j = \theta_j \text{ (Kroese and Smagt, 1996).} \\ \text{undefined}, & net_j < \theta_j \end{cases} \)
3.4.2 Numerical Example

A four-layered Multilayer Perceptron (MLP) with two hidden layers is one of the most widely implemented neural network topologies and is a universal pattern classifier (Sut and Senocak, 2007). In this example, such an MLP which is trained with a back-propagation learning algorithm was used to classify the patients into the most probable state of severity. Considering the $N = 4$ screening tests case and the same data set, the input layer had four input nodes identifying each test and the TanhAxon transfer function was used in the nodes of the hidden and output layers of the ANN. The training data had 51 samples. After the network was established via Neurosolutions, multiple trainings were conducted and the best one (with the minimum mean square error (MSE)) was used. Thus, the same patient in Section 3.2.2 was given a final diagnosis of “Moderate”.

Figure 3-8 ANN network established in Neurosolutions for $N = 4$ screening test example

3.4.3 Model Accuracy

Using the same evaluation data set as above, Figure 3-9 compares the actual and predicted states over time with 81% of patients correctly classified after all 4 tests and a correlation coefficient of $R = 0.848659$, which is slightly better than the fuzzy and MLR categorical models.
3.5 Model Comparison and Optimization

Table 3-6 and 3-7 summarize the results of all three approaches, indicating that each method provides a fairly strong agreement with patients’ known health states.

<table>
<thead>
<tr>
<th>Time</th>
<th>Screening test 1</th>
<th>Screening test 2</th>
<th>Screening test 3</th>
<th>Screening test 4</th>
<th>Final diagnosis</th>
</tr>
</thead>
<tbody>
<tr>
<td>( t_1 )</td>
<td>45</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>Moderate, Moderate, Moderate</td>
</tr>
<tr>
<td>( t_2 )</td>
<td>45</td>
<td>56</td>
<td>N/A</td>
<td>N/A</td>
<td>Severe, Moderate, Severe</td>
</tr>
<tr>
<td>( t_3 )</td>
<td>45</td>
<td>56</td>
<td>63</td>
<td>N/A</td>
<td>Severe, Severe, Severe</td>
</tr>
<tr>
<td>( t_4 )</td>
<td>45</td>
<td>56</td>
<td>63</td>
<td>20</td>
<td>Moderate, Moderate, Moderate</td>
</tr>
</tbody>
</table>

FL: Fuzzy logic, MLR: Multinomial logistic regression, ANN: Artificial neural networks
### Table 3-7 Overall comparison of model performances

<table>
<thead>
<tr>
<th>Method</th>
<th>Correct diagnosis %</th>
<th>MSE</th>
<th>Correlation coefficient ($R$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>FL</td>
<td>77</td>
<td>0.26</td>
<td>0.837047</td>
</tr>
<tr>
<td>MLR</td>
<td>73</td>
<td>0.30</td>
<td>0.799164</td>
</tr>
<tr>
<td>ANN</td>
<td>81</td>
<td>0.25</td>
<td>0.848659</td>
</tr>
</tbody>
</table>

FL: Fuzzy logic, MLR: Multinomial logistic regression, ANN: Artificial neural networks

In the fuzzy categorical model, intermediate and final diagnoses may be fairly dependent on all $4N$ membership function shapes and $16N$ parameters ($a$, $b$, $c$, $d$), $N$ intermediate and final rule-bases, and $N$ defuzzification scales (all potentially decision variables). Thus, model performance could be improved or optimized by searching out the values of some or all of the above decision variables to optimize each of the above performance criteria and to minimize the number of undetected or misdiagnosed cases. Several models could be developed to optimize membership function values, rule-bases, and defuzzification scales – the optimal values of which might be different for each intermediate and final number of screening tests ($N$). For example, an optimal rule-base might be developed using a quadratic mathematical model that minimizes the square errors, as shown in Eq.(3-8); even in this simple case the number of decision variables may be as many as $4^n$.

Minimize $z = \sum_{j|O_{ij}=k,O_{lj}=l}(x_{kl} - r_j)^2 \quad \forall \ k, l$

Subject to

- $x_{kl} \leq 4$
- $x_{kl} \geq 1$ and integer. \hspace{1cm} (3-8)

The different solutions that are based on different criteria also can be weighted or traded-off against one another using various multi-criteria or desirability approaches. Note that, depending
on how exhaustive the model is, the number of decision variables may grow exponentially and special heuristics may be necessary for finding optima.

3.6 Discussion

TBI is a major military, civilian, and public health concern and one of the leading causes of disability and death in the U.S. Although improvements in protective equipment and battlefield care have reduced the incidence of penetrating head injuries and death among military personnel, there has been a significant increase in the incidence of closed-head TBI, which often occurs without any visible symptoms and goes unrecognized and untreated indefinitely. Proper and timely diagnosis is essential in effectively treating a TBI and preventing any medical, behavioral, and social consequences that untreated patients may experience. Currently available methods to assess the existence and severity of TBI may be insufficient, especially in detecting mild cases. Additionally, the initial categorization of a TBI does not necessarily correspond to the ultimate outcome as symptoms may get worse over time, eventually leading to severe dysfunction and disability.

This study presents a comprehensive screening procedure, which combines different assessment tools and consists of multiple screenings over time, in order to reduce the number of unrecognized and misdiagnosed cases. Three types of mathematical methods – fuzzy logic, multinomial logistic regression, and neural networks – were used to categorize TBI patients into the one of four health states, namely, none, mild, moderate, or severe. A numerical example was developed for illustration purposes and preliminary results for each method indicated a fairly strong agreement with the patients’ known health states.
Future work should be conducted to further develop the proposed modeling framework, which includes identifying the appropriate screening tests and their characteristics, determining the predicted diagnoses for patients whose actual health states are known, and finally evaluating the overall or severity-specific performance. Additional research could be useful to better identify the model parameters, such as fuzzy logic membership functions, final rule-bases, and defuzzification scales. Also, overall or severity-specific performance of the proposed model could be optimized by developing several optimization models, such as illustrated in Section 3.5.
Chapter 4 - Network Optimization Models

Optimal location of specialty care services and allocation of patients to these services within any healthcare network are increasingly important for balancing costs, access to care, and patient-centeredness. Typical long-range planning efforts attempt to address a myriad of quantitative and qualitative issues, including within-network access within reasonable travel distances, space capacity constraints, costs, politics, and community commitments. To help inform these decisions, this chapter presents several mathematical integer programs that minimize total procedure, travel, non-coverage, and start-up costs to increase network capacity subject to access constraints. These models were used across a range of specialty care services within the Veterans Health Administration (VHA) to help decision makers explore relationships and tradeoffs between costs, coverage, service location, and capacity and to inform larger strategic planning discussions.

4.1 Background

Over a decade ago, the Institute of Medicine (IOM) released two landmark reports outlining significant deficiencies in the quality and safety of medical care provided in the United States. To Err is Human (1999) documented the nearly 100,000 lives lost each year due to systematic problems in the design and execution of healthcare delivery. Crossing the Quality Chasm (2001)
provided a broad vision of how health care could be effectively transformed and improved, and recommended the use of systems and industrial engineering techniques to assist with understanding and optimizing healthcare processes. In response to this call, the National Academy of Engineering (NAE) and the IOM initiated a study in 2002 to identify engineering tools and technologies that could help the health system improve and deliver care that is safe, effective, timely, patient-centered, efficient, and equitable – the six health system aims defined in *Crossing the Quality Chasm*.

The result of this study was published as *Building a Better Delivery System* (National Academy of Engineering and Institute of Medicine, 2005), which concluded that the U.S. healthcare industry has neglected use of systems engineering methods and concepts that have revolutionized quality, productivity, and performance in many other industries, and that this “collective inattention” has contributed to serious consequences within health care – huge amounts of preventable harm and deaths, outdated procedures, an approximately half-trillion dollars wasted annually through inefficiency, costs rising at roughly three times the rate of inflation, and 43 million people uninsured. Today, health care is still an area of crucial concern in the United States, spending substantially more than any other developed country in the Organization for Economic Co-operation and Development (OECD). The use of healthcare services in the U.S., however, is below the OECD median by almost all measures, which means that Americans are receiving fewer real resources than people in over 50% of OECD countries. These facts suggest that the difference in spending is caused mostly by higher prices for the delivery of healthcare goods and services in the U.S. (Anderson *et al.*, 2003). Rising costs and inadequate levels of care underscore the importance of making strategic decisions within this industry.
The use of systems engineering models to help make decisions in many key areas, in fact, can be very effective at helping design both cost-efficient and qualified healthcare services, especially in cases involving complex and competing considerations. Network planning is one of several areas that can be improved via systems engineering methods, with the optimal location of health services across geographic care networks having significant potential to improve costs, patient-centeredness, and care continuity (3 of the IOM dimensions).

Location-allocation models seek to simultaneously determine optimal facility locations and the assignment of customers (here, patients) to facilities. These models have been applied widely in many industries to reduce costs and increase access, with examples spanning retail, telecommunication, energy, and manufacturing. In health care, poor location decisions also can result in increased mortality and morbidity, with optimal facility location therefore having even greater importance (Daskin and Dean, 2005). Historical healthcare applications have included locating ambulances (Adenso-Díaz and Rodríguez, 1997; Sasaki et al., 2010) and hospitals in rural regions (Mehrez et al., 1996), trauma care resources (Branas et al., 2000), organ transplant services (Bruni et al., 2006), blood facilities (Jacobs et al., 1996), emergency medical service vehicles (Eaton et al., 1985), hospital networks (Santibanez et al., 2009), and specialized care services such as traumatic brain injury (TBI) treatment (Côté et al., 2007; Syam and Côté, 2010). Models that account for transient populations who change their locations seasonally (Ndiaye and Alfares, 2008) and continuously changing nature of health systems (Harper et al., 2005) also have been developed for specific applications.

From a modeling perspective, location problems can be classified as discrete and continuous. Discrete models assume a finite number of candidate locations exist at which facilities can be
sited, whereas continuous models assume facilities can be located anywhere within a geographic area. All models aim simultaneously to determine both the best locations for facilities and the best assignment (“allocation”) of individuals to these facilities, where “best” usually is defined by some combination of total travel distance, facility costs that include fixed locating and variable operating costs, and demand coverage. The work in this chapter focuses on discrete location models, which also have been used more extensively in healthcare location problems (Daskin and Dean, 2005), because typically the VHA has been interested in considering specific existing or candidate locations. Table 4-1 summarizes the three basic types of discrete facility location models and the performance measures they seek to optimize.

Table 4-1 Representation of basic discrete facility location models

<table>
<thead>
<tr>
<th>Model</th>
<th>Description</th>
<th>Assumptions</th>
<th>Decision variables</th>
<th>Performance measures</th>
</tr>
</thead>
</table>
| Location set covering model  | Minimize total cost of providing service | - Specify potential facility locations  
- Specify maximum acceptable distance (S)  
- Iterate on S | Location, number | Cost (dollars) |
| Maximal covering location model | Maximize demand coverage | - Specify potential facility locations  
- Specify maximum acceptable distance (S)  
- Specify total number of facilities (P)  
- Iterate on S and P | Location, number | Demand coverage (percent) |
| P-median model               | Minimize total weighted travel distance | - Specify potential facility locations  
- Specify total number of facilities (P)  
- Iterate on P | Location | Travel distance (miles) |

The key concept in covering models is the notion of coverage, with a maximum acceptable travel distance specified such that a “demand node” (here, a patient’s 3-digit zip code home location) can be served only by a facility within this distance. Covering models have been studied within two major classes in the literature (Owen and Daskin, 1998). The first class is the location set covering problem (LSCP), introduced by Toregas (1970), which aims to minimize the cost of
facility location while ensuring that all demand nodes are covered. These models are useful for cases in which all demand must be covered and one wishes to minimize the total cost to accomplish this. The second class is the maximal covering location problem (MCLP), introduced by Church and ReVelle (1974), which aims to maximize the level of coverage (i.e., the number of covered demand nodes) using only a specified number of facilities and a maximum acceptable travel distance. These models are useful in cases for which the cost of covering all demand nodes is prohibitive. Finally, \( P \)-median models (ReVelle and Swain, 1970) seek to minimize the average travel distance and are useful in cases for which a specific maximum acceptable distance is less clear (Daskin and Dean, 2005). Many single and multiple objective extensions of these three basic ideas have been developed over the past four decades to address particular problems specific in manufacturing (Melachrinoudis and Min, 2000), distribution networks (Klose and Drexl, 2005), web services (Aboolian et al., 2009), and other contexts. It also should be noted that all of these types of models have continuous counterparts – for example planar LSCP and MCLP (Church, 1984; Current and O'Kelly, 1992; Murray, 2001) and multi-source Weber formulations (Cooper, 1963) – where facilities can be located anywhere in the plane.

4.2 Optimal Location and Use of In-Person and Video-Based PTSD Treatment Services within the VHA

Given the vast medical and social costs of untreated or undertreated posttraumatic stress disorder (PTSD), there is a critical need to improve access to evidence-based treatments (Murdoch et al., 2005; Cohen et al., 2010; Ivanova et al., 2011). In the Veterans Affairs (VA), this need has become of even greater importance with the development of effective treatments (Smith et al., 2005; Davis et al., 2012; Schnurr and Lunney, 2012). However, access to these effective
treatments, particularly for those residing in rural areas, is a practical issue. Innovative treatment modalities, such as telehealth or video-based services, hold great promise to increase access to these important mental health treatment services (Hailey et al., 2008; Wilson and Wells, 2009; Mohr et al., 2010; Tutty et al., 2010), particularly in rural areas where access to in-person mental health treatment is limited (Barnwell et al., 2012). Large integrated health systems such as VHA can benefit from systems engineering tools to help plan the most appropriate ways to meet mental health needs of patient populations. This study aims to apply longitudinal systems engineering care location models to help determine the combined optimal geographic locations and capacities for in-person care for veterans with PTSD across the New England VA network and where instead video-based care would be advantageous over expensive within-system capacity expansion or fee-based external care, if even feasible.

4.2.1 VA New England PTSD Treatment Services

The VA New England healthcare system is one of 21 Veterans Integrated Service Networks (VISNs) within the Department of Veterans Affairs, with each VISN being regionally managed and operating as a somewhat autonomous decision-making system. The VA New England VISN (VISN 1) provides comprehensive medical services (including primary, mental health, specialty, and hospital-based care) to more than 240,000 veterans residing in any of the six New England states (Maine, New Hampshire, Vermont, Massachusetts, Rhode Island, and Connecticut) annually. With an annual operating budget of approximately $2.2 billion, VISN 1 has 8 medical centers (a total of 11 facilities as Connecticut and Boston healthcare system has 2 and 3 campuses, respectively), 40 community-based outpatient clinics, 5 outpatient clinics, 6 nursing homes, and 2 domiciliaries – institutional houses for aged and disabled veterans (U.S.)
Department of Veterans Affairs, 2013). Nationwide, the VA spends more than $2.5 billion annually to provide specialty care services, of which $125 million spent by VISN 1 (Veterans Service Support Center, 2011a). While the VA has strived to develop integrated and coordinated care, veterans who also are eligible for Medicare increasingly use non-VA specialty services, resulting in increased fragmentation in care continuity (Liu et al., 2011). The fees for these out-of-network services are reimbursed by the VA (“fee basis” visits) and can represent large amounts of monies for some specialties. Better availability of specialty services within the VA could decrease this fragmentation as well as result in lower overall costs.

Since the number of veterans diagnosed with PTSD is a function of access to care and may under-represent true demand, overall PTSD prevalence among VA users rather than diagnostic codes in VISN 1 was used to estimate the true need for PTSD care services. In November 2011, the VA Office of Public Health reported that of the cumulative total of 741,954 Iraq and Afghanistan veterans who utilized VA health services from October 2001 through September 2011, a total of 223,609 (30.14%) were diagnosed with PTSD (VA Office of Public Health, 2011a, 2011b). Given that the rate of PTSD diagnosis has been increasing over time (Rosenheck and Fontana, 2007; Hermes et al., 2012; Shiner et al., 2012), a prevalence of 25% among veterans who utilized VISN 1 VHA facilities from January 2009 through May 2010 was assumed and true demand for PTSD care services over this time period was estimated based on this prevalence. Figure 4-1 summarizes the estimated number of veterans with PTSD in VISN 1 by state, with the majority residing in Massachusetts and Connecticut. Figure 4-2 depicts the heterogeneous geographic distribution of veterans with PTSD in (a) raw counts and (b) per square mile across VISN 1, with the locations of VA facilities.
Within the VA, veterans with PTSD are treated in specialized PTSD clinics, general mental health clinics, or primary care clinics. Specialized PTSD clinics currently are located only at some of the medical centers across New England. General mental health services are provided at medical centers and large community-based outpatient clinics, while smaller community-based outpatient clinics typically provide PTSD care either by referral to their affiliated medical center or through the use of video-based telemental health services (Kussman, 2008). In VISN 1, six of eight medical centers (Boston MA, Brocton MA, West Haven CT, Providence RI, Togus ME, and White River Junction VT) have specialized PTSD care teams. Data between October 2009 and September 2010 shows that 35.15% of patients who received a primary diagnosis of PTSD used specialized PTSD services across VISN 1, while more than half of those received care in general mental health services (Veterans Service Support Center, 2011b), as presented in Figure 4-3.
Figure 4-2 Geographic distribution of veterans with PTSD in (a) raw counts and (b) per square mile by 3-digit zip code and all VA facilities across VISN 1 (January 2009 - May 2010)
4.2.2 Methodology

4.2.2.1 Model Overview and Assumptions

The below model determines the optimal assignment of PTSD patients to the existing VISN 1 VHA facilities that minimizes overall total costs and calculates the total amount of mental health manpower needed at each facility. A maximum acceptable travel distance (S) to a VA facility, beyond which a patient is directed to an out-of-network provider at the VA’s expense, is specified. Travel distances exceeding 30 miles are fully reimbursed by the VA at a current rate of $0.70 per mile to cover gas, maintenance, vehicle depreciation, and inconvenience. Patients at demand nodes (3-digit zip codes) assigned to an available VA facility within this acceptable distance incur treatment and travel costs, while others are directed to an out-of-network provider at a higher cost to the VA (i.e., non-coverage cost). Given that most private practice providers lack knowledge of PTSD and other mental health disorders prevalent among veterans and that they are unfamiliar with the VA’s treatment resources for such conditions (Boscarino et al., 2010; Kizer, 2012), demand nodes that are sent to a non-VA provider are considered candidates.
for video-based services as a more effective alternative both in terms of quality and cost of care.

Other model assumptions include

(1) Each patient receives care at the facility to which he is assigned,

(2) Demand is deterministic, equally distributed over time, and with no seasonality effects,

(3) Care cost variations between patients, medical centers, and geographic regions are negligible,

(4) PTSD treatment preferences among VA users are assumed to be a) 60% want evidence-based psychotherapy only, b) 25% want medication only, and c) 15% want both (Watts et al., 2012),

(5) Assumptions on evidence-based psychotherapy included: a) Conducted by psychotherapists (with at minimum a master’s degree), b) consisted of one evaluation, 12 treatment sessions, and three phone calls, and c) a total of 13.75 clinical hours needed per patient per year (i.e., one hour for evaluation, one hour for each treatment session, and 15 minutes for each phone call) (Kussman, 2008; U.S. Department of Veterans Affairs and Department of Defense, 2010),

(6) Medication assumptions included: a) Prescribed by psychiatrists, b) consisted of one evaluation session, six follow-up visits, and three phone calls, and c) a total of 4.75 clinical hours needed per patient per year (i.e., one hour for evaluation session, 30 minutes for each follow-up visit, and 15 minutes for each phone call),
(7) Mental health providers located in specialized PTSD clinics only focus on patients with PTSD, while those in general mental health and primary care clinics spend 75% of their time addressing disorders other than PTSD.

### 4.2.2.2 Integer Programming Model

The following notation and input parameters are used in the mathematical model, with

- $i$ = index of geographic demand nodes,
- $j$ = index of facilities,
- $h_i^P$ = total number of patients who need psychotherapy at demand node $i$,
- $h_i^M$ = total number of patients who need medication at demand node $i$,
- $v^P$ = total number of psychotherapy visits needed per year per patient,
- $v^M$ = total number of medication visits needed per year per patient,
- $k^P$ = total clinical hours needed for psychotherapy per year per patient,
- $k^M$ = total clinical hours needed for medication per year per patient,
- $S$ = maximum acceptable travel distance,
- $t_{ij}$ = distance between demand node $i$ and facility $j$,
- $C_i^P$ = total cost of psychotherapy per year per patient,
- $C_i^M$ = total cost of medication per year per patient,
- $C_2$ = travel cost per mile,
- $C_3$ = non-coverage cost per year per patient,
- $n_j^P$ = total number of psychotherapists needed at facility $j$ per year,
- $n_j^M$ = total number of psychiatrists needed at facility $j$ per year,
- $T$ = total working hours of a therapist/psychiatrist per year,
- $O$ = set of medical centers that have specialized PTSD clinics,
- $l_j$ = rate of time that a therapist/psychiatrist spends for treating patients with PTSD in facility $j$, equal to 1 if $j \in O$, and 0.25 otherwise,
- $D$ = total demand (for both psychotherapy and medication) in the corresponding planning horizon, and
\[d_{ij} = \text{reimbursed travel distance, equal to } t_{ij} \text{ if } t_{ij} > 30, \text{ and } 0 \text{ otherwise.}\]

The decision variable is

\[Y_{ij} = 1 \text{ if patients in demand node } i \text{ are treated at facility } j, \text{ and } 0 \text{ otherwise.}\]

Using the above notation, an integer programming model can be written as

\[
\begin{aligned}
\text{Minimize} & \quad \sum_i \sum_j Y_{ij} \left( h_i^p (C_i^p + v^p d_{ij} C_2) + h_i^M (C_i^M + v^M d_{ij} C_2) \right) \\
& \quad + C_3 \left( D - \sum_i \sum_j Y_{ij} (h_i^p + h_i^M) \right)
\end{aligned}
\]

Subject to

\[
\begin{align*}
\sum_j Y_{ij} & \leq 1 & \forall i \quad (4-1) \\
\sum_{j\mid t_{ij} > S} Y_{ij} & \leq 0 & \forall i \quad (4-2) \\
n_j^p & = \sum_i Y_{ij} h_i^p k^p / (l_j T) & \forall j \quad (4-3) \\
n_j^M & = \sum_i Y_{ij} h_i^M k^M / (l_j T) & \forall j \quad (4-4) \\
Y_{ij} & \in \{0, 1\} & \forall i, j \quad (4-5)
\end{align*}
\]

The objective function minimizes the total of treatment, travel and non-coverage costs. In the first term, treatment (i.e., psychotherapy and medication) and travel costs are calculated for the patients who receive care within-network. The second term represents the non-coverage cost calculated for the remaining patients that are not covered by any VA facility. Constraint (4-1) assigns every demand node to at most one VA facility. Constraint (4-2) ensures that a demand node can be assigned to a facility if and only if the travel distance is less than or equal to the acceptable maximum. The total number of psychotherapists and psychiatrists needed in each
facility is calculated by constraints (4-3) and (4-4). Constraint (4-5) defines the decision variable $Y_{ij}$ to be binary.

4.2.3 Application and Results

4.2.3.1 Application Details

The mathematical model was used to determine for each facility the optimal amount of each type of PTSD care service to provide and for each PTSD patient whether service could be economically provided within a specified travel distance to any existing PTSD treatment facility. Assignments of care type to facilities were made regardless of current services at them, with assignment of PTSD patients to sites used to determine which patients ideally should be seen where. VA goals for maximum acceptable travel distances (currently 30 miles for PTSD care) were used to determine these service locations, patient destinations, and if a particular patient could receive within-network in-person care. Those that could not receive in-person care within the VISN 1 network were considered candidates for video-based care rather than outside the VA system in a manner so as to minimize total care and delivery costs. This analysis was also repeated for other maximum travel distances in order to investigate how results might be impacted by this somewhat arbitrary service threshold.

4.2.3.2 Results

Figure 4-4 shows the tradeoff between maximum acceptable driving distance, total minimal cost, and access for the estimated 2010 PTSD care demand. Also shown is the current situation for comparison, estimated based on the current patient allocations, which has an 11.08% ($10.9 million) higher cost and 56.26% lower access than even the optimized case with no telehealth
service due to current sub-optimal patient-to-facility allocation. Utilizing telehealth services where demand cannot be covered by the VA within 30 miles, rather than providing care at non-VA facilities, would result in further savings of $4,001,664 (4.04%) given the same access coverage, or alternately a roughly 50% reduction in maximum driving distance to 30 miles given the same cost. Note that all demand can be covered by the VA if a maximum acceptable driving distance of more than 60 miles is allowed. This, however, causes a 45% increase in average driving distance per patient that exceeds the VA’s maximum travel distance policy.

Figure 4-4 Tradeoff between acceptable distance and (a) total annual cost, and (b) coverage percentage assuming estimated care demand in 2010
Table 4-2 summarizes the resulting optimal locations of in-person PTSD services and the number of psychotherapist and psychiatrist FTEs required at each site to provide this care. As shown, rural pockets of Maine, Vermont, and New Hampshire would be better served with video-based mental health services since demand in these areas cannot be covered by any of the VA facilities. In Massachusetts, Connecticut, and Rhode Island, there is sufficient need and access for in-person PTSD services, with the possible exception of the medical center located in Jamaica Plain, MA which might be treated in the next closest location instead since only an estimated 17 patients here will need care during the year.

Figure 4-5 illustrates how optimal within-network and video-based care locations change if some other maximum travel distances were allowed. Finally, Figure 4-6 summarizes future projected needs for PTSD treatment by therapists and psychiatrists, along with the estimated number of veterans needing PTSD care. Future PTSD care demand was estimated via the VA’s veteran population projections (www.va.gov/VAEKGAR/Demographics/Demographics.asp) and geographically distributed among all 59 VISN 1 zip codes based on current relative proportions. As expected, the number of providers required decreases over the next 10 years as a result of shrinking veteran population projections, while zip codes that would be better served with video-based services remain the same.
### Table 4-2 Optimal allocation of patient zip codes to VISN 1 facilities for a 30-mile maximum acceptable distance assuming estimated care demand in 2010

<table>
<thead>
<tr>
<th>VISN 1 VA facilities</th>
<th>Type of care provided</th>
<th>Patient 3-digit zip codes</th>
<th>Total number of patients assigned</th>
<th>Total number of psychotherapists needed (FTEs)</th>
<th>Total number of psychiatrists needed (FTEs)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Brocton, MA</td>
<td>PCT</td>
<td>023, 027</td>
<td>1687</td>
<td>16.12</td>
<td>2.61</td>
</tr>
<tr>
<td>Jamaica Plain, MA</td>
<td>PCT</td>
<td>022</td>
<td>17</td>
<td>0.16</td>
<td>0.03</td>
</tr>
<tr>
<td>West Roxbury, MA</td>
<td>MHS</td>
<td>020, 024</td>
<td>801</td>
<td>30.61</td>
<td>4.95</td>
</tr>
<tr>
<td>Newington, CT</td>
<td>MHS</td>
<td>060, 061, 064</td>
<td>4570</td>
<td>174.71</td>
<td>28.24</td>
</tr>
<tr>
<td>West Haven, CT</td>
<td>PCT</td>
<td>065, 066</td>
<td>1336</td>
<td>12.76</td>
<td>2.06</td>
</tr>
<tr>
<td>Manchester, NH</td>
<td>MHS</td>
<td>030, 031</td>
<td>1257</td>
<td>48.04</td>
<td>7.76</td>
</tr>
<tr>
<td>Providence, RI</td>
<td>PCT</td>
<td>028, 029</td>
<td>3722</td>
<td>35.57</td>
<td>5.75</td>
</tr>
<tr>
<td>Togus, ME</td>
<td>PCT</td>
<td>043</td>
<td>598</td>
<td>5.71</td>
<td>0.92</td>
</tr>
<tr>
<td>White River Junction, VT</td>
<td>PCT</td>
<td>037, 050</td>
<td>711</td>
<td>6.79</td>
<td>1.10</td>
</tr>
<tr>
<td>Fitchburg, MA</td>
<td>MHS</td>
<td>014</td>
<td>471</td>
<td>17.99</td>
<td>2.91</td>
</tr>
<tr>
<td>Gloucester, MA</td>
<td>Rf/TH</td>
<td>019</td>
<td>832</td>
<td>31.80</td>
<td>5.14</td>
</tr>
<tr>
<td>Dorchester, MA</td>
<td>Rf/TH</td>
<td>021</td>
<td>1769</td>
<td>67.62</td>
<td>10.93</td>
</tr>
<tr>
<td>Framingham, MA</td>
<td>MHS</td>
<td>017</td>
<td>353</td>
<td>13.47</td>
<td>2.18</td>
</tr>
<tr>
<td>Lowell CBOC, MA</td>
<td>MHS</td>
<td>018</td>
<td>1157</td>
<td>44.22</td>
<td>7.15</td>
</tr>
<tr>
<td>Worcester, MA</td>
<td>MHS</td>
<td>015, 016</td>
<td>1085</td>
<td>41.47</td>
<td>6.70</td>
</tr>
<tr>
<td>New London, CT</td>
<td>MHS</td>
<td>063</td>
<td>973</td>
<td>37.19</td>
<td>6.01</td>
</tr>
<tr>
<td>Stamford, CT</td>
<td>MHS</td>
<td>068, 069</td>
<td>737</td>
<td>28.17</td>
<td>4.55</td>
</tr>
<tr>
<td>Waterbury, CT</td>
<td>MHS</td>
<td>067</td>
<td>948</td>
<td>36.24</td>
<td>5.86</td>
</tr>
<tr>
<td>Windham, CT</td>
<td>MHS</td>
<td>062</td>
<td>430</td>
<td>16.42</td>
<td>2.65</td>
</tr>
<tr>
<td>Somersworth CBOC, NH</td>
<td>MHS</td>
<td>038, 039</td>
<td>1037</td>
<td>39.65</td>
<td>6.41</td>
</tr>
<tr>
<td>Tilton CBOC, NH</td>
<td>MHS</td>
<td>032, 033</td>
<td>737</td>
<td>28.17</td>
<td>4.55</td>
</tr>
<tr>
<td>Greenfield, MA</td>
<td>MHS</td>
<td>013</td>
<td>275</td>
<td>10.51</td>
<td>1.70</td>
</tr>
<tr>
<td>Pittsfield, MA</td>
<td>Rf/TH</td>
<td>012</td>
<td>181</td>
<td>6.91</td>
<td>1.12</td>
</tr>
<tr>
<td>Springfield, MA</td>
<td>MHS</td>
<td>010, 011</td>
<td>1974</td>
<td>75.44</td>
<td>12.19</td>
</tr>
<tr>
<td>Hyannis, MA</td>
<td>MHS</td>
<td>026</td>
<td>657</td>
<td>25.09</td>
<td>4.05</td>
</tr>
<tr>
<td>Martha’s Vineyard, MA</td>
<td>Rf/TH</td>
<td>025</td>
<td>267</td>
<td>10.19</td>
<td>1.65</td>
</tr>
<tr>
<td>Bangor CBOC, ME</td>
<td>MHS</td>
<td>044</td>
<td>822</td>
<td>31.43</td>
<td>5.08</td>
</tr>
<tr>
<td>Portland CBOC, ME</td>
<td>Rf/TH</td>
<td>040, 041</td>
<td>1052</td>
<td>40.22</td>
<td>6.50</td>
</tr>
<tr>
<td>Rumford CBOC, ME</td>
<td>MHS</td>
<td>042</td>
<td>963</td>
<td>36.80</td>
<td>5.95</td>
</tr>
<tr>
<td>Bennington CBOC, VT</td>
<td>MHS</td>
<td>052</td>
<td>176</td>
<td>6.73</td>
<td>1.09</td>
</tr>
<tr>
<td>Brattleboro CBOC, VT</td>
<td>Rf/TH</td>
<td>053</td>
<td>137</td>
<td>5.24</td>
<td>0.85</td>
</tr>
<tr>
<td>Colchester CBOC, VT</td>
<td>MHS</td>
<td>054</td>
<td>771</td>
<td>29.47</td>
<td>4.76</td>
</tr>
<tr>
<td>Newport CBOC, VT</td>
<td>Rf/TH</td>
<td>058</td>
<td>320</td>
<td>12.22</td>
<td>1.98</td>
</tr>
<tr>
<td>Rutland CBOC, VT</td>
<td>Rf/TH</td>
<td>057</td>
<td>321</td>
<td>12.25</td>
<td>1.98</td>
</tr>
<tr>
<td>Non-covered (candidates for video services)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| Non-covered (candidates for video services) | 034, 035, 036 (NH); 045, 046, 047, 048, 049 (ME); 051, 056, 059 (VT) |

PCT: Specialized PTSD care teams, MHS: General mental health services, Rf/TH: By referral or telemental health
Figure 4-5 Optimal areas for video-based services for a (a) 15-mile, (b) 30-mile, (c) 45-mile, and (d) 60-mile maximum acceptable distance assuming estimated care demand in 2010
4.2.4 Discussion

This study highlights the need for appropriate allocation of PTSD services for veterans. Results may be useful as a planning tool for PTSD resource allocation within the U.S. Department of Veteran Affairs. Access to care is a particular concern for veterans who are unable to readily access PTSD treatment, such as those living in rural environments. In such cases, new treatment modalities such as telehealth (i.e., via video conference) can be considered. Such treatment actually may be preferred by some veterans given that it increases convenience and reduces privacy concerns.

For the VA population, most clinic sites have sufficient volume to justify in-person staff for face-to-face PTSD treatment. There are a small number of community-based clinics, however, whose estimated demand is insufficient to justify in-person staff. For these sites, this analysis suggests that PTSD treatment needs would be better met by video-based services. In general, these sites were in rural or highly rural regions of New England, primarily in central Vermont, southwestern
New Hampshire, and much of northern Maine. All other areas have sufficient need to support in-person services, assuming availability.

Study limitations include logic, data, and preference assumptions in the mathematical model. For example, it is assumed that providers trained in treating PTSD are available or could be relocated to each clinic site, which may not be possible immediately, although in such cases additional or temporary reliance on telemental health services may be possible. Assumptions regarding patient treatment preferences also were based on a pilot sample (Watts et al., 2012) and were assumed to be equally distributed across the patient population. Rural patients instead, for example, may have higher preference for medication treatment. It also is not clear that all VA users have a preference for or an understanding of evidence-based psychotherapy. While the VA follows guidelines that promote evidence-based treatment (Kussman, 2008; U.S. Department of Veterans Affairs and Department of Defense, 2010; Cloitre et al., 2012), furthermore, the definition of evidence-based psychotherapy is not universal and instead could be defined as any of those listed in the National Registry of Evidence-based Programs and Practices (Miller et al., 2006; Sorensen, 2011). In addition to psychiatrists, medications also can be prescribed by nurse practitioners, physician assistants, or primary care physicians and therefore the estimate for medications prescribed may be conservative. The recent RESPECT-PTSD trial (Schnurr et al., 2013) also showed the difficulty of implementing evidence-based care for PTSD by primary care providers. Finally, results have based only on the New England VA network, although the general value of this approach seems generalizable to all VISNs.

Additional research could examine if tele treatment modalities affect treatment retention among veterans or the active duty population. More generally, use of operations research and systems
engineering methods appears useful to help inform these types of macro system design and policy issues. Similar analysis for the location of other evidence-based mental health services therefore also may be beneficial.

4.3 Single and Multi-Period Location-Allocation Models for Sleep Apnea Testing Services within the VHA

Along with the care services for combat-related conditions such as TBI and PTSD, VA also is responsible for providing comprehensive medical services to the veterans with other chronic health problems. There has been, for instance, an increased demand for sleep apnea testing and treatment services among Vietnam veterans in their 50s and 60s given the high prevalence of sleep disorders in this age group (Ancoli-Isreal et al., 1991; Young et al., 1993; Punjabi et al., 2000; Rosenheck and Fontana, 2007). In the New England VA network (VISN 1), it is reported that there is insufficient access to sleep apnea testing services, causing a significant increase in the number of “fee basis” visits and thus cost. This study aims to illustrate the use of mathematical optimization models to help improve the geographic locations and capacities of sleep apnea testing services across the New England VA network in a manner so as to provide appropriate access at minimal cost considering patient needs and preferences.

4.3.1 VA New England Sleep Apnea Services

Sleep apnea is a disorder characterized by abnormal pauses in breathing or instances of abnormally low breathing during sleep and has been linked with high blood pressure, heart problems, stroke, fatigue, headaches, snoring, daytime drowsiness, and memory problems (Ancoli-Isreal and Avalon, 2006; Erman, 2006). These health problems occur at a much higher
rate among veterans than in the general population due to long-term exposure to dust, smoke, chemicals, and other environments (U.S. Department of Health and Human Services, 2012). Up to 20% of all U.S. veterans currently suffer from sleep apnea, with an annual treatment cost of $225 million to the VA (Veterans Service Support Center, 2011a).

Figure 4-7 summarizes the number of treated sleep apnea patients in VISN 1 by state, with the majority residing in Connecticut and Massachusetts as expected given greater overall populations. Since observed utilization is a function of capacity and may under- or over-represent true demand, also shown is an estimate of the greater true need over this time period, based on clinical input and patient-specific predictors, such as their body mass index (since it is well-known that the risk for sleep apnea is higher for people who are overweight (U.S. Department of Health and Human Services, 2012)). Figure 4-8 shows estimated demand by 3-digit zip code relative to VA facilities across VISN 1.

Figure 4-7 Sleep apnea care demand by state across VISN 1 (January 2009 - May 2010)
A definitive diagnosis of sleep apnea is made through use of overnight polysomnographic testing that involves having a patient sleep in a polysomnogram bed (referred to as a “sleep bed”) with continuous monitoring of their blood oxygen levels, electroencephalogram, and vital signs. A general belief has been that inadequate access to overnight polysomnographic testing in VISN 1 is causing a large number of veterans to be referred to out-of-network providers. The fees for
these out-of-network services are reimbursed by the VA (“fee basis” visits) and can represent large amounts of monies for some specialties. In the sleep apnea case, data between January 2009 and May 2010 shows that of 4712 tested patients from VISN 1, 3927 were tested in one of the VA facilities at a total operation and travel cost of $2,431,762, with 785 tested outside the network at a total cost of $667,250 or 9% of the overall cost.

Currently, only four of eight VA medical centers in VISN 1 (i.e., West Haven CT; West Roxbury MA; Manchester NH; and Providence RI) have sleep laboratories that provide polysomnographic testing. Closing these laboratories or decreasing their capacities is not feasible within the VA given policy and other considerations. Table 4-3 summarizes the number of sleep beds and visits at each facility.

Table 4-3 Number of sleep beds, within-network and outside-network sleep apnea patients in VISN 1 (January 2009 - May 2010)

<table>
<thead>
<tr>
<th>Facilities</th>
<th>Current number of sleep beds</th>
<th>Number of within-network visits</th>
<th>Number of outside-network visits</th>
</tr>
</thead>
<tbody>
<tr>
<td>West Haven, CT</td>
<td>4</td>
<td>1,738</td>
<td>255</td>
</tr>
<tr>
<td>West Roxbury, MA</td>
<td>6</td>
<td>1,119</td>
<td>-</td>
</tr>
<tr>
<td>Manchester, NH</td>
<td>4</td>
<td>476</td>
<td>-</td>
</tr>
<tr>
<td>Providence, RI</td>
<td>2</td>
<td>434</td>
<td>-</td>
</tr>
<tr>
<td>Togus, ME*</td>
<td>-</td>
<td>97*</td>
<td>286</td>
</tr>
<tr>
<td>White River Junction (WRJ), VT</td>
<td>-</td>
<td>-</td>
<td>204</td>
</tr>
<tr>
<td>Total</td>
<td>16</td>
<td>3,767</td>
<td>745</td>
</tr>
</tbody>
</table>

*In 2009, there were two sleep beds in Togus.

The West Haven medical center generated outside-network fee visits over this time period, even though it has sleep beds, due to demand at times exceeding capacity as shown in Figure 4-9. However, a relatively large number of fee type visits occurred in months in which utilization (measured as the percent of capacity usage) was low. Fee type utilization associated with other
facilities with sleep beds suggests regional demand does not exceed regional capacities, as shown in Figure 4-10. Due to the fact that patient records are based on billing rather than visit dates, utilizations in some months are over 100%, which may not reflect actual utilization levels. Those observations underscore data inaccuracies and idiosyncrasies, the need for sensitivity analyses, and the value of working in an integrated analyst-clinician research team.

Figure 4-9 Fee visits for West Haven, (a) number of fee visits versus monthly utilization, (b) weekly fee visits versus in-house visits (months with relatively low utilization (< 50%) and high fee visits are circled)
Figure 4-10 Utilization of sleep apnea capacity by medical center over time (utilizations over 100% are circled)

Figure 4-11 summarizes the distances patients travelled to receive within-network care over the same time period, with a mean of 44.5 miles but a high variability ranging from 2 to 420 miles. Approximately 70% of patients travelled more than 30 miles and 25% more than 60 miles, typical acceptable distance thresholds. Roughly 5% of patients traveled more than 100 miles. As mentioned earlier, travel distances exceeding 30 miles are fully reimbursed by the VA at a current rate of $0.70 per mile.

Figure 4-11 Travel distance distribution, VISN 1 (January 2009 - May 2010)
4.3.2 Methodology

4.3.2.1 Model Overview and Assumptions

The below models determine the optimal sleep bed capacity in each existing VISN 1 VHA facility that minimizes overall total costs over a single or multiple periods, relative to geographic demand. The single-period model is useful for short term planning or for cases in which future demand is unlikely to change, whereas the multi-period model is important in cases for which demand or patient volume may significantly change over time, as can be the case with veteran populations.

Similar to the above model developed for the case of PTSD treatment services, a maximum acceptable travel distance ($S$) to a VA facility is specified, beyond which a patient is directed to an out-of-network provider at the VA’s expense. Patients at demand nodes (3-digit zip codes) assigned to an available VA facility within this acceptable distance are referred to as in-house type visits and incur procedure costs as well as travel costs, as described above. Other model assumptions include

1. Each patient receives his care at the facility to which he is assigned and spends one night there,
2. Annual capacity of a sleep bed is equal to the number of working days in a year,
3. Facilities that currently provide this care continue to do so and capacity down-sizing is not allowed,
(4) Procedure costs include overhead and labor costs independent from a patient’s health
state,

(5) Demand is deterministic, equally distributed over time, and with no seasonality effects,
and

(6) Care cost variations between medical centers and geographic regions are negligible.

The first four assumptions reflect the current circumstances of the VA sleep apnea testing
services. The fifth and sixth assumptions also are realistic since seasonality is not indicated in the
literature as a major factor in sleep apnea cases and there is not a significant cost variation
between VISN 1 facilities all of which operate under the same local administration. These two
assumptions, in fact, decrease the complexity of the model.

The multi-period model considers an $n$-year planning horizon, with a user-specified maximum
number of facilities $P$ that can be added each year (assumed the same for each period). An
underutilization cost is included in the multi-period model since expansion decisions may
produce over capacity in some following years. Facilities in which service is added or expanded
also are assumed to continue to operate at this capacity level over the remainder of the planning
horizon.

### 4.3.2.2 Single-Period Model

The following notation and input parameters are used in the single-period model, with

\[ i \quad \text{=} \quad \text{index of geographic demand nodes,} \]
The decision variables are

\[ X_j = \begin{cases} 1 & \text{if facility } j \text{ provides sleep apnea service and 0 otherwise,} \\ 0 & \text{otherwise} \end{cases} \]

\[ Y_{ij} = \begin{cases} 1 & \text{if patients in demand node } i \text{ are treated at facility } j \text{ and 0 otherwise, and} \\ 0 & \text{otherwise} \end{cases} \]

\[ a_j = \text{number of additional sleep beds to add in facility } j. \]

Using the above notation, an integer programming mathematical model can be written as

\[
\text{Minimize } \left( \sum_i \sum_j Y_{ij} h_i \left( C_1 + d_{ij} C_2 \right) + C_3 \left( D - \sum_i \sum_j Y_{ij} h_i \right) + C_4 \sum_j a_j \right)
\]

Subject to

\[
\sum_{j \in O} X_j \leq P \quad (4-6)
\]

\[
\sum_j Y_{ij} \leq 1 \quad \forall i \quad (4-7)
\]

\[
Y_{ij} - X_j \leq 0 \quad \forall i, j \quad (4-8)
\]
The objective function minimizes the total of in-house, fee, and set-up cost for additional beds. In the first term of the objective function, treatment and travel costs are calculated for the patients who receive in-house type service. The second term represents the fee type service cost calculated for the remaining patients that are not covered by any VA facility. The weighted sum of non-covered demand nodes is multiplied by the fixed external visit cost to calculate the total cost of fee type visits. It should be noted that the term “C3D” in the fee type cost equation could be excluded from the objective function since it is constant and non-optimizable, but is included here for completeness. Finally, the last term multiplies the total number of additional sleep beds in all facilities by a fixed bed set-up cost. Constraint (4-6) requires that service is added to at most \( P \) additional facilities, while constraint (4-7) assigns every demand node to at most one VA facility (those unassigned receive their care outside the network). Constraints (4-8) and (4-9) ensure that a demand node can be assigned to a facility if and only if this facility provides this type of care service and the travel distance is less than or equal to the acceptable maximum, respectively. The number of additional beds at each facility is determined by constraint (4-10). Constraint (4-11) ensures that the maximum capacity of each facility is not exceeded. Constraint (4-12) ensures that all facilities currently providing service remain open. Constraints (4-13)
through (4-15) define the location \( (X_j) \), allocation \( (Y_{ij}) \), and capacity expansion \( (a_j) \) decision variables to be binary or non-negative integers.

### 4.3.2.3 Multi-Period Model

The multi-period model expands the above to include the additional notation and parameters, with most inputs and decision variables now having a time period index

\[
t = \text{index of period},
\]

\[
h_i^t = \text{total number of patients at demand node } i \text{ during period } t,
\]

\[
C_5 = \text{unused sleep bed cost per day, and}
\]

\[
D^t = \text{total demand during period } t,
\]

with the decision variables now being

\[
X_{jt} = \begin{cases} 
1 & \text{if facility } j \text{ provides sleep apnea service during period } t, \\
0 & \text{otherwise,}
\end{cases}
\]

\[
Y_{ijt} = \begin{cases} 
1 & \text{if demand node } i \text{ is covered by facility } j \text{ during period } t, \\
0 & \text{otherwise,}
\end{cases}
\]

\[
a_{jt} = \text{number of additional sleep beds in facility } j \text{ during period } t, \text{ and}
\]

\[
b_{jt} = \text{number of total sleep beds in facility } j \text{ at the beginning of period } t \text{ (an auxiliary decision variable).}
\]

The mathematical model then extends to

Minimize

\[
\sum_i \sum_j Y_{ij} h_i^t (C_1 + d_{ij} C_2) + \sum_i \sum_j Y_{ij} h_i^t
\]

\[
+ C_4 \sum_j a_j + C_5 \sum_i \sum_j \left( T(b_j^t + a_j^t) - \sum_i Y_{ij} h_i^t \right)
\]

Subject to

\[
\sum_{j \in O} X_{jt} \leq P \quad \forall \ i
\]

\[
X_{jt} \geq X_{jt}^{t-1} \quad \forall \ j, t \in \{2, 3, 4, 5, 6\}
\]
The objective function now includes a fourth term that represents the total underutilization cost, where the number of empty beds at each facility is the difference between its total capacity and total allocated patients. Constraint (4-16) ensures that at most $P$ number of sleep laboratories can be located in each period. Closing a service over the remaining planning horizon per VISN 1 policy is prevented by constraint (4-17), although this could be considered in future models. Constraints (4-18) and (4-19) calculate the number of beds in each facility at the beginning of each period, and constraint (4-20) determines the number of needed beds in each facility at each period. Constraints (4-21)-(4-28) function much as their counterparts in the single-period model (constraints (4-7)-(4-9) and (4-11)-(4-15)).
4.3.3 Application and Results

4.3.3.1 Application Details

The above models were solved using both the observed and predicted geographic demand patterns described earlier. Future demand for the next five years was estimated via the VA’s veteran population projections, summarized in Figure 4-12, with significant estimated decreases over the next 20 years due to aging veteran populations and smaller modern military sizes. The implication of these trends is that in the short term it may be optimal to source a large percent of care outside the network so as to not overbuild long-term excess capacity. These future annual demand estimates were geographically distributed across the 59 zip code demand nodes based on the current relative proportions.

![Figure 4-12 Total projected sleep apnea demand by year for entire VISN 1 (New England)](image)

Based on discussions with VISN 1 leadership, for both types of models three sets of facilities were considered for expansion:
- **Case 1**: All seven existing VA medical centers in New England,
- **Case 2**: All seven existing VA medical centers and the five largest outpatient clinics (Burlington VT, Springfield MA, Worcester MA, Portland ME, Bangor ME),
- **Case 3**: All seven existing VA medical centers and all 45 outpatient clinics across New England, regardless of size.

Cost parameters provided by VISN 1 are summarized in Table 4-4.

<table>
<thead>
<tr>
<th>Notation</th>
<th>Explanation</th>
<th>Cost ($)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$C_1$</td>
<td>In-house procedure cost per patient</td>
<td>$588</td>
</tr>
<tr>
<td>$C_2$</td>
<td>Travel cost per mile</td>
<td>$0.70</td>
</tr>
<tr>
<td>$C_3$</td>
<td>Non-coverage cost per patient</td>
<td>$850</td>
</tr>
<tr>
<td>$C_4$</td>
<td>Cost of adding a sleep bed (capacity expansion)</td>
<td>$21,000</td>
</tr>
<tr>
<td>$C_5$</td>
<td>Unused sleep bed (underutilization) cost per day</td>
<td>$385</td>
</tr>
</tbody>
</table>

The mathematical models were solved using the LINGO 11 software on a Dell desktop computer with 3 GB of RAM running a 1.8 GHz Intel Core 2 Duo processor. Both models were run for almost 100 scenarios in order to examine different coverage distance policies ($S$), numbers of allowable additional facilities ($P$), and candidate locations. With larger coverage distances or candidate locations, the feasible region and thus run times significantly increase (up to 44 minutes), as illustrated in Table 4-5 for the multi-period model.
Table 4-5 Multi-period model solution times based on different scenarios for case 1, 2, and 3

<table>
<thead>
<tr>
<th>Case</th>
<th>Acceptable travel distance ($S$) (miles)</th>
<th>Solution times (in seconds)</th>
<th>Case</th>
<th>Acceptable travel distance ($S$) (miles)</th>
<th>Solution times (in seconds)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>20</td>
<td>$&lt;1$</td>
<td></td>
<td>20</td>
<td>$&lt;1$</td>
</tr>
<tr>
<td></td>
<td>30</td>
<td>$&lt;1$</td>
<td></td>
<td>30</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>40</td>
<td>1</td>
<td></td>
<td>40</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>2</td>
<td></td>
<td>50</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>60</td>
<td>12</td>
<td></td>
<td>60</td>
<td>12</td>
</tr>
<tr>
<td>2</td>
<td>20</td>
<td>1</td>
<td></td>
<td>20</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>30</td>
<td>3</td>
<td></td>
<td>30</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>40</td>
<td>2</td>
<td></td>
<td>40</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>2</td>
<td></td>
<td>50</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>60</td>
<td>34</td>
<td></td>
<td>60</td>
<td>34</td>
</tr>
<tr>
<td>3</td>
<td>20</td>
<td>2</td>
<td></td>
<td>20</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>30</td>
<td>6</td>
<td></td>
<td>30</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>40</td>
<td>10</td>
<td></td>
<td>40</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>9</td>
<td></td>
<td>50</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>60</td>
<td>71</td>
<td></td>
<td>60</td>
<td>71</td>
</tr>
</tbody>
</table>

### 4.3.3.2 Single-Period Results

For the single-period model, Figure 4-13 summarizes the optimized tradeoff between maximum acceptable driving distance, total minimal cost, and access assuming current observed demands and that sleep apnea care can be added at zero through five additional facilities. Also shown is the current situation for comparison, which has a 2.3% higher cost and 4.5% lower access than even the “zero” case with no added capacity nor facilities, due to sub-optimal patient-to-facility allocation. Increasing the number of new facilities to three ($P = 3$) would result in a savings of $145,188 (4.7\%)$ given the same access coverage, a roughly 50% reduction in maximum driving distance to 50 miles given the same cost, or any pairwise cost-distance point on the shown curve. After three new facilities, the cost-access tradeoff starts to become negligible, indicating expansion beyond $P > 3$ facilities may not be useful for practical purposes.
Table 4-6 summarizes general results and savings for the single-period model assuming estimated true demands and based on a 60-mile maximum travel distance, showing the costs of fee and in-house visits. The rows with $P = 0$ correspond to the cases of modified capacities in existing facilities but not additional sleep care anywhere else.
Table 4-6 Results for single-period model with 60-mile maximum acceptable travel distance

<table>
<thead>
<tr>
<th>Case</th>
<th>Num. add. facilities ($P$)</th>
<th>Total cost</th>
<th>In-house type visit cost</th>
<th>Fee type visit cost</th>
<th>% in-house type visits</th>
<th>Optimal facility locations (total number of beds)**</th>
<th>Savings</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 (7 candidate locations)</td>
<td>Est. obs.*</td>
<td>$7,455,189</td>
<td>$3,252,421</td>
<td>$4,202,768</td>
<td>51%</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>$7,274,793</td>
<td>$3,494,030</td>
<td>$3,759,762</td>
<td>53%</td>
<td>-</td>
<td>$110,085</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>$7,020,768</td>
<td>$4,342,393</td>
<td>$2,573,375</td>
<td>70%</td>
<td>Northampton(4)</td>
<td>$434,421</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>$6,874,504</td>
<td>$4,896,599</td>
<td>$1,809,905</td>
<td>79%</td>
<td>Togus(3)-Northampton(4)</td>
<td>$580,685</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>$6,810,171</td>
<td>$5,102,943</td>
<td>$1,518,228</td>
<td>82%</td>
<td>Togus(3)-Northampton(4)-WRJ(1)</td>
<td>$645,018</td>
</tr>
<tr>
<td>2 (12 candidate locations)</td>
<td>Est. obs.*</td>
<td>$7,455,189</td>
<td>$3,252,421</td>
<td>$4,202,768</td>
<td>51%</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>$7,274,793</td>
<td>$3,494,030</td>
<td>$3,759,762</td>
<td>53%</td>
<td>-</td>
<td>$110,085</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>$7,007,062</td>
<td>$4,296,089</td>
<td>$2,605,972</td>
<td>70%</td>
<td>Springfield(4)</td>
<td>$448,127</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>$6,860,798</td>
<td>$4,850,296</td>
<td>$1,842,502</td>
<td>79%</td>
<td>Togus(3)-Springfield(4)</td>
<td>$594,391</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>$6,775,736</td>
<td>$5,134,086</td>
<td>$1,452,650</td>
<td>83%</td>
<td>Togus(3)-Springfield(4)-Portland(2)</td>
<td>$679,453</td>
</tr>
<tr>
<td>3 (52 candidate locations)</td>
<td>Est. obs.*</td>
<td>$7,455,189</td>
<td>$3,252,421</td>
<td>$4,202,768</td>
<td>51%</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>$7,274,793</td>
<td>$3,494,030</td>
<td>$3,759,762</td>
<td>53%</td>
<td>-</td>
<td>$110,085</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>$7,007,062</td>
<td>$4,296,089</td>
<td>$2,605,972</td>
<td>70%</td>
<td>Springfield(4)</td>
<td>$448,127</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>$6,860,798</td>
<td>$4,850,296</td>
<td>$1,842,502</td>
<td>79%</td>
<td>Togus(3)-Springfield(4)</td>
<td>$594,391</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>$6,774,937</td>
<td>$5,153,385</td>
<td>$1,411,553</td>
<td>84%</td>
<td>Togus(3)-Springfield(4)-Saco(2)</td>
<td>$680,252</td>
</tr>
</tbody>
</table>

*Estimated observed case
**For all cases, optimum solution also requires a capacity expansion of one bed at the current facility in Providence.

Figure 4-14 illustrates similar information and conclusions now using the estimated true demands. Table 4-7 further summarizes optimal locations and their corresponding capacities, using true demand, for 27 different scenarios based on coverage policy, additional number of facilities, and set of candidates. Patterns in this table provided important insight for decision makers. For example, capacity expansion at the current sleep laboratory in Providence always is suggested in all scenarios, removing most debate about the accuracy of various assumptions and demand estimates. Several other candidates also are frequently suggested as optimal expansion facilities, namely, Newington, Togus, Northampton, and Springfield.
Figure 4-14 Optimized tradeoff between acceptable distance and (a) total cost, and (b) coverage percentage for single-period model assuming estimated true demand and candidate facilities under case 3.
Table 4-7 Optimal facility locations for single-period model based on estimated true demand (numbers in parentheses represent the number of additional beds in corresponding facilities)

<table>
<thead>
<tr>
<th>$S$ (miles)</th>
<th>$P$</th>
<th>Case 1</th>
<th>Case 2</th>
<th>Case 3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Facility</td>
<td>Beds</td>
<td>Capacity expansion</td>
</tr>
<tr>
<td>30</td>
<td>1</td>
<td>Newington</td>
<td>4</td>
<td>Providence (+2)</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>Newington</td>
<td>4</td>
<td>Providence (+2)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Northampton</td>
<td>2</td>
<td>Providence (+2)</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>Jamaica Plain</td>
<td>4</td>
<td>Providence (+2)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Northampton</td>
<td>1</td>
<td>Providence (+2)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>WRJ</td>
<td>1</td>
<td>Providence (+2)</td>
</tr>
<tr>
<td>60</td>
<td>1</td>
<td>Newington</td>
<td>4</td>
<td>Providence (+1)</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>Togus</td>
<td>3</td>
<td>Providence (+1)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Northampton</td>
<td>4</td>
<td>Providence (+1)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>WRJ</td>
<td>1</td>
<td>Providence (+1)</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>Togus</td>
<td>3</td>
<td>Providence (+1)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Northampton</td>
<td>4</td>
<td>Providence (+2)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>WRJ</td>
<td>1</td>
<td>Providence (+2)</td>
</tr>
<tr>
<td>90</td>
<td>1</td>
<td>Northampton</td>
<td>4</td>
<td>Providence (+2)</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>Togus</td>
<td>3</td>
<td>Providence (+2)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Northampton</td>
<td>4</td>
<td>Providence (+2)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>WRJ</td>
<td>1</td>
<td>Providence (+2)</td>
</tr>
</tbody>
</table>

Table 4-8 summarizes the corresponding “allocation” results; that is, the assignment of patients to facilities (or fee basis) based on home zip code. As shown, even though the optimal locations change for each scenario, zip codes are clustered in similar manners; that is, similar groups of demand nodes are assigned to the same facilities. Most patients living in Vermont also become fee type visits, which follows intuition given that demand is relatively low in Vermont as compared to other regions. Figure 4-15 visualizes the demand nodes covered within-network (in-house type) and outside-network (fee type) as well as the VA facility locations that provide service assuming a 60-mile acceptable maximum travel distance and the candidate facilities given by case 3.
### Table 4-8 Allocation of demand nodes to facilities for single-period model

<table>
<thead>
<tr>
<th>Facilities</th>
<th>Assigned demand node 3-digit zip codes</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Case 1</strong></td>
<td></td>
</tr>
<tr>
<td>West Haven</td>
<td>064-065-066-067-068-069 (CT)</td>
</tr>
<tr>
<td>Togus</td>
<td>041-042-043-045-048-049 (ME)</td>
</tr>
<tr>
<td>West Roxbury</td>
<td>015-016-017-019-020-021-022-023-024-027 (MA), 029 (RI)</td>
</tr>
<tr>
<td>Northampton</td>
<td>060-061 (CT), 010-011-012-013-053 (MA)</td>
</tr>
<tr>
<td>Manchester</td>
<td>014-018 (MA), 030-031-032-033-034 (NH), 039 (ME)</td>
</tr>
<tr>
<td>Providence</td>
<td>062-063 (CT), 028 (MA)</td>
</tr>
<tr>
<td>WRJ</td>
<td>037 (NH), 050-051-057 (VT)</td>
</tr>
<tr>
<td>Fee type</td>
<td>025-026 (MA), 035-036-038 (NH), 040-044-046-047 (ME) 052-054-056-058-059 (VT)</td>
</tr>
<tr>
<td><strong>Case 2</strong></td>
<td></td>
</tr>
<tr>
<td>West Haven</td>
<td>064-065-066-067-068-069 (CT)</td>
</tr>
<tr>
<td>Togus</td>
<td>042-045-048-049 (ME)</td>
</tr>
<tr>
<td>West Roxbury</td>
<td>015-016-017-019-020-021-022-023-024-027 (MA), 029 (RI)</td>
</tr>
<tr>
<td>Manchester</td>
<td>014-018 (MA), 030-031-032-033-034 (NH)</td>
</tr>
<tr>
<td>Providence</td>
<td>062-063 (CT), 028 (RI)</td>
</tr>
<tr>
<td>Springfield</td>
<td>060-061 (CT), 010-011-012-013 (RI)</td>
</tr>
<tr>
<td>Portland</td>
<td>038-039-040-044 (ME)</td>
</tr>
<tr>
<td>Fee type</td>
<td>025-026 (MA), 035-036-037 (NH), 041-044-046-047 (ME), 050-051-052-053-054-056-057-058-059 (VT)</td>
</tr>
<tr>
<td><strong>Case 3</strong></td>
<td></td>
</tr>
<tr>
<td>West Haven</td>
<td>064-065-066-067-068-069 (CT)</td>
</tr>
<tr>
<td>Togus</td>
<td>042-043-045-048-049 (ME)</td>
</tr>
<tr>
<td>West Roxbury</td>
<td>015-016-017-019-020-021-022-023-024-027 (MA), 029 (RI)</td>
</tr>
<tr>
<td>Manchester</td>
<td>014-018 (MA), 030-031-032-033-034 (NH)</td>
</tr>
<tr>
<td>Providence</td>
<td>062-063 (CT), 028 (MA)</td>
</tr>
<tr>
<td>Springfield</td>
<td>060-061 (CT), 010-011-012-013 (MA)</td>
</tr>
<tr>
<td>Saco</td>
<td>038-039-040-041 (ME)</td>
</tr>
<tr>
<td>Fee type</td>
<td>025-026 (MA), 035-036-037 (NH), 044-046-047 (ME), 050-051-052-053-054-056-057-058-059 (VT)</td>
</tr>
</tbody>
</table>
4.3.3.3 Multi-Period Results

The multi-period model was solved for five, ten, and twenty-five year planning horizons. For the 5-year case, Table 4-9 summarizes general results, again assuming a 60-mile acceptable maximum travel distance. Since demand is expected to decrease over time, the optimal solution
adds additional facilities only in the first year, although this must not necessarily be the case in
general.

Table 4-9 Results for multi-period model with 60-mile maximum acceptable travel distance

<table>
<thead>
<tr>
<th>Case</th>
<th>Num. add. facilities (P)</th>
<th>Average annual cost</th>
<th>In-house type visit cost</th>
<th>Fee type visit cost</th>
<th>% in-house type visits</th>
<th>Optimal facility locations (total number of beds)**</th>
<th>Savings</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 (7 candidate locations)</td>
<td>Est. obs.* $7,455,189</td>
<td>$3,252,421</td>
<td>$4,202,768</td>
<td>51%</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>0</td>
<td>$6,960,730</td>
<td>$3,305,465</td>
<td>$3,392,237</td>
<td>58%</td>
<td>-</td>
<td>-</td>
<td>$494,459</td>
</tr>
<tr>
<td>1</td>
<td>$6,734,157</td>
<td>$3,929,943</td>
<td>$2,513,087</td>
<td>69%</td>
<td>Northampton(3)</td>
<td>$721,032</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>$6,587,880</td>
<td>$4,360,938</td>
<td>$1,920,172</td>
<td>76%</td>
<td>Togus(2)-Northampton(3)</td>
<td>$867,309</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>$6,517,772</td>
<td>$4,562,840</td>
<td>$1,635,157</td>
<td>80%</td>
<td>Togus(2)-Northampton(3)-WRJ(1)</td>
<td>$937,417</td>
<td></td>
</tr>
<tr>
<td>2 (12 candidate locations)</td>
<td>Est. obs.* $7,455,189</td>
<td>$3,252,421</td>
<td>$4,202,768</td>
<td>51%</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>1</td>
<td>$6,723,166</td>
<td>$3,914,317</td>
<td>$2,516,278</td>
<td>69%</td>
<td>Springfield(3)</td>
<td>$732,023</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>$6,576,890</td>
<td>$4,345,313</td>
<td>$1,923,362</td>
<td>76%</td>
<td>Togus(2)-Springfield(3)</td>
<td>$878,299</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>$6,489,340</td>
<td>$4,709,745</td>
<td>$1,416,678</td>
<td>82%</td>
<td>Togus(2)-Springfield(3)-Portland(2)</td>
<td>$965,849</td>
<td></td>
</tr>
<tr>
<td>3 (52 candidate locations)</td>
<td>Est. obs.* $7,455,189</td>
<td>$3,252,421</td>
<td>$4,202,768</td>
<td>51%</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>0</td>
<td>$6,960,730</td>
<td>$3,305,465</td>
<td>$3,392,237</td>
<td>58%</td>
<td>-</td>
<td>-</td>
<td>$494,459</td>
</tr>
<tr>
<td>1</td>
<td>$6,723,166</td>
<td>$3,914,317</td>
<td>$2,516,278</td>
<td>69%</td>
<td>Springfield(3)</td>
<td>$732,023</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>$6,574,090</td>
<td>$4,346,015</td>
<td>$1,920,952</td>
<td>76%</td>
<td>Togus(2)-Springfield(3)</td>
<td>$878,299</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>$6,489,340</td>
<td>$4,709,745</td>
<td>$1,416,678</td>
<td>82%</td>
<td>Togus(2)-Springfield(3)-Portland(2)</td>
<td>$965,849</td>
<td></td>
</tr>
</tbody>
</table>

*Estimated observed case
**For all cases, optimum solution also requires a capacity expansion of one bed at the current facility in Providence.

As above, Figure 4-16 shows the tradeoff frontier between acceptable travel distance, average
annual cost, and coverage percentage assuming sleep beds can be added to zero through four
additional facilities and the candidate facilities given by case 3. Due to excessive computational
times, the multi-period model was solved only for maximum coverage distances of 60 miles or
less. Similar to the single-period results, the current situation also is shown for comparison. Even
with the “zero” case, an improvement of 6.6% in annual cost and 12.1% in access is possible.
Increasing the number of new facilities in which sleep beds could be added to three could result
in a savings of $965,849 (12.9%) given the same access coverage, a reduction in maximum
driving distance to 42 miles given the same cost, or any pairwise cost-distance point on the
shown curve. After service capacity is added in three new facilities, again, the cost-access tradeoff starts to become negligible.

![Optimized tradeoff between acceptable distance and cost](image)

*Figure 4-16 Optimized tradeoff between acceptable distance and (a) average annual cost, and (b) coverage percentage for multi-period model assuming case 3*

Table 4-10 summarizes the optimal locations and their corresponding capacities for 27 different scenarios as previously. Similar to the single-period case, capacity expansion at the current Providence sleep laboratory always is suggested regardless of assumptions about acceptable
travel distances (greater than 20 miles), with facilities Newington, Togus, Bedford, and Springfield again being common additional expansion candidates. The broad similarity of these results under a variety of assumptions and to those of the single-period model has been useful for management to better understand obvious and robust decisions.

Table 4-10 Optimal facility locations for multi-period model (numbers in parentheses represent the number of additional beds in corresponding facilities)

<table>
<thead>
<tr>
<th>S (miles)</th>
<th>P</th>
<th>Case 1</th>
<th>Case 2</th>
<th>Case 3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Facility</td>
<td>Beds</td>
<td>Capacity expansion</td>
</tr>
<tr>
<td>20</td>
<td>1</td>
<td>Newington</td>
<td>2</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>Newington, Bedford</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>Newington, Bedford, Northampton</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>40</td>
<td>1</td>
<td>Newington, Providence (+1)</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>Newington, Providence (+1)</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>Newington, Providence (+1)</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>60</td>
<td>1</td>
<td>Northampton, Providence (+1)</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>Togus, Northampton, Providence (+1)</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>Togus, Northampton, WRJ, Providence (+1)</td>
<td>2</td>
<td></td>
</tr>
</tbody>
</table>

Table 4-11 summarizes patient-facility assignments (based on home zip code) for the multi-period case, which again are very similar to those for the single-period model. One noticeable difference is that the capacity expansions in the multi-period model are relatively lower than those in the single-period model, which makes intuitive sense given shrinking veteran population projections. This also results in the optimal number of patients that receive fee type service being
higher especially in earlier years, contrary to current policy, with more patients who live outside of Vermont now also being sent to non-VA facilities.

Table 4-11 Allocation of demand nodes to facilities for multi-period model

<table>
<thead>
<tr>
<th>Case</th>
<th>Facilities</th>
<th>Assigned demand node 3-digit zip codes</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>West Haven</td>
<td>061-064-065-066-067-068-069 (CT)</td>
</tr>
<tr>
<td></td>
<td>Togus</td>
<td>041-042-043-045-048-049 (ME)</td>
</tr>
<tr>
<td></td>
<td>West Roxbury</td>
<td>015-016 (year 1-3)-017-018 (year 5-6)-019-020-021-022-023-024-027 (MA); 029 (RI)</td>
</tr>
<tr>
<td></td>
<td>Northampton</td>
<td>010-011-012-013 (MA); 053 (VT); 060 (CT)</td>
</tr>
<tr>
<td></td>
<td>Manchester</td>
<td>014-018 (year 1-4) (MA); 030-031-032-033-034 (NH); 039 (ME)</td>
</tr>
<tr>
<td></td>
<td>Providence</td>
<td>016 (year 4-6) (MA); 028 (RI); 062-063 (CT)</td>
</tr>
<tr>
<td></td>
<td>WRJ</td>
<td>036-037 (NH); 050-051-057 (VT)</td>
</tr>
<tr>
<td></td>
<td>Fee type</td>
<td>025-026 (MA); 035-038 (NH); 040-044-046-047 (ME); 052-054-056-058-059 (VT)</td>
</tr>
<tr>
<td>2</td>
<td>West Haven</td>
<td>060 (year 3)-061-064 (year 1-2,4-6)-065-066-067-068-069 (CT)</td>
</tr>
<tr>
<td></td>
<td>Togus</td>
<td>042-043 (year 4-6)-045 (year 1-3)-048-049 (ME)</td>
</tr>
<tr>
<td></td>
<td>West Roxbury</td>
<td>015-016 (year 1-3)-017-018 (year 5-6)-019-020-021-022-023-024-027 (MA); 029 (RI)</td>
</tr>
<tr>
<td></td>
<td>Manchester</td>
<td>014-018 (year 1-4) (MA); 030-031-032-033-034 (NH); 039 (year 2) (ME)</td>
</tr>
<tr>
<td></td>
<td>Providence</td>
<td>016 (year 4-6) (MA); 028 (RI); 062-063 (CT)</td>
</tr>
<tr>
<td></td>
<td>Springfield</td>
<td>010-011-012-013 (MA); 060 (year 1-2,4-6)-064 (year 3) (CT)</td>
</tr>
<tr>
<td></td>
<td>Portland</td>
<td>038 (NH); 039 (year 1,3-6)-040-041-043 (year 1-3)-045 (year 4-6) (ME)</td>
</tr>
<tr>
<td></td>
<td>Fee type</td>
<td>025-026 (MA); 035-036-037 (NH); 044-046-047 (ME); 050-051-052-053-054-056-057-058-059 (VT)</td>
</tr>
<tr>
<td>3</td>
<td>West Haven</td>
<td>060 (year 3)-061-064 (year 1-2,4-6)-065-066-067-068-069 (CT)</td>
</tr>
<tr>
<td></td>
<td>Togus</td>
<td>042-043 (year 4-6)-045 (year 1-3)-048-049 (ME)</td>
</tr>
<tr>
<td></td>
<td>West Roxbury</td>
<td>015-016 (year 1-3)-017-018 (year 5-6)-019-020-021-022-023-024-027 (MA); 029 (RI)</td>
</tr>
<tr>
<td></td>
<td>Manchester</td>
<td>014-018 (year 1-4) (MA); 030-031-032-033-034 (NH); 039 (year 2) (ME)</td>
</tr>
<tr>
<td></td>
<td>Providence</td>
<td>016 (year 4-6) (MA); 028 (RI); 062-063 (CT)</td>
</tr>
<tr>
<td></td>
<td>Springfield</td>
<td>010-011-012-013 (MA); 060 (year 1-2,4-6)-064 (CT)</td>
</tr>
<tr>
<td></td>
<td>Portland</td>
<td>038 (NH); 039 (year 1,3-6)-040-041-043 (year 1-3)-045 (year 4-6) (ME)</td>
</tr>
<tr>
<td></td>
<td>Fee type</td>
<td>025-026 (MA); 035-036-037 (NH); 044-046-047 (ME); 050-051-052-053-054-056-057-058-059 (VT)</td>
</tr>
</tbody>
</table>

In addition to the 5 year planning horizon, 10 and 25 year horizons also were considered. To reduce execution times for the 10 and 25 year horizon cases, increments of two-year and five-year periods were employed respectively, with the projected demand for each $n$-year period aggregated accordingly (that is, each model still used 5 periods, for example with the 25 year...
model treating the first 5 years as one period, the second 5 years as a second period, and so on). The average annual cost and coverage percentage for each planning horizon assuming sleep beds can be added to at most three additional facilities and the candidate facilities given by case 3 are shown in Figure 4-17. Of note, as the planning horizon increases the average annual cost decreases, primarily because of long-term decreases in demand.

Figure 4-17 Optimal (a) annual cost and (b) coverage percentages considering different planning horizons (both for case 3 and $P = 3$)
4.3.4 Sensitivity Analysis on Demand Variability

The above mathematical models assume that demand for VA specialty care services is deterministic and equally distributed over time. Although assumption of deterministic demand significantly decreases the complexity of a mathematical model, it may not always be justified. Given the future veteran population uncertainty due to unpredictable military conflicts, additional research was conducted to study the impact of stochastic demand on the deterministic model results. Two simulation models, as explained in Figure 4-18, were developed that consider demand variability on a yearly and monthly basis. Assuming the optimal locations and capacities that resulted from the “single period model for case 3 and $P = 3$”, the changes in the overall system results were examined assuming variable annual and monthly demands.

```
Prompt user for desired coefficient of variation (CV) and half-width values (desired HW)
Initializations (outside-replications)
Do While HW > desired HW
  Initializations (within-replications)
    FOR i = 1 to 59  !for each demand node
      Generate demand; ~ NORMAL($\mu_i$, $\sigma_i$)
    NEXT
    FOR j = 1 to 7  !for each facility
      FOR i = 1 to 59
        IF demand node $i$ is assigned to facility $j$
          IF capacity $j$ is not exceeded
            Let house_cost = house_cost + unit_house_cost*demand;
          ENDIF
        ENDIF
        IF capacity $j$ is exceeded
          Let fee_cost = fee_cost + unit_fee_cost*demand;
          Let unanticipated_fee = unanticipated_fee + demand;
          Let noncovered = noncovered + demand;
        ENDIF
      NEXT
    NEXT
  Calculate HW
ENDDO
Calculate average and standard deviation of costs, unanticipated fee patients, coverage percentage
Print results
END
```

If monthly demand considered, add:
```
  For k = 1 to 12  !for each month
    Generate demand$_k$ ~ NORMAL($\mu_i/12$, $\sigma_i^2/12$)
```

Figure 4-18 Pseudo code for the simulation models with demand variability
Figure 4-19 illustrates the changes in average total, house, and fee costs and number of unanticipated fee patients (i.e., patients that are sent to a VA facility by the deterministic model but cannot be served there due to capacity constraints) assuming variable annual demands. As expected, the number of unanticipated fee patients significantly increases as the coefficient of variation (CV) gets larger, resulting in higher fee cost and thus total cost. The average house cost, on the other hand, decreases as the CV increases, mainly because of the capacity constraint. Due to the limited capacity, as demand varies, house cost cannot increase more than an upper limit (i.e., where all facilities are fully utilized), while there is no lower limit. Since all facilities are almost fully utilized in the no-variation experiments (i.e., CV = 0), as the CV increases, positive deviations have no/limited effects on the house cost while negative deviations pull it down drastically. This causes the average house cost to be lower than the one with the deterministic demand.

As shown in Figure 4-20, the number of unanticipated fee patients assuming variable monthly demands is significantly higher than those with the annual demand variability. This is because the monthly variation is higher than the annual variation assuming the same annual CV values, which results in less ability to meet all demands within-network. Accordingly, increases in the average total cost, number of unanticipated fee patients, and related fee cost and decreases in the average house cost are more significant in the monthly demand variability case when compared to those with the variable annual demand.

Overall, this analysis suggests that accounting for demand variability makes significant changes in the performance and may change the system design completely. Thus, stochastic modeling approaches may be beneficial even though they increase mathematical complexity significantly.
Figure 4-19 Changes in average (a) total cost, (b) number of unanticipated fee patients, (c) house cost, and (d) fee cost assuming variable annual demands and single period model results for case 3 and $P = 3$.
Figure 4-20 Changes in average (a) total cost, (b) number of unanticipated fee patients, (c) house cost, and (d) fee cost assuming variable monthly demands and single period model results for case 3 and $P = 3$. 
4.3.5 Discussion

As the U.S. veteran population ages, effective delivery of health services becomes more important in terms of optimal within-network coverage at minimum total cost and travel distance. This study investigated the optimal short and long term location of sleep apnea testing services across New England VA facilities. Results indicate that significant cost savings and improvements in patient-centeredness can be achieved simply by optimizing capacity in existing VA facilities. These potential savings demonstrate this is a useful approach to help decision makers make more deliberate and tactical healthcare service location and capacity optimization decisions. In this particular application, significant savings are possible by opening an additional sleep laboratory to serve unmet need in northern Connecticut and western Massachusetts. If closing or relocating underutilized facilities were allowed, even greater improvements could be achieved, especially in the multi-period case. These results also have implications on the optimal amount of out-of-network “fee” care that run contrary to current practice and beliefs. Counter to the VA’s desire to eliminate almost all fee basis care, in almost all specialty care applications and scenarios examined to-date, roughly 10 to 25% of all care is optimal to occur outside of the VA network. To better understand the costs of out-of-network care, the above models were modified to require that all patients must be covered by VA facilities, resulting in a 34% increase in average travel distance while decreasing the total cost only by 4% (versus roughly 5% when compared with the $S = 60$ and $P = 3$ case).

Possible limitations of this study include the assumption that out of network capacity always is available within desired distances, which may not reflect the actual market. Additional savings might be achieved by adding or relocating existing facilities. Expansion of the system is perhaps
less disruptive than facility relocation, the latter potentially forcing staff to relocate or find work elsewhere. Since the disruptive nature of change may limit a health system’s ability to immediately implement an optimized network design, some thought might be put into multi-period transition or minimally disruptive models.

These models could be expanded in several manners. These include accounting for demand variability naturally occurring over time and future veteran population uncertainty due to such things as the unpredictability of future military conflicts. Stochastic modeling approaches might be appropriate here, such as recourse, chance-constraint, and dependent-chance programs. Multi-period capacity scale-down models also might be considered, although implementation initially may be limited by “permanent” contracts and the ability to retrain staff. Finally, implementation of these models in user-friendly decision support tools also would be useful to enable decision makers to explore what-if scenarios more readily.
Chapter 5 - Conclusions and Future Work

This study focuses on developing systems engineering models to improve the effectiveness and cost of and access to screening, diagnostic, and care services for several mental/general health conditions within large integrated healthcare systems. The use of several systems engineering/operations research modeling techniques were illustrated within the Veterans Health Administration (VHA) to discover potential improvement opportunities in providing effective and timely care to the veterans with significant health problems while considering overall system cost. Six mathematical models – probability modeling, Monte Carlo, fuzzy logic, logistic regression, artificial neural networks, and integer programming – were developed and illustrated across a range of specialty care services, namely, traumatic brain injury (TBI), posttraumatic stress disorder (PTSD), and sleep apnea screening, diagnostic, and treatment services. Results indicate that these models could make a significant contribution to the safety and long-term well-being of U.S. servicemen. For each specific model, detailed results, limitations, and possible extensions are provided at the end of each chapter.

As a summary, the first chapter of this study investigated the current and proposed screening processes for PTSD within the VHA and further explored the effects of several model parameters on the number of false-diagnoses and overall cost. Results indicate that lower false diagnosis
rates, predictable performance, and reduced costs can be achieved by using a more intentionally
designed system. Also developed was a sequential screening model for mild TBI to illustrate
how the general approach can be applied to other disorders. The biggest limitation of this study
is that the results are very dependent on the model inputs, most of which are not known with
certainty. One possible extension, therefore, is to conduct further research to better estimate these
inputs. Additionally, the proposed screening process for mild TBI could be further developed
and evaluated by better identifying the model parameters and by conducting a comprehensive
sensitivity analysis on the inputs that are subject to uncertainty.

In the second chapter of this study, a comprehensive diagnostic model which aims to ensure
proper and timely diagnosis of TBI was developed. Three types of mathematical methods were
used to categorize TBI patients into the most probable severity states, with preliminary results
indicating a fairly strong agreement with the patients’ known health states for each method. This
study could be extended in a number of ways, one of them being to develop mathematical
models to optimize overall or severity-specific performance of the proposed model.

Finally, the last chapter illustrated the use of location-allocation modeling approach across a
range of specialty care services, namely, PTSD treatment and sleep apnea testing services, within
the VHA. In both applications, results indicate that significant cost savings and improvements in
patient-centeredness can be achieved. These potential savings demonstrate that this is a useful
approach to help decision makers make more informed and tactical decisions. General
limitations include the assumptions regarding model inputs and deterministic demand. The
mathematical models could be extended in a manner so as to account for demand variability and
uncertainty. Another extension could be to develop a co-location model which simultaneously
determines the optimal locations and capacities for multiple mental health services (e.g., PTSD and depression care services), given that the mental health conditions often co-occur and interact with each other.

More generally, the results of this study underscore three important issues: (1) current care systems are insufficient to meet the mental health and cognitive needs of U.S. service members and produce extra costs, (2) significant opportunities for mental health care process improvements and cost reductions are possible with the use of systems engineering methods such as statistical modeling and mathematical optimization, and (3) further improvements in the long-term care of U.S. servicemen can be achieved by applying these models over a range of other chronic health conditions. However, there exist several implementation challenges given that the results do not consider the potential for reduced care continuity and that the disruptive nature of change may limit a health system’s ability to immediately implement an optimized system design.

The mathematical models described in this study all share another related potential limitation which is poor “trialability”, especially given that healthcare practitioners have become more accustomed to gradual iterative PDSA improvement cycles than to instantaneous reengineering. Small cycles of change allow process improvement personnel to obtain safe feedback to minor changes before proceeding with broader implementation. Solutions such as those in this study, conversely, may require more of a leap of faith. Unfortunately, since few healthcare practitioners and leaders understand modeling, simulation, and optimization methods, it may be difficult for them to fully trust in results and make disruptive changes.
Although the potential improvements that can be achieved with the use of each mathematical model developed in this study are illustrated separately, it is important to emphasize that these models and their results highly interact with each other. For instance, the screening and diagnostic models determine the number of patients in need of using care services in a population, which actually constitutes the true demand for care services – one of the inputs in the network optimization models. Integration of these models in a manner so as to use one model’s output as another’s input, therefore, could help identify the correct inputs for each model and thus provide more accurate results.

In conclusion, there are a number of possibilities to extend this study, all of which are discussed throughout the dissertation. These recommendations are summarized as follows:

1. Given that the results are very dependent on the model inputs, additional research could be conducted either to confirm the estimates used in this study or to make better estimates. Integration of mathematical models and using one’s output as another’s input also could help identify the inputs more accurately.

2. Mathematical models could be extended in order to account for (i) demand variability naturally occurring over time and (ii) future veteran population uncertainty due to the unpredictability of future military conflicts, using stochastic modeling approaches such as recourse, chance-constraint, and dependent-chance programs.

3. In order to address the difficulties in implementation and the poor trialability, multi-period transition or minimally disruptive models could be considered. The development
of graphical and interactive tools also could be useful for building initial acceptance, understanding, support, and trust.

(4) Finally, implementation of these models in user-friendly decision support tools would be useful to enable decision makers to explore what-if scenarios more readily.
Bibliography


Bibliography


Colton, C.W. and Manderscheid, R.W., 2006, "Congruencies in increased mortality rates, years of potential life lost, and causes of death among public mental health clients in eight states", *Preventing Chronic Disease*, 3(2): A42.


Hoge, C.W., Auchterlonie, J.L., and Milliken, C.S., 2006, "Mental health problems, use of mental health services, and attrition from military service after returning from deployment to Iraq or Afghanistan", *Journal of the American Medical Association (JAMA)*, 295(9): 1023-1032.


Bibliography


Seal, K.H., Bertenthal, D., Miner, C.R., Sen, S., and Marmar, C., 2007, "Bringing the war back home: Mental health disorders among 103,788 US veterans returning from Iraq and Afghanistan seen at Department of Veterans Affairs facilities", Archives of Internal Medicine, 167(5): 476-482.


Bibliography


Appendix A - Overall Sensitivity of PTSD Screening Process

In order to introduce some notation and motivate the general approach, the one-year analysis horizon is considered first. In this process, a truly positive patient can be given a final correct positive diagnosis in the following two mutually exclusive and collectively exhaustive manners.

Note that PC-PTSD, PCL, and CAPS tests are represented by S, P, and C, respectively.

(S′, S+, P′, P\bar{C}, \bar{P}):
- Patient takes PC-PTSD,
- PC-PTSD correctly identifies patient as positive,
- Patient is sent to PCL for review and takes it, and
- PCL makes a definite diagnosis and confirms patient as positive.

(S′, S+, P′, P\bar{C}, C\bar{C}, C^+):
- Patient takes PC-PTSD,
- PC-PTSD correctly identifies patient as positive,
- Patient is sent to PCL for review and takes it,
- PCL cannot make a definite diagnosis,
- Patient is sent to CAPS for further review and takes it, and
- CAPS confirms patient as positive.

The probability of detecting a truly positive patient in one year, \( p_1 \), then is the sum of the probabilities of these events:
Appendix A - Overall Sensitivity of PTSD Screening Process

\[ p_j = \text{Probability of detecting a truly positive patient in one year} \]

\[ = P\left\{ \left( S^\prime, S^+, P^\prime, P^+ \right) \text{ or } \left( S^\prime, S^+, P^\prime, P^+, C^\prime, C^+ \right) \right\} \]

\[ = P\{s^\prime\} P\{s^+\} P\{p^\prime\} \left( 1 - P\{p^+\} \right) P\{p^+\} + P\{s^\prime\} P\{s^+\} P\{p^\prime\} P\{p^+\} P\{c^\prime\} P\{c^+\} \]

\[ = r_s(1 - \alpha_s)r_p(1 - u)(1 - \alpha_p) + r_s(1 - \alpha_s) r_p u \epsilon (1 - \alpha_c) \]

\[ = r_s(1 - \alpha_s) r_p \left( 1 - \alpha_p - u \left( 1 - \alpha_p - r_c (1 - \alpha_c) \right) \right). \quad \text{(A-1)} \]

More generally, a correct positive determination in a \( j \)-year horizon can occur in the following manners:

\[ (((S_1^\prime, S_1^-) \text{ or } S_1^\prime L \text{ or } S_1^\prime, S_1^+, P^\prime) \text{ or } (S_1^\prime, S_1^+, P^\prime, P^+, C^\prime)), \ldots, ((S_i^\prime, S_{i-1}^-) \text{ or } S_i^\prime L \text{ or } S_i^\prime, S_i^+, P^\prime, P^+, C^\prime)), \ldots, ((S_i^\prime, S_{i-1}^-) \text{ or } S_i^\prime L \text{ or } S_i^\prime, S_i^+, P^\prime, P^+, C^\prime)), S_i^\prime, S_i^+, P^\prime, P^+, C^\prime), \ldots), i \leq j): \]

- PC-PTSD in year \( i, 1 \leq i \leq j \), correctly identifies patient as positive after \( i-1 \geq 0 \) years being determined as negative, skipping PC-PTSD, or being determined as positive by PC-PTSD but not following up with confirmatory testing,
- Patient is sent to PCL for review and takes it, and
- PCL makes a definite diagnosis and confirms patient as positive.

\[ (((S_1^\prime, S_1^-) \text{ or } S_1^\prime L \text{ or } S_1^\prime, S_1^+, P^\prime) \text{ or } (S_1^\prime, S_1^+, P^\prime, P^+, C^\prime)), \ldots, ((S_i^\prime, S_{i-1}^-) \text{ or } S_i^\prime L \text{ or } S_i^\prime, S_i^+, P^\prime, P^+, C^\prime)), S_i^\prime, S_i^+, P^\prime, P^+, C^\prime, C^+), i \leq j): \]

- PC-PTSD in year \( i, 1 \leq i \leq j \), correctly identifies patient as positive after \( i-1 \geq 0 \) years being determined as negative or not skipping PC-PTSD, or being determined as positive by PC-PTSD but not following up with confirmatory testing,
- Patient is sent to PCL for review and takes it,
- PCL cannot make a definite diagnosis,
- Patient is sent to CAPS for further review and takes it, and
- CAPS confirms patient as positive.
Appendix A - Overall Sensitivity of PTSD Screening Process

Note that both events above can occur in \( j \) manners, as the patient can be identified as positive by either PC-PTSD test in year \( i = 1, 2, \ldots, j \). Now, the generalized analog to Eq. (A-1) is

\[ p_j = \text{Probability of detecting a truly positive patient in } j \text{ years} \]

\[ = P\left\{ \left( \left( S_1^-, S_1^- \right) \text{ or } S_1^- \text{ or } \left( S_1^-, S_1^+ \text{ or } P^\prime \text{ or } P^\prime, C^\prime \right) \right), \ldots, \left( \left( S_i^-, S_i^- \right) \text{ or } S_i^- \text{ or } \left( S_i^-, S_i^+ \text{ or } P^\prime \text{ or } P^\prime, C^\prime \right) \right), \ldots, \left( \left( S_{i-1}^-, S_{i-1}^- \text{ or } S_{i-1}^- \text{ or } \left( S_{i-1}^-, S_{i-1}^+ \text{ or } P^\prime \text{ or } P^\prime, C^\prime \right) \right), \ldots, \left( \left( S_j^-, S_j^- \right) \text{ or } S_j^- \text{ or } \left( S_j^-, S_j^+ \text{ or } P^\prime \text{ or } P^\prime, C^\prime \right) \right), \right\} \times \left( P\{P^\prime, P^\prime, C^\prime, C^\prime\} + P\{P^\prime, P^\prime, C^\prime, C^\prime\} \right) \]

\[ = P\left\{ \left( \left( S_1^-, S_1^- \right) \text{ or } S_1^- \text{ or } \left( S_1^-, S_1^+ \text{ or } P^\prime \text{ or } P^\prime, C^\prime \right) \right), \ldots, \left( \left( S_i^-, S_i^- \right) \text{ or } S_i^- \text{ or } \left( S_i^-, S_i^+ \text{ or } P^\prime \text{ or } P^\prime, C^\prime \right) \right), \ldots, \left( \left( S_{i-1}^-, S_{i-1}^- \right) \text{ or } S_{i-1}^- \text{ or } \left( S_{i-1}^-, S_{i-1}^+ \text{ or } P^\prime \text{ or } P^\prime, C^\prime \right) \right), \ldots, \left( \left( S_{i-1}^-, S_{i-1}^- \right) \text{ or } S_{i-1}^- \text{ or } \left( S_{i-1}^-, S_{i-1}^+ \text{ or } P^\prime \text{ or } P^\prime, C^\prime \right) \right), \right\} \times P\{P^\prime\} \times \left( P\{P^\prime\} P\{C^\prime\} P\{C^\prime\} \right) \]

\[ = \sum_{i=0}^{j-1} r_s \alpha_s + 1 - r_s + r_s(1 - \alpha_s)(1 - r_p) + r_s(1 - \alpha_s)r_p u(1 - r_c) r_s(1 - \alpha_s) \times r_p \left( (1 - u)(1 - \alpha_p) + u r_c(1 - \alpha_c) \right) \]

\[ = \sum_{i=0}^{j-1} \left( 1 - r_s \alpha_s \frac{1 - u(1 - r_c)}{r_p(1 - \alpha_p - u (1 - \alpha_p - r_c(1 - \alpha_c)))} \right) \]
Using the finite series

\[ \sum_{i=0}^{k-1} ax^i = a \frac{1-x^k}{1-x}, \]  

(A-3)

Eq. (A-2) becomes

\[ p_j = \frac{1-(1-r_j(1-a_o)r_p(1-u(1-r_c)))}{1-(1-r_j(1-a_o)r_p(1-u(1-r_c)))} r_s(1-a_s)r_p \left( 1 - \alpha_p - u \left( 1 - \alpha_p - r_c(1-a_c) \right) \right) \]

\[ = \frac{1-(1-r_j(1-a_o)r_p(1-u(1-r_c)))}{r_s(1-a_s)r_p(1-u(1-r_c))} \left( 1 - \alpha_p - u \left( 1 - \alpha_p - r_c(1-a_c) \right) \right) \]

\[ = \frac{\frac{1-(1-r_j(1-a_o)r_p(1-u(1-r_c)))}{1-u(1-r_c)}}{1-u(1-r_c)}. \]  

(A-4)
Appendix B - Overall Specificity of PTSD Screening Process

Following the same logic as for the deviation of overall system sensitivity in Appendix A, a correct negative determination in \( j \) years can occur in the following manners:

\[
((S_1^\prime, S_1^-) \lor \overline{S_1^\prime} \lor (S_1^\prime, S_1^+, P^\prime, \overline{C^\prime})) \lor ((S_1^\prime, S_1^+, P^\prime, P^?, C^\prime)), \ldots, ((S_j^\prime, S_j^-) \lor \overline{S_j^\prime} \lor (S_j^\prime, S_j^+, P^\prime, \overline{P^?}, \overline{C^\prime}))
\]

- PC-PTSDs during \( j \) years either classify patient as negative, give positive result but not confirmed by confirmatory testing, or are skipped.

\[
(((S_1^\prime, S_1^-) \lor \overline{S_1^\prime} \lor (S_1^\prime, S_1^+, P^\prime, \overline{C^\prime})) \lor ((S_1^\prime, S_1^+, P^\prime, P^?, C^\prime)), \ldots, ((S_{i-1}^\prime, S_{i-1}^-) \lor \overline{S_{i-1}^\prime} \lor (S_{i-1}^\prime, S_{i-1}^+, P^\prime, \overline{P^?}, \overline{C^\prime})), S_i^\prime, S_i^+, P^\prime, P^?, C^\prime), i \leq j)
\]

- PC-PTSD in year \( i \), \( 1 \leq i \leq j \), incorrectly identifies patient as positive after \( i-1 \geq 0 \) years either being determined as negative, skipping PC-PTSD, or being determined as positive by PC-PTSD but not following up with confirmatory testing.
- Patient is sent to PCL for review and takes it, and
- PCL makes a definite diagnosis and correctly classifies patient as negative.

\[
(((S_1^\prime, S_1^-) \lor \overline{S_1^\prime} \lor (S_1^\prime, S_1^+, P^\prime, \overline{C^\prime})) \lor ((S_1^\prime, S_1^+, P^\prime, P^?, C^\prime)), \ldots, ((S_{i-1}^\prime, S_{i-1}^-) \lor \overline{S_{i-1}^\prime} \lor (S_{i-1}^\prime, S_{i-1}^+, P^\prime, \overline{P^?}, \overline{C^\prime})), S_i^\prime, S_i^+, P^\prime, P^?, C^\prime, C^?, C^-, i \leq j)
\]

- PC-PTSD in year \( i \), \( 1 \leq i \leq j \), incorrectly identifies patient as positive after \( i-1 \geq 0 \) years being determined as negative, skipping PC-PTSD, or being determined as positive by PC-PTSD but not following up with confirmatory testing.
Appendix B - Overall Specificity of PTSD Screening Process

- Patient is sent to PCL for review and takes it,
- PCL cannot make a definite diagnosis,
- Patient is sent to CAPS for further review and takes it, and
- CAPS correctly classifies patient as negative.

All events above except for the first one can occur in \( j \) manners, as the patient can be identified as positive by PC-PTSD test in either year \( i = 1, 2, \ldots, j \). Now considering these events one at a time, the probability of detecting a truly negative patient in \( j \) years is

\[
q_j = \text{Probability of identifying a truly negative patient as negative in } j \text{ years} \\
= P\left\{ \left( S_1^-, S_1^- \right) \text{ or } S_1^- \text{ or } \left( S_1^-, S_1^+, P^-, P^2, C^- \right) \right\}, \ldots, \\
\left( S_j^-, S_j^- \right) \text{ or } S_j^- \text{ or } \left( S_j^-, S_j^+, P^-, P^2, C^- \right) \right\} \\
+ P\left\{ \left( S_1^-, S_1^- \right) \text{ or } S_1^- \text{ or } \left( S_1^-, S_1^+, P^-, P^2, C^- \right) \right\}, \ldots, \\
\left( S_{i-1}^-, S_{i-1}^- \right) \text{ or } S_{i-1}^- \text{ or } \left( S_{i-1}^-, S_{i-1}^+, P^-, P^2, C^- \right) \right\}, \ldots, \\
\left( S_j^-, S_j^- \right) \text{ or } S_j^- \text{ or } \left( S_j^-, S_j^+, P^-, P^2, C^- \right) \right\}, i \leq j \}
\]

\[
+ P\left\{ \left( S_1^-, S_1^- \right) \text{ or } S_1^- \text{ or } \left( S_1^-, S_1^+, P^-, P^2, C^- \right) \right\}, \ldots, \\
\left( S_{i-1}^-, S_{i-1}^- \right) \text{ or } S_{i-1}^- \text{ or } \left( S_{i-1}^-, S_{i-1}^+, P^-, P^2, C^- \right) \right\}, \ldots, \\
\left( S_j^-, S_j^- \right) \text{ or } S_j^- \text{ or } \left( S_j^-, S_j^+, P^-, P^2, C^- \right) \right\}, i \leq j \}
\]
\[ q_j = \left( 1 - r_s \beta s r_p (1 - u(1 - r_c)) \right)^j + \frac{1 - r_s \beta s r_p (1 - u(1 - r_c))}{1 - r_s \beta s r_p (1 - u(1 - r_c))} \times r_s \beta s r_p \left( 1 - \beta p - u \left( 1 - \beta p - r_c (1 - \beta c) \right) \right), \]
Appendix C - Expected Total Cost of PTSD Screening Process

The expected cost per \( n \) patients after \( j \) years using PC-PTSD, PCL, and CAPS tests, \( EC_j \), is equal to the following sum:

\[
EC_j = \text{Expected cost of all PC-PTSD tests per } n \text{ patients} + \\
\text{Expected cost of all PCL tests per } n \text{ patients} + \\
\text{Expected cost of all CAPS tests per } n \text{ patients} + \\
\text{Expected cost of all false-positives per } n \text{ patients} + \\
\text{Expected cost of all false-negatives per } n \text{ patients}
\]

\[
= (k_s \times \text{Expected number of PC-PTSD tests per } n \text{ patients}) + \\
(k_p \times \text{Expected number of PCL tests per } n \text{ patients}) + \\
(k_c \times \text{Expected number of CAPS tests per } n \text{ patients}) + \\
(k_{FP} \times \text{Expected number of false-positives per } n \text{ patients}) + \\
(k_{FN} \times \text{Expected cumulative number of false-negatives per } n \text{ patients}). 
\]  

(C-1)

For readability, it is convenient first to consider each of the above four expectations separately and then to sum them at the end of this appendix.

**Expected number of PC-PTSD tests per \( n \) patients:**

The expected number of PC-PTSD tests per \( n \) patients, \( E[I_{s,n}] \), is a sum of the expected number of tests per single positive patient, \( E[I_s \mid \text{Positive}] \), and the expected number of tests per single negative patient, \( E[I_s \mid \text{Negative}] \), conditioned as follows:
\[ E[I, n] = n(P(\text{Patient is positive})E[I, Positive] + P(\text{Patient is negative})E[I, Negative]). \]  

(C-2)

The expected number of PC-PTSD tests for a positive patient is

\[
E[I, Positive] = \sum_{i=1}^{j} i P[I_s = i] \\
= 1P[I_s = 1] + 2P[I_s = 2] + \cdots + (j - 1)P[I_s = j - 1] + jP[I_s = j] \\
= 1 \left( \sum_{a=0}^{j-1} (1 - r_s)^a p^+_{con} + j p^+_{s, nc} (1 - r_s)^{j-1} \right) \\
+ 2 \left( \sum_{a=0}^{j-2} \frac{(a+1)!}{a!(j-2)!} (1 - r_s)^a p^+_{s, nc} p^+_{con} + \frac{j}{1!(j-2)!} p^+_{s, nc} (1 - r_s)^{j-2} \right) + \cdots \\
+ (j - 1) \left( \sum_{a=0}^{j-1} \frac{(a+j-2)!}{a!(j-2)!} (1 - r_s)^a p^+_{s, nc} \right)^{j-2} p^+_{con} + \frac{j}{1!(j-2)!} p^+_{s, nc} (1 - r_s)^{j-1} (1 - r_s) \\
+ j \left( p^+_{s, nc} \right)^{j-1} p^+_{con} + \left( p^+_{s, nc} \right)^j \right) \\
= \sum_{i=1}^{j} i \left( p^+_{s, nc} \right)^{j-1} p^+_{con} \sum_{a=0}^{j-i} \frac{(a+j-2)!}{a!(j-2)!} (1 - r_s)^a + \frac{j}{1!(j-2)!} p^+_{s, nc} (1 - r_s)^{j-i}, \quad \text{(C-3)}
\]

and similarly the expected number of PC-PTSD tests for a negative patient is

\[
E[I, Negative] = \sum_{i=1}^{j} i P[I_s = i] \\
= 1P[I_s = 1] + 2P[I_s = 2] + \cdots + (j - 1)P[I_s = j - 1] + jP[I_s = j] \\
= 1 \left( \sum_{a=0}^{j-1} (1 - r_s)^a p^-_{con} + j p^-_{s, nc} (1 - r_s)^{j-1} \right) \\
+ 2 \left( \sum_{a=0}^{j-2} \frac{(a+1)!}{a!(j-2)!} (1 - r_s)^a p^-_{s, nc} p^-_{con} + \frac{j}{1!(j-2)!} p^-_{s, nc} (1 - r_s)^{j-2} \right) + \cdots \\
+ (j - 1) \left( \sum_{a=0}^{j-1} \frac{(a+j-2)!}{a!(j-2)!} (1 - r_s)^a p^-_{s, nc} \right)^{j-2} p^-_{con} + \frac{j}{1!(j-2)!} p^-_{s, nc} (1 - r_s)^{j-1} (1 - r_s) \\
+ j \left( p^-_{s, nc} \right)^{j-1} p^-_{con} + \left( p^-_{s, nc} \right)^j \right) \\
= \sum_{i=1}^{j} i \left( p^-_{s, nc} \right)^{j-1} p^-_{con} \sum_{a=0}^{j-i} \frac{(a+j-2)!}{a!(j-2)!} (1 - r_s)^a + \frac{j}{1!(j-2)!} p^-_{s, nc} (1 - r_s)^{j-i}, \quad \text{(C-4)}
\]

where the probabilities \( p^+_{con}, p^-_{con}, p^+_{s, nc}, \) and \( p^-_{s, nc} \) are given as
Appendix C - Expected Total Cost of PTSD Screening Process

\[ p_{\text{con}}^+ = P\{\text{Result is confirmed by either PCL or CAPS|Patient is positive}\} \]
\[ = r_s(1 - \alpha_s)r_p(1 - u) + ur_c = r_s(1 - \alpha_s)r_p(1 - u(1 - r_c)), \quad (C-5) \]

\[ p_{\text{con}}^- = P\{\text{Result is confirmed by either PCL or CAPS|Patient is negative}\} \]
\[ = r_s\beta sr_p(1 - u(1 - r_c)), \quad (C-6) \]

\[ p_{s, \text{nc}}^+ = P\{\text{PC-PTSD is taken but its result remains unconfirmed|Patient is positive}\} \]
\[ = 1 - p_{\text{con}}^+ - P\{\text{PC-PTSD is not taken}\} = 1 - r_s - (1 - r_s) = r_s - p_{\text{con}}^+, \quad (C-7) \]

\[ p_{s, \text{nc}}^- = P\{\text{PC-PTSD is taken but its result remains unconfirmed|Patient is negative}\} \]
\[ = r_s - p_{\text{con}}^-. \quad (C-8) \]

Thus, by substituting Eq. (C-3) and (C-4) into Eq. (C-2), and then by using the binomial theorem

\[ (a + b)^m = \sum_{y=0}^{m} \binom{m}{y} a^y b^{m-y}, \quad (C-9) \]

and after some algebra, the expected number of PC-PTSD tests per \( n \) patients is calculated as

\[ E[I_{s,n}] = n \left( p \left( \sum_{i=1}^{j} \binom{i}{j} p_{s,\text{nc}}^+ \right) \right)^{i-1} p_{\text{con}}^+ \sum_{a=0}^{j-i} \binom{(a+i-1)!}{a!(i-1)!} (1 - r_s)^a + j p_{s,\text{nc}}^+ (1 - p_{\text{con}}^+)^{j-1} \\
+ (1 - p) \left( \sum_{i=1}^{j} \binom{i}{j} p_{s,\text{nc}}^- \right) \right)^{i-1} p_{\text{con}}^- \sum_{a=0}^{j-i} \binom{(a+i-1)!}{a!(i-1)!} (1 - r_s)^a + j p_{s,\text{nc}}^- (1 - p_{\text{con}}^-)^{j-1} \right). \quad (C-10) \]

Note that the variance in number of PC-PTSD tests per patient can be found in a similar manner, using the relations that \( E[X^2] = \sum_{x=1}^{\infty} x^2 P(X = x) \), \( Var[X] = E[X^2] - E^2[X] \), \( Var[X + Y] = \)
Var[X] + Var[Y] + 2Cov[X, Y], and Cov[X, Y] = 0 if X and Y are independent. After some algebra, by using the same approach above, the following result for \( Var[I_{s,n}] \) can be obtained as

\[
Var[I_{s,n}] = \sum_{i=1}^{j} \left( \left( \frac{p^+}{p_{nc}} \right)^{i-1} p_{con} \sum_{a=0}^{i-j} \left( \frac{(a+j-1)!}{a!(j-1)!} \right) (1 - r_s)^a \right) + \left( \frac{p^-}{p_{nc}} \right)^i (1 - r_s)^j
\]

\[
\times \left( \left( \frac{p^+}{p_{con}} \right)^{i-1} p_{nc} \sum_{a=0}^{i-j} \left( \frac{(a+j-1)!}{a!(j-1)!} \right) (1 - r_s)^a + \left( \frac{p^-}{p_{con}} \right)^{i-1} \right)
\]

\[
+ (1 - p) \left( \sum_{i=1}^{j} i \left( \left( \frac{p^+}{p_{nc}} \right)^{i-1} p_{con} \sum_{a=0}^{i-j} \left( \frac{(a+j-1)!}{a!(j-1)!} \right) (1 - r_s)^a + \left( \frac{p^-}{p_{con}} \right)^{i-1} \right) \right)
\]

\[
\times \left( \left( \frac{p^+}{p_{nc}} \right)^i (1 - r_s)^j + \left( \frac{p^-}{p_{con}} \right)^{i-1} \right) \right)^2.
\]

(C-11)

**Expected number of PCL tests per n patients:**

The expected number of PCL tests per n patients, \( E[I_{s,n}] \), similar to the number of PC-PTSD tests, is a sum of the expected number of tests per single positive patient, \( E[I_p | Positive] \), and the expected number of tests per single negative patient, \( E[I_p | Negative] \), conditioned as follows:

\[
E[I_{p,n}] = n \{ P\{ Patient \ is \ positive \} E[I_p | Positive] + P\{ Patient \ is \ negative \} E[I_p | Negative] \}.
\]

(C-12)

The expected number of PCL tests for a positive patient is

\[
E[I_p | Positive] = \Sigma_{i=1}^{j} iP[I_p = i]
\]
and similarly the expected number of PCL tests for a negative patient is

\[
E[I_p|\text{Negative}] = \sum_{i=1}^{j} i P\{I_p = i\} \\
= 1P\{I_p = 1\} + 2P\{I_p = 2\} + \cdots + (j-1)P\{I_p = j-1\} + jP\{I_p = j\} \\
= 1 \left( \sum_{a=0}^{j-1} (1 - r_s(1 - \alpha_s) r_p)^a p_{\text{con}}^+ + j p_{\text{nc}}^- (1 - r_s(1 - \alpha_s) r_p)^{j-1} \right) \\
+ 2 \left( \sum_{a=0}^{j-2} \frac{(a+1)!}{a!1!} (1 - r_s(1 - \alpha_s) r_p)^a p_{\text{nc}}^- p_{\text{con}}^+ + \left( \frac{j}{2((j-2)!)} \right) \left( p_{\text{nc}}^+ \right)^2 (1 - r_s(1 - \alpha_s) r_p)^{j-2} \right) + \cdots \\
+(j-1) \left( \sum_{a=0}^{j-1} \frac{(a+j-2)!}{a!(j-2)!} (1 - r_s(1 - \alpha_s) r_p)^a p_{\text{nc}}^- p_{\text{con}}^+ + \left( \frac{j}{(j-1)!1!} \right) \left( p_{\text{nc}}^- \right)^j (1 - r_s(1 - \alpha_s) r_p)^{j-1} \right) \\
+j \left( \left( p_{\text{nc}}^- \right)^{j-1} p_{\text{con}}^- + \left( p_{\text{nc}}^- \right)^j \right) \\
= \sum_{i=1}^{j} i \left( (p_{\text{nc}}^-)^{i-1} p_{\text{con}}^- \sum_{a=0}^{i-1} \frac{(a+i-1)!}{a!(i-1)!} (1 - r_s(1 - \alpha_s) r_p)^a + \left( \frac{j}{(j-1)!} \right) \left( p_{\text{nc}}^- \right)^i (1 - r_s(1 - \alpha_s) r_p)^{j-i} \right), \tag{C-14}
\]

where the probabilities \( p_{\text{nc}}^+ \) and \( p_{\text{nc}}^- \) are given as

\[
p_{\text{nc}}^+ = P\{\text{PCL is taken but its result remains unconfirmed} | \text{Patient is positive}\} \\
= r_s (1 - \alpha_s) r_p u(1 - r_c), \tag{C-15}
\]
\[ p_{p_{nc}}^- = P\{\text{PCL is taken but its result remains unconfirmed}\mid \text{Patient is negative}\} \]
\[ = r_s \beta_s r_p u (1 - r_c). \quad (C-16) \]

Thus, by substituting Eq. (C-13) and (C-14) into Eq. (C-12), and then by using the binomial theorem given in Eq. (C-9) and after some algebra, the expected number of PCL tests per \( n \) patients can be calculated as

\[
E[I_{p, n}] = n \left( p \left( \sum_{i=1}^j i \binom{p_{p_{nc}}^+}{p_{p_{nc}}^+} \binom{p_{con}^+}{p_{con}^+} \sum_{a=0}^{(a+i-1)!} \binom{a+i-1}{a} (1 - r_s (1 - \alpha_s) r_p)^a + p_{p_{nc}}^- (1 - p_{con}^+)^{j-1} \right) \right. 
\]
\[
+ (1 - p) \left( \sum_{i=1}^j i \binom{p_{p_{nc}}^-}{p_{p_{nc}}^-} \binom{p_{con}^-}{p_{con}^-} \sum_{a=0}^{(a+i-1)!} \binom{a+i-1}{a} (1 - r_s \beta_s r_p)^a + p_{p_{nc}}^- (1 - p_{con}^-)^{j-1} \right) \right). 
\]

(C-17)

The variance in number of PCL tests per patient can be found in a similar manner, as explained above. After some algebra, the following result for \( Var[I_{p, n}] \) can be obtained as

\[
Var[I_{p, n}] = n \left( E[I_p^2] - E[I_p]^2 \right) 
\]
\[
= n \left( p \sum_{i=1}^j i^2 \left( p_{p_{nc}}^+ \binom{p_{con}^+}{p_{con}^+} \sum_{a=0}^{(a+i-1)!} \binom{a+i-1}{a} (1 - r_s (1 - \alpha_s) r_p)^a \right. 
\]
\[
+ \left. (1 - p) \sum_{i=1}^j i^2 \left( p_{p_{nc}}^- \binom{p_{con}^-}{p_{con}^-} \sum_{a=0}^{(a+i-1)!} \binom{a+i-1}{a} (1 - r_s \beta_s r_p)^a \right) \right) 
\]
\[
+ \left( p \sum_{i=1}^j i \binom{p_{p_{nc}}^+}{p_{p_{nc}}^+} \binom{p_{con}^+}{p_{con}^+} \sum_{a=0}^{(a+i-1)!} \binom{a+i-1}{a} (1 - r_s (1 - \alpha_s) r_p)^a \right. 
\]
\[
\left. + p_{p_{nc}}^- (1 - p_{con}^+)^{j-1} \right) \right) 
\]
\[
- \left( p \sum_{i=1}^j i \binom{p_{p_{nc}}^-}{p_{p_{nc}}^-} \binom{p_{con}^-}{p_{con}^-} \sum_{a=0}^{(a+i-1)!} \binom{a+i-1}{a} (1 - r_s \beta_s r_p)^a \right. 
\]
\[
\left. + j p_{p_{nc}}^+ (1 - p_{con}^+)^{j-1} \right) \right) 
\]
\[
+ \left( p \sum_{i=1}^j i \binom{p_{p_{nc}}^-}{p_{p_{nc}}^-} \binom{p_{con}^-}{p_{con}^-} \sum_{a=0}^{(a+i-1)!} \binom{a+i-1}{a} (1 - r_s \beta_s r_p)^a \right. 
\]
\[
\left. + j p_{p_{nc}}^- (1 - p_{con}^-)^{j-1} \right) \right) \right)^2. 
\]

(C-18)
While Eq. (C-10), (C-11), (C-17), and (C-18), are rather lengthy expressions, special-purpose software greatly facilitates their evaluation. Also note that the computer simulation results presented elsewhere in this dissertation in Table 2-3 independently arrive at similar results, in this sense confirming above derivations.

**Expected number of CAPS tests per n patients:**

The expected number of CAPS tests per \( n \) patients, \( E[I_{c,n}] \), simply is a function of the probability that any single patient is checked using the CAPS test. As every patient is checked once using the CAPS test or not checked using this test, this constitutes a Bernoulli trial and the distribution of the number of CAPS tests per \( n \) patients therefore is binomial with binomial expectation

\[
E[I_{c,n}] = n \times P\{\text{Patient is checked using CAPS test}\} \\
= n \times P\{\text{Result is confirmed by CAPS after } 0 \leq i < j - 1 \text{ years not having a confirmed result}\} \\
= n \left[ \left( r_s(1 - \alpha_s)r_purc + (1 - p^+_r)\right)\left( r_s(1 - \alpha_s)r_purc + \cdots + (1 - p^+_r)^{j-1} r_s(1 - \alpha_s)r_purc \right) \right] \\
+ (1 - p) \left[ r_s\sum_{i=0}^{j-1} (1 - p^+_r)^i + \left( r_s\sum_{i=0}^{j-1} (1 - p^-_r)^i \right) \right] \\
= n \left[ \left( r_s(1 - \alpha_s)r_purc \sum_{i=0}^{j-1} (1 - p^+_r)^i \right) + (1 - p) \left( r_s\sum_{i=0}^{j-1} (1 - p^-_r)^i \right) \right] \\
= n \left[ \left( r_s(1 - \alpha_s)r_purc \sum_{i=0}^{j-1} \left( \frac{1-(1-p^+_r)^i}{1-(1-p^+_r)} \right) \right) + (1 - p) \left( r_s\sum_{i=0}^{j-1} \left( \frac{1-(1-p^-_r)^i}{1-(1-p^-_r)} \right) \right) \right] \\
= n \left[ \left( r_s(1 - \alpha_s)r_purc \sum_{i=0}^{j-1} \left( \frac{1-(1-p^+_r)^i}{r_s(1-\alpha_s)r_p(1-u(1-r_c))} \right) \right) + (1 - p) \left( r_s\sum_{i=0}^{j-1} \left( \frac{1-(1-p^-_r)^i}{r_s\sum_{i=0}^{j-1} \left( \frac{1-(1-p^-_r)^i}{r_s\sum_{i=0}^{j-1} \left( \frac{1-(1-p^-_r)^i}{1-u(1-r_c)} \right)} \right) \right) \right) \right] \\
= n \left[ \left( r_s(1 - \alpha_s)r_purc \sum_{i=0}^{j-1} \left( \frac{1-(1-p^+_r)^i}{1-u(1-r_c)} \right) \right) + (1 - p) \left( r_s\sum_{i=0}^{j-1} \left( \frac{1-(1-p^-_r)^i}{1-u(1-r_c)} \right) \right) \right], \quad (C-19)
\]
and with binomial variance $\text{Var}[I_{c,n}] = n\{\cdot\}(1 - \{\cdot\})$, where $\{\cdot\}$ denotes the term in square brackets in Eq. (C-19) and is equal to the probability that any patient is checked using the CAPS test:

$$\text{Var}[I_{c,n}] = n \left( \frac{ur_c}{1-u(1-r_c)} \right) \left( p \left( 1 - (1 - p^+_{con})^j \right) + (1 - p) \left( 1 - (1 - p^-_{con})^j \right) \right) \times \left( 1 - \left( \frac{ur_c}{1-u(1-r_c)} \right) \left( p \left( 1 - (1 - p^+_{con})^j \right) + (1 - p) \left( 1 - (1 - p^-_{con})^j \right) \right) \right).$$  \hspace{1cm} (C-20)

**Expected number of false-positives per $n$ patients:**

The expected number of false-positives per $n$ patients, $E[FP]$, follows directly from the probability of detecting a truly negative patient, $q_j$, found in Eq. (B-2) of Appendix B:

$$E[FP] = n \times P\{\text{Patient is negative}\} \times P\{\text{Positive determination|Patient is negative}\}$$
$$= n(1 - p)(1 - P\{\text{Negative determination|Patient is negative}\})$$
$$= n(1 - p) \left( 1 - q_j \right)$$
$$= n(1 - p) \left( 1 - \left( \frac{(1-r_c)p_{p}(1-u(1-r_c))^j(1-u)p_{p}+ur_c\beta_p}{1-u(1-r_c)} \right) \right)$$
$$= n(1 - p) \left( \frac{(1-(1-r_c)p_{p}(1-u(1-r_c))^j(1-u)p_{p}+ur_c\beta_p)}{1-u(1-r_c)} \right).$$  \hspace{1cm} (C-21)

**Expected number of false-negatives per $n$ patients:**

Similarly, the expected number of false-negatives per $n$ patients, $E[FN]$, follows directly from the probability of detecting a truly positive patient, $p_j$, found in Eq. (A-4) of Appendix A:
Appendix C - Expected Total Cost of PTSD Screening Process

\[ E[FN] = n \times P\{\text{Patient is positive}\} \times P\{\text{Negative determination}\mid \text{Patient is positive}\} \]
\[ = np(1 - P\{\text{Positive determination}\mid \text{Patient is positive}\}) \]
\[ = np\left(1 - \frac{1 - (1-r_c(1-a_c)r_p(1-u(1-r_c)))}{1-u(1-r_c)}\right)^j \left(1-a_p-u\left(1-a_p-r_c(1-a_c)\right)\right) \]
\[ = np\left(\sum_{j=1}^{\infty} \frac{(1-r_c(1-a_c)r_p(1-u(1-r_c)))}{1-u(1-r_c)}^j \left(1-a_p-u\left(1-a_p-r_c(1-a_c)\right)\right) + (1-u)a_p+ur_c.a_c\right). \tag{C-22} \]

**Expected overall cost per n patients:**

Now summing the products of each of the above five terms with their corresponding costs as shown in Eq. (C-1) yields the expected total cost per n patients:

\[ EC_j = k_i E[I_{s,n}] + k_p E[I_{p,n}] + k_c E[I_{c,n}] + k_{fp} E[FP] + k_{FN} j E[FN] \]
\[ = k_i n \left(\sum_{i=1}^{\infty} \frac{(p_{s,nc}^+)^{i-1} p_{con}^+ \sum_{a=0}^{(a+i-1)} \frac{(a+i-1)!}{(a)!i!} (1-r_s)^a + j p_{s,nc}^- (1-p_{con}^-)^{i-1}}{1-u(1-r_c)} \right) \]
\[ + (1-p) \left(\sum_{i=1}^{\infty} \frac{(p_{s,nc}^-)^{i-1} p_{con}^- \sum_{a=0}^{(a+i-1)} \frac{(a+i-1)!}{(a)!i!} (1-r_s)^a + j p_{s,nc}^+ (1-p_{con}^+)^{i-1}}{1-u(1-r_c)} \right) \]
\[ + k_p n \left(\sum_{i=1}^{\infty} \frac{(p_{p,nc}^+)^{i-1} p_{con}^+ \sum_{a=0}^{(a+i-1)} \frac{(a+i-1)!}{(a)!i!} (1-r_s(1-a_c)r_p)^a + j p_{p,nc}^- (1-p_{con}^-)^{i-1}}{1-u(1-r_c)} \right) \]
\[ + (1-p) \left(\sum_{i=1}^{\infty} \frac{(p_{p,nc}^-)^{i-1} p_{con}^- \sum_{a=0}^{(a+i-1)} \frac{(a+i-1)!}{(a)!i!} (1-r_s(1-a_c)r_p)^a + j p_{p,nc}^+ (1-p_{con}^+)^{i-1}}{1-u(1-r_c)} \right) \]
\[ + k_n \left(\frac{ur_c}{1-u(1-r_c)} \left(p - (1-p_{con}^-)^j\right) \right) + (1-p) \left(1 - (1-p_{con}^-)^j\right) \]
\[ + k_{fp} n(1-p) \left(\frac{1-\beta_u(1-u(1-r_c))}{1-u(1-r_c)}\right) \]
\[ + k_{FN} j np \left(\frac{1-r_c(1-a_c)r_p(1-u(1-r_c))}{1-u(1-r_c)}\right)\left(1-a_p-u\left(1-a_p-r_c(1-a_c)\right)\right) + (1-u)a_p+ur_c.a_c\). \tag{C-23} \]
Appendix D - Overall Accuracy and Cost of TBI Screening Process

**Notation and Assumptions:**

\[
\begin{align*}
\alpha_i &= \text{probability that any screening test } i \text{ will determine a truly positive patient as negative}, \\
\alpha_c &= \text{probability that the confirmation test will determine a truly positive patient as negative}, \\
\beta_i &= \text{probability that any screening test } i \text{ will determine a truly negative patient as positive}, \\
\beta_c &= \text{probability that the confirmation test will determine a truly negative patient as positive}, \\
r_i &= \text{probability that a truly positive patient has at least one teammate diagnosed with TBI}, \\
r_2 &= \text{probability that a truly negative patient has at least one teammate diagnosed with TBI}, \\
k_s &= \text{average cost of a screening test}, \\
k_c &= \text{average cost of the confirmation test}, \\
k_{FP} &= \text{average cost of a false-positive final result}, \\
k_{FN} &= \text{average cost of a false-negative final result}, \text{ and} \\
j &= \text{number of screening tests}.
\end{align*}
\]

In the proposed screening process (given in Figure 2-18), after a negative determination by any screening test \( i < j \), the patient is scheduled for the next screening test \( i+1 \). If any screening test \( i \leq j \) determines the patient as positive, he/she is sent to the confirmation test. If the patient, on the other hand, is determined as negative by all \( j \) screening tests, it is checked whether or not any of his teammates has diagnosed with TBI recently. If so, the patient is sent to the confirmation test. Otherwise, no further action is taken and the patient is determined as negative. Note that this decision stage is incorporated into the mathematical model using the rescreening rates of \( r_1 \) and
$r_2$ (explained above), and that $r_1$ is assumed to be greater than $r_2$. Other model assumptions include

(1) Test sensitivities and specificities are assumed to remain constant over time.

(2) All tests are assumed to be independent of each other.

(3) The actual number of screening tests per any given patient, the expected number of screening tests per positive patient and per negative patient are denoted, respectively, as $I_s$, $E[I_s \mid \text{Positive}]$ and $E[I_s \mid \text{Negative}]$. Similarly, the actual number of confirmation tests per any given patient, the expected number of confirmation tests per positive patient and per negative patient are denoted, respectively, as $I_c$, $E[I_c \mid \text{Positive}]$ and $E[I_c \mid \text{Negative}]$.

(4) Screening and false diagnosis costs are assumed to remain constant over time.

Overall System Sensitivity:

A correct positive determination using $j \geq 1$ screening tests can occur in the following manners:

$$(S_1^-, S_2^-, \ldots, S_j^-, R, C^+): \quad \text{– All } j \text{ screening tests incorrectly classify patient as negative,}$$

$$\quad \text{– Patient is sent to confirmation test, and}$$

$$\quad \text{– Confirmation test correctly revises diagnosis to positive.}$$

$$((S_1^-, S_2^-, \ldots, S_{i-1}^-, S_i^+, C^+), i \leq j): \quad \text{– Some screening test } i, 1 \leq i \leq j, \text{ correctly identifies patient as positive after } i-1 \geq 0 \text{ incorrect negative determinations,}$$

$$\quad \text{– Patient is sent to confirmation test for review (always), and}$$

$$\quad \text{– Confirmation test confirms diagnosis as positive.}$$
The probability of detecting a truly positive patient using $j$ screening tests, $p_j$, then is the sum of the probabilities of these events:

\[
p_j = \text{Probability of detecting a truly positive patient using } j \text{ screening tests} \\
= P\{S_1^-, S_2^-, \ldots, S_j^-, R, C^+\} + P\{S_1^-, S_2^-, \ldots, S_{j-1}^-, S_j^+, C^+\} i \leq j \\
= a_1a_2 \ldots a_j r_j (1 - a_c) + (1 - a_j)(1 - a_c) + a_j(1 - a_2)(1 - a_c) + \ldots \\
\ldots + a_ja_2 \ldots a_{j-1}(1 - a_j)(1 - a_c) \\
= a_1a_2 \ldots a_j r_j (1 - a_c) + (1 - a_ja_2 \ldots a_{j-1}a_j)(1 - a_c) \\
= \left(1 - \prod_{i=1}^{j} a_i (1 - r_j)\right)(1 - a_c). \quad (D-1)
\]

Note that if all screening tests are assumed to have the same sensitivity, $(1 - a_c)$, the above expression in Eq. (D-1) simplifies to closed form as

\[
p_j = \left(1 - a_j(1 - r_j)\right)(1 - a_c). \quad (D-2)
\]

**Overall System Specificity:**

Following the same logic above, a correct negative determination using $j \geq 1$ screening tests can occur in the following manners:

\((S_1^-, S_2^-, \ldots, S_j^-, \bar{R})\):  
- All $j$ screening tests correctly classify patient as negative, and  
- Patient is not sent to confirmation test.

\((S_1^-, S_2^-, \ldots, S_j^-, R, C^-)\):  
- All $j$ screening tests correctly classify patient as negative,  
- Patient is sent to confirmation test, and  
- Confirmation test confirms patient as negative.
Appendix D - Overall Accuracy and Cost of TBI Screening Process

((S_1^-, S_2^-, ..., S_{i-1}^-, S_i^+, C^-), i \leq j): – Some screening test \(i\), \(1 \leq i \leq j\), incorrectly identifies patient as positive after \(i-1 \geq 0\) correct negative determinations,
– Patient is sent to confirmation test for review (always), and
– Confirmation test correctly revises diagnosis to negative.

Considering these events one at a time, the probability of detecting a truly negative patient using \(j\) screening tests is

\[
q_j = \text{Probability of identifying a truly negative patient as negative using } j \text{ screening tests}
\]

\begin{align*}
&= P\{S_1^-, S_2^-, ..., S_j^-, \bar{R}\} + P\{S_1^-, S_2^-, ..., S_j^-, R, C^-\} + P\{(S_1^-, S_2^-, ..., S_{i-1}^-, S_i^+, C^-) i \leq j\} \\
&= \prod_{i=1}^{j}(1-\beta_i)(1-r_2) + \prod_{i=1}^{j}(1-\beta_i) r_2(1-\beta_c) + (1 - \prod_{i=1}^{j}(1-\beta_i))(1-\beta_c) \\
&= 1 - \beta_c \left(1 - \prod_{i=1}^{j}(1-\beta_i)(1-r_2)\right). \quad (D-3)
\end{align*}

As above, if all \(j\) screening tests are assumed to have similar specificities, \((1-\beta_s)\), the above expression in Eq. (D-3) simplifies to closed form as

\[
q_j = 1 - \beta_c \left(1 - (1-\beta_i)(1-r_2)\right). \quad (D-4)
\]

**Expected Total Cost:**

The expected cost per \(n\) patients using \(j\) screening tests, \(EC_j\), is equal to the following sum:

\[
EC_j = \text{Expected cost of all screening tests per } n \text{ patients} + \\
\text{Expected cost of all confirmation tests per } n \text{ patients} + \\
\text{Expected cost of all false-positives per } n \text{ patients} + \\
\text{Expected cost of all false-negatives per } n \text{ patients}
\]
where each of the above four expectations are given separately below.

The expected number of screening tests per \( n \) patients, \( E[I_{s,n}] \), is the sum of the expected number of tests per single positive patient, \( E[I_s | \text{Positive}] \), and the expected number of tests per single negative patient, \( E[I_s | \text{Negative}] \), conditioned as

\[
E[I_{s,n}] = n(P(\text{Patient is positive})E[I_s | \text{Positive}] + P(\text{Patient is negative})E[I_s | \text{Negative}]), \quad (D-6)
\]

where the expected number of screening tests for a positive patient is

\[
E[I_s | \text{Positive}] = \sum_{i=1}^{j} iP(I_s = i) = IP(I_s = 1) + 2P(I_s = 2) + \cdots + (j-1)P(I_s = j-1) + jP(I_s = j) = 1(1 - \alpha_1) + 2\alpha_1(1 - \alpha_2) + \cdots + (j-1)\alpha_1\alpha_2 \cdots \alpha_{j-2}(1 - \alpha_{j-1}) + j(\alpha_1\alpha_2 \cdots \alpha_{j-1}(1 - \alpha_j) + \alpha_1\alpha_2 \cdots \alpha_j)
\]

\[
= (1 - \alpha_1) + \sum_{j=2}^{j-1} \left( i \prod_{k=1}^{j-1} \alpha_k (1 - \alpha_i) \right) + j \prod_{k=1}^{j-1} \alpha_k, \quad (D-7)
\]

and similarly the expected number of screening tests for a negative patient is

\[
E[I_s | \text{Negative}] = \sum_{i=1}^{j} iP(I_s = i) = IP(I_s = 1) + 2P(I_s = 2) + \cdots + (j-1)P(I_s = j-1) + jP(I_s = j) = +1\beta_1 + 2(1 - \beta_1)\beta_2 + \cdots + (j-1)(1 - \beta_j)(1 - \beta_2) \cdots (1 - \beta_{j-2})\beta_{j-1} + j\left( (1 - \beta_1)(1 - \beta_2) \cdots (1 - \beta_{j-1}) \beta_j + (1 - \beta_1)(1 - \beta_2) \cdots (1 - \beta_j) \right)
\]
\begin{align*}
\beta_1 + \sum_{i=2}^{j} (i \prod_{k=1}^{i-1} (1 - \beta_k) \beta_i) + j \prod_{k=1}^{j-1} (1 - \beta_k). \quad (D-8)
\end{align*}

The expected number of confirmation tests per \( n \) patients, \( E[I_{c,n}] \), given below, simply is a function of the probability that any single patient is checked using the confirmation test. As every patient is screened once using the confirmation test or not screened using this test, this constitutes a Bernoulli trial and the distribution of the number of confirmation tests per \( n \) patients therefore is binomial with binomial expectation.

\begin{align*}
E[I_{c,n}] &= n \times P \{ \text{Patient is screened by confirmation test} \} \\
&= n (P \{ \text{Some screening test determines patient as positive} \} \\
&\quad + P \{ \text{All screening tests determine patient as negative} \} P \{ \text{Patient is sent to confirmation test} \}) \\
&= n \left( p (1 - \prod_{i=1}^{j} \alpha_i + \prod_{i=1}^{j} \alpha_i r_i) + (1 - p) (1 - \prod_{i=1}^{j} (1 - \beta_i) + \prod_{i=1}^{j} (1 - \beta_i) r_2) \right) \\
&= n \left( 1 - \left( p \prod_{i=1}^{j} \alpha_i (1 - r_i) + (1 - p) \prod_{i=1}^{j} (1 - \beta_i) (1 - r_2) \right) \right). \quad (D-9)
\end{align*}

The expected number of false-positives per \( n \) patients, \( E[FP] \), follows directly from the probability of detecting a truly negative patient, \( q_j \), found in Eq. (D-3):

\begin{align*}
E[FP] &= n \times P \{ \text{Patient is negative} \} \times P \{ \text{Positive determination|Patient is negative} \} \\
&= n (1 - p) \left( 1 - q_j \right) \\
&= n (1 - p) \beta_c \left( 1 - \prod_{i=1}^{j} (1 - \beta_i) (1 - r_2) \right). \quad (D-10)
\end{align*}

Similarly, the expected number of false-negatives per \( n \) patients, \( E[FN] \), follows directly from the probability of detecting a truly positive patient, \( p_j \), found in Eq. (D-1):

\begin{align*}
E[FN] &= n \times P \{ \text{Patient is positive} \} \times P \{ \text{Negative determination|Patient is positive} \}
\end{align*}
\[ = np \left( 1 - p_j \right) \]
\[ = np \left( \alpha_c + (1 - \alpha_c) \prod_{i=1}^{l} \alpha_i \left( 1 - r_i \right) \right). \quad \text{(D-11)} \]

Now summing the products of each of the above terms with their corresponding costs as shown in Eq. (D-5) yields the expected total cost per \( n \) patients:

\[ EC_j = k_s E[I_s, n] + k_c E[I_c, n] + k_{FP} E[FP] + k_{FN} E[FN] \]
\[ = k_s n \left( p \left( (1 - \alpha_j) + \sum_{i=1}^{j-1} \left( i \prod_{k=1}^{i} \alpha_k \left( 1 - \alpha_i \right) \right) + j \Pi_{k=1}^{i} \alpha_k \right) \right) \]
\[ + \left( 1 - p \right) \left( \beta_j + \sum_{i=1}^{j-1} \left( i \Pi_{k=1}^{i} \left( 1 - \beta_i \right) \right) + j \Pi_{k=1}^{i} \left( 1 - \beta_k \right) \right) \]
\[ + k_c n \left( 1 - \left( p \prod_{i=1}^{j} \alpha_i \left( 1 - r_j \right) + \left( 1 - p \right) \prod_{i=1}^{l} \left( 1 - \beta_i \right) \left( 1 - r_2 \right) \right) \right) \]
\[ + k_{FP} n \left( 1 - p \right) \beta_c \left( 1 - \prod_{i=1}^{j} \left( 1 - \beta_i \right) \left( 1 - r_2 \right) \right) \]
\[ + k_{FN} n p \left( \alpha_c + (1 - \alpha_c) \prod_{i=1}^{l} \alpha_i \left( 1 - r_i \right) \right). \quad \text{(D-12)} \]

If all screening tests can be assumed to have identical sensitivities \((1 - \alpha_i) = (1 - \alpha_s)\) and identical specificities \((1 - \beta_i) = (1 - \beta_s)\) for all \( i \), then the four expectations above in Eq. (D-12) simplify to closed form and the total expected cost expression becomes

\[ EC_j = k_s n \left( p \frac{1 - \alpha_j}{1 - \alpha_s} + \left( 1 - p \right) \frac{1 - (1 - \beta_j)}{\beta_s} \right) \]
\[ + k_c n \left( 1 - \left( p \alpha_j \left( 1 - r_j \right) + \left( 1 - p \right) \left( 1 - \beta_s \right) \left( 1 - r_2 \right) \right) \right) \]
\[ + k_{FP} n \left( 1 - p \right) \beta_c \left( 1 - \left( 1 - \beta_s \right) \left( 1 - r_2 \right) \right) \]
\[ + k_{FN} n p \left( \alpha_c + (1 - \alpha_c) \alpha_j \left( 1 - r_j \right) \right). \quad \text{(D-13)} \]