3-21-2013

Traumatic Brain Injury Recovery Care: Demand Forecasting, Staffing, and Treatment Planning

Mitchell R. Kieffer

Recommended Citation

Kieffer, Mitchell R., "Traumatic Brain Injury Recovery Care: Demand Forecasting, Staffing, and Treatment Planning" (2013). *Theses and Dissertations*. 967.
https://scholar.afit.edu/etd/967

TRAUMATIC BRAIN INJURY RECOVERY CARE: DEMAND FORECASTING, STAFFING, AND TREATMENT PLANNING

THESIS

Mitchell R. Kieffer, Captain, USAF

AFIT-ENS-13-M-08

DEPARTMENT OF THE AIR FORCE
AIR UNIVERSITY

AIR FORCE INSTITUTE OF TECHNOLOGY

Wright-Patterson Air Force Base, Ohio

AFIT-ENS-13-M-08

TRAUMATIC BRAIN INJURY RECOVERY CARE: DEMAND FORECASTING, STAFFING, AND TREATMENT PLANNING

THESIS

Presented to the Faculty

Department of Operational Sciences

Graduate School of Engineering and Management

Air Force Institute of Technology

Air University

Air Education and Training Command

In Partial Fulfillment of the Requirements for the

Degree of Master of Science in Operations Research

Mitchell R. Kieffer, MS

Captain, USAF

March 2013

AFIT-ENS-13-M-08

TRAUMATIC BRAIN INJURY RECOVERY CARE: DEMAND FORECASTING, STAFFING, AND TREATMENT PLANNING

Mitchell R. Kieffer, MS
Captain, USAF

Approved:

Jeffery K. Cochran, Dr., Professor (Advisory)	Date
Jeffery D. Weir, Dr., Associate Professor (Reader)	Date

AFIT-ENS-13-M-08

Abstract

Improvised Explosive Device attacks have skyrocketed since the start of the War on Terror. Many troops wounded by these tactics receive long-lasting unseen wounds including Traumatic Brain Injury (TBI). TBI sufferers are treated along with other casualties. This has created an increasing, and varying, demand for ongoing post operative recovery care for troops returning from deployments. Diagnosis and treatment for TBI wounded troops is costly. This thesis is motivated by the recognition that budgets are constrained yet quality of care should not be compromised. Additive Holtz–Winters smoothing is used to forecast overall patient care demand, a regression based on queueing theory determines care consultant staffing levels, and reliability theory quantifies the idea of reducing cost by reducing parallel treatment planning. The scope is the Warfighter Rehabilitation Centers and AF Warrior and Survivor Care with data from SMEs, the Brookings Institution, and icasualties.org. This thesis provides a step-by-step methodology and analyzes the actual situation that leadership encountered from 2010-2012.

I would like to dedicate this work to all of those who've been down-range and could benefit from improved recovery care, and those who may need it in the future.

Acknowledgments

I'd like to thank my advisor, Dr. Jeffery Cochran, for his wonderful enthusiasm and expert guidance throughout the research process. I'd like to thank specific subject matter experts, Timothy Townes (USAF/A1SZ, Pentagon), Michelle Lindsey (Traumatic Brain Injury Center, Brooke Army Medical Center, TX), and Rachel Morgan (Recovery Care Coordinator, Wright Patterson Medical Center, OH), for taking the time to be interviewed in order to understand current practices. I'd also like to thank all Recovery Care Coordinators and Air Force Wounded Warrior consultants that have helped me through my own recovery after being wounded in 2011 while on deployment with the US Army Corps of Engineers, Baghdad, Iraq.

Mitchell R. Kieffer, CPT, USAF

Table of Contents

Page

List of Figures

Page

List of Tables

List of Equations

TRAUMATIC BRAIN INJURY RECOVERY CARE: DEMAND FORECASTING, STAFFING, AND TREATMENT PLANNING

I. Introduction

General Issue

Since the beginning of the War on Terror, the number of Improvised Explosive Device (IED) attacks has skyrocketed. This ambush tactic has become a staple for the recent fighting in countries like Iraq and Afghanistan. IED attacks often wound troops from flying debris and/or blast overpressure. Many troops that are wounded by these tactics receive long-lasting problems due to the high number of Post Traumatic Stress Disorder (PTSD) and Traumatic Brain Injury (TBI) cases. “Among veterans of the wars in Iraq and Afghanistan, traumatic brain injury is the leading cause of disability, labeled, ‘The signature wound of the war on terror’.”(Harch, et al., 2012).

According to a newspaper article from USA Today, Pentagon officials have stated that up to 360,000 Iraq and Afghanistan veterans may have sustained TBI, with 45,000 to 90,000 needing continuing specialty care (Zoroya, 2009). This has created a great demand for ongoing post operative recovery care for troops returning from deployments. I have seen this first hand, as I was wounded myself, receiving a Purple Heart for a Traumatic Brian Injury while deployed in Iraq. During my own TBI recovery process, I spoke with a number of employees and subject matter experts (SME) in this healthcare realm. Many of the SMEs had indicated that there were a number of shortfalls in the current system. The cost involved with diagnosis and treatment for wounded troops is extensive, and there are many issues, inconsistencies, waiting patients that could benefit if able to enter the system, and budget constraints to combat in this healthcare environment.

Problem Statement

As the fiscal climate of the US Government becomes more volatile and budget cuts are expected, there is a need to determine "what if" scenarios if or when certain resources are minimized. Healthcare is known to have escalating costs. "For instance, the United States spends more than $2 trillion or 16% of it GDP on Healthcare…" (Viana & Rais, 2010). For the military, the Department of Veterans Affairs (VA) has the brunt of these costs and is constantly hard-pressed for extra budget and resources to meet the needs of recovering veterans from war wounds, illness, and/or injuries. Furthermore, with many clients waiting to enter the recovery system, the VA needs to work with the WRCs and the active duty component of recovery care, the US Air Force Warrior and Survivor Care (USAF/A1SZ) personnel section to determine the best ways to minimize service time and cost in order to provide service to more wounded, ill and injured (WII), while continuing to provide a high quality and valuable service.

There are different phases of recovery that can be affected. When troops are wounded in the Middle East, they are usually sent to a Deployed Warrior Medical Center (DWMC) such as Landstuhl Regional Medical Center in Germany. Once received via Medical Evacuation (MEDEVAC) from down range, troops will receive initial care toward recovery, and possibly surgical procedures for those in need. When the troops are ready to redeploy back to the Continental United States (CONUS), they arrive via MEDEVAC to a Warrior Rehabilitation Center (WRC), such as Brooke Army Medical Center in San Antonio, TX, or Walter Reed Army Medical Center located in Washington D.C.

At the WRC, clients will be treated for their respective wounds/illness/injuries. These wounded, ill, or injured (WII) troops will be diagnosed and receive treatment at specialty clinics. Once they have completed the necessary diagnosis and initial treatment, they are sent back to their home base for their latter stages of recovery when deemed not needing specialty care.

Research Focus

This thesis will concentrate primarily on the two phases of recovery that happen in CONUS; at WRC and Home Base Medical Centers (HBMC). When troops are receiving recovery care, they are followed and helped by the USAF/A1SZ. This personnel section provides the services of Air Force Wounded Warrior (AFW2) Program and the Recovery Care Program (RCP). Both will be detailed later in the thesis, but in essence, the AFW2 will handle most issues arising with military pay, benefits, entitlements, transition to civilian at medical board, etc. The RCP handles the medical treatment aspects of recovery. Specialty clinics at the WRC are separate from the USAF/A1SZ, and are classified in different categories. Specifically, the Traumatic Brain Injury clinic's classification is based upon the "Army's Office of Reintegration and Rehabilitation (PR&R) published operation order" in 2007 (Lindsay, 2012). It indicates that the TBI Center at BAMC is Category (Cat) 1, where it can, "Provide inpatient and outpatient care for the full spectrum of TBI Severity (mild, moderate, severe)." Other facilities have different ratings. "Category 2: Provide inpatient and outpatient care for mild and moderate TBI (typically these clients don't require surgical intervention or ICU care). Category 3: Provide outpatient medical and rehab care to mild TBI and refer out for additional services as needed (these facilities may or may not have neurology consultation available). Category 4: Provide outpatient care for mild TBI (no rehab i.e.: Occupational Therapy, Speech Language Pathology) and refer for additional services." The main goal is for troops to benefit from a Cat 1 facility and transition back to their home station to complete their recovery at most likely a Cat 4 medical facility.

The specialty clinics at the WRC (and healthcare in general) have high costs stemming from treatment, staffing (people), techniques, and technology. Currently, funding for the WRC and the TBI Center in particular stems from the operations order from 2007 stated previously where it dictates, "…what level of TBI services would be offered at each facility…[and] provide[s] comprehensive outpatient rehab and inpatient consult service. Our staffing was dictated by this memo. We are budgeted under special funds from Office of the Surgeon General (OTSG)" (Lindsay, 2012).

On the other hand, the USAF/A1SZ programs (RCP, AFW2) are federally funded. "Staffing [is]…based on case load and acuity levels" (Townes, 2012). Rachel Morgan, a currently operating Recovery Care Coordinator (RCC), has mentioned volatility in recent staffing due to budget constraints. She stated that workforce was let go only to be rehired to maintain guidance set by the Department of Defense (DoD) and Air Force Instruction (AFI) (Morgan, 2012). Mr. Townes at USAF/A1SZ Headquarters confirmed stating, "We have recently almost doubled our RCC numbers because of the case load and acuity levels for our RCCs. In order to stay in compliance with the OSD mandated case load, we determined the need to add to the RCC total."

These are different funding methods, but the golden thread found with both the WRC and USAF/A1SZ programs is that staffing is based upon current and past levels of clients (military, retired, dependents). They do not forecast future demand. They also have not researched how to streamline the process to allow more clients to enter the system, or research what-if scenarios of mission impact if budget constraints come into play. There are many military, retired, and dependents waiting to enter to the recovery system, while the fiscal climate tightens, it will be imperative for these organizations to have mathematical reasoning behind their staffing to prove to budget officials their worth and impact to their important mission if not subsidized. It is also imperative to have backup plans for environments where funding cannot keep up with demand. Leadership will have to make tough decisions, and mathematical analysis on operations can be vital to finding the correct path.

Hence, the main focus of this thesis is directed toward the post-operative recovery care staffing and assessing reliability of treatment decisions (Figure 1).

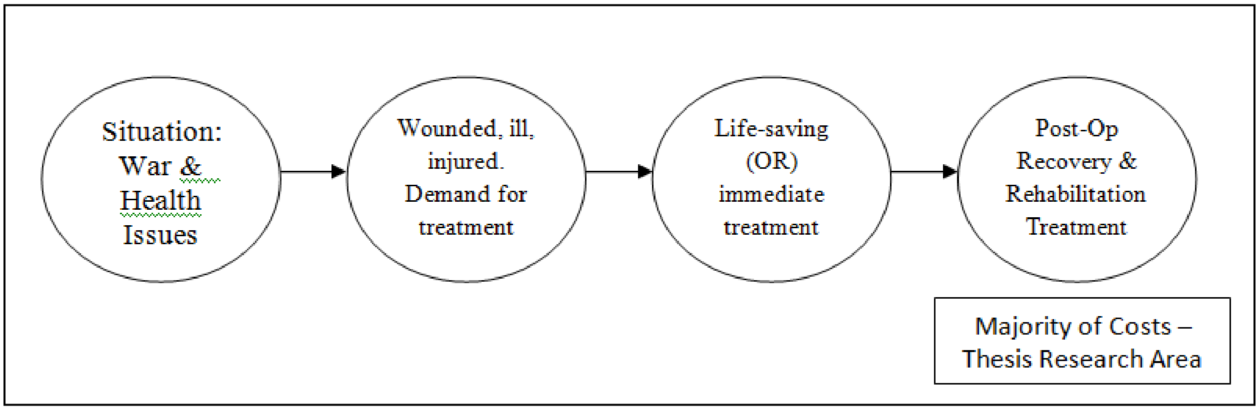


Figure 1: Recovery Care Demand Diagram

The goal of the Air Force Warrior and Survivor Care Program is to provide, "Support to families of the fallen and Airmen wounded, ill and injured and their families" (Demmons, 2011). As seen below (Figure 2), the RCP and AFW2 are interrelated. Staffing scenarios will be related to the RCC and AFW2 consultants as they are the primary drivers of the RCP and AFW2 programs.

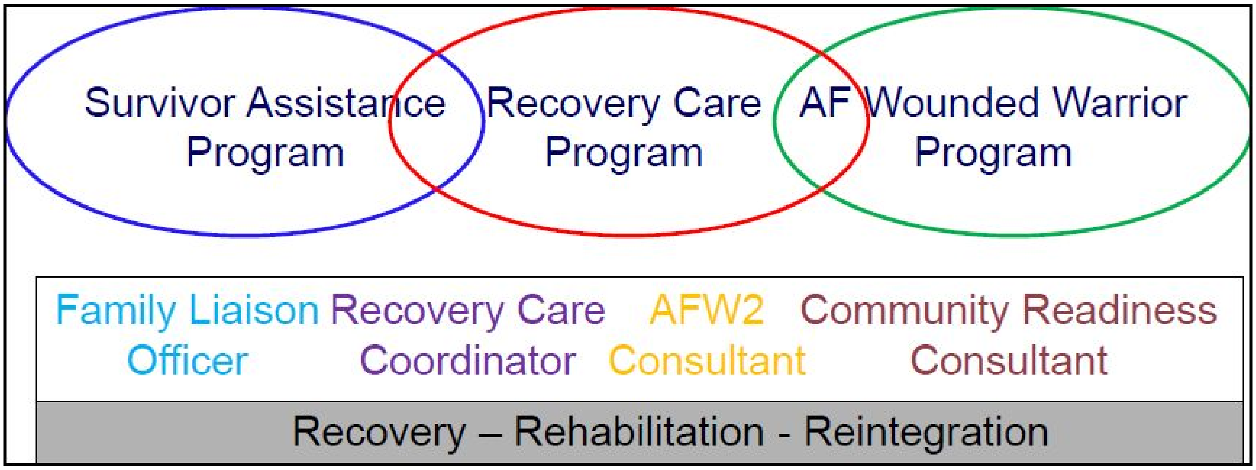


Figure 2: Recovery Care System Overview

Research Objectives

The research questions hope to contribute to the ability to advise, inform and/or recommend courses of action for decision makers in warrior and survivor care. Investigative questions attempting to be answered using Operations Research (OR) techniques are:

1) What constraints are present in current staffing techniques and how can they be remedied?
2) What variability is there in staffing and the number of WII needing treatment? Is there variability based upon seasons and/or milestones in war?
3) How many RCC/AFW2 consultants are needed given workload thresholds using different forecasting scenarios?
4) What techniques help to improve staffing efficiency while maintaining quality?
5) What techniques can be used to maintain reliability in treatment system effectiveness when constraints limit the treatment options at WRC's?

Assumptions/Limitations

There are a few major assumptions and limitations to the analysis in this thesis. The warrior and survivor care system has drastically evolved in the recent years, and there are many wounded warriors from previous campaigns that are not actively in the system. The arrival rates into the system are strictly based on expected future wounded USAF troops, and every instance is assumed automatically entered into the system. There is minimal knowledge about the average time in system as well, so any analysis in consultant staffing based upon queueing methods should utilize in-house statistics known about the average client service and departure rates. Accuracy of all data and confidence of estimates provides extensive limitations to the analysis used in this thesis. Wide confidence intervals should be used in decision making when results follow an in-line path of estimates. Specific assumptions to each analytical investigation will be presented in its respective section.

II. Literature Review

The balance of the literature review is divided into four main sections. The literature review is followed by the methodology, technical analysis, and conclusions with recommendations for future work.

Defining Air Force Warrior and Survivor Care System (USAF/A1SZ)

Airmen are enrolled in the USAF/A1SZ programs when the personnel section has been notified of wounded, or relocated ill/injured Airmen. Throughout the process of MEDEVAC and treatment at a DWMC, WRC, or HBMC, the USAF/A1SZ provides personal representatives to the WII and their families for a variety of assistance and support services.

The RCC leads the WII through their respective medical Comprehensive Recovery Plan (CRP) as part of the RCP. They are the focal point for non-clinical case management, helping to establish career goals and timelines for accomplishment, and document non-clinical support to the WII. They are the direct link to Wounded Airmen and their families. According to Lt Col Beth Demmons (2011), there were 32 RCCs and one program manager staffed at 30 major installations as of 23 February 2011 (Figure 3). As of 31 January 2011, the "RCC program is currently providing service to and/or tracking 805 total WII with 574 [71.3%] falling into the injured and ill category. [There are] 228 of total are combat wounded [clients that are] also tracked by AFW2."

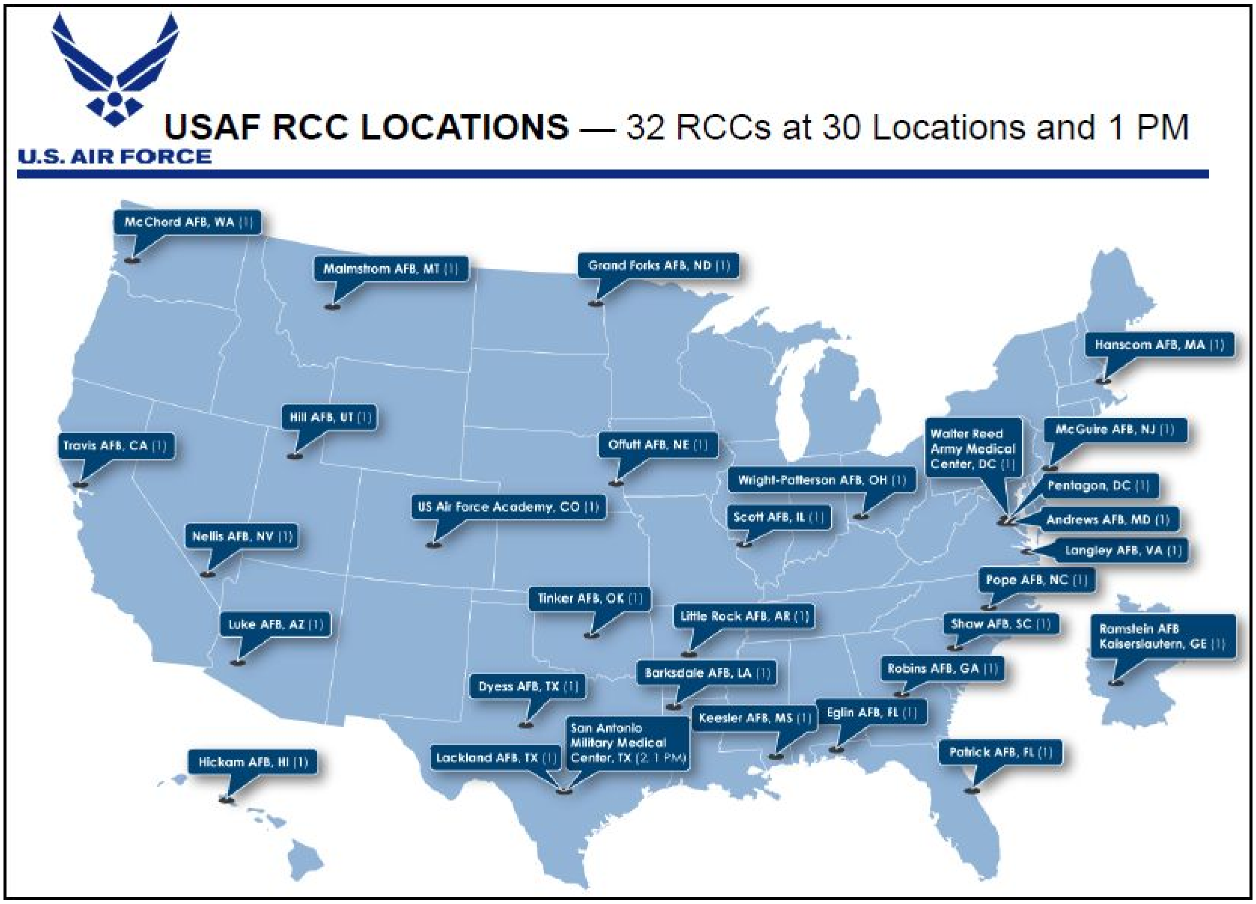


Figure 3: RCC Locations

The RCP services all WII personnel, while the AFW2 program mainly services the combat-related injured/ill troops. The AFW2 Care Management system has 17 AFW2 consultant care managers with varied backgrounds along with four administrative personnel. They are considered non-medical care-management and are located at the Air Force Personnel Center at Randolph Air Force Base, TX. They make monthly contact with AFW2 members, family members or caretakers, help with reintegration and transitioning through medical boards, and provide counseling for financial planning, employment, benefits, compensation, relocation, and so on (AFI 34-1101: Air Force Warrior and Survivor Care, 2012). Both programs aim to ensure first rate care and to do as much as possible prior to separation.

As of 31 January 2011, there were 975 AFW2 clients in the program: 706 from regular Air Force, 143 from Air National Guard, and 126 from the Air Force Reserve. There has been non-linear growth of service to combat wounded veterans from year to year (Figure 4). The prevalence of "unseen" injuries has been the primary driver, with a 66.5% increase in TBI/PTSD cases from 2009 to 2010 (Demmons, 2011). However, this exponential increase in client load is not sustainable and fitting an exponential curve would not be an accurate way to forecast future clientele.

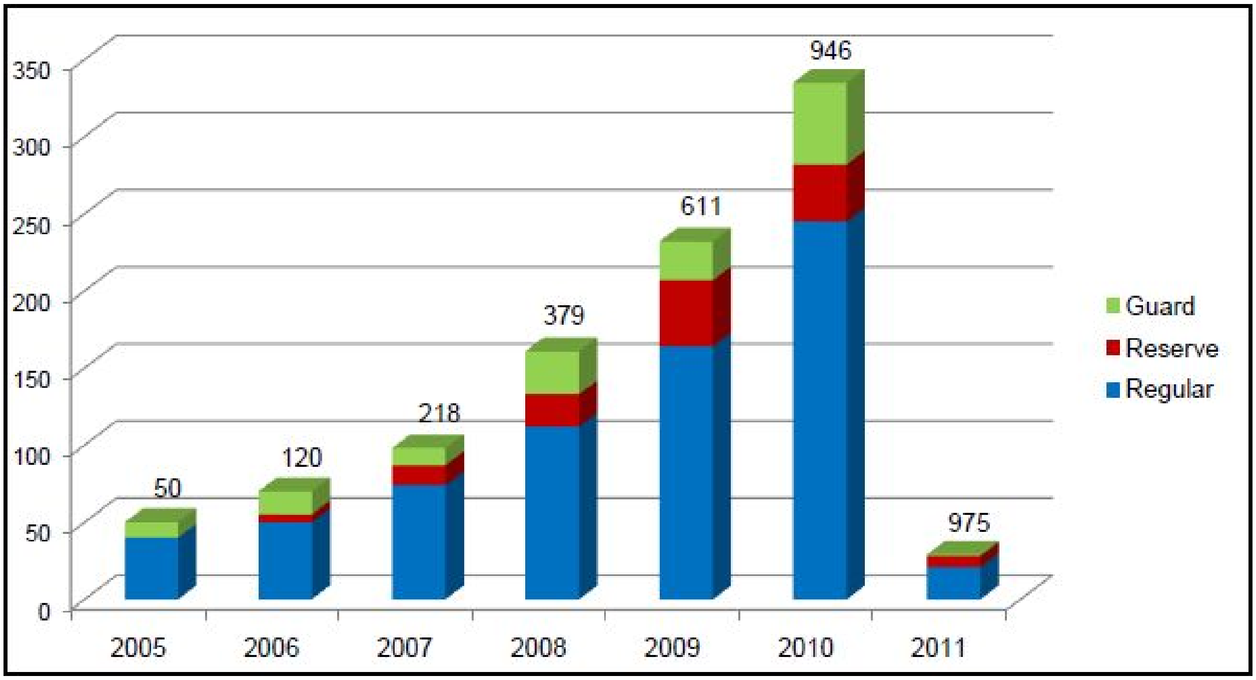
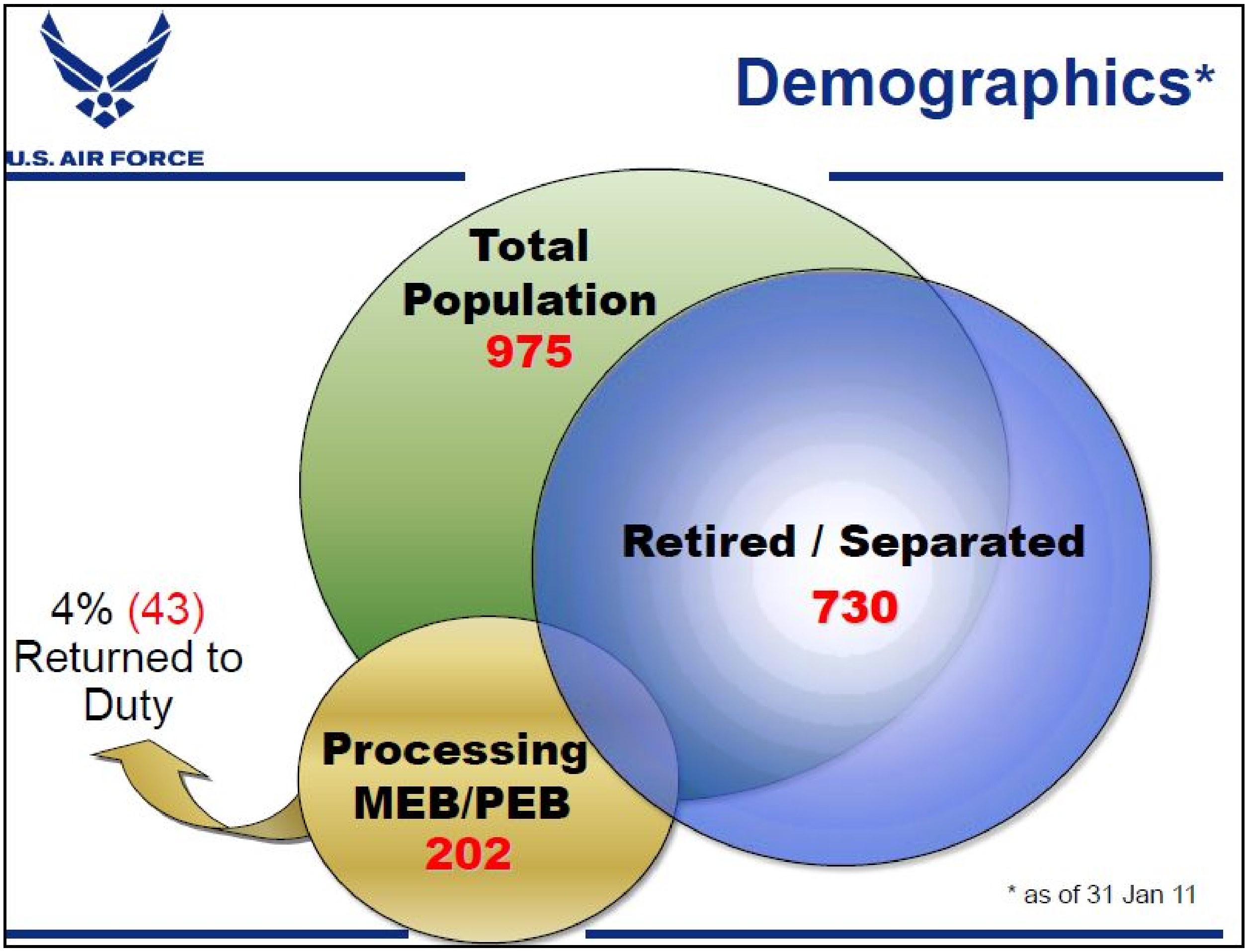


Figure 4: Times Series of AFW2 Clientele

Of the total population at the beginning of 2011, 61% had primary injuries/illness of a psychological nature such as PTSD while the remaining 39% had physical injuries from combat which includes TBI. There were 335 newly identified AFW2 members in 2010, which is an average of approximately 28 new Air Force wounded warriors per month. Injuries take a heavy toll on the members, where only 4% of the total AFW2 members have been able to recover to a full duty status (Figure 5).



U.S. AIR FORCE

Demographics*

Total Population 975

Retired / Separated 730

Processing MEB/PEB 202

4% (43) Returned to Duty

* as of 31 Jan 11

Figure 5: AFW2 Demographics

The programs have many challenges to include classification. PTSD makes up approximately 61% of wounded cases and is identified via disability evaluation instead of MEDEVAC. TBI and PTSD are very closely related and it is difficult to distinguish which wound causes different symptoms. Dr. Lisa Brenner (2011) indicated a number of challenges for treatment and classification of TBI, along with the plausible synergy with PTSD. Many have attempted to standardize and streamline an identification system. Wojcik et al. (2010) examined a new “Barell Matrix” classification system that includes data on the presence or absence of intracranial injury, level of consciousness, and allows cross analysis with military TBI and civilian population TBI. The VA has taken great interest in the prevalence of TBI and PTSD as well. Many studies have researched the accuracy of current diagnostic tests, psychological or pharmacological therapies used for treatment, and therapies for treatment of PTSD and TBI (Carlson, et al., 2009). With this new demand for post operative recovery care, a new database has been designed to help solve some of the current issues. The Internal Mission for Prognosis and Clinical Trial (IMPACT) database of TBI includes complete datasets from most clinical trials of the past 20 years (Marmarou, et al., 2007). The effort is funded by the US National Institute of Health. This database will come online in March 2013. This database could be very useful for verification/validation of work done in this thesis or be used for recommended future research.

Forecasting Demand Relevant Research

The type of attacks against US troops has changed from conventional war to ambush tactics where heavy Improvised Explosive Device use allows enemy personnel to avoid meeting US/Allied forces head on. Even though the main mechanism in combat causing injury has changed, the wounding pattern has not. “The wounding patterns currently seen in Iraq and Afghanistan resemble the patterns from previous conflicts” (Owens, Kragh, Wenke, Macaitis, Wade, & Holcomb, 2008). Hence, displaying the wounding sample of the Afghanistan/Iraq wars can provide insight to forecasting future conflicts. Historically, staffing for RCC has been simply a moving average of the past three years of clientele in the system to provide a potential case load for the future (Townes, 2012). This method of forecasting does not take into account trend, seasonality, or discrete changes (Bowerman, O'Connell, & Koehler, 2005). It merely portrays that past rate of client addition into the program, instead of actually trying to predict the number of new clients that will be needing care.

Most of the current medical forecasting literature focuses on short term hospital planning. Broyles, Cochran, and Montgomery (2011) have attempted to forecast demand by using a Markov decision process for dynamic hospital inpatient staffing. “Appropriate inpatient staffing levels minimize hospital cost and increase client safety.” The service process and transient inventory are estimated in this method. “[Relevant] literature’s primary goal is to minimize waiting and service cost, and to maximize service completion rewards. The majority of the literature assumes constant arrival and service processes and requires explicit knowledge of the service rates and non-stationary profile.” However, the topic of this thesis does not assume a constant arrival or service process and there is not explicitly known service rates.

Other ways the medical community has tried to predict hospital inventory levels is by using Erlang distributions when clients have more than one phase of treatment (Millard & McClean, 1994). Empirical statistical models on length of hospital stay has been attempted to forecast demand (Littig & Isken, 2007). However, this method requires extensive hospital data records which are usually proprietary and confidential due to patient privacy.

These short-term hospital forecasting methods become hard to adapt to long-term care, “Throughput complexities exist in both the arrival and service process of the client flow” (Broyles, Cochran, & Montgomery, 2011). The actual service rates for each client will have much more variability for the WII in comparison to what inpatient treatment centers see in hospitals due to the nature of injury. Also, there is an obvious correlation with client inventory due to wartime operations, which must be taken into account in the forecast, while hospitals that serve the general population do not have congruent seasonal trends. There have been attempts to forecast ground force casualties by the Naval Health Research Center (1997). They use the FORECAS ground casualty forecasting system model to simulate wounded-in-action, killed-in-action, and disease/non-battle injury incidence for U.S. forces. Historical casualty data from four combat operations was used for their baseline. “These battle intensity-specific baselines are then adjusted to reflect changes in weapons parity, troop motivation, environmental factors, and battlefield awareness between the past operations and the hypothesized future scenarios.” This forecasting method helps medical personnel plan resource needs/allocation for specific adversary combat operations. This technique is better utilized at the engagement or operational level while the purpose of this thesis’ forecasting is at the strategic level where a cheaper, faster, and simpler evaluation technique is needed. Instead of the simulation technique, a smoothing method is better utilized because of its simplicity and familiarity to those that will do the demand (casualties) forecasting.

The current simple moving average smoothing method utilized by USAF/A1SZ is demonstrated by Equation 1 (Bowerman, O'Connell, & Koehler, 2005).

$$F_{t+1} = \frac{1}{k} \sum_{i=t-k+1}^{t} Y_i$$

$$where\ F = forecast, t = time, k = order, Y = time\ series\ data.$$

Equation 1

The simple moving average is easy to compute and understand. A single exponential smoothing method is another option. It can create a weighted moving average that might improve the fit to the data while maintaining simplicity. This method provides very similar results, formulated by Equation 2.

$$F_{t+1} = F_t + \alpha(Y_t - F_t)$$

$$where\ \alpha = exponential\ weight.$$

Equation 2

With regard to war, there are obvious trends and seasonality in casualty data, hence, a naïve method or moving average smoothing method would not be wise to implement as they do not take into account these features in the time series. Therefore, the Additive Holtz-Winters smoothing method provides an adequate tool. It assumes a constant (additive) seasonal variation with a slowly changing linear trend. This method will be compared to the simple exponential smoothing method as used for a current baseline in the analysis section.

Qualitative historical/political information can help shed light on different trends, seasonality, and randomness of the quantitative series to help build the Holtz-Winters method. The Brookings Indexes provide important background information regarding past, current, and future anti-terrorism operations to help form the analysis of trends and seasonality indicators. Once the final model is built and provides expected casualties based upon different scenarios, AFW2 and RCC consultant staffing can be analyzed based on these forecasts.

Medical Staffing Relevant Research

There has been limited research to model hospital staffing. Current RCC staffing is done by looking strictly at the previous inventory levels and a service capacity set by the Air Force and DoD Instructions, AFI 34-1101 and DoDI 1300.24 (DoDI 1300.24: Recovery Coordination Program, 2009; AFI 34-1101: Air Force Warrior and Survivor Care, 2012). "We don't really do much as far as long term forecasting. We hope long term numbers go down, but all we can do is use the average of the last three years to give us an idea of potential case load increases or decreases in the future…The current case load for our RCCs tops out, by DoDI, at

40:1"(Townes, 2012). As stated earlier, this method can pose problems in forecasting. The RCC numbers were recently doubled to stay in DoDI compliance. RCC Rachel Morgan (2012) stated that prior to the hiring, RCCs were actually let go due to budget constraints. Following the realization of increase demand, the personnel that were laid off were rehired along with new hires. This portrays an inconsistent staffing process. Queueing theory can alleviate some of this inconsistency.

Queueing methods utilized in literature control the arrival and service rates using a threshold service capacity. Most of the literature has been done for inpatient hospital bed allocation which is not the nature of this thesis, however there are still lessons to be learned and can be applied to medical staffing in general. In all cases, queueing theory is used as a function of demand. There are *N-type*, *D-type,* and *T-type* policies. The leaders in *N-type* research switch servers on and off relying upon the number in the system and a constant threshold of service capacity per server (Takagi, 1991; Yadin & Naor, 1967). *D-type* methods utilize an amount of workload in the system for each server and a capacity of workload threshold (Artalejo, 2002; Lillo & Martin, 2000). "*T-type* policies also make server on and off decisions based on current entity inventory levels but can only be made at discrete times(Federgruen & So, 1991; Wang, 1996; Okamura, Dohi, & Osaki, 2000; Tadj, 2003)"(Broyles, Cochran, & Montgomery, 2011). With the different types of threshold policies, queueing theory has been used, "…to prove that a hospital units should maintain at least 10% emptiness in order to run efficiently due to stochastic arrival and service processes" (Gorunescu, McClean, & Millard, 2002). This 10% rate is a bottom line figure, emphasizing that a server utilization rate over 90% will rapidly increase the number of entities in the system and will reap havoc on serviceability. "Right now the average [RCC case load] is 35:1" (Townes, 2012). Hence, the USAF/A1SZ section has been utilizing a similar approach by staffing based upon a N-Type policy with 12.5% emptiness to total allowed service capacity given by the DoDI.

Cochran and Roche (2008) demonstrate that, "Queuing theory uses analytical expressions rather than statistical experiments to evaluate hospital client flows and determine bed allocation which has the appeal of ease of use." Simulation has also provided a flexible tool for determining bed allocation. However, "…the widespread application of solution techniques in hospitals requires a simpler and faster foundation than simulation." To remedy the complexity that can be found in simulation alone, joint applications in simulation and queueing have been attempted. Cochran and Bharti (2006) use a joint application in order to balance utilizations in bed allocation. However these joint methods also incur long development times in creating the simulation models.

Queueing theory also demonstrates that increasing demand does not necessarily imply a linear increase in staffing. Whitt (2007) uses a square-root-staffing formula to determine entity delay probability (the time a customer must wait before starting service). The author emphasizes this model use for a system with time-varying demand.

RCC and AFW2 staffing should also be analyzed with a time-varying demand in forecasting client load due to changes in the troop casualties time series. In this case, it is an $M/M/c$ multi-server stochastic queueing system. The number of entities in an $M/M/c$ system depends upon the number of servers (c), with an assumed Poisson process rate of arrival (λ), and service rate (μ) following an exponential distribution. The assumed time between successive arrivals is an independent exponential random variable with mean $1/\lambda$. If a server is busy, then the customer waits in the queue. When the server is no longer busy, they begin service. The successive service times are assumed independent exponential random variables with mean 1/μ.

When managing a queueing system, the appropriate number of servers is paramount. A large number will allow shorter waiting times and improve customer satisfaction, yet it can be very costly. Having a low number of servers, though cheap to operate, increases customer wait time and lowers customer satisfaction. Decision makers for these systems pay close attention to measures of effectiveness (MOE). The MOEs can describe a system and predict how efficient and the level of quality it will have. A few typical MOEs are the server utilization rate (ρ), mean number in the system (L), expected number in the queue (L_q), the mean waiting time in the queue (W_q), and the mean waiting time in the entire system (W). In a single server capacity, the derived formulas for these measures of effectiveness are given by Equation 3-Equation 7.

$$\rho = \lambda/\mu$$

Equation 3

$$L = \lambda W = \frac{\rho}{1-\rho} = \frac{\lambda}{\mu - \lambda}$$

Equation 4

$$L_q = \lambda W_q = \frac{\rho^2}{1-\rho} = \frac{\lambda^2}{\mu(\mu - \lambda)}$$

Equation 5

$$W = W_q + \frac{1}{\lambda} = \frac{L}{\lambda} = \frac{1}{\mu - \lambda}$$

Equation 6

$$W_q = \frac{L_q}{\lambda} = \frac{\rho}{\mu - \lambda}$$

Equation 7

There is a need to find an optimal balance of server utilization and adequate flow through the system. Using multiple servers, “When ρ>1 (λ>cμ), the mean number of arrivals into the system exceeds the maximum mean service rate of the system, and we would expect, as time goes on, the queue to get bigger and bigger, unless, at some point, customers were not allowed to join…Therefore, if one knows the mean arrival rate and mean service rate, the minimum number of parallel servers required to guarantee a steady-state [long run] solution can be calculated immediately by finding the smallest c such that λ/cμ < 1”(Gross, Shortle, Thompson, & Harris, 2008). In fact, even when λ/cμ = 1, the queue will lead to excessive congestion. In steady state, the number of servers must be greater than the load ω to ensure the queue is stable. Thus, the number of servers is given by Equation 8.

$$c = \omega + \Delta,$$
$$where\ \Delta > 0\ (number\ of\ additional\ servers\ used).$$

Equation 8

Now suppose that the offered load quadruples to 4ω. By analyzing the MOEs, they can help shed light in determining the appropriate number of servers for an efficient system. To maintain constant traffic intensity, it would seem reasonable to also quadruple the number of servers. This is referred to as the Quality Domain. Another technique is to have a constant padding of servers so that the actual value of c is something like ‘ω + 3’ (Efficiency Domain). One could also attempt to maintain a constant measure of congestion by controlling the probability of a customer having a delay before service (Quality and Efficiency Domain).

“In the quality domain, there is an emphasis on providing a high level of service at the expense of server cost…Yet, as the offered load increases, the probability of delay in the queue decreases” (Gross, Shortle, Thompson, & Harris, 2008). This might seem like a good thing.

However, a decreasing probability of delay causes an abundance of unutilized servers, which causes high cost and waste. This approach often over-estimates the number of additional servers needed.

The efficiency domain technique minimizes cost at the expense of service quality. “As the offered load increases, congestion in the queue increases.” Therefore, the efficiency domain approach often leads to high congestion and customer dissatisfaction by underestimating the number of additional servers required.

The quality and efficiency domain approach provides a balance between the two previous domains. The object is to maintain a fixed quality of service. The formula to find *c* in this approach is very complex compared to the other domains in which they are very simple and intuitive. However, there is an approximation formula that minimizes the complexity. “The basic idea is that the number of excess servers should increase with the *square root* of the offered load.” From Equation 8, the appropriate number of servers is given by Equation 9.

$$c \approx \omega + \beta\sqrt{\omega} \ or \ \Delta = \beta\sqrt{\omega}$$

Equation 9

The constant, β, represents the quality of service, having a relationship to the probability of nonzero delay in the queue (α). β can be approximated by taking the (1 – α) quantile of the standard normal distribution. “In summary, the sequence of queues has approximately the same quality of service provided that the excess number of servers grows with the square root of the offered load.” Therefore, this ‘Square Root Law’ is a new approach to determining the proper number of consultants for the RCC and AFW2 programs.

It is difficult to estimate an arrival rate for these programs because it is time dependent, and the rate of service is unstable due to the new prominence of long-lasting unseen injuries. Therefore, it is increasingly complex to calculate respective queueing theory measures of effectiveness for this topic. Deriving these measures (Equation 3-Equation 7) require steady-state probabilities from an infinite series that is computationally intensive. The square root law becomes very simple and applicable in this case to provide insight into matching consultant staffing to demand because; “The square-root law can be used in a relative sense without specifying the precise values of these constants” (Gross, Shortle, Thompson, & Harris, 2008). This method for staffing can help minimize cost, waste, and variability in warrior and survivor care while maintaining optimal performance.

Reliability Theory Relevant Research

Another way to improve cost and efficiency in WII medical care is to minimize the number of parallel treatments, and keeping the most reliable. The type of wounds received by US troops in battle has shifted in the past decade due to the heavily exploited ambush tactics. “[There have been] a greater proportion of head and neck wounds, and a lower proportion of thoracic wounds [as seen in previous wars]. An explosive mechanism accounted for 78% of injuries [in the war on terror], which is the highest proportion seen in any large-scale conflict”(Owens, Kragh, Wenke, Macaitis, Wade, & Holcomb, 2008). “Primary blast injuries to the brain include concussion as well as barotrama caused by acute gas embolism…Serious late effects of traumatic brain injuries, such as central nervous system residua, have brought attention to the need for rehabilitation of the central nervous system after blast exposure”(DePalma, Burris, Champion, & Hodgson, 2005). This new ‘staple’ injury requires long-term treatment planning. "Most people who have had a significant brain injury will require rehabilitation. They may need to relearn basic skills, such as walking or talking" (Mayo Clinic Staff, 2012).

"Several excellent sets of review articles on TBI have recently been published…[there are] review articles developed by the Institute of Medicine in response to the VA's request for an examination of the strength of the evidence for potential long-term health outcomes related to TBI" (Taber & Hurley, 2010). The US Army Medical Research and Materiel Command has submitted research regarding brain vulnerability to repeated blast overpressure and polytrauma, as the risks associated with multiple concussions displays a direct need for post operative recovery care (Long, 2010).

Numerous studies have been attempted to adopt certain system of treatment for such blast injuries. An article from the New England Journal of Medicine focused on examination. Once initial evaluation has concluded there are no life-threatening injuries present, treating injuries as indicated is the path taken (DePalma, Burris, Champion, & Hodgson, 2005). There has been no proven magical remedy to rehabilitate from a TBI, therefore rehabilitation takes many forms. "The goal [of rehabilitation] is to improve their abilities to perform daily activities. Therapy usually begins in the hospital and continues at an inpatient rehabilitation unit, a residential treatment facility or through outpatient services. The type and duration of rehabilitation varies by individual, depending on the severity of the brain injury and what part of the brain was injured" (Mayo Clinic Staff, 2012). "The TBI Services encompasses Physical Medicine and Rehabilitation, Neuro-psychology, Psychology, Clinical Pharmacology, Occupational Therapy, Physical Therapy, Speech-Language Pathology, Recreational Therapy, Case Management, and Defense & Veterans Brain Injury Center" (Traumatic Brain Injury, 2011). This philosophy is adopted by the Traumatic Brain Injury Service clinic at the WRC, and will provide the example in the analysis section of how to maintain treatment reliability when constraints may hinder resources available to the clinic.

The TBI clinic implements their treatment plan using specialists in speech and language pathology, occupational therapy, psychologists, physical therapy, and rehab nurses. Treatment is based upon symptoms of the client. “We treat headaches like headaches, etc.” (Lindsay, 2012). Clients will be seen by each specialist and they will determine the type and level of treatment required. Obviously, the more treatment providers necessary for recovery will be a direct cause of higher cost. “[There are] resources including staffing, service capacity (beds, treatment areas, etc.), medical devices, and medication…Poor resource matching leads to increased hospital operations cost (Runy, 2005), decreased client safety, and business loss” (Broyles, Cochran, & Montgomery, 2011).

In budget constrained environments, staffing and treatments might be an item on the chopping block. Rehabilitation paths are constantly being researched, with new studies showing evidence proving and/or disproving success reliability of treatments. Therefore, it is not uncommon to add, minimize, or exchange different treatments within a WRC.

It should be noted that losing specified treatment(s) will most likely not return a similar loss in probability of success. Reliability theory can be used in conjunction with known treatments’ probability of success to help gauge influence on performance and success rates when budget decisions are to be made. Painton (1995) uses reliability algorithms to optimize system performance subject to cost constraints and identifies possible improvements with needed effort. Yet, his system involves uncertain component failure rates instead of using treatment effectiveness rates, which is the concentration of this thesis.

Currently, when a wounded troop arrives at the WRC, they receive a set protocol of treatments at a specialty clinic respective to their injury. Every client will be seen by a specialist for each treatment type to assess the level of care needed for the symptoms that persist. The typical treatment plan follows a parallel system (Figure 6).

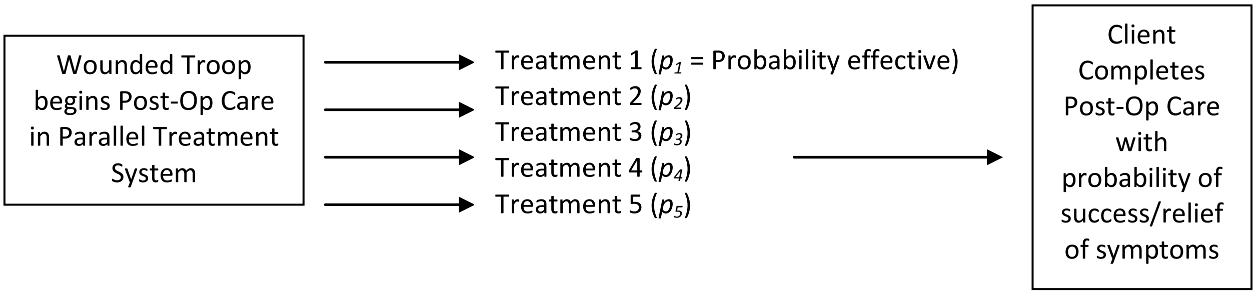


Figure 6: WRC Parallel System of Treatments

The clients go through an '*n-out-of-n*' system of treatments. The system has multiple components (treatments) that work simultaneously, each which functions with some known probability of effectiveness (recovery/relief of symptoms) with the assumption that they are independent from the other system components. The system functions if at least one of the system components is implemented; however, will not have the same measure of effectiveness than if all components were functioning. Redundancy and serviceability is a prominent issue, as defense budgets continually seem to be decreasing for all DoD related services. Every treatment has costs in clinic space, equipment, salary paid to technicians, and the client's time. The reliability function of a parallel system with *n* components assumed to be independent, is given by Equation 10 (Ross, 2007).

$$r(p) = P\{\emptyset(X) = 1\} \text{ where } \mathrm{X} = (\mathrm{X}_1, \dots, \mathrm{X}_\mathrm{n}) = \text{ state of system components}$$

$$= 1 - \prod_{i=1}^{n} (1 - p_i)$$

Equation 10

Budget Reduction/Sequestration (Fiscal Year 2013)

According to a memorandum outlining the current budget Sequestration (2013), the Wounded Warrior programs will be fully protected to the $12B reductions. However, it is imperative to have a plan of action for the possibility of mandated budget limitations, as well as finding ways to allow more clients to enter the system when there is a budget freeze, which is very realistic due to the Sequestration.

A parallel protocol of treatments will most likely have the highest efficacy to improve patients' symptoms when all treatments are given. If it is possible to keep all treatments in the protocol, it is recommended to do so for the highest probability of effectiveness. A possible solution to combat constraints may call for changing treatment protocol such as reducing the number of treatments.

The protocol chosen could be based upon medical effectiveness of treatments, patient's choice of what they would be willing to participate, availability of services, or perhaps some other clinical reasons. Reducing the treatment protocol would reduce the total time it takes to service each client. This would increase the overall μ, as seen in the previous section on queueing theory. Recalling Equation 3(server utilization), a faster rate of service (larger μ) decreases the server utilization rate. Therefore, the cost in server's and patient's time is reduced. That would decrease the number of required servers for the system. Additionally, it would allow more of the clients that are waiting at the moment to enter the system and receive care.

As stated in the introduction, approximately 45,000-90,000 veterans are in need of ongoing recovery care for TBI alone. There are thousands of other casualties with different wounds from previous military conflicts that would gain benefit from treatment as well. The solution provided could allow many of those veterans to gain access to the system and be helped.

Decision makers can use reliability theory for a *'k-out-of-n'* system, where *'k'* is the number of treatments still used in the recovery protocol. If the probability of effectiveness for each treatment was equal ($p_i = p$; $i = 1,\ldots, n$), then the reliability function for the system is given by Equation 11.

$$
\begin{aligned}
r(p) &= P\{\emptyset(X) = 1\} \text{ where } \mathrm{X} = (\mathrm{X}_1, \ldots, \mathrm{X}_\mathrm{n}) \; = \; \text{state of system components} \\
&= P\left\{\sum_{i=1}^{n} X_i \geq k\right\} \\
&= \sum_{i=k}^{n} \binom{n}{i} p^i (1-p)^{n-i}
\end{aligned}
$$

Equation 11

However, without all treatments having the same rate of effectiveness, it leads to a different way of calculating reliability. For simplicity, assume there is a 'Two-out-of-Three System' where at least two of three treatments are utilized. The reliability function is now given by Equation 12. It can be modified for almost any treatment protocol scenario. The theory used in the simple two-out-of-three system is easily applied to any *k-out-of-n* system. Further scenario refinement is added when inserting respective success probabilities for each treatment.

$$
\begin{aligned}
& r(p) = P\{\emptyset(X) = 1\} \\
& = P\{X = (1,1,1)\} + P\{X = (1,1,0)\} + P\{X = (1,0,1)\} + P\{X = (0,1,1)\} \\
& = p_1 p_2 p_3 + p_1 p_2 (1 - p_3) + p_1 (1 - p_2)\, p_3 + (1 - p_1) p_2 p_3 \\
& = p_1 p_2 + p_1 p_3 + p_2 p_3 - 2(p_1 p_2 p_3)
\end{aligned}
$$

Equation 12

Summary

The analytical methods stated in this section provide ways for decision makers to assess what-if scenarios. They provide evidence to help leadership understand possible effect on system performance and/or give a foresight ‘*battle damage assessment*’ to decisions they could make. Literature from treatment reliability theory coupled with forecasting and queueing theory lay the groundwork for efficient consultant staffing that leads to minimizing waste, and maximizing quality and efficiency in recovery care. The next section is dedicated to providing specific methodology to analyze actual scenarios that leadership may encounter.

III. Methodology

Figure 7 shows a methodology to evaluate the WRC and USAF/A1SZ warrior and survivor care program operations. Three main steps are proposed to minimize cost while maximizing efficiency and performance.

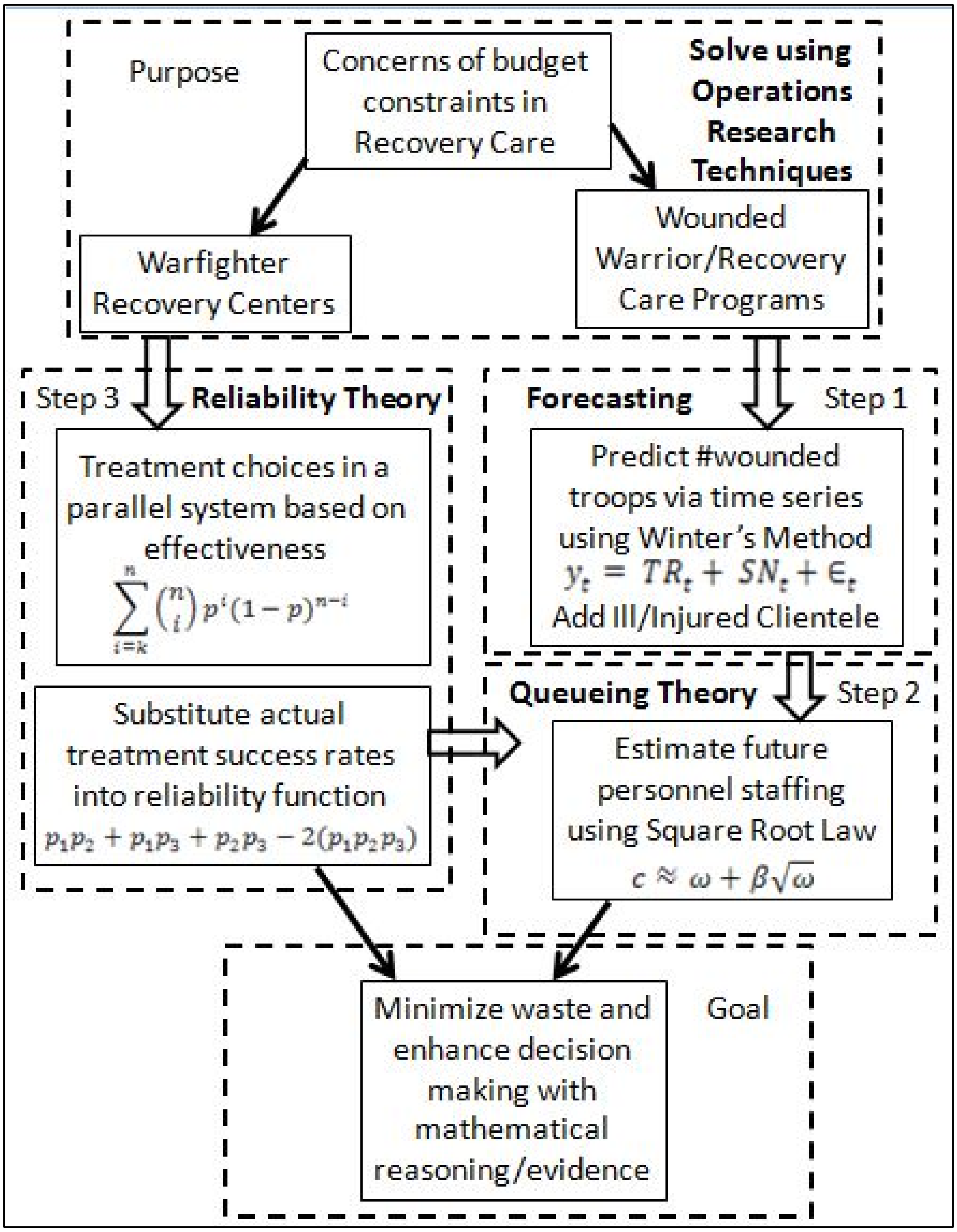


Figure 7: Methodology Diagram

Methodological Steps

Step 1: Forecast the future number of wounded USAF troops given total number deployed to symbolize workload demand. This proportion of wounded troops to total deployed allows multiple scenarios to be compared. Wartime operations, draw-down, and peacetime scenarios are able to be analyzed and compared due to using this proportion instead of outright number of casualties. Because of trend and seasonal correlation of time to the number of troops wounded, Additive Holtz-Winters smoothing method provides the framework for the future forecasts and is compared to the baseline moving average method. This demand output feeds into the next step.

Step 2: Use the forecasted proportion in efforts to estimate future workload of RCC and AFW2 consultants. Given DoDI regulation of no more than 40 clients per consultant, staffing decisions are made based upon forecasted future clientele. The square root law utilizes the Quality and Efficiency Domain to estimate an appropriate number of consultants needed in each program system to maintain quality, improve efficiency and server utilization, thereby minimizing cost and waste. Choosing treatments to remove from a WRC parallel treatment plan (*Step 3*) also provides an avenue to reduce cost. Staffing techniques used in *Step 2* should also be utilized at the WRC to reduce cost and minimize waste. USAF/A1SZ program efficiency coupled with WRC reliability drives toward the main goal of an effective post-operative medical care system with *Step 3*.

Step 3: This step provides a framework for choosing which treatments to be included in the rehabilitation protocol. In a constrained budget scenario, reliability theory is used to evaluate the expected loss in system performance when required to remove treatment(s) in the protocol. System reliability is calculated in the analysis section based upon scenarios with arbitrary treatment reliabilities to demonstrate possible outcomes. The actual treatment reliabilities that are known from research by an internal organization can be substituted for making applicable reliability calculations.

Forecasting Demand Using Smoothing Methods

Additive Holtz-Winters smoothing provides a forecasting method for a time series shown in Equation 13 (Bowerman, O'Connell, & Koehler, 2005).

$$y_t = TR_t + SN_t + \in_t$$

Equation 13

This implies the permanent and trend components TR_t=$(\beta_0 + \beta_1 t)$; seasonal component, SN_t; and error that is inherent to the model, $\in_t$.

When the time series is observed, the updated estimates of the permanent component (level), trend component (slope) and seasonal factor are, shown in Equation 14-Equation 16, respectively.

$$l_t = \alpha\,(y_t - sn_{t-L}) + (1-\alpha)\,(l_{t-1} + b_{t-1})$$

Equation 14

$$b_t = \gamma\,(l_t - l_{t-1}) + (1-\gamma)\,b_{t-1}$$

Equation 15

$$sn_t = \delta\,(y_t - l_t) + (1-\delta)\,sn_{t-L}$$

Equation 16

The values of smoothing constants, α, γ, and δ, are between 0 and 1, and L is the number of periods in the seasonal component.

The point forecast, ‘τ’ periods after time ‘t’, is calculated via Equation 17.

$$Forecast = \hat{y}_{t+\tau} = l_t + \tau\,(b_t) + sn_{t+\tau-L}$$

Equation 17

It takes a number of resources in order to determine the components and variables. Through iCasualites.org, the US combat casualties (by state, month) are modeled as a time dependent series, including breaking out IED casualties in Iraq and Afghanistan (Operation Enduring Freedom). The Congressional Research Service Report for Congress and Brookings Indexes also provides extensive “Boots on Ground” figures. The ratio of USAF casualties over total troop levels in-theatre estimates the targeted proportion that will be used in time series forecasting (Belasco, 2011; Livingston & O'Hanlon, 2012; O'Hankes & Livingston, 2011). With these time series, the Holtz-Winters method estimates the number and proportion of casualties that can be used by commanders involved in planning troop levels in-theatre. This method allows toggle of the troop levels based upon different wartime scenarios to estimate future casualties. These time series will help the AFW2 program forecast their possible new client load.

To forecast future client load for the RCP program and help provide insight into RCC staffing, the forecasting models can be used as an active entity into the ‘wounded’ proportion of the total client load. This proportion is given by Lieutenant Colonel Demmons (2011) as the ratio of wounded troops over the total WII serviced clientele. The injured/ill portion will be assumed constant in the analysis section, also assuming the absence of a large scale attack on US soil or natural disaster(s).

Both the proposed Holtz-Winters method and the baseline simple exponential smoothing method can be compared via root mean squared error (RMSE). The RMSE provides a measure of error from the predicted forecast value to the actual data. This statistic is considered the best way to compare different models. The statistic is calculated by Equation 18 (Ross, 2007).

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2}$$

$$where\ N = number\ of\ data\ points, y = actual\ data, \hat{y} = proposed\ model\ data$$

Equation 18

After choosing the best model it will provide the expected casualties to be used in AFW2 consultant staffing analysis. It will also be a portion of the workload used in analyzing RCC staffing levels, which also depends upon the non-combat injured and ill clientele.

Medical Staffing Using Simple Markovian Queueing Models & The Square Root Law

In the literature review, the quality and efficiency domain was identified as the appropriate way to maintain a fixed quality of service. This method disproves simple intuition that an arbitrary increasing multiple of workload should lead to a congruent increase in the number of servers in the system. Utilizing the square root law in this case demonstrates the number of consultants needed for recovery care based upon future forecasted workload.

AFW2 Consultant Staffing

In order to ensure a certain quality of service for the given system, the equations for the measures of effectiveness must accommodate multiple servers. The steady state probability at zero is needed to calculate these values (Equation 19).

$$p_0 = \left(\frac{{}^{\lambda}/_{\mu}{}^{c}}{c!\,(1-\rho)} + \sum_{n=0}^{c-1}\frac{{}^{\lambda}/_{\mu}{}^{c}}{n!}\right)^{-1}$$

Equation 19

The MOEs, server utilization rate (ρ), mean number in the system (L), expected number in the queue (L_q), the mean waiting time in the queue (W_q), and the mean waiting time in the entire system (W) formulas are derived in Equation 20-Equation 24.

$$\rho = \lambda/c\mu$$

Equation 20

$$L = \frac{\lambda}{\mu} + L_q$$

Equation 21

$$L_q = \frac{{}^{\lambda}\!/_{\mu}{}^{c} * \rho}{c!\,(1-\rho)^2} * p_0$$

Equation 22

$$W = \frac{1}{\mu} + \left(\frac{{}^{\lambda}\!/_{\mu}{}^{c} * \rho}{c!\,(c\mu)(1-\rho)^2} \right) * p_0$$

Equation 23

$$W_q = \frac{L_q}{\lambda}$$

Equation 24

A table of these values can be evaluated to test how the system would react to management decisions, such as changing the number of servers, the client load, and/or the service rate. To determine the appropriate number of servers that would maintain the baseline measure of congestion, 1-$W_q(0)$ is used (Equation 25). This is the probability that a customer is delayed in the queue. Within the view of the thesis, it is construed as the probability that a client would need to wait for a server.

$$1 - W_q(0) = \left(\frac{{\lambda/\mu}^{c}}{c!\,(1-\rho)}\right) \Big/ \left(\frac{{\lambda/\mu}^{c}}{c!\,(1-\rho)} + \sum_{n=0}^{c-1} \frac{{\lambda/\mu}^{n}}{n!}\right)$$

Equation 25

This process is labor intensive and requires complex spreadsheet manipulation to calculate and analyze the MOEs.

The square root law provides a fast and simple tool to approximate the required number of servers to use based upon a forecasted load while maintaining a congruent measure of congestion. Recall Equation 8 for the hypothesized number of servers to work in a system at steady state. Workload (ω) is the ratio of total clientele divided by the DoDI mandated maximum 40 clients per consultant. Hence, the steady state equation is seen below (Equation 26).

$$c = \omega + \Delta\,, where\ \omega = \frac{\#\ of\ Clients}{40}\ and\ recall\ \Delta > 0$$

Equation 26

Therefore, consultant workload must be less than the current number of servers to be in DoDI regulation and to ensure a workable server utilization rate. Now recall the square root law Equation 9 where $\Delta = \beta\sqrt{\omega}$. When workload increases by a multiple 'ν', instead of increasing servers by that same multiple, the proper number of servers for the system is calculated by Equation 27.

$$c = \nu\omega + \beta\sqrt{\nu\omega}$$

Equation 27

The workload multiple is a ratio of the added forecast of future new clientele (casualties) with respect to the current number of clients in the system (Equation 28).

$$\nu = \frac{(Expected\ \#wounded\ in\ forecasted\ year) + (\#\ Current\ Clients)}{(\#\ Current\ Clients)}$$

Equation 28

RCC Staffing

The theory is the same for RCC staffing. However, there is an added component for non-combat ill and injured workload (Equation 29).

$$c = \omega_{ii} + (\omega_w + \Delta)\ , where\ \omega_{ii} = \frac{\#injured\ \&\ ill\ clients}{40}\ \ and\ \omega_w = \frac{\#wounded\ clients}{40}$$

Equation 29

Rachel Morgan (2012) stated that the number of injured and ill clients is currently increasing due to the addition of more retirees and dependents into the system. Yet, the number of injured and ill clients will be assumed constant for the analysis of forecasted time period because of assumed budget constraints. In that environment, it is possible to imagine that combat-wounded personnel would take priority and the program would be unable to undertake more ill and injured clients. Each workload component will have a multiple specific to its client type, and will follow Equation 28. The server approximation square root formula evolves into Equation 30.

$$c = \eta\omega_{ii} + \nu\omega_w + \beta\sqrt{(\eta\omega_{ii} + \nu\omega_w)}\ , where\ \eta = 1$$

Equation 30

The workload multiple for the injured and ill clients (η) will equal one so that it assumes a constant load as defined in the previous paragraph. This value can easily accommodate changes to injured and ill client workload that may be known by the internal organization if required. These formulas will be used for a variety of scenarios in Chapter 4 of this thesis to display how to maximize staffing efficiency in efforts to control costs and quality in warrior and survivor care.

Reliability Theory Analyzing System of Treatments

The main scenario that leadership at the WRC encounters is changing the treatment protocol for patients. This allows a WRC clinic to improve efficiency/efficacy, and possibly cut cost in space, equipment, salary paid to technicians, and the servers/client's time. Similarly to the previous section on queueing based staffing, maximizing technician and consultant efficiency can provide benefits in reducing the budget. The same methods can be used to staff clinical personnel based upon workload changes. However, the WRC requires much more attention to the system effectiveness on relieving symptoms of the patients, so there is more to it than simply matching a certain number of servers to patients. There are many possible scenarios to cut cost at the WRC clinics, but the effect on performance may differ. The reliability function scenarios in the analysis section provide a tool to exemplify treatment protocol changes to cut costs while maintaining high performance.

Using the same method applied in Equation 12 to the WRC TBI center's five treatment protocol, the reliability function for a *Four-out-of-Five* system is seen in Equation 31.

$$
\begin{aligned}
r(p) &= P\{\emptyset(X) = 1\} \\
&= P\{X = (1,1,1,1,1)\} + P\{X = (0,1,1,1,1)\} + P\{X = (1,0,1,1,1)\} + P\{X = (1,1,0,1,1)\} \\
&\qquad + P\{X = (1,1,1,0,1)\} + P\{X = (1,1,1,1,0)\} \\
&= p_1p_2p_3p_4p_5 + (1-p_1)p_2p_3p_4p_5 + \ p_1(1-p_2)p_3p_4p_5 + p_1p_2(1-p_3)\,p_4p_5 \\
&\qquad + p_1p_2p_3(1-p_4)p_5 + p_1p_2p_3p_4(1-p_5) \\
&= p_2p_3p_4p_5 + p_1p_3p_4p_5 + p_1p_2p_4p_5 + p_1p_2p_3p_5 + p_1p_2p_3p_4 \ - 4(p_1p_2p_3p_4p_5)
\end{aligned}
$$

Equation 31

In a different scenario, assume that two of the treatments were needed to be cut. However, it is given that two out of these five treatments *must* be part of the protocol but only one of the remaining three treatments is able to be funded. The reliability function is now given by Equation 32.

$$
\begin{aligned}
r(p) &= P\{X_1 = 1, X_2 = 1, \max(X_3, X_4, X_5) = 1\} \\
&= P\{X_1 = 1\}\,P\{X_2 = 1\}\,P\{\max(X_3, X_4, X_5) = 1\} \\
&= p_1p_2[1 - (1-p_3)(1-p_4)(1-p_5)]
\end{aligned}
$$

Equation 32

Summary

Equation 31 and Equation 32 are useful in a parallel system that can have arbitrary parts fail. However, in the case of WRC clinics, decisions on which treatments will be removed for cost cutting reasons are not arbitrary and would need to be carefully planned. Hence, the scenarios in the following section will be based upon manipulating Equation 10 since management would directly choose which treatment(s) to remove. Hypothetical treatment success probabilities will be substituted into the reliability equation to compare mission effectiveness when choosing specific treatments to remove from the system. Leadership gains insight into the magnitude of performance degradation from scenarios in the next chapter. They also see how the use of forecasting and queueing theory can provide insight into staffing decisions that are optimal in a budget constrained environment.

IV. Analysis, Results, and Evaluation

The current operating procedure to forecast future workload in the RCP and AFW2 programs is by a simple moving average of the past three years of clientele already in the system. With thousands of Airmen from previous conflicts waiting for system entry (Vietnam, Gulf War, Iraq, Afghanistan, etc), these programs are already running behind in participation. This method will always lag behind the exponential growth of clientele that is seen in Figure 8, and does not account for trend or seasonality in clientele arrival data.

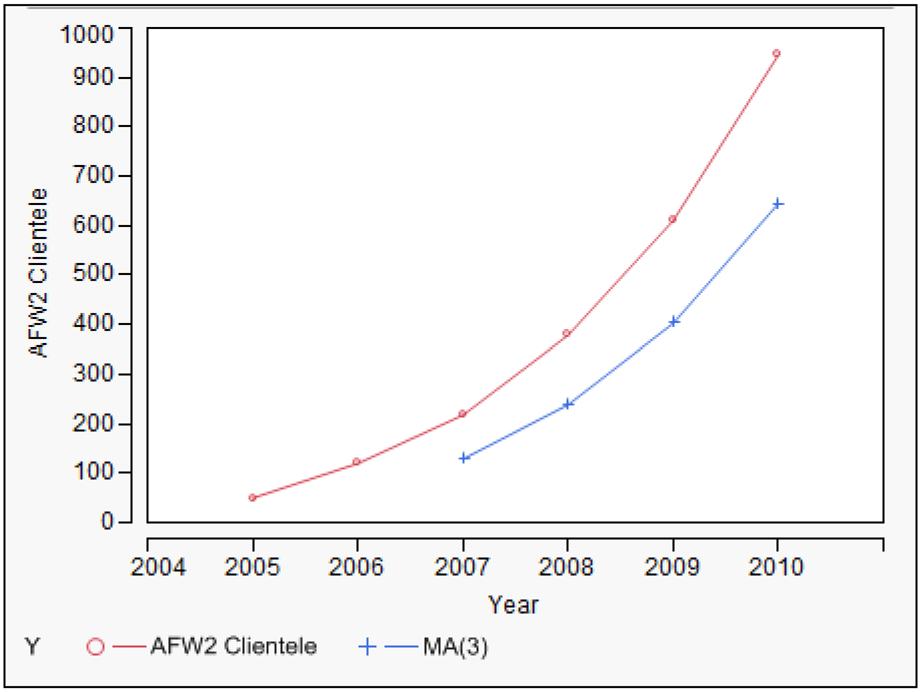


Figure 8: MA Forecast on Previous AFW2 Clientele

Forecasting Demand Analysis

Forecasting Casualties

Instead of using the past clientele in efforts to forecast the future demand, it is recommended to forecast the predicted future casualties from military conflicts. This statistic provides a better forecast of workload demand.

Figure 9 displays a time series of the total number of USAF wounded personnel from Operation Iraqi Freedom (OIF) and Operation Enduring Freedom (OEF). There is a decrease in wounded totals going into every winter, and a spike at the start of every spring, noted as the Taliban "Spring Offensive" where fighting dormancy ceases after the thaw of winter. This implies a hypothesized seasonal component in the casualty data. The average number of wounded Airmen per month would not suffice in describing this time series, so there is an obvious trend component to this data as well.

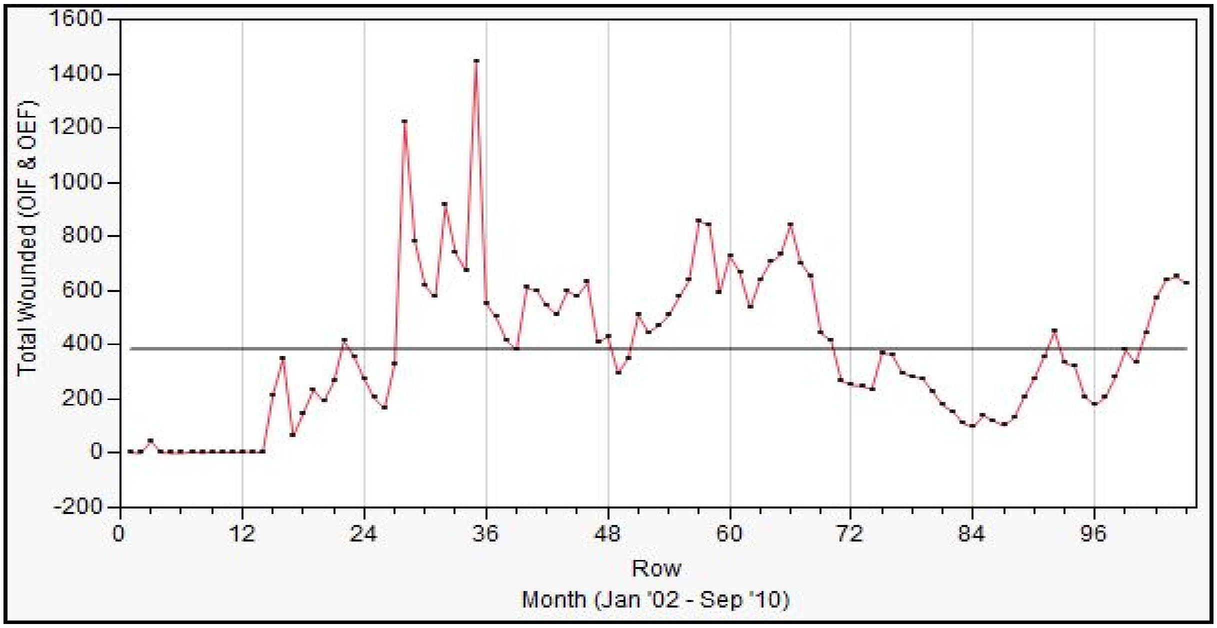
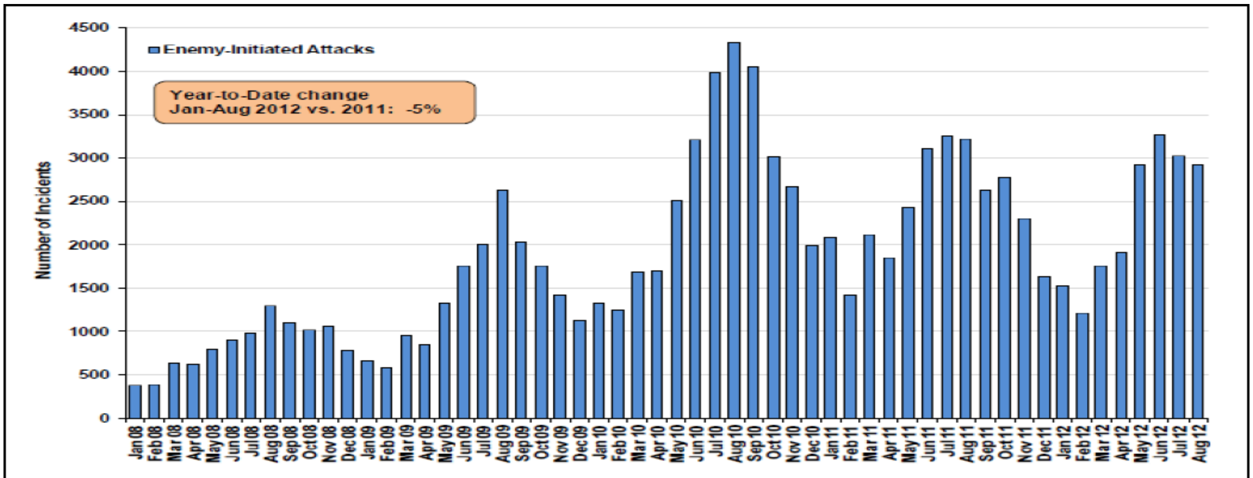


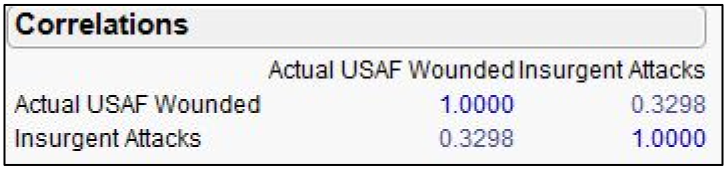
Figure 9: Time Series – Total Wounded US Troops (Afghanistan & Iraq)

Figure 10 displays the pattern of insurgent attacks in Afghanistan. The obvious seasonal factor was also noted in the wounded Airmen data. There is a trend component to the number of attacks as time progresses. The correlation between the wounded Airmen time series (Figure 9) and insurgent attack time series (Figure 10) is believed to be significant at 33% (0.3298) (Figure 11). Hence, the method to forecast the expected number of future USAF casualties should be able to account for this correlated seasonality.



NOTE: Attacks from April-June 2012 were 11% higher than during the same period of 2011, with roughly 110 attacks per day in June. That number per day was the highest seen in June since the start of the war.

Figure 10: Times Series – Number of Incidents/Attacks (Afghanistan & Iraq)



Correlations

	Actual USAF Wounded	Insurgent Attacks
Actual USAF Wounded	1.0000	0.3298
Insurgent Attacks	0.3298	1.0000

Figure 11: Correlation between USAF Casualties and Insurgent Attack Pattern

The trend component in the casualty time series (Figure 9) correlates significantly with the number of total troops that were deployed at the time. As seen in Figure 12, the total number of troops deployed in the Middle East (Iraq & Afghanistan) displays evident trend changes. There was a dramatic increase in deployed troops when Operation Iraqi Freedom began around April 2003. The Iraq troop surge in January 2007 (month 61) starts an increasing linear trend. In April 2010 (month 100), the trend reverses, and decreasing as Iraq begins to draw down. There is also a definitive decreasing trend from October to December 2011 (Months 118-120) as all troops from Iraq were redeployed elsewhere or back to their home station. The correlation between the wounded Airmen time series (Figure 9) and deployed troop levels time series (Figure 12) is believed to be significant at 40% (0.3989) (Figure 13). Therefore, the method to forecast the expected number of future USAF casualties should be able to account for trend in the data.

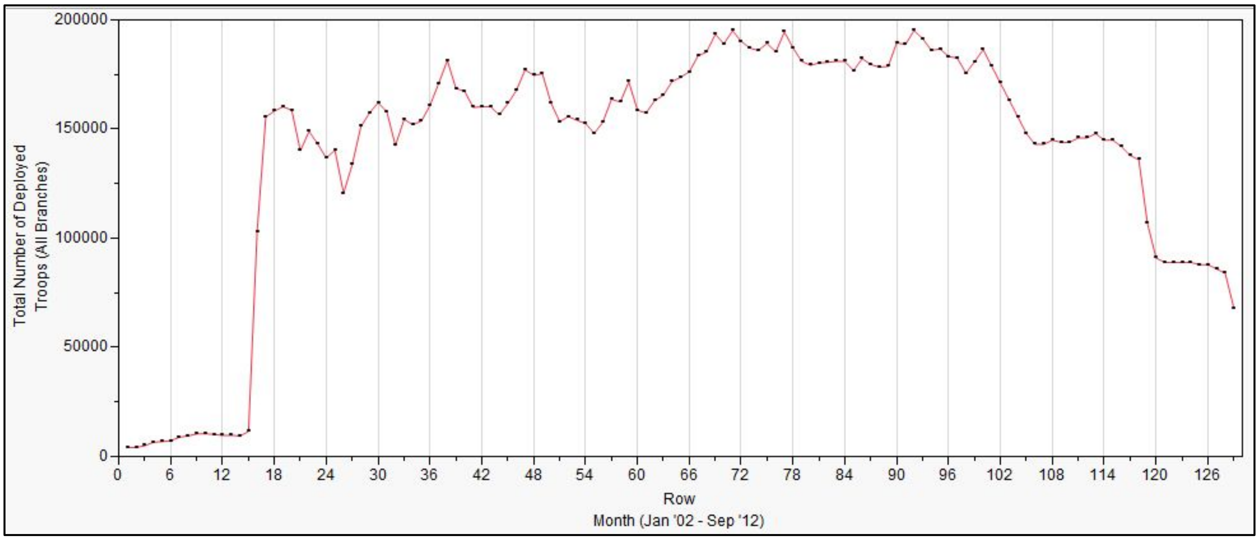
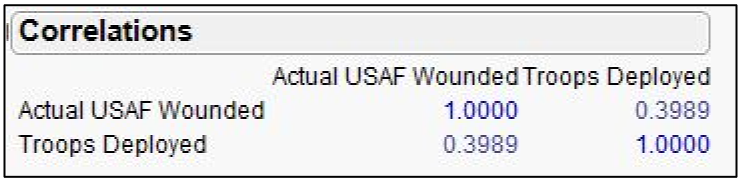


Figure 12: Time Series - Number of Total US Troops Deployed (Afghanistan & Iraq)



Correlations

	Actual USAF Wounded	Troops Deployed
Actual USAF Wounded	1.0000	0.3989
Troops Deployed	0.3989	1.0000

Figure 13: Correlation between USAF Casualties and Deployed Troop Levels

To be able to forecast USAF casualties, the number of troops deployed must be taken into account. Forecasting casualties alone will not be accurate during tempo changes because of this dependence (to be outlined later in this analysis section). Including the total troops deployed allows analysis of any military scenario (ramp up, operational war, drawdown, peacetime). Therefore, the forecast statistic of choice is the proportion of USAF wounded over the total number of troops deployed.

Once the ratio has been forecasted on a monthly basis, the planned deployed number of troops will be multiplied by this ratio to predict that number of USAF casualties used in later analysis. Figure 14 demonstrates the time series proportion based upon the data from OIF and OEF from January 2002 through September 2012.

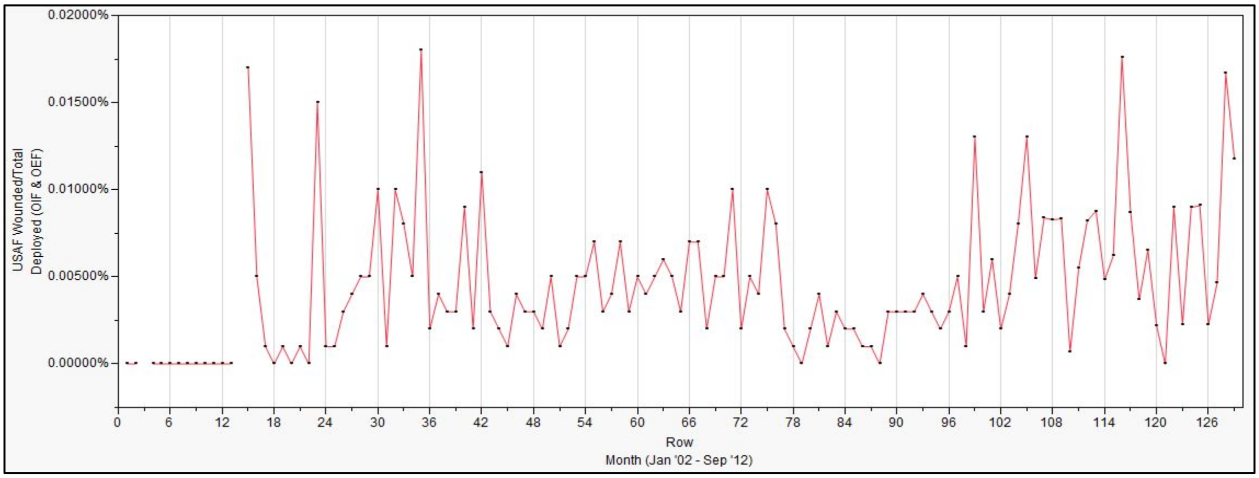


Figure 14: Times Series – Proportion of USAF Wounded to Total US Troops Deployed (Afghanistan & Iraq)

To analyze the forecasting methods, the last 24 months of actual data was not used in configuring the forecast. This allowed two full years to see how well the forecasts stood in comparison with the actual data that was held out. The forecasts attempted to “learn” from the data in order to project an estimated proportion of USAF wounded personnel, given deployed troop levels. The first forecasted year was from October 2010 to September 2011, when there was a heavy military presence in both Iraq and Afghanistan. The second forecasted year was from October 2011 to September 2012, which encompasses the Iraq troop drawdown. Comparison of the actual wounded data to the forecasted data will help to validate model adequacy in any war/peacetime environment.

Moving Average

The first method used in forecasting was a simple exponential moving average. This method is cheap, quick, simple, and well known. Using this method to forecast future casualties is already an improvement from the moving average method currently used on previous clientele.

As seen in Figure 15, this method does not account for trend or seasonality, and produces a relaxed fit to the actual ratio of wounded/troops deployed. The first forecasted year is displayed visually with the 95% confidence interval. The model and the forecast stay close to the grand average throughout the entire time domain.

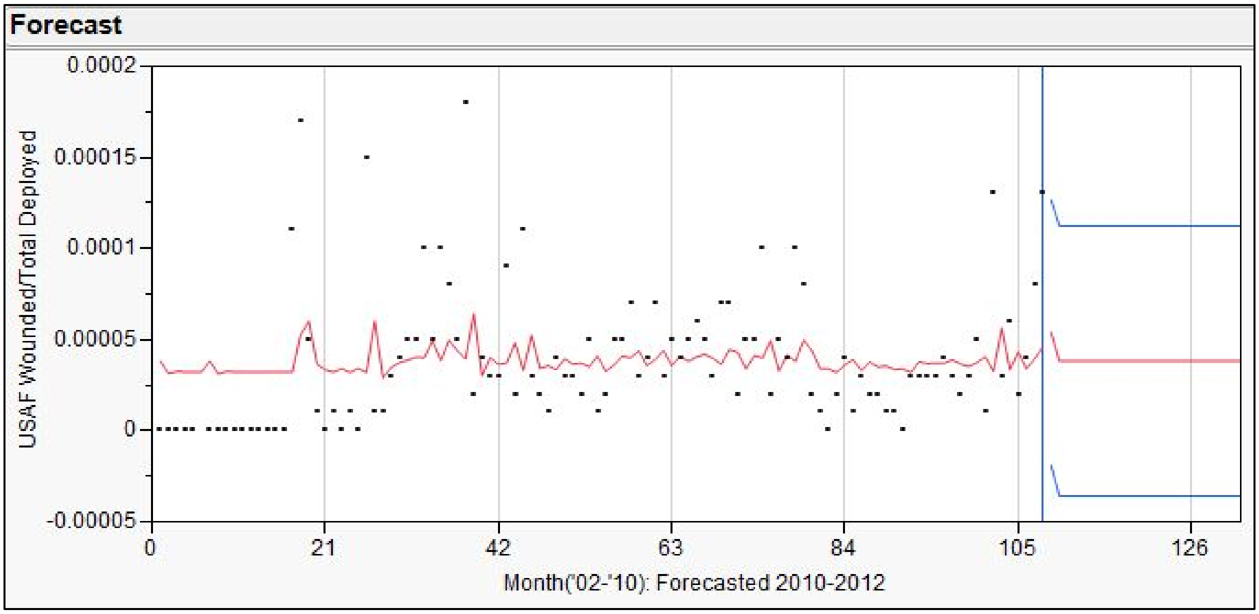


Figure 15: MA Ratio Forecast

Now that the forecast of the ratio has been computed on a monthly basis (24 months forecasted), each ratio is multiplied by the actual number of troops that were deployed during that respective month. When forecasting into the future, analysts should use the planned/projected number of troops that will be deployed each month for their investigation. The time series forecast along with the 95% confidence levels are seen in Figure 16, overlaid with the actual number of casualties. Even though a constant ratio for all 24 months was forecasted, there is a decreasing predicted number of casualties. This is due to less troops being down-range due to the drawdown in Iraq. The MA forecast mostly underestimates the actual number of casualties, but is rather accurate with only two data points breaking outside the 95%

confidence limits. The RMSE for this method was calculated as 6.203. This value will be compared to the following method.

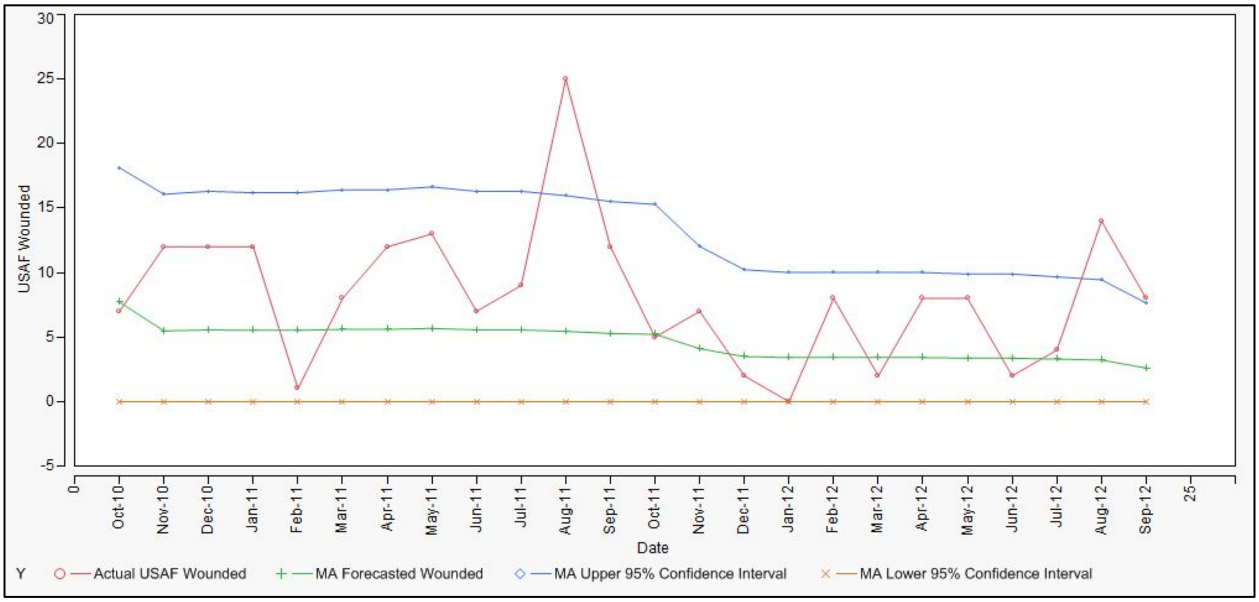


Figure 16: MA Casualty Forecast

Additive Holtz-Winters

The Holtz-Winters method (Figure 17) appears to fit the actual ratio better than the moving average method, as it accounts for the trend and seasonality gyrations in the data. The forecast includes growth in the trend, indicating a higher proportion of USAF personnel are expected to be wounded. This could be for any hypothesized purpose and is outside of the scope of this thesis. The forecast also includes seasonal spikes as seen in the fitted model.

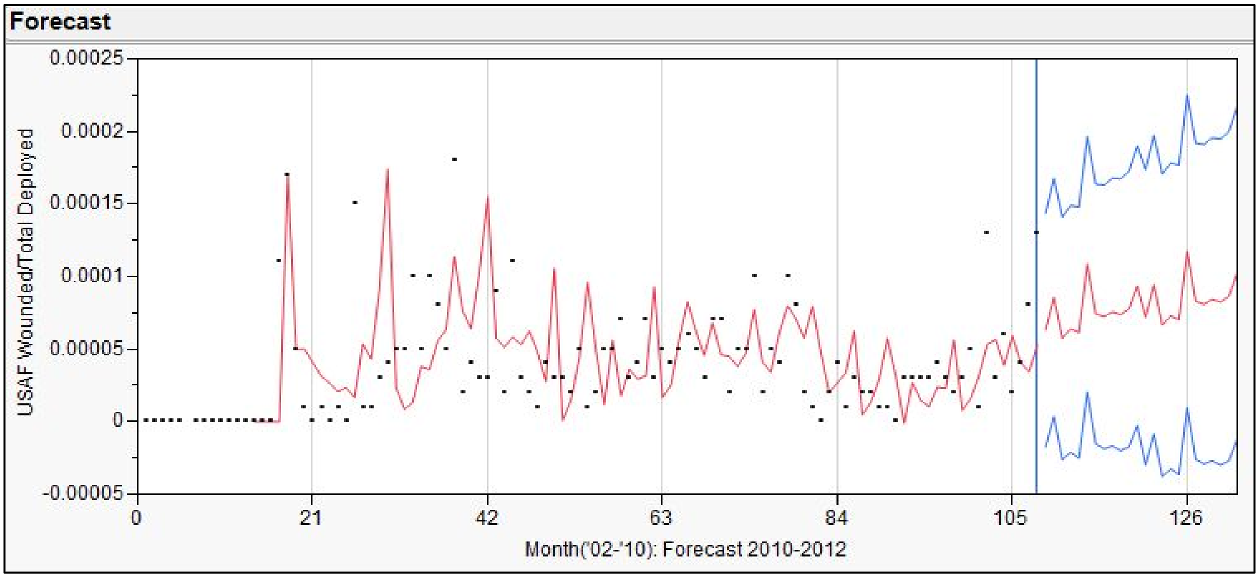


Figure 17: Holtz-Winters Ratio Forecast

This ratio is then multiplied by the number of troops deployed for the respective month, to create the predicted casualty time series (Figure 18). The forecast attempts to follow the gyrations in the actual data. There was an increasing ratio forecasted, yet the number of casualties is seen to be decreasing. Again, this is due to Iraq drawdown. The forecasted number of casualties is accurate and all actual data points are contained within the 95% confidence limits. The RMSE calculated for this method was 4.862. That is a 22% improvement over the MA method. Therefore, the Holtz-Winters method seems superior. However, the MA method to forecast casualties could still be a valid option if needing computational simplicity.

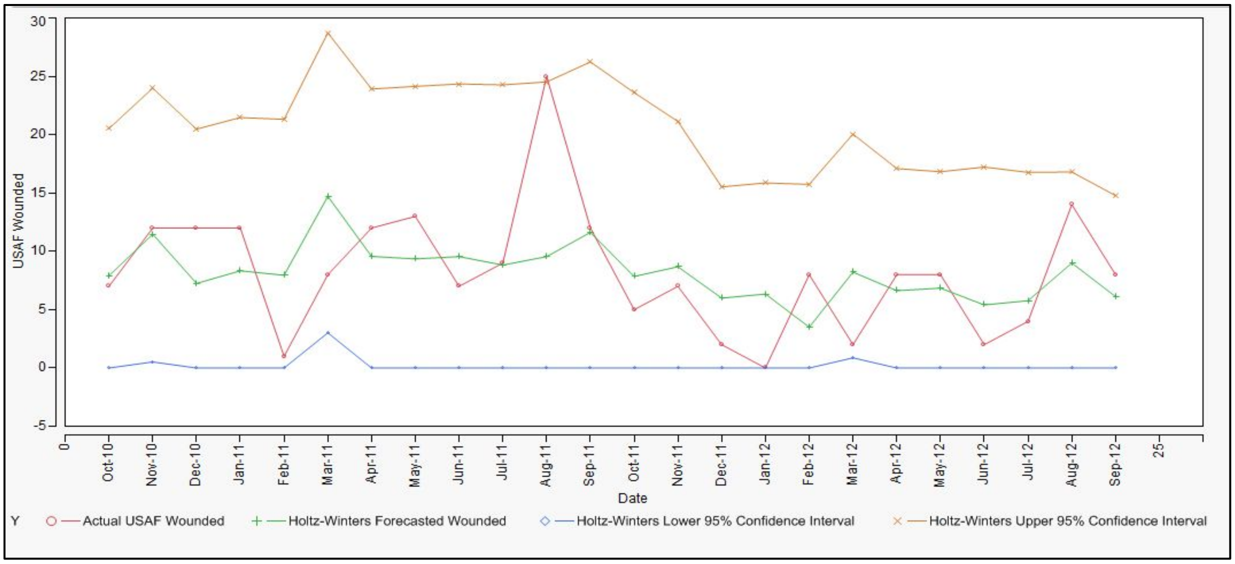


Figure 18: Holtz-Winters Casualty Forecast

After choosing the Holtz-Winters method as the superior smoothing technique, it is important to validate using the ratio of USAF wounded/troops deployed instead of the raw casualty data in order to forecast. For comparison purposes, the casualty data alone is used to forecast future casualties using the Holtz-Winters method. As seen in Figure 19, using only the casualty data leads to a rather level forecast with minimal trend changes and some seasonal spikes. Recall that forecasting the ratio which includes the number of troops deployed (Figure 17) displayed an increasing trend.

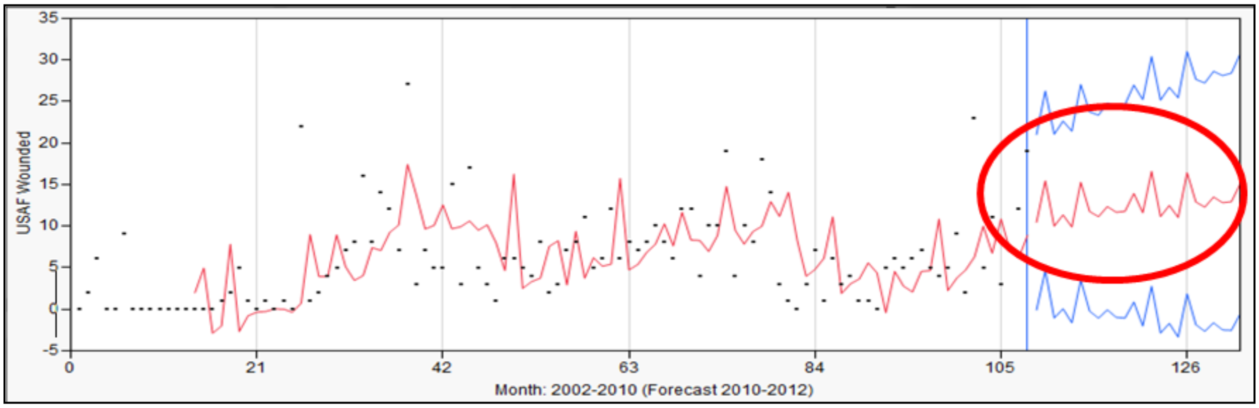


Figure 19: Holtz-Winters Casualty Alone Forecast (no ratio)

After the ratio forecast is multiplied by the deployed troops for each respective month, the resultant Holtz-Winters ratio method casualty forecast is overlaid with the casualty alone forecast, and the actual data (Figure 20). Note that the casualty alone forecast overestimates the actual number of USAF casualties during the second year, while the ratio forecast adapted to the change in trend. As stated earlier, this reduction in casualties is correlated with the drawdown from Iraq. To further ratio validation, the RMSE of the casualty alone forecast is 7.299. With the Holtz-Winters RMSE equal to 4.862, it is a 33% improvement over the casualty alone forecast. Even the MA method using the ratio displays an improvement over the casualty alone forecast (15% improvement). Hence, the ratio method to predict casualties is able to accommodate all war scenarios, while trying to forecast casualties from that data alone can only be used when operational tempo remains constant.

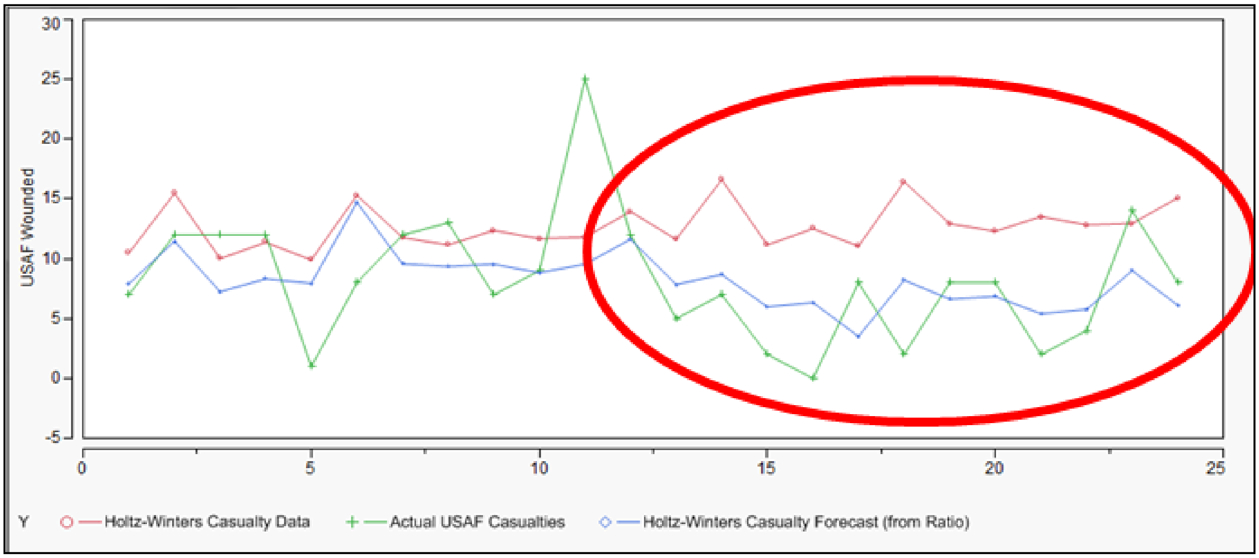


Figure 20: Casualty Alone Versus Ratio Forecast

After choosing the Holtz-Winters ratio method as the superior model for the methodology step 2 input, we must ensure the assumptions are met. The residual plot (Figure 21) and the Basic Diagnostics Chart (Figure 22) imply an adequate model that assumes the errors are random, are normally distributed, and are independent from each other. There is no significant autocorrelation and a slight violation of homoscedasticity has a negligible effect on the resulting forecast. The non-constant variance is most apparent at the start of the two campaigns in which variation may come heavily from the '*Fog of War*'. The smoothing techniques learn most from the data leading closely up to the forecasting period. One should take extra precaution if applying these methods at the start of a new military campaign.

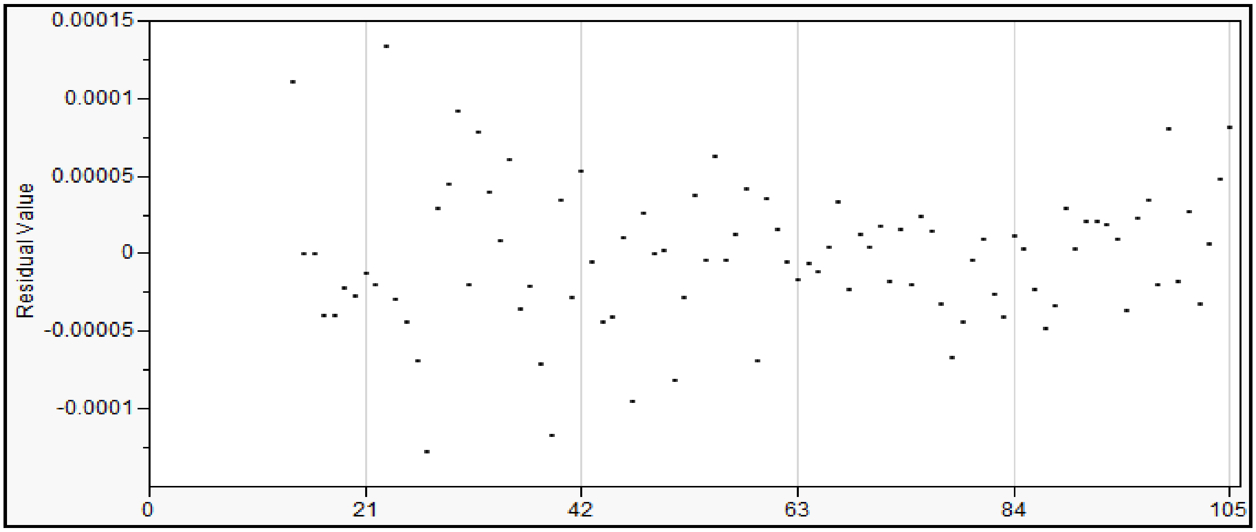
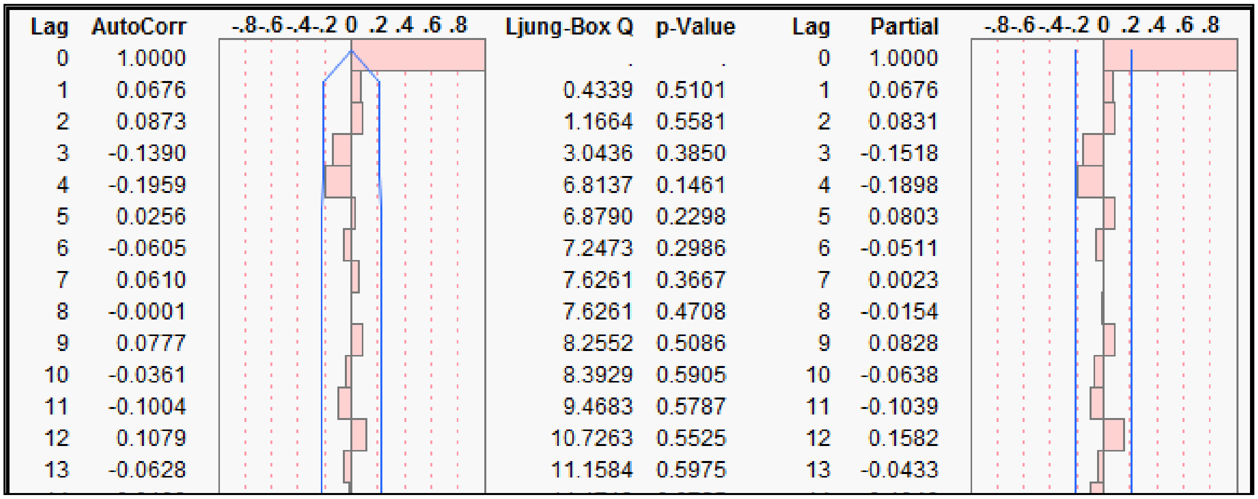


Figure 21: Holtz-Winters Forecast Residual Plot



Lag	AutoCorr	Ljung-Box Q	p-Value	Lag	Partial
0	1.0000	.	.	0	1.0000
1	0.0676	0.4339	0.5101	1	0.0676
2	0.0873	1.1664	0.5581	2	0.0831
3	-0.1390	3.0436	0.3850	3	-0.1518
4	-0.1959	6.8137	0.1461	4	-0.1898
5	0.0256	6.8790	0.2298	5	0.0803
6	-0.0605	7.2473	0.2986	6	-0.0511
7	0.0610	7.6261	0.3667	7	0.0023
8	-0.0001	7.6261	0.4708	8	-0.0154
9	0.0777	8.2552	0.5086	9	0.0828
10	-0.0361	8.3929	0.5905	10	-0.0638
11	-0.1004	9.4683	0.5787	11	-0.1039
12	0.1079	10.7263	0.5525	12	0.1582
13	-0.0628	11.1584	0.5975	13	-0.0433

Figure 22: Holtz-Winters Forecast Basic Diagnostics Chart

Therefore, the Additive Holtz-Winters ratio smoothing method is the best forecast for determining the expected number of USAF casualties. This forecast data will be the input used in the following section in Step 2 (Medical Staffing Analysis).

Medical Staffing Analysis

All newly wounded warriors are automatically entered into the recovery care system. Therefore, the demand of new clientele is estimated using the expected casualty forecast from step 1 of the methodology. Queueing theory can provide specific measures and statistics that are helpful in determining the appropriate number of servers that can maintain a high level of quality and efficiency. Different queueing scenarios can provide evidence to decision makers of possible solutions to operating in the current and future environment. The variables and MOEs of interest for these scenarios are defined in Table 1.

c (# of consultants)	**ρ - Server utilization rate**	**Lq - Mean expected # of people waiting**	**1-Wq(0): Probability of waiting for server**

Table 1: Medical Staffing Variables & MOEs

AFW2 Consultant Staffing

As of 31 January 2011, the number of AFW2 consultants was 17 with a total clientele equaling 975. Hence, the steady state equation is seen in Equation 33.

$$c = \omega + \Delta, \qquad where\ \omega = \frac{975}{40} = 24.375\ and\ recall\ \Delta\ must\ be\ greater\ than\ zero$$

Equation 33

Consultant workload at this point is not within regulation, as the right side of the equation must equal the left. The AFW2 consultants are servicing many more wounded warriors than they are supposed to and that they have time for. In this situation, the server utilization is well over one, the queue is unstable and growing exponentially, and no new clients will enter the system because the probability of waiting for a server is greater than one. Hence, the consultants will be unable to keep up with the workload to provide high quality of service to the wounded warriors. At this point in time, AFW2 consultants were very understaffed.

AFW2	**# of clients**	**# of clients per consultant**	**c (# of consultants)**	**ρ - Server utilization rate**	**Lq - Mean expected # of people waiting**	**1-Wq(0): Probability of waiting for server**
Current Ops	975	40	17	1.43	UNSTABLE	4.220

Table 2: Current AFW2 Consultant Workload Statistics

Therefore, the current situation should not be the baseline to analyze. For explanatory purposes, consider a baseline number of consultants equal to 30. The baseline load and server workload capacity are given.

$$30 = \omega + \Delta, \qquad where\ \omega = \frac{975}{40} = 24.375$$

In this case, Δ = 6 additional servers are needed for the system when rounding workload to the nearest integer (24). After the baseline system has been defined, a series of scenarios are analyzed. The baseline along with all other scenario variations are solved using the techniques developed earlier in this thesis to explore the effect when changing inputs of the number of clients in the system, the number of consultants, and the number of clients that each consultant can provide service to.

Efficiency Domain

Recall the Efficiency Domain method that recommends server increase based upon a specified buffer. In the baseline case, this buffer is 6 consultants. When applying this method with a constant increase in workload, the percentage of this buffer to the total number of servers required for the given load starts to shrink. As seen in Table 3, server utilization, the number in the queue awaiting system entry, and the probability that a new client will need to wait for a server is increasing with the increased load. Therefore, this method leads to a level of high congestion and customer dissatisfaction.

AFW2	**% Change in Client Load**	**# of clients per consultant**	**c**	**ρ**	**Lq**	**1-Wq(0)**
Efficiency Domain Method	-30%	40	23	0.74	0.35	0.123
	-20%	40	26	0.75	0.34	0.114
	-10%	40	28	0.78	0.56	0.154
	0%	40	30	0.81	0.86	0.199
	10%	40	33	0.81	0.78	0.180
	20%	40	35	0.84	1.14	0.224
	30%	40	38	0.83	1.02	0.202
	40%	40	40	0.85	1.42	0.244
	50%	40	43	0.85	1.25	0.221
	60%	40	45	0.87	1.70	0.261
	70%	40	47	0.88	2.26	0.304
	80%	40	50	0.88	1.97	0.275
	90%	40	52	0.89	2.58	0.316
	100%	40	55	0.89	2.24	0.288
	110%	40	57	0.90	2.88	0.327
	120%	40	60	0.89	2.51	0.298
	130%	40	62	0.90	3.18	0.336
	140%	40	65	0.90	2.77	0.307
	150%	40	67	0.91	3.46	0.344
	160%	40	69	0.92	4.31	0.383
	170%	40	72	0.91	3.73	0.351
	180%	40	74	0.92	4.61	0.388
	190%	40	77	0.92	4.00	0.357
	200%	40	79	0.93	4.90	0.393

Table 3: AFW2 Staffing Efficiency Domain Method

Quality Domain

Intuitively, it would seem appropriate to increase an organization's staff based upon the same rate of new clientele entering the system. This describes the Quality Domain method, which attempts to maintain the server utilization in the system. This is done by simply multiplying the baseline number of servers by the percent increase in client load. This method bodes well for simplicity in calculating the optimal solution. However, this method often overestimates the optimal number of servers. As seen in Table 4, system quality improves in that the expected number in the queue and the probability that a new client must wait for a server decreases. Yet, when the probability of delay decreases much past 15-20%, the system yields an abundance of unutilized servers. Hence, while holding quality alone as the goal, the organization may incur excess cost with negligible system improvement. This problem is not a luxury that an organization can afford in a budget constrained environment.

AFW2	% Change in Client Load	# of clients per consultant	c	ρ	Lq	1-Wq(0)
Quality Domain Method	-30%	40	21	0.81	1.19	0.274
	-20%	40	24	0.81	1.06	0.246
	-10%	40	27	0.81	0.96	0.221
	0%	40	30	0.81	0.86	0.199
	10%	40	33	0.81	0.78	0.180
	20%	40	36	0.81	0.71	0.164
	30%	40	39	0.81	0.65	0.149
	40%	40	42	0.81	0.59	0.136
	50%	40	45	0.81	0.54	0.124
	60%	40	48	0.81	0.49	0.114
	70%	40	51	0.81	0.45	0.104
	80%	40	54	0.81	0.41	0.095
	90%	40	57	0.81	0.38	0.088
	100%	40	60	0.81	0.35	0.081
	110%	40	63	0.81	0.32	0.074
	120%	40	66	0.81	0.30	0.068
	130%	40	69	0.81	0.27	0.063
	140%	40	72	0.81	0.25	0.058
	150%	40	75	0.81	0.23	0.054
	160%	40	78	0.81	0.21	0.050
	170%	40	81	0.81	0.20	0.046
	180%	40	84	0.81	0.18	0.042
	190%	40	87	0.81	0.17	0.040
	200%	40	90	0.81	0.17	0.039

Table 4: AFW2 Staffing Quality Domain Method

Quality and Efficiency Domain

The Quality and Efficiency Domain method attempts to maintain the baseline congestion and provides the best method to find optimal staffing solutions. The following scenarios that will be analyzed utilizing the Quality and Efficiency Domain represent real-world situations that leadership making staffing decisions has faced from 2010-2012:

- A change in the server workload capacity to analyze affect on server utilization and system MOEs:
    - inferior efficiency
    - improved efficiency.
- An increase in total client load in the system based upon forecasted wounded totals to determine the optimal number of servers to match baseline system congestion.
    - First forecasted year using Holtz-Winters Method, and
    - Through the second forecasted year using Holtz-Winters Method.
- An increase in total client load in the system based upon forecasted wounded totals through the second year, with improved server workload capacity. Attempts to analyze improved workload efficiency with increasing demand to determine the optimal number of servers to match baseline congestion.

The inputs for all the scenarios can be seen below in Table 5.

AFW2	**# of clients**	**# of clients per consultant**	**c**	**ρ**
Baseline	975	40	30	0.81
Inferior Efficiency	975	33-39	30	0.83-0.98
Improved Efficiency	975	41-47	30	0.69-0.79
1yr Forecast Load	1091	40	28-39	0.70-0.97
2yr Forecast Load	1171	40	30-41	0.71-0.98
2yr Forecast & Improved Efficiency	1171	45	27-37	0.70-0.96

Table 5: AFW2 Staffing Scenario Inputs

The in-regulation baseline MOE values of interest based upon 30 consultants are as follows in Table 6.

AFW2	**# of clients**	**# of clients per consultant**	**c**	**ρ**	**Lq**	**1-Wq(0)**
Baseline	975	40	30	0.81	0.86	0.199

Table 6: AFW2 Baseline Statistics

The most appropriate MOEs that indicates congestion in this case is the probability of having to wait for a server [1-W_q(0)] and the average expected number of clients waiting in the queue (L_q). In the baseline case, the busy/utilization rate of a server is 81% (19% emptiness). The probability that a client will need to wait for a server is just under 0.2, or a 20% chance. The expected number of clients waiting for a server is less than 1. When analyzing different load, service rate, and number of server scenarios, the best solution(s) should maintain the baseline figures to ensure quality of care and to avoid waste.

Results of AFW2 Consultant Staffing Scenarios

In analyzing the baseline scenario, it would be beneficial to understand the effect of increasing the workload on each server if regulation was changed and management found ways to streamline processes/treatments or consolidate the way clients are handled. As seen below in Table 7, server utilization decreases when servers can assume a higher client workload. It would be imperative for a reduction in workload per client via management changes for this to take effect. If client workload increases and servers are not able to handle as many as 40 clients, the server utilization will rise and cause a high probability waiting for a consultant. When consultants are only able to have 35 clients, their emptiness rate falls below 10%, and the queue starts to fill with an expected 8 clients waiting to join the system. When consultants have a reduction in workload per client and can have 45 clients, the emptiness approaches 30% with a server utilization of 0.72. This utilization rate is rather low and the ample supply of servers for a low demand of clients would waste resources.

AFW2	**# of clients**	**# of clients per consultant**	**c**	**ρ**	**Lq**	**1-Wq(0)**
Change in server workload capacity	975	33	30	0.98	58.74	0.904
	975	35	30	0.93	7.76	0.597
	975	37	30	0.88	2.81	0.389
	975	39	30	0.83	1.25	0.250
	975	41	30	0.79	0.61	0.159
	975	43	30	0.76	0.31	0.100
	975	45	30	0.72	0.16	0.062
	975	47	30	0.69	0.09	0.038

Table 7: AFW2 Staffing Change in Server Workload Capacity

The future client load can be estimated by taking the forecasted ratio of wounded USAF personnel over the number of troops deployed for each month from Holtz-Winters method, multiplied by the actual number of troops deployed for that month. Adding the first twelve months together of predicted USAF wounded personnel yields an expected 116. Hence, as the Iraq war was winding down and redeploying troops back to CONUS from October 2010 to September 2011, the Holtz-Winters method predicted 116 wounded Airmen. Applying the same procedure for the second year (October 2011-September 2012), the forecast method predicts 80 additional USAF wounded personnel after the Iraq drawdown and continued Afghanistan military presence.

Therefore, an increase in load of 116 clients is expected for the first forecasted year (October 2010 – September 2011). Leadership could plan their consultant staffing based upon this number and determine how many additional servers to add to their workforce by analyzing Table 8. As seen in Figure 18, the 95% confidence interval for the Holtz-Winters forecasting method captures the actual wounded totals in all 24 forecasted months. Therefore, planners can be confident in using the upper 95% confidence interval limit as the increase in load if they would like to create a staffing buffer. The change in load would also have to take into account the expected number of clients that will complete AFW2 service, which is internally known to the AFW2 organization. Since this rate is unknown to the author, for analysis purposes, the change in load will be configured as the expected forecast for the first year. This is approximately 1.12 times the baseline load, or a 12% increase.

Server workload is kept constant at 40 clients, the mandated maximum via DoD regulation. If budget constraints required cutting servers, AFW2 would still be in regulation with 28 servers. However, over 84% of clients would expect to wait for consultants as the server utilization rate is dangerously close to one. This would entail approximately 32 clients waiting in the queue for entry into the system. Given the increase in load and keeping the same number of servers as the baseline (30), server utilization will rise from 0.81 to 0.91, falling under 10% emptiness. At this rate, the quality of service will undoubtedly diminish. Table 8 indicates that 33-34 servers is the best solution that would maintain the baseline's probability of waiting and expected number in the queue. More than 34 servers would not be an effective solution for the minimal quality increase and excess cost in server supply.

AFW2	**# of clients**	**# of clients per consultant**	**c**	**ρ**	**Lq**	**1-Wq(0)**
Change in total client load: First forecasted year	1091	40	28	0.97	31.77	0.844
	1091	40	29	0.94	10.45	0.661
	1091	40	30	0.91	5.10	0.510
	1091	40	31	0.88	2.83	0.387
	1091	40	32	0.85	1.67	0.289
	1091	40	33	0.83	1.01	0.213
	1091	40	34	0.80	0.62	0.154
	1091	40	35	0.78	0.39	0.109
	1091	40	36	0.76	0.24	0.076
	1091	40	37	0.74	0.15	0.052
	1091	40	38	0.72	0.09	0.035
	1091	40	39	0.70	0.05	0.023

Table 8: AFW2 Staffing Change in Total Client Load (First Yr Forecast)

The second forecasted year (October 2011 – September 2012) adds another 80 clients to the predicted program clientele, which now has increased by approximately 20% from the baseline. As seen in Table 9, if the baseline number of servers is maintained through the second year, the quality of service would suffer. The probability of waiting is approximately 85%. In this case, the optimal number of consultants to have on staff is approximately 35-36 to match baseline congestion.

AFW2	**# of clients**	**# of clients per consultant**	**c**	**ρ**	**Lq**	**1-Wq(0)**
Change in total client load: Second forecasted year	1171	40	30	0.98	34.29	0.849
	1171	40	31	0.94	11.38	0.670
	1171	40	32	0.91	5.61	0.522
	1171	40	33	0.89	3.15	0.401
	1171	40	34	0.86	1.88	0.303
	1171	40	35	0.84	1.15	0.226
	1171	40	36	0.81	0.72	0.165
	1171	40	37	0.79	0.45	0.119
	1171	40	38	0.77	0.28	0.084
	1171	40	39	0.75	0.18	0.059
	1171	40	40	0.73	0.11	0.040

Table 9: AFW2 Staffing Change in Total Client Load (Second Yr Forecast)

For the next scenario, assume that management was able to decrease replication of work and consultants were able to increase their workload capacity to 45 clients instead of the regulated 40 clients. Table 10 illustrates this scenario and implies that 31-32 servers could maintain baseline congestion. Hence, further investigation into improving work efficiency could be very cost effective as four fewer servers are required.

AFW2	**# of clients**	**# of clients per consultant**	**c**	**ρ**	**Lq**	**1-Wq(0)**
Change in total client load with increased server workload capacity	1171	45	27	0.96	21.04	0.790
	1171	45	28	0.93	8.06	0.613
	1171	45	29	0.90	4.09	0.468
	1171	45	30	0.87	2.30	0.351
	1171	45	31	0.84	1.36	0.259
	1171	45	32	0.81	0.82	0.188
	1171	45	33	0.79	0.50	0.134
	1171	45	34	0.77	0.31	0.094
	1171	45	35	0.74	0.19	0.064
	1171	45	36	0.72	0.11	0.043
	1171	45	37	0.70	0.07	0.028

Table 10: AFW2 Staffing Change in Total Client Load & Server Workload Capacity

Figure 23 indicates the appropriate number of servers based upon client load changes for the ‘Quality Domain’ and ‘Quality and Efficiency Domain’. The Quality Domain matches a change in workload to the appropriate change in servers, while the Quality and Efficiency Domain changes the number of servers in effort to maintain baseline congestion. The slope of the Quality Domain is greater than the slope of the Quality and Efficiency Domain. Hence, the organization does not need to increase their supply of staff to match the incoming demand of clientele when following the Quality and Efficiency Domain method. This provides a cheaper and less wasteful option than intuitively matching workload to servers.

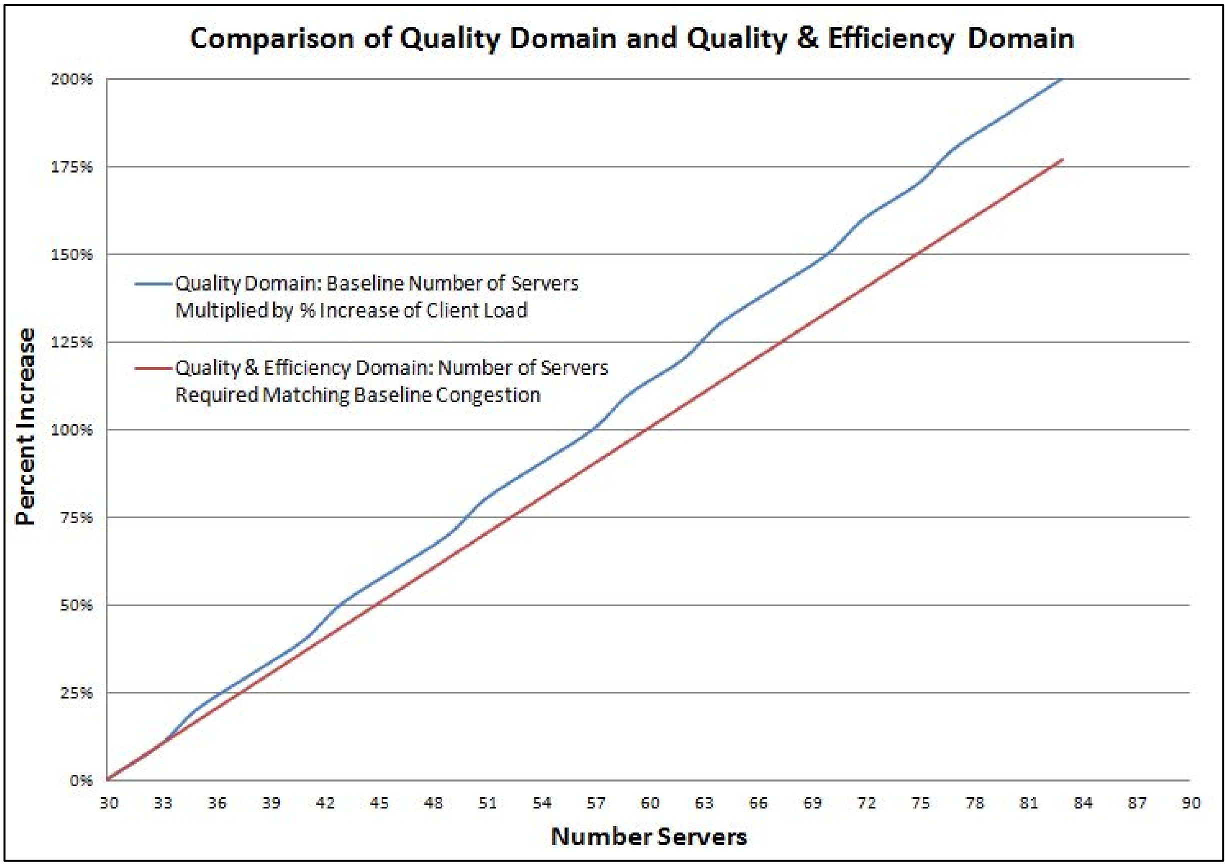


Figure 23: Comparison of Slopes - 'Quality Domain' & 'Quality and Efficiency Domain'

This process to find optimal solution(s) in the previous scenarios requires complex functions and time consuming methods. Hence, having the ability to approximate the appropriate number of additional servers without the need of calculating respective MOEs will simplify this multifaceted problem.

Using Equation 28, the workload for the first forecasted year increases by the multiple '1.12'. This value is substituted into Equation 27 as $'\nu'$. Recall that the value for β is the (1-α) quantile of the normal distribution. Setting $\alpha = 0.1$, the (1-0.1) quantile of the normal distribution is 1.2816. Hence, the appropriate number of servers for the system based upon the approximation formula is 34 consultants (Equation 34). This answer coincides with the answer found from Table 8, when baseline MOE values were maintained.

$$c \approx 1.12\left(\frac{975}{40}\right) + 1.2816\sqrt{1.12\left(\frac{975}{40}\right)} \approx 33.996 \sim 34\, consultants$$

Equation 34

Keeping with the same procedure, the forecast for the second year yields a total of 36 consultants (Equation 35), which also coincides with the answer found from Table 9, when baseline MOE values were maintained.

$$c \approx 1.20\left(\frac{975}{40}\right) + 1.2816\sqrt{1.20\left(\frac{975}{40}\right)} \approx 36.181 \sim 36\, consultants$$

Equation 35

Therefore, an additional six consultants are needed to be staffed within the next two years. It is not cost effective to hire all of them 24 months in advance. Since the data suggests a seasonal variation in wounded forecasting, it is wise to determine on a monthly basis when a new consultant would be needed. The load increase can be broken down by month to determine when an additional consultant would acquire close to a full client workload. It is recommended to staff prior to this point to avoid congestion. As seen in Figure 24, the third and fourth consultants are needed more quickly than the other consultants due to the increased demand from the “Spring Offensive”. There also seems to be decreasing demand throughout the second year due to the absence of troops in Iraq. This shows the importance of using a forecasting method that takes into account changes in trend and seasonality. Without such, staffing would be based upon equal time intervals across the forecasted two year period. However, it is evident that additional consultants are needed at non-linear times throughout the year.

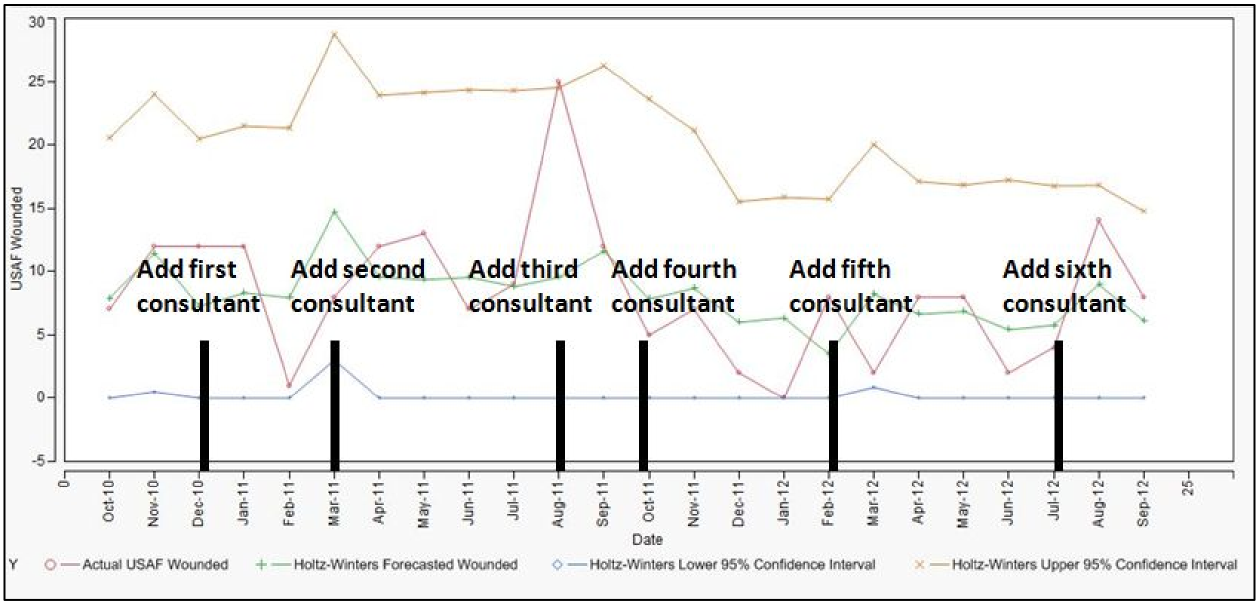


Figure 24: Recommended AFW2 Timing Staffing Decisions Based Upon Forecast

The square root approximation function can easily be modified to analyze any load change that may be concluded from forecasting future clientele. It can also adapt to changes in regulation, if each server is required to take on more clientele. This method is simple, quick, and reliable for any staffing planner to utilize.

RCC Staffing

As of 31 January 2011, the number of RCC consultants was 32. The program contained 231 wounded troops, 574 injured or ill clients, equaling a total of 805 clients. These values are substituted into the load variables in Equation 29 in order to find the optimal number of consultants.

$$c = 32 = \omega_{ii} + (\omega_w + \Delta)\,, where\ \omega_w = \frac{231}{40} \approx 5.8\ and\ \omega_{ii} = \frac{574}{40} \approx 14.4$$

Consultant workload is in regulation, requiring at least 21 total RCC consultants. In the baseline, Δ = 9 additional servers are needed for the system. The baseline MOE values with 32 RCC consultants are seen in Table 11.

RCC	**# of clients**	**# of clients per consultant**	**c**	**ρ**	**Lq**	**1-Wq(0)**
Baseline	805	40	32	0.63	0.02	0.010

Table 11: RCC Baseline Statistics

In the baseline case, the utilization rate of each server is 0.63. The probability that a client will need to wait for a server is very low at 0.01, or a 1% chance. The expected number of clients waiting for a server is also very low, a value of less than 0.02 which would be equivalent to zero in reality. In comparison to the baseline AFW2 values, it seems that RCC consultants are underutilized. However, since the RCC is the primary focal point for the medical recovery care plan, they need to be geographically separated to cover most regions of the US in order to meet with clients and have an impact in appointment/referral management. This is in contrast to AFW2 consultants, as most of their duties can be performed via phone call and email.

Therefore, the RCC staffing system may need to take into account the need for geographical coverage. This may entail lowering the number of clients per RCC in order to accommodate this variable. Nevertheless, the following staffing scenarios will attempt to match the server utilization rate and the quality of care from the AFW2 baseline for simplicity. Hence, based upon 805 total clients there should be approximately 25 servers. When analyzing different scenarios, the best solution(s) should maintain the baseline figures found in Table 12, to ensure quality of care and to avoid wasting resources.

RCC	**# of clients**	**# of clients per consultant**	**c**	**ρ**	**Lq**	**1-Wq(0)**
Baseline	805	40	25	0.81	0.91	0.22

Table 12: Recommended RCC Baseline Statistics

Quality and Efficiency Domain

Having been proven the superior method, the Quality and Efficiency Domain will be used to analyze the RCC consultant staffing scenarios. The following scenarios that will be analyzed represent real-world situations that leadership making staffing decisions has faced from 2010-2012:

- A change in the server workload capacity to analyze affect on server utilization and system MOEs:
    - inferior efficiency
    - improved efficiency.
- An increase in total client load in the system based upon forecasted wounded totals to determine the optimal number of servers to match baseline system congestion.
    - First forecasted year using Holtz-Winters Method, and
    - Through the second forecasted year using Holtz-Winters Method.
- An increase in total client load in the system based upon forecasted wounded totals through the second year, with improved server workload capacity. Attempts to analyze improved workload efficiency with increasing demand to determine the optimal number of servers to match baseline congestion.

Results of RCC Staffing Scenarios

Changing the regulated workload of each server can help to understand the effect on the system. As seen in Table 13, server utilization decreases when servers can assume a higher client workload. If client workload increases and servers are not able to handle 40 clients, the utilization rate will rise and cause a high probability of clients waiting for consultants. This appears to happen when each consultant can only serve. An increase in workload could come from the clients needing more attention, or consultants needing to perform training or additional duties. It is vital for leadership to note this effect on the system. If consultants are able to serve 50 or more clients, given a reduction in workload per client, the system would ultimately see no wait time and an empty queue.

RCC	**# of clients**	**# of clients per consultant**	**c**	**ρ**	**Lq**	**1-Wq(0)**
Change in server workload capacity	805	34	25	0.95	12.77	0.714
	805	36	25	0.89	4.12	0.487
	805	38	25	0.85	1.82	0.328
	805	40	25	0.81	0.91	0.220
	805	42	25	0.77	0.48	0.146
	805	44	25	0.73	0.26	0.097
	805	46	25	0.70	0.15	0.064
	805	48	25	0.67	0.08	0.042
	805	50	25	0.64	0.05	0.027
	805	52	25	0.62	0.03	0.018
	805	54	25	0.60	0.02	0.012

Table 13: RCC Staffing Change in Server Workload Capacity

Similarly to the AFW2 program, leadership needs to plan for a change in future workload. Without knowledge of an explicit departure from system rate, it is assumed negligible on the load for the following scenarios. The injured/ill load will be constant for reasons stated in the methodology. The first forecasted year, with an expected 116 additional USAF wounded troops that will be added to the program, and yields the MOE values seen in Table 14.

RCC	# of clients	# of clients per consultant	c	ρ	Lq	1-Wq(0)
Change in total client load: First forecasted year	921	40	24	0.96	18.41	0.780
	921	40	25	0.92	6.93	0.594
	921	40	26	0.89	3.45	0.445
	921	40	27	0.85	1.90	0.328
	921	40	28	0.82	1.09	0.237
	921	40	29	0.79	0.65	0.167
	921	40	30	0.77	0.38	0.116
	921	40	31	0.74	0.23	0.079
	921	40	32	0.72	0.13	0.052
	921	40	33	0.70	0.08	0.034
	921	40	34	0.68	0.05	0.022

Table 14: RCC Staffing Change in Total Client Load (First Yr Forecast)

Regulation is still met with 24 servers; however, extremely high server utilization will cause poor client satisfaction because of congestion. To maintain the baseline congestion, there should be at least a total of 28-29 consultants with this load increase.

After the second forecasted year with an expected addition of 80 wounded clients, targeted MOE values indicate a need for at least 30-31 RCCs (Table 15).

RCC	# of clients	# of clients per consultant	c	ρ	Lq	1-Wq(0)
Change in total client load: Second forecasted year	1001	40	26	0.96	20.21	0.787
	1001	40	27	0.93	7.69	0.607
	1001	40	28	0.89	3.88	0.461
	1001	40	29	0.86	2.16	0.344
	1001	40	30	0.83	1.27	0.252
	1001	40	31	0.81	0.76	0.181
	1001	40	32	0.78	0.46	0.128
	1001	40	33	0.76	0.28	0.089
	1001	40	34	0.74	0.17	0.060
	1001	40	35	0.72	0.10	0.040
	1001	40	36	0.70	0.06	0.026
	1001	40	37	0.68	0.03	0.017
	1001	40	38	0.66	0.02	0.010

Table 15: RCC Staffing Change in Total Client Load (Second Yr Forecast)

Equation 30 can be used to simplify the analysis and get an approximate result quickly and with minimal computation. Using Equation 28, the workload for the first forecasted year increases by the multiple ‘1.502’. This value is substituted into Equation 27 as $'\nu'$. Recall that the value for β is the (1-α) quantile of the normal distribution. Setting α = 0.1, the (1-0.1) quantile of the normal distribution is 1.2816. Hence, the appropriate number of servers for the system is 29 consultants (Equation 36). This answer coincides with the answer found from Table 14, when baseline MOE values were maintained.

$$c \approx 1\left(\frac{574}{40}\right) + 1.502\left(\frac{231}{40}\right) + 1.2816\sqrt{1\left(\frac{574}{40}\right) + 1.502\left(\frac{231}{40}\right)} \approx 29.17 \sim 29\ consultants$$

Equation 36

Following the same procedure to forecast into the second year, the approximation formula yields a total of 36 consultants (Equation 37), which coincides with the answer found from Table 15, when baseline MOE values were maintained.

$$c \approx 1\left(\frac{574}{40}\right) + 1.848\left(\frac{231}{40}\right) + 1.2816\sqrt{1\left(\frac{574}{40}\right) + 1.848\left(\frac{231}{40}\right)} \approx 31.43 \sim 31\ consultants$$

Equation 37

It is understandable that the main driver to RCC staffing will be the injured/ill clientele. The baseline scenario exhibits 2½ times as many injured/ill clients to wounded clients. Hence, a change in the injured/ill clientele demand would have a greater effect on RCC staffing than a change in wounded clientele. As the scenarios progress into the second forecasted year, injured/ill clientele remains constant, and the ratio reduces to $1\,{}^{1}\!/_{3}$ times as many wounded clients. Hence, the wounded demand becomes a bigger driver to RCC staffing. The approximation formula can easily be modified to analyze any arbitrary load change in both types of clientele seen from the forecast. This staffing technique can improve efficiency at WRCs, yet treatment planning can have an even larger impact on system efficiency solutions.

Reliability Theory Analysis

At the Cat 1 TBI center at BAMC, they provide treatment five main treatments via speech/language pathology, occupational therapy, psychology, physical therapy, and rehab nursing. In Figure 25 there are five hypothetical treatments. Each treatment has a different reliability (probability of effectiveness in relieving symptoms). In house studies should be done to determine actual treatment reliability values. Changing this parallel treatment protocol could improve efficiency to allow better quality in a budget freeze and/or allow more clients to receive treatment. Without specific treatment efficacy known, reliability theory used in different scenarios can help understand possible trade-off solutions.

For the following scenarios, it is assumed that the treatment effectiveness rates are independent probabilities. It is understood that this may not be the case in the real world and that some treatments may lessen the effectiveness or synergize with others. However, the effect on total system reliability is still assumed negligible. Based upon several scenarios, it will be demonstrated how much removing certain treatments could reduce cost to the care facilities without losing significant system performance.

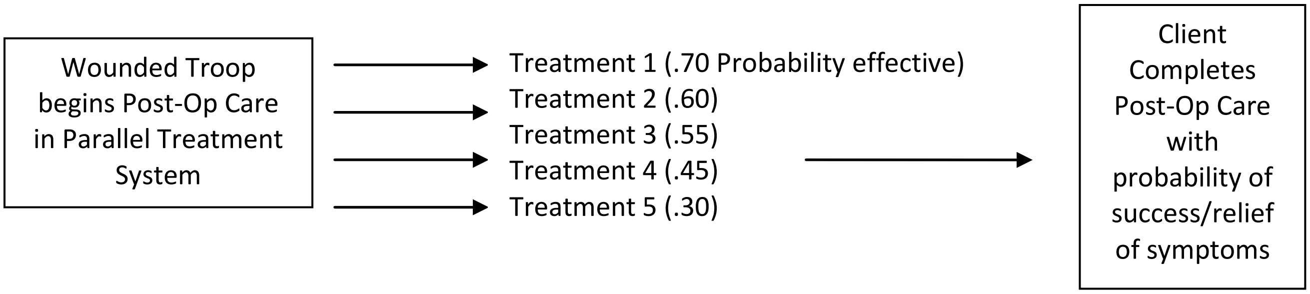


Figure 25: Non-Equal Treatment Reliability System Scenario

Utilizing Equation 10, the reliability of this system scenario is 0.979 (calculated below).

$$r(p) = 1 - \prod_{i=1}^{n}(1 - p_i) = 1 - (1 - 0.7)(1 - 0.6)(1 - 0.55)(1 - 0.45)(1 - 0.3) = \mathbf{0.979}$$

Manipulating Equation 10, one treatment is removed from the system yielding new system reliabilities (Table 16).

Treatment Removed	System Reliability r(p)	Effectiveness Loss
1	0.931	0.931/0.979= **4.9%**
2	0.948	**3.2%**
3	0.954	**2.6%**
4	0.962	**1.7%**
5	0.970	**0.9%**

Table 16: Non-Equal System Reliability Analysis

As seen in Table 16, even if the most successful treatment (Treatment #1) is removed from the system, the reliability only drops by approximately 4.9%, still yielding total system reliability above 0.93. If the least successful treatment (Treatment #5) was removed, the total system reliability drops less than 1%.

While each treatment costs in clinic space, equipment, salary paid to technicians, and the client's time; removing one treatment may be a viable option for management. If 100% of the monetary and time cost is variable (treatment) cost, removing one of these treatments could cut costs by 20% or more, while only reducing system effectiveness by 1-5%. Even in the case that there are 50% fixed costs (administrative) and 50% variable costs, there could be 10% or more cost reduction that yields improved efficiency. That could free 10% of technician's time to serve new clients, or free 10% more budget to enhance current rehab equipment.

To further analyze the effect of removing treatments in a parallel system, four other scenarios are investigated:

1. All treatments are highly successful where $p_i = 0.75$,
2. All treatments are fairly successful where $p_i = 0.5$,
3. All treatments are somewhat successful where $p_i = 0.25$,
4. All treatments are minimally successful where $p_i = 0.1$.

The total system reliability for each scenario is seen in Table 17.

Scenario	System Reliability r(p)
1	0.999
2	0.969
3	0.763
4	0.410

Table 17: Equal Treatment Reliability System Scenarios

If the probability of success for every treatment is greater or equal to 0.5, the system reliability with five parallel treatments will be approximately 0.97 or greater. Even when all treatments maintain a 0.25 effectiveness rate, the total system reliability is greater than 0.76.

Results of System of Treatments Reliability Scenarios

The resulting system reliability for possible treatment protocol changes for each scenario is seen in Table 18.

Scenario	System Reliability r(p) (all 5 treatments)	System Reliability r(p) (1 treatment removed)	System Reliability r(p) (2 treatments removed)	System Reliability r(p) (3 treatments removed)
1	0.999	0.996	0.984	0.938
2	0.969	0.938	0.875	0.750
3	0.763	0.684	0.578	0.438
4	0.410	0.344	0.271	0.190

Table 18: Equal System Reliability Analysis

In Scenario 1, removing up to three treatments will have little impact on the treatment success, still yielding greater than 0.93 system reliability. Scenario 2 displays a minimal drop in performance when losing one to two treatments, and still produces system reliability of 0.75 when only two treatments remain. When analyzing the tradeoff analysis for this scenario, recall that removing one treatment from the protocol yields about a 20% reduction in cost/time in a 100% variable cost system, and a 10% reduction in cost/time in a 50% fixed/50% variable cost system, with the treatment protocol's loss in effectiveness being a meager 3%. Removing two treatments from the protocol would double the cost/time savings, while reducing protocol effectiveness by only 9%. Removing three treatments would triple the reduction in time/cost, while reducing protocol effectiveness by 22%. Hence, cases could be made to change treatment protocol in this treatment efficacy scenario based upon the benefits received in cost and time reduction in order to use those resources more efficiently.

Scenario 3 and 4 are more significantly impacted when losing treatments. Consequently, significance of removing treatments increases as the individual treatment probabilities of effectiveness decreases. The magnitude of cost/time reduction may/may not outweigh the magnitude of performance degradation. Therefore, changing treatment protocol in a parallel system will be most effective with a more reliable system such as scenario 1 and 2.

Management can use tables like this to weigh system performance they are willing to give up versus the benefits of cost and time reduction. Coupled with queueing theory to make staffing changes based upon load changes, these methods provide incredibly effective cost/system management analysis for WRCs.

V. Conclusions and Recommendations

Conclusions of Research

This paper presents a Holtz-Winters Method of forecasting to predict clientele demand changes in US Air Force Warrior and Survivor Care. Current methods rely solely upon smoothing methods on the past levels of clientele. They do not forecast the expected number of new clients that will need treatment, nor does the smoothing method account for trend or seasonality, which has been demonstrated to be analytically naïve. Forecasting future clientele is necessary for effective resource planning. Utilizing the forecasted demand of future casualties, the staffing of recovery care consultants and AF Wounded Warrior consultants is modeled via queueing theory. Numerous SMEs reported recent staffing variability that caused hardship on consultants and clientele. Quality and efficiency methods in this paper provide procedures to minimize staffing variability and that adapt to changes in demand. They improve cost effectiveness while providing a continual high quality service. The described decision process occurs constantly in USAF/A1SZ programs, and has become increasingly important due to the recent DoD constrained budget environment and the plethora of clients waiting that could benefit from treatment. It also provides leadership with mathematical methods to discover understaffing and how to adjust efficiently. Due to the same constraints, Warfighter Rehabilitation Centers, such as the Traumatic Brain Injury Center, also requires quality and efficiency analysis and could staff based upon this queuing theory. Through stochastic reliability theory, this paper also implies possible cost/time saving methods through treatment protocol changes in a parallel system that retains a high total system success rate. Increasing the rate of service will reduce the cost of a server's and patient's time, and reduce the server utilization rate. A streamlined treatment protocol will minimize waste. It can provide leadership with mathematical reasoning behind decisions to conform to a fixed or reduced budget. The analytical methods could also reduce the number of servers required, and/or allow more clients waiting to enter the system and receive treatment. These Operations Research techniques provide budget solutions via abundant cost reduction (in time and monetary value) without a significant reduction in performance.

Recommendations for Action

Research questions were provided viable and executable options in the analysis section. Current staff planning, based upon a basic moving average of past clientele, constrains the ability to effectively staff for the future. Forecasting future casualties based upon trend and seasonal factors in wounded troop totals remedy the complexity of predicting future workload. Current fighting in the Middle East displays seasonal attack and correlated wounding patterns. Trend changes can be seen through operational aggressiveness in deployed troop totals. The method required to forecast casualties should accommodate such trend and seasonality.

The analysis section provided wounded troop forecast examples during different war scenarios from 2010-2012. Forecasting analysis should be done often and will be the most accurate for a few short periods following the most recent actual data. Therefore, staffing based upon workload forecasting is recommended to be updated bi-annually or even quarterly. Forecasting in wartime and drawdown scenarios minimize RCC and AFW2 staffing variation. Both consultant positions require extensive training. Temporary workers and rapid staffing changes are detrimental to the process, as well as client satisfaction. Hence, forecasting in order to plan consultant staffing ensures a steady system for consultants and clients. The most recent data should be applied in these methods to configure their respective staffing figures. The internal organization should research their client departure rates to determine if they need to account for clientele leaving the system, as this was assumed negligible in this thesis. This will ensure proper service rates of clientele in the system, which helps the accuracy in consultant staffing analysis.

It was demonstrated in the analysis section that the appropriate increase in consultants does not have to be of the same magnitude of increased workload demand. To maintain an efficient and high quality system that is analogous to the baseline, the suitable number of servers can be estimated based upon the square root of the modified workload, via the Quality and Efficiency Domain. Consequently, the increase in cost for hiring new consultants is less significant on the organization, than when attempting to match staff totals to the expected change in workload.

This staffing technique should also be used at the WRC to maintain efficiency and quality of their treatment providers. When treatment protocol changes are required due to budget constraints, the cost reduction in the server's and patient's time, along with any monetary cost reduction may outweigh the loss in treatment plan efficacy. It was shown that removing one of five treatments in a parallel system may cut at least 20% of the cost in space, equipment, salary paid to technicians, and the client's time. This cutback only leads to approximately a 1-5% reduction in total system performance. Reliability of the system will most likely be the highest when all treatments are included. It is recommended whenever possible to retain all treatments in a parallel system to do so. However, reliability theory provides possible alternative optimal solutions to accommodate budget freeze/reduction and allow more clients to receive treatment.

Recommendations for Future Research

I recommend further research into refining the arrival and departure rates accuracy for clientele in the USAF/A1SZ programs. The WRC should analyze the performance of treatments in their system, and perform cost/trade-off analysis upon each treatment. Researchers could use the IMPACT Database (Marmarou, et al., 2007) statistics or results from TBI rehab research in fields such as hockey, football, and car accidents to analyze the performance of experimental rehab techniques.

These experiment techniques may provide better performance efficacy than current treatments in the protocol. Hyperbaric Oxygen Therapy (HOT) has gained recent media attention and has shown impressive symptom improvement in patients with TBI (Harch, et al., 2012). Cranial Sacral Therapy is an experimental treatment that gave the most relief to me during my recovery process. Both HOT and Cranial Sacral Therapy are not part of the current protocol. Yet, I believe a parallel system with solely these two treatments could perform better than the current system of five treatments. Hence, further investigation into one or two of the most promising treatments may be able to provide a system that is superior in service time, cost, and performance.

The analytical improvements via Operations Research techniques to the USAF/A1SZ and WRC systems outlined in this thesis can provide immediate time and cost saving impact with improved quality and efficiency over the long term. As the author, I am a wounded client in the system, and the last thing that I want is to minimize the resources available for recovery in efforts to simply save money. Dr. James Bales, AFW2 Adaptive Sports Program (2013), indicated that there are over 6,000 USAF WII veterans from current and previous military conflicts that the AFW2 program is trying to reach out to for participation. Without improving the efficiency of the recovery care process as outlined in this thesis, many of these individuals will not benefit from the great service provided. The recommended solutions and future research investigations in this thesis could help re-write the future of TBI and casualty recovery care.

VI. Appendix: PowerPoint Thesis Defense Slides & Quad Chart

Air Force Institute of Technology

The AFIT of Today is the Air Force of Tomorrow.

Traumatic Brain Injury Recovery Care: Demand Forecasting, Staffing, and Treatment Planning

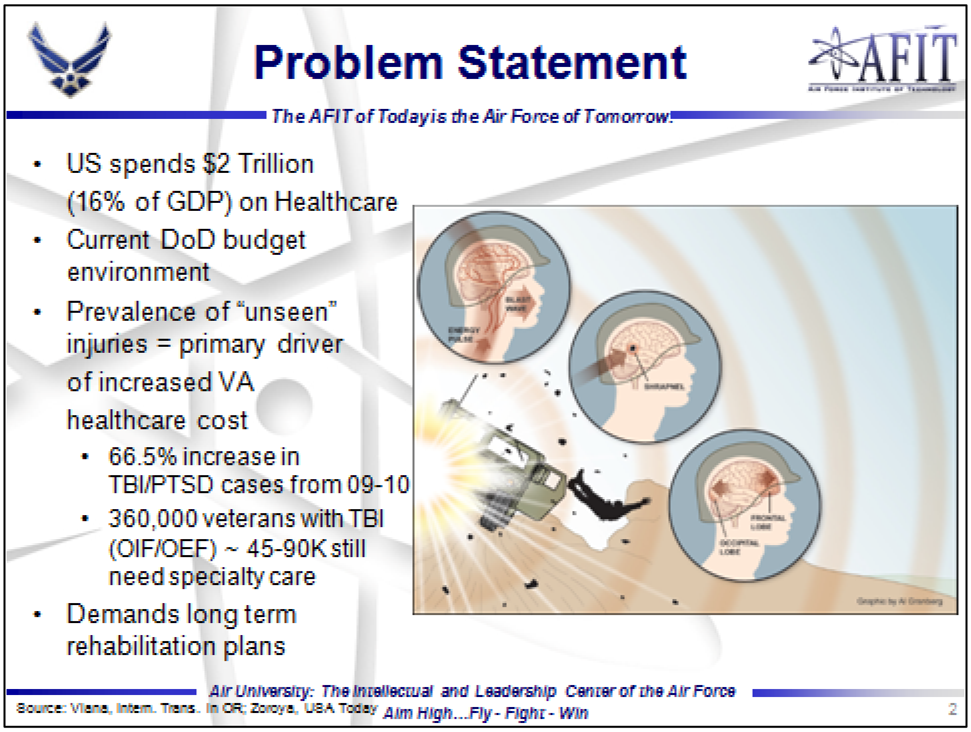
Mitchell Kieffer

CPT, USAF
ENS GOR 13M

8 March 2013

THESIS DEFENSE

Air University: The Intellectual and Leadership Center of the Air Force
Aim High…Fly - Fight - Win



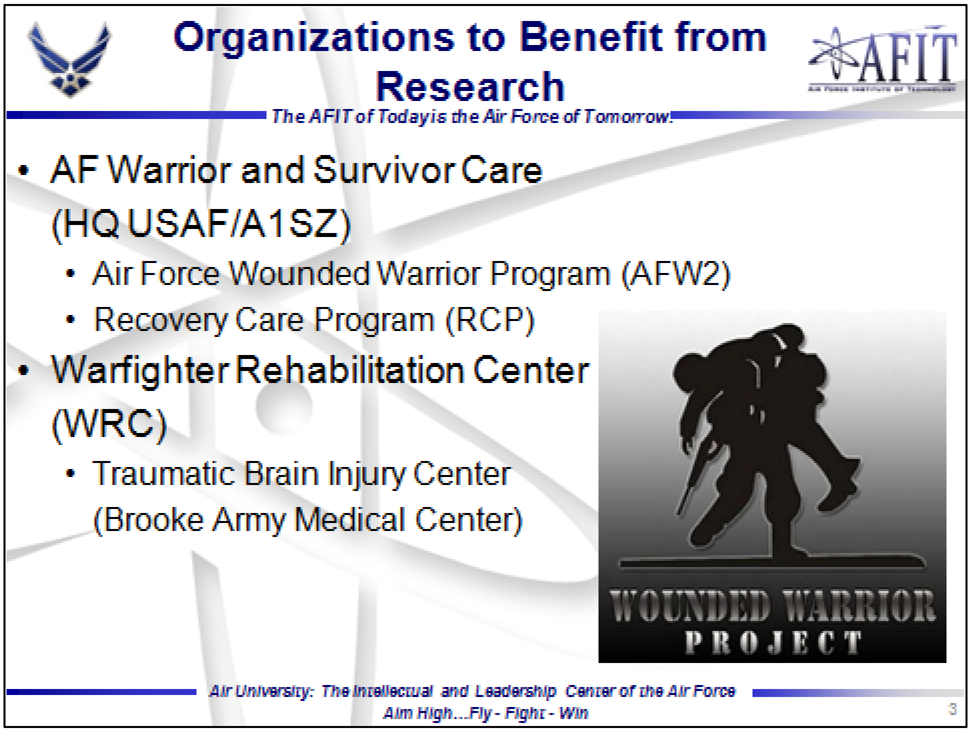
Problem Statement

The AFIT of Today is the Air Force of Tomorrow.

- US spends $2 Trillion (16% of GDP) on Healthcare
- Current DoD budget environment
- Prevalence of "unseen" injuries = primary driver of increased VA healthcare cost
  - 66.5% increase in TBI/PTSD cases from 09-10
  - 360,000 veterans with TBI (OIF/OEF) ~ 45-90K still need specialty care
- Demands long term rehabilitation plans

Air University: The Intellectual and Leadership Center of the Air Force

Source: Viana, Intern. Trans. in OR; Zoroya, USA Today *Aim High…Fly - Fight - Win*



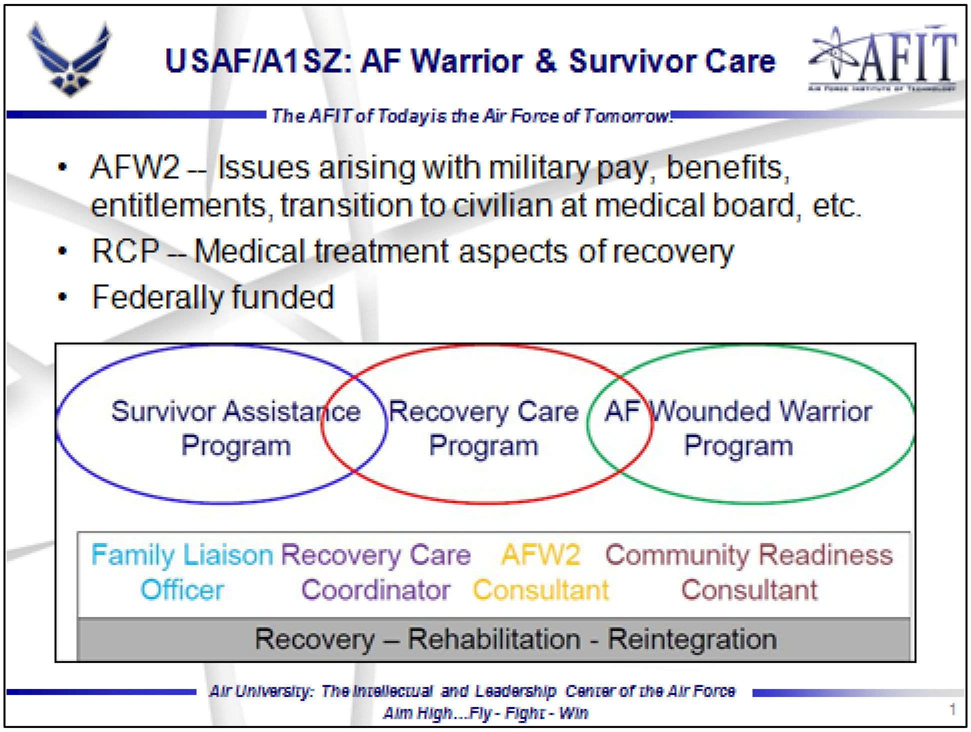
Organizations to Benefit from Research

The AFIT of Today is the Air Force of Tomorrow!

- AF Warrior and Survivor Care (HQ USAF/A1SZ)
  - Air Force Wounded Warrior Program (AFW2)
  - Recovery Care Program (RCP)
- Warfighter Rehabilitation Center (WRC)
  - Traumatic Brain Injury Center (Brooke Army Medical Center)

Air University: The Intellectual and Leadership Center of the Air Force
Aim High...Fly - Fight - Win

3



USAF/A1SZ: AF Warrior & Survivor Care

The AFIT of Today is the Air Force of Tomorrow!

- AFW2 -- Issues arising with military pay, benefits, entitlements, transition to civilian at medical board, etc.
- RCP -- Medical treatment aspects of recovery
- Federally funded

Survivor Assistance Program

Recovery Care Program

AF Wounded Warrior Program

Family Liaison Officer

Recovery Care Coordinator

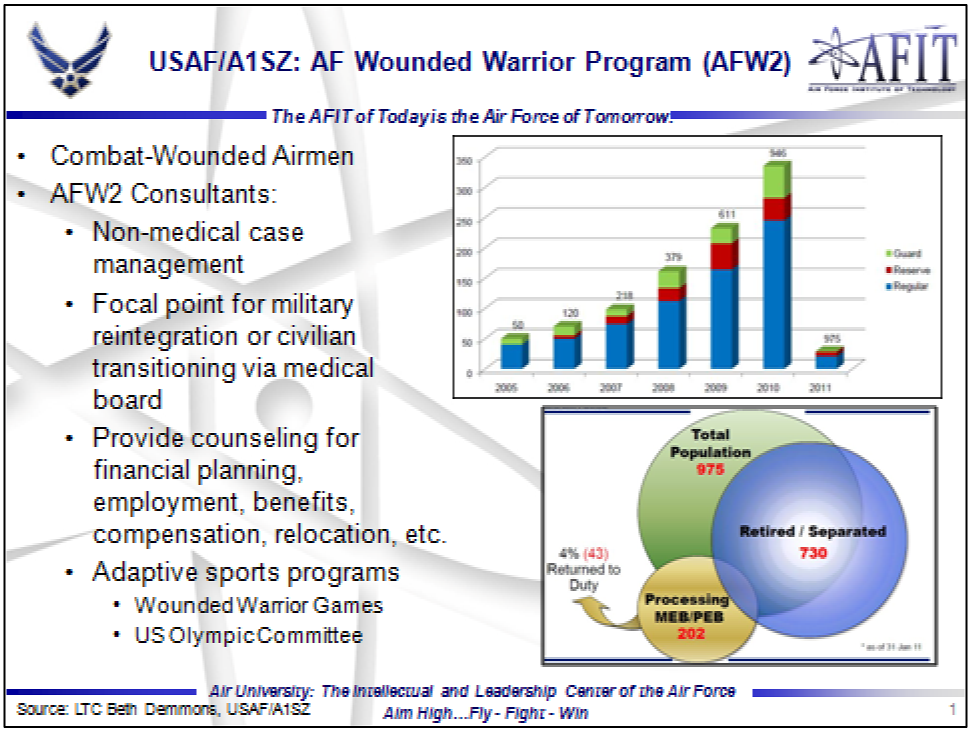
AFW2 Consultant

Community Readiness Consultant

Recovery – Rehabilitation - Reintegration

Air University: The Intellectual and Leadership Center of the Air Force
Aim High...Fly - Fight - Win

1



USAF/A1SZ: AF Wounded Warrior Program (AFW2)

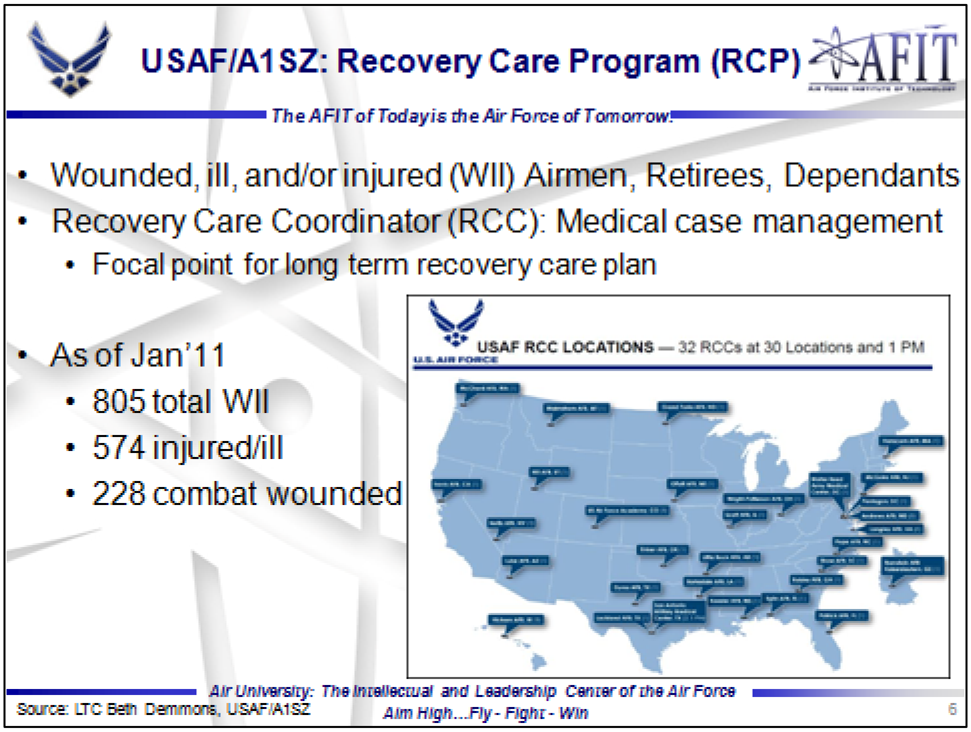
The AFIT of Today is the Air Force of Tomorrow.

- Combat-Wounded Airmen
- AFW2 Consultants:
  - Non-medical case management
  - Focal point for military reintegration or civilian transitioning via medical board
  - Provide counseling for financial planning, employment, benefits, compensation, relocation, etc.
  - Adaptive sports programs
    - Wounded Warrior Games
    - US Olympic Committee

Air University: The Intellectual and Leadership Center of the Air Force

Source: LTC Beth Demmons, USAF/A1SZ

Aim High…Fly - Fight - Win



USAF/A1SZ: Recovery Care Program (RCP)

The AFIT of Today is the Air Force of Tomorrow.

- Wounded, ill, and/or injured (WII) Airmen, Retirees, Dependants
- Recovery Care Coordinator (RCC): Medical case management
  - Focal point for long term recovery care plan
- As of Jan'11
  - 805 total WII
  - 574 injured/ill
  - 228 combat wounded

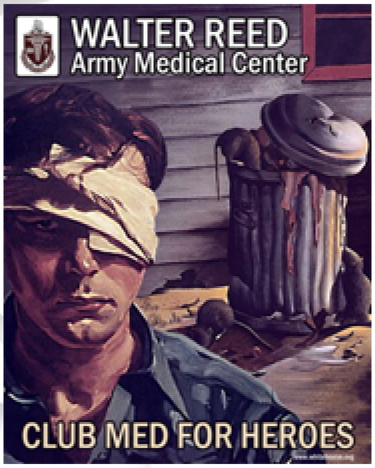
Air University: The Intellectual and Leadership Center of the Air Force

Source: LTC Beth Demmons, USAF/A1SZ

Aim High…Fly - Fight - Win

Warfighter Rehabilitation Center

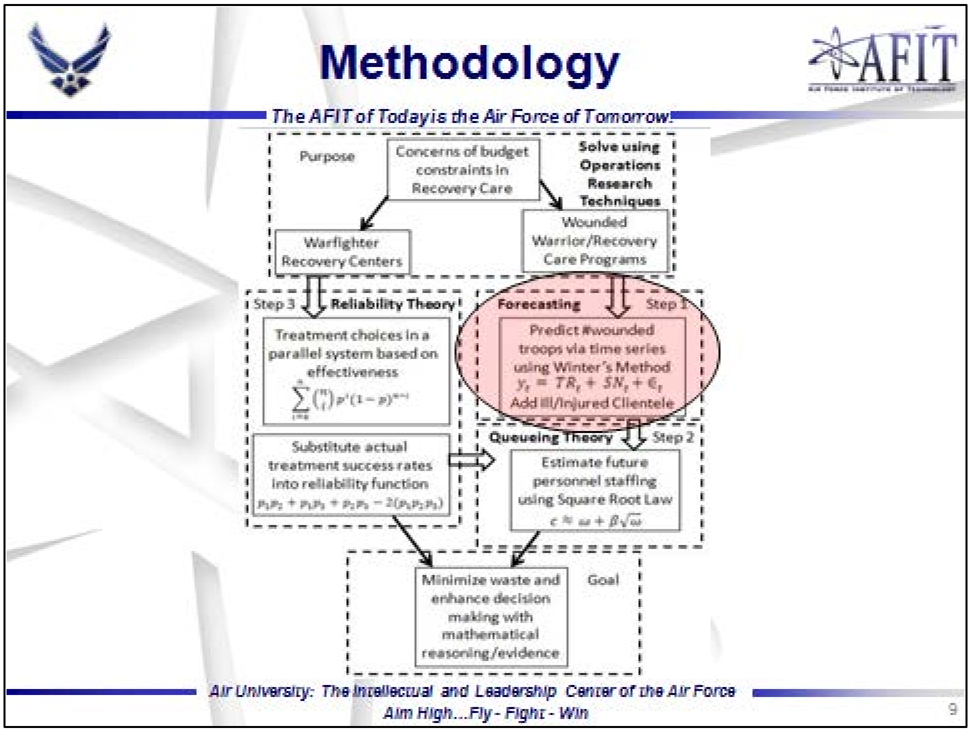
- Walter Reed Army Medical Center (Washington D.C.)
- Brooke Army Medical Center (San Antonio, TX)
- Specialty clinic rehab
  - Often inpatient or outpatient TDY
  - Category I: Services to full spectrum of severity
  - Goal: Transition to home base medical facility
- Budgeted under special funds from Office of the Surgeon General



Source: Michelle Lindsey, TBI Center

Research Questions

- What constraints are present in current staffing techniques and how can they be remedied?
- What variability is there in staffing and the number of WII needing treatment? Is there variability based upon seasons and/or milestones in war?
- How many RCC/AFW2 consultants are needed given workload thresholds using different forecasting scenarios?
- What techniques help to improve staffing efficiency while maintaining quality?
- What techniques can be used to maintain reliability in treatment system effectiveness when constraints limit the treatment options at WRC's?



Methodology

The AFIT of Today is the Air Force of Tomorrow!

Purpose

Concerns of budget constraints in Recovery Care

Warfighter Recovery Centers

Solve using Operations Research Techniques

Wounded Warrior/Recovery Care Programs

Step 3 **Reliability Theory**

Treatment choices in a parallel system based on effectiveness

$$\sum_{i=k}^{n}\binom{n}{i}p^i(1-p)^{n-i}$$

Substitute actual treatment success rates into reliability function

$$p_1p_2 + p_1p_3 + p_2p_3 - 2(p_1p_2p_3)$$

Forecasting Step 1

Predict #wounded troops via time series using Winter's Method

$$y_t = TR_t + SN_t + \epsilon_t$$

Add Ill/Injured Clientele

Queueing Theory Step 2

Estimate future personnel staffing using Square Root Law

$$c \approx \omega + \beta\sqrt{\omega}$$

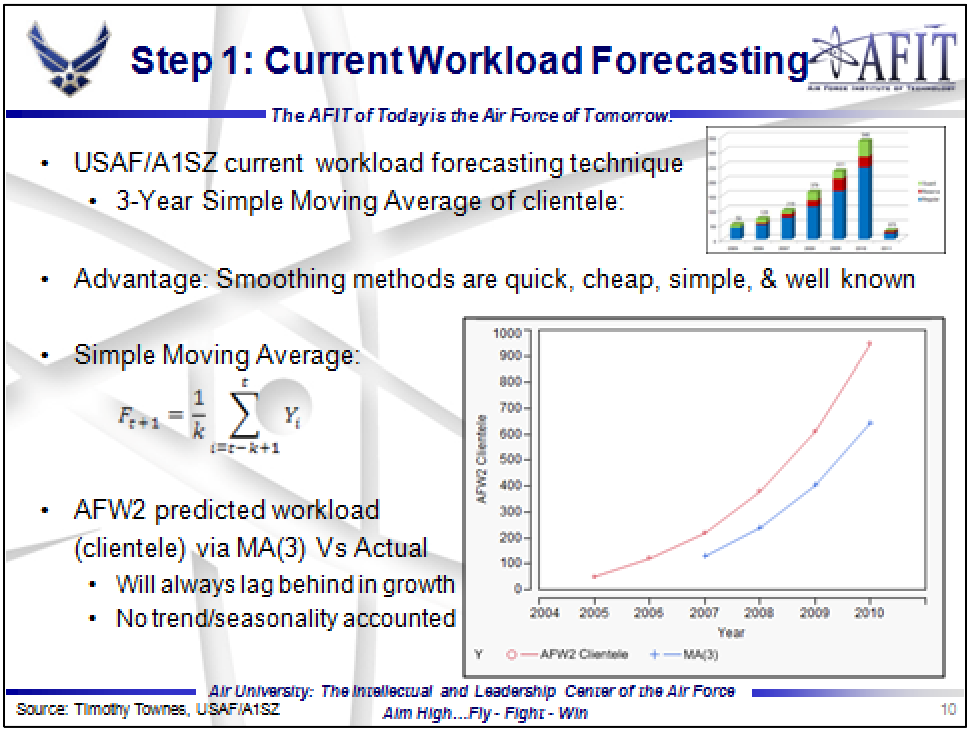
Minimize waste and enhance decision making with mathematical reasoning/evidence

Goal

Air University: The Intellectual and Leadership Center of the Air Force

Aim High…Fly - Fight - Win

9



Step 1: Current Workload Forecasting

The AFIT of Today is the Air Force of Tomorrow!

- USAF/A1SZ current workload forecasting technique
  - 3-Year Simple Moving Average of clientele:
- Advantage: Smoothing methods are quick, cheap, simple, & well known
- Simple Moving Average:

$$F_{t+1} = \frac{1}{k}\sum_{i=t-k+1}^{t} Y_i$$

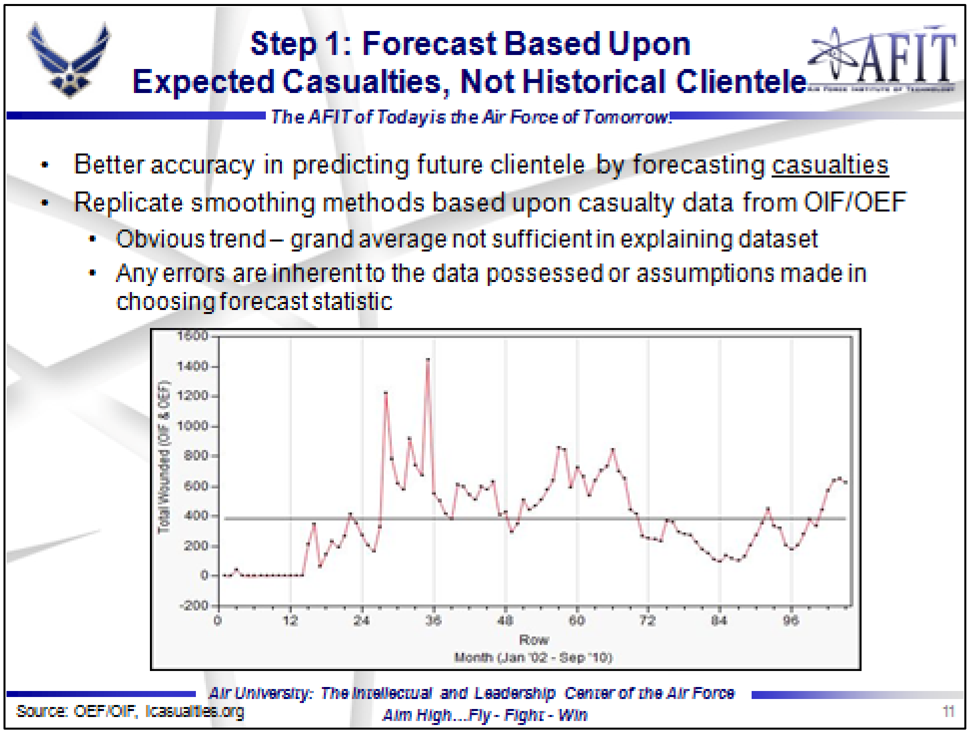
- AFW2 predicted workload (clientele) via MA(3) Vs Actual
  - Will always lag behind in growth
  - No trend/seasonality accounted

Source: Timothy Townes, USAF/A1SZ

Air University: The Intellectual and Leadership Center of the Air Force

Aim High…Fly - Fight - Win

10



Step 1: Forecast Based Upon Expected Casualties, Not Historical Clientele

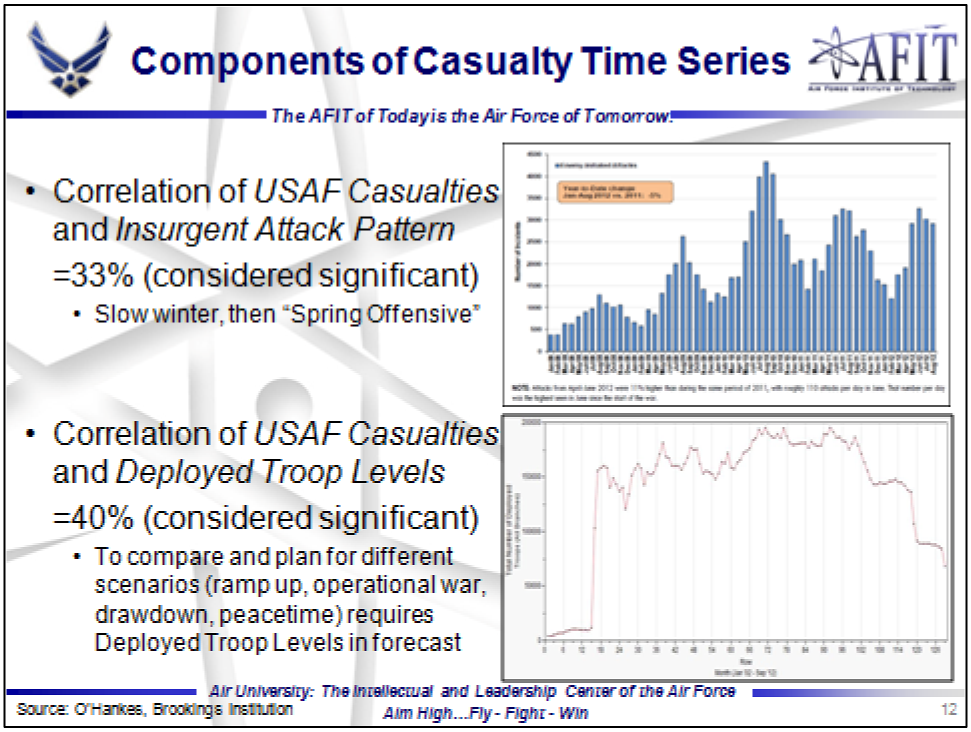
The AFIT of Today is the Air Force of Tomorrow!

- Better accuracy in predicting future clientele by forecasting <u>casualties</u>
- Replicate smoothing methods based upon casualty data from OIF/OEF
  - Obvious trend – grand average not sufficient in explaining dataset
  - Any errors are inherent to the data possessed or assumptions made in choosing forecast statistic

Air University: The Intellectual and Leadership Center of the Air Force

Source: OEF/OIF, Icasualties.org

Aim High…Fly - Fight - Win



Components of Casualty Time Series

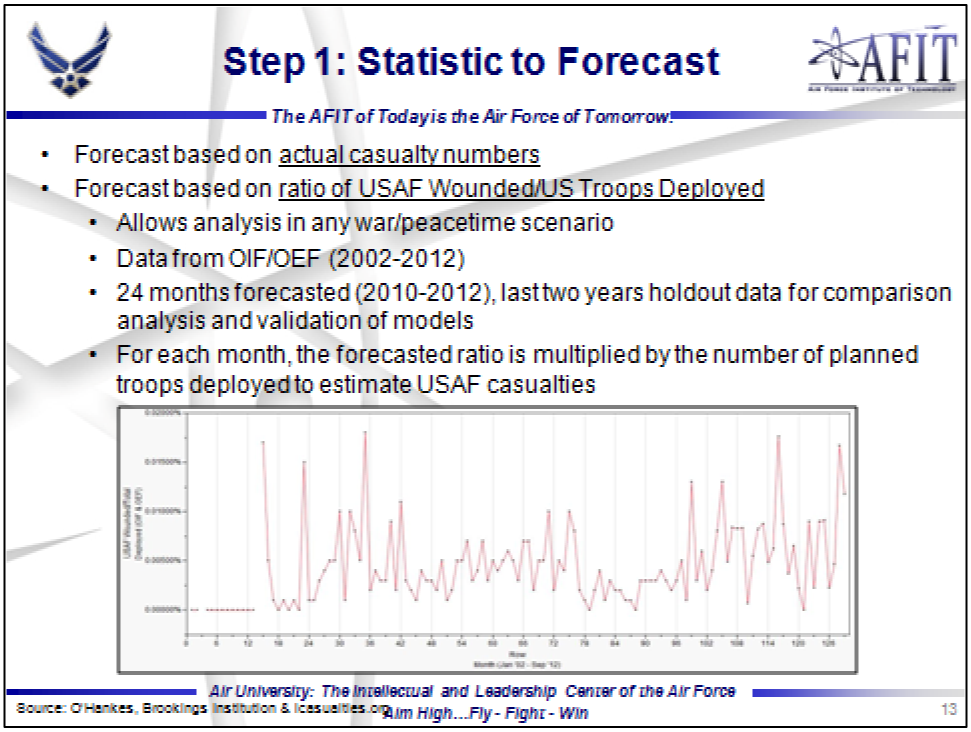
The AFIT of Today is the Air Force of Tomorrow!

- Correlation of *USAF Casualties* and *Insurgent Attack Pattern* =33% (considered significant)
  - Slow winter, then "Spring Offensive"
- Correlation of *USAF Casualties* and *Deployed Troop Levels* =40% (considered significant)
  - To compare and plan for different scenarios (ramp up, operational war, drawdown, peacetime) requires Deployed Troop Levels in forecast

Air University: The Intellectual and Leadership Center of the Air Force

Source: O'Hankes, Brookings Institution

Aim High…Fly - Fight - Win



Step 1: Statistic to Forecast

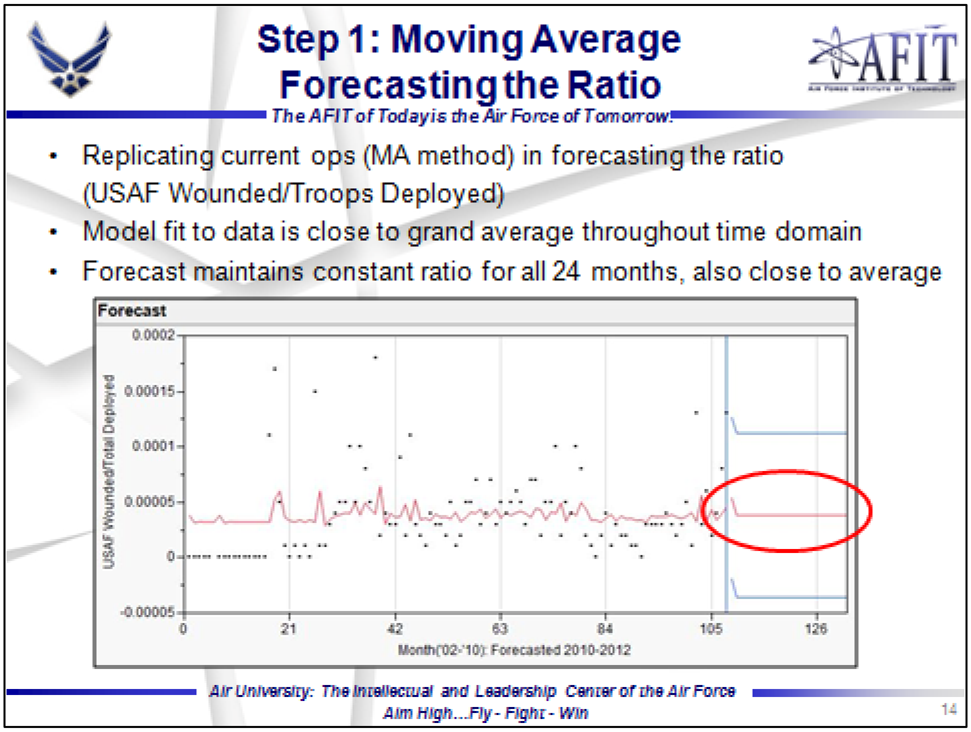
The AFIT of Today is the Air Force of Tomorrow!

- Forecast based on actual casualty numbers
- Forecast based on ratio of USAF Wounded/US Troops Deployed
  - Allows analysis in any war/peacetime scenario
  - Data from OIF/OEF (2002-2012)
  - 24 months forecasted (2010-2012), last two years holdout data for comparison analysis and validation of models
  - For each month, the forecasted ratio is multiplied by the number of planned troops deployed to estimate USAF casualties

Air University: The Intellectual and Leadership Center of the Air Force

Source: O'Hankes, Brookings Institution & icasualties.org

Aim High…Fly - Fight - Win



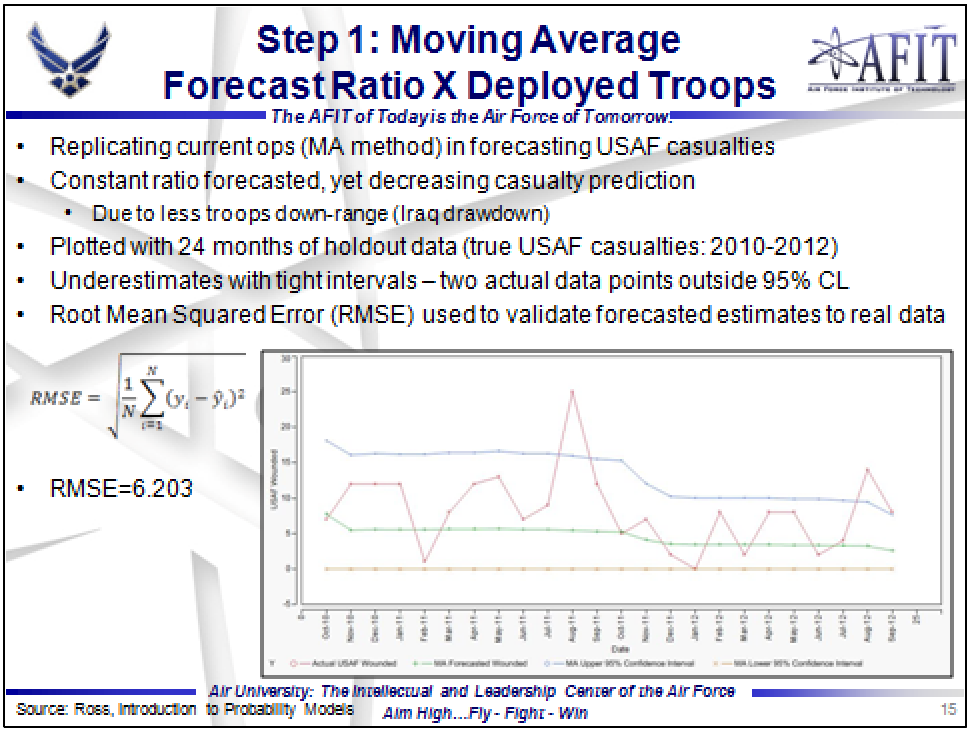
Step 1: Moving Average Forecasting the Ratio

The AFIT of Today is the Air Force of Tomorrow!

- Replicating current ops (MA method) in forecasting the ratio (USAF Wounded/Troops Deployed)
- Model fit to data is close to grand average throughout time domain
- Forecast maintains constant ratio for all 24 months, also close to average

Air University: The Intellectual and Leadership Center of the Air Force

Aim High…Fly - Fight - Win



Step 1: Moving Average Forecast Ratio X Deployed Troops

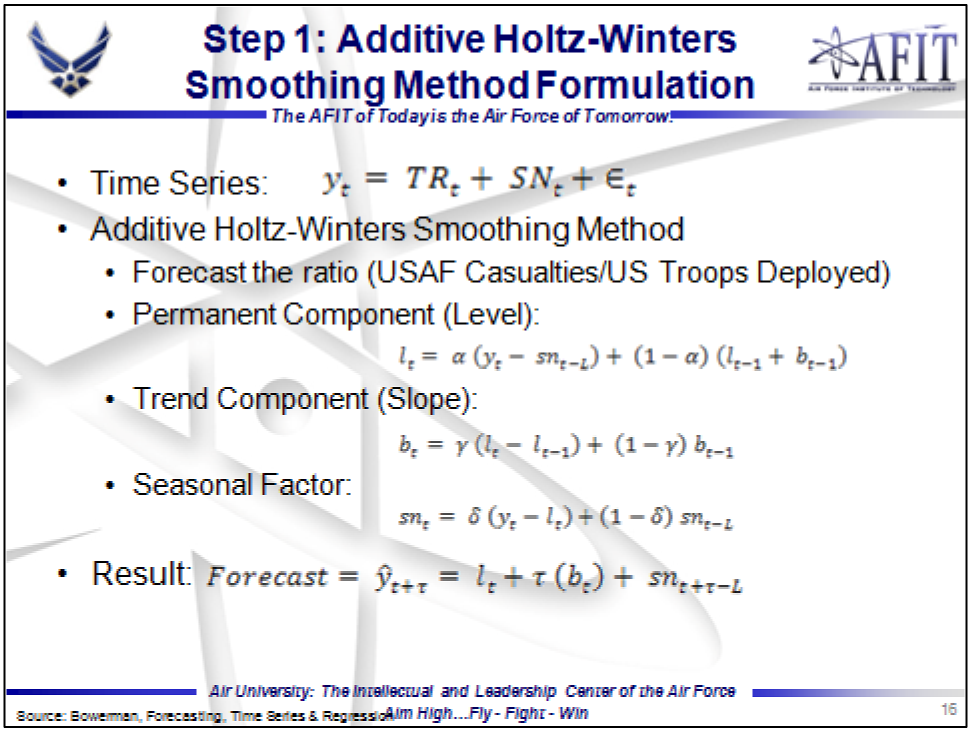
The AFIT of Today is the Air Force of Tomorrow!

- Replicating current ops (MA method) in forecasting USAF casualties
- Constant ratio forecasted, yet decreasing casualty prediction
  - Due to less troops down-range (Iraq drawdown)
- Plotted with 24 months of holdout data (true USAF casualties: 2010-2012)
- Underestimates with tight intervals – two actual data points outside 95% CL
- Root Mean Squared Error (RMSE) used to validate forecasted estimates to real data

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2}$$

- RMSE=6.203

Source: Ross, Introduction to Probability Models



Step 1: Additive Holtz-Winters Smoothing Method Formulation

The AFIT of Today is the Air Force of Tomorrow!

- Time Series: $y_t = TR_t + SN_t + \epsilon_t$
- Additive Holtz-Winters Smoothing Method
  - Forecast the ratio (USAF Casualties/US Troops Deployed)
  - Permanent Component (Level):

    $$l_t = \alpha\,(y_t - sn_{t-L}) + (1-\alpha)\,(l_{t-1} + b_{t-1})$$

  - Trend Component (Slope):

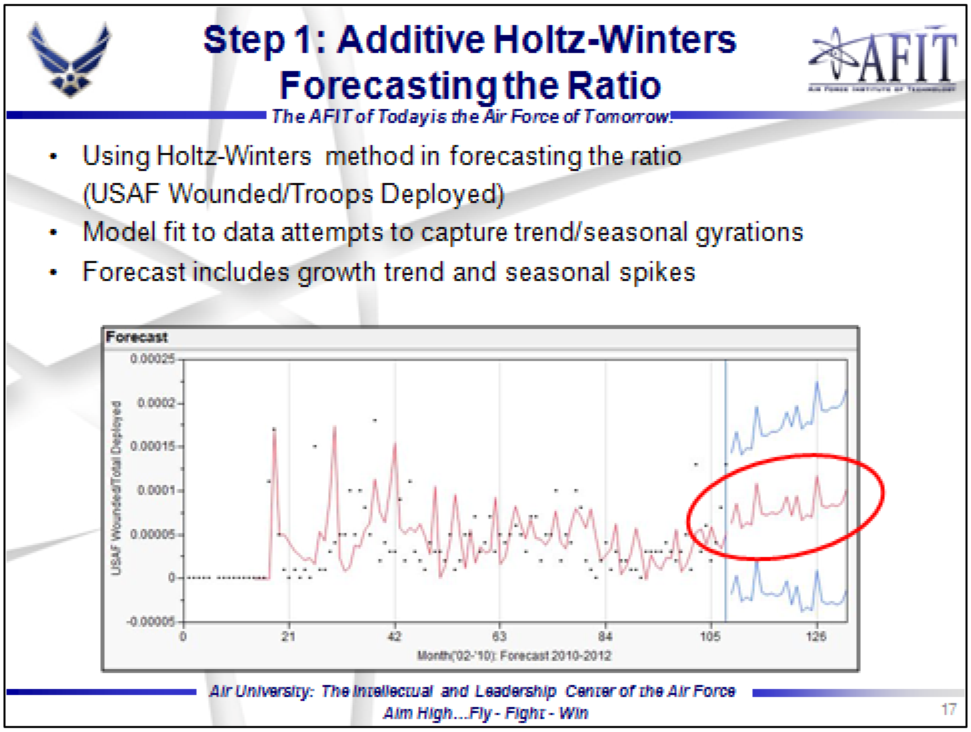
    $$b_t = \gamma\,(l_t - l_{t-1}) + (1-\gamma)\,b_{t-1}$$

  - Seasonal Factor:

    $$sn_t = \delta\,(y_t - l_t) + (1-\delta)\,sn_{t-L}$$

- Result: $Forecast = \hat{y}_{t+\tau} = l_t + \tau\,(b_t) + sn_{t+\tau-L}$

Source: Bowerman, Forecasting, Time Series & Regressio

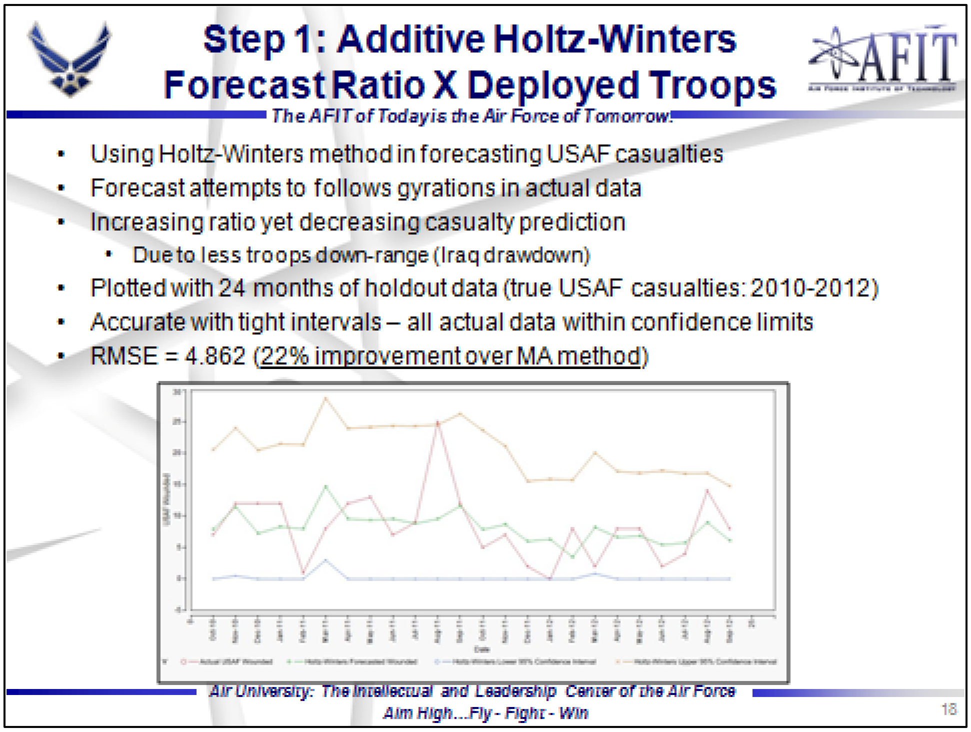


Step 1: Additive Holtz-Winters Forecasting the Ratio

The AFIT of Today is the Air Force of Tomorrow!

- Using Holtz-Winters method in forecasting the ratio (USAF Wounded/Troops Deployed)
- Model fit to data attempts to capture trend/seasonal gyrations
- Forecast includes growth trend and seasonal spikes

Air University: The Intellectual and Leadership Center of the Air Force
Aim High…Fly - Fight - Win

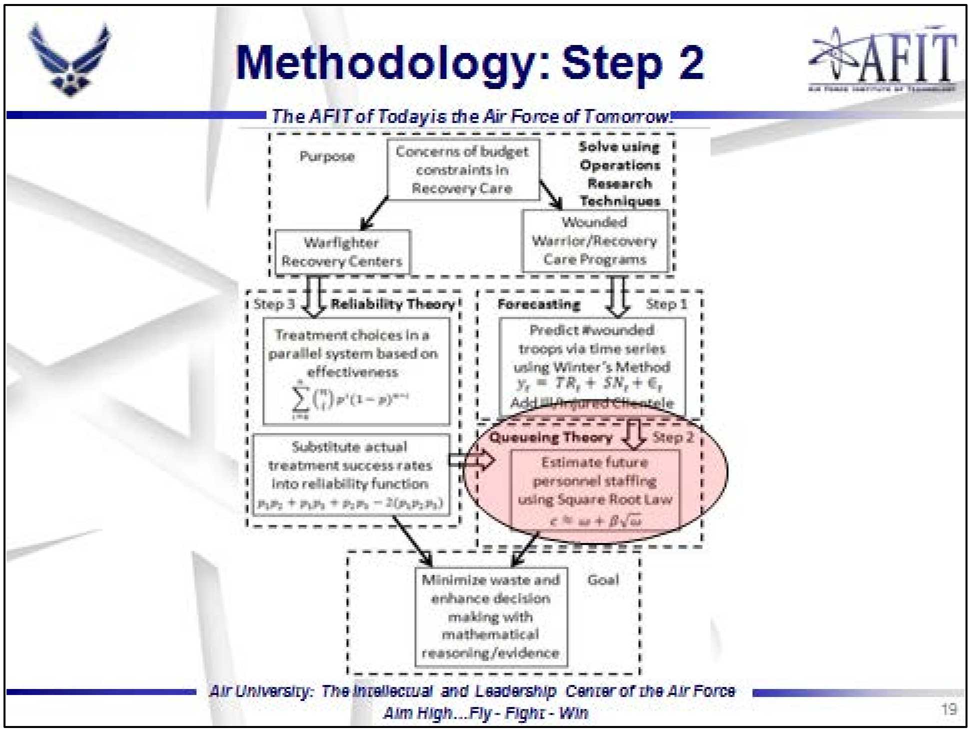


Step 1: Additive Holtz-Winters Forecast Ratio X Deployed Troops

The AFIT of Today is the Air Force of Tomorrow!

- Using Holtz-Winters method in forecasting USAF casualties
- Forecast attempts to follows gyrations in actual data
- Increasing ratio yet decreasing casualty prediction
  - Due to less troops down-range (Iraq drawdown)
- Plotted with 24 months of holdout data (true USAF casualties: 2010-2012)
- Accurate with tight intervals – all actual data within confidence limits
- RMSE = 4.862 (22% improvement over MA method)

Air University: The Intellectual and Leadership Center of the Air Force
Aim High…Fly - Fight - Win



Methodology: Step 2

Purpose

Concerns of budget constraints in Recovery Care

Solve using Operations Research Techniques

Warfighter Recovery Centers

Wounded Warrior/Recovery Care Programs

Step 3 **Reliability Theory**

Treatment choices in a parallel system based on effectiveness

$\sum_{i=k}^{n}\binom{n}{i}p^i(1-p)^{n-i}$

Substitute actual treatment success rates into reliability function

$p_1p_2+p_1p_3+p_2p_3-2(p_1p_2p_3)$

Forecasting Step 1

Predict #wounded troops via time series using Winter's Method

$y_t = TR_t + SN_t + \epsilon_t$

Add Ill/Injured Clientele

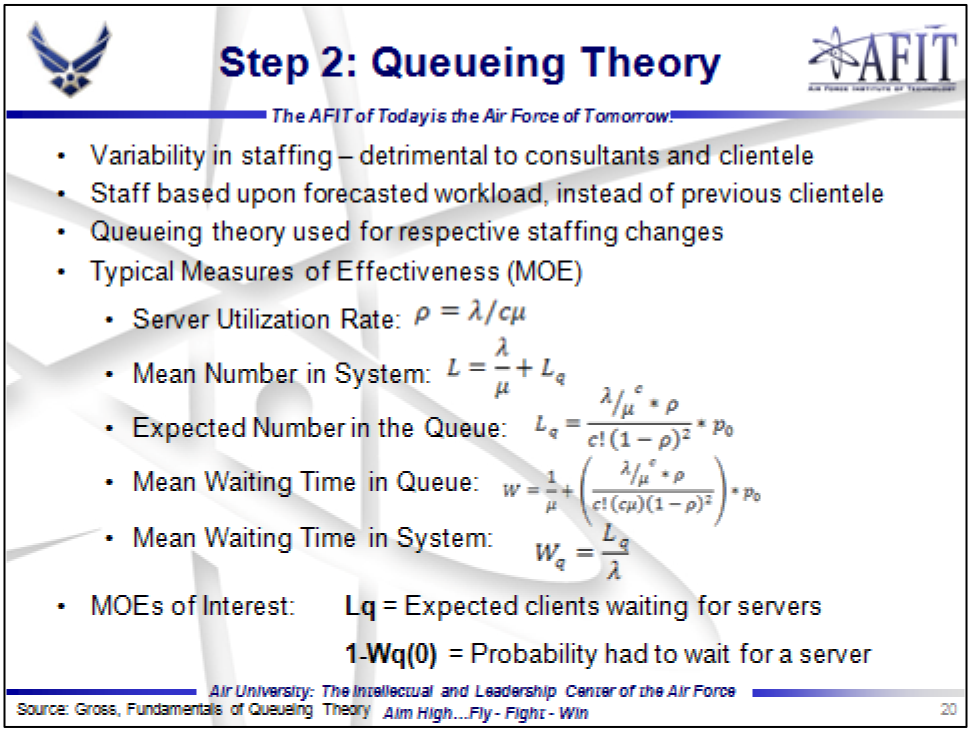
Queueing Theory Step 2

Estimate future personnel staffing using Square Root Law

$c \approx \omega + \beta\sqrt{\omega}$

Minimize waste and enhance decision making with mathematical reasoning/evidence

Goal



Step 2: Queueing Theory

- Variability in staffing – detrimental to consultants and clientele
- Staff based upon forecasted workload, instead of previous clientele
- Queueing theory used for respective staffing changes
- Typical Measures of Effectiveness (MOE)
  - Server Utilization Rate: $\rho = \lambda / c\mu$
  - Mean Number in System: $L = \frac{\lambda}{\mu} + L_q$
  - Expected Number in the Queue: $L_q = \frac{\lambda/\mu^{c} * \rho}{c!\,(1-\rho)^2} * p_0$
  - Mean Waiting Time in Queue: $W = \frac{1}{\mu} + \left(\frac{\lambda/\mu^{c} * \rho}{c!\,(c\mu)(1-\rho)^2}\right) * p_0$
  - Mean Waiting Time in System: $W_q = \frac{L_q}{\lambda}$
- MOEs of Interest: **Lq** = Expected clients waiting for servers

  1-Wq(0) = Probability had to wait for a server

Source: Gross, Fundamentals of Queueing Theory

Step 2: Matching Servers to Workload Demand

The AFIT of Today is the Air Force of Tomorrow!

- Appropriate number of servers for stable system: $c = \omega + \Delta,$

  $c = \#servers\ , \omega = load\ and\ \Delta > 0\ (number\ of\ additional\ servers\ used).$
- Quality Domain: match staffing to demand (constant utilization)

  $Load\ quadruples\ (4\omega){:}\ Quadruple\ the\ number\ of\ servers$
- Efficiency Domain: constant padding/flex factor

  $Load\ quadruples\ (4\omega){:}\ Appropriate\ servers\ for\ new\ load + 3$
- Quality & Efficiency Domain: maintain level of congestion
  - Line search for appropriate number of servers maintaining baseline values of Lq & 1-Wq(0) -- (computationally intensive)
  - Square Root Formula (quick approximation)

    $Load\ quadruples\ (4\omega){:}\ \ Adjust\ servers\ by\ square\ root\ of\ new\ load$

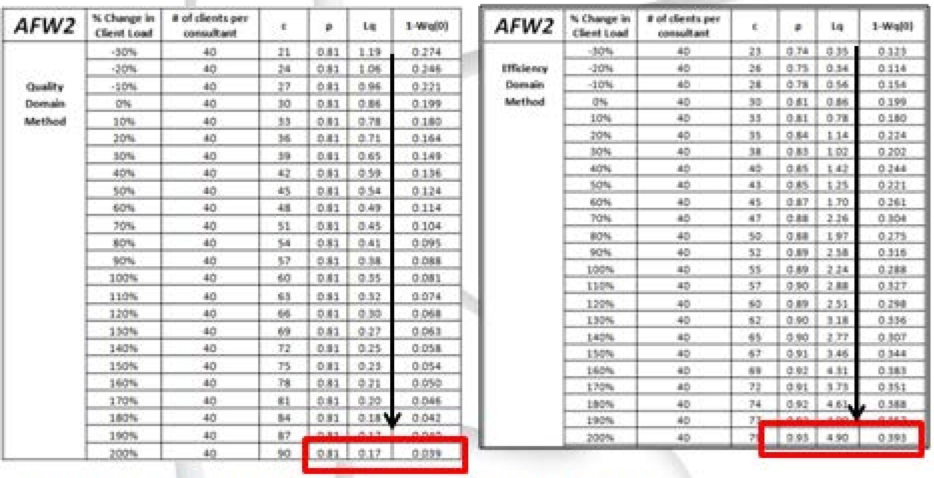
$$c \approx 4\omega + \beta\sqrt{4\omega}$$

Air University: The Intellectual and Leadership Center of the Air Force

Source: Gross, Fundamentals of Queueing Theory *Aim High...Fly - Fight - Win*

Step 2: Quality Vs. Efficiency

The AFIT of Today is the Air Force of Tomorrow!



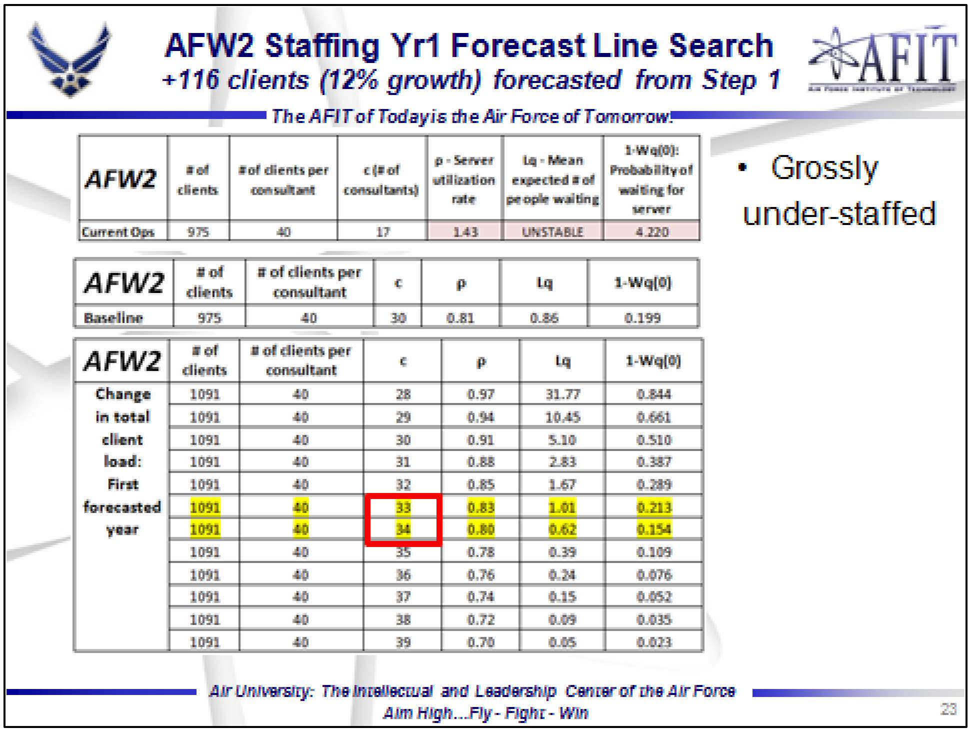
AFW2	% Change in Client Load	# of clients per consultant	c	ρ	Lq	1-Wq(0)
Quality Domain Method	-30%	40	21	0.81	1.19	0.274
	-20%	40	24	0.81	1.06	0.246
	-10%	40	27	0.81	0.96	0.221
	0%	40	30	0.81	0.86	0.199
	10%	40	33	0.81	0.78	0.180
	20%	40	36	0.81	0.71	0.164
	30%	40	39	0.81	0.65	0.149
	40%	40	42	0.81	0.59	0.136
	50%	40	45	0.81	0.54	0.124
	60%	40	48	0.81	0.49	0.114
	70%	40	51	0.81	0.45	0.104
	80%	40	54	0.81	0.41	0.095
	90%	40	57	0.81	0.38	0.088
	100%	40	60	0.81	0.35	0.081
	110%	40	63	0.81	0.32	0.074
	120%	40	66	0.81	0.30	0.068
	130%	40	69	0.81	0.27	0.063
	140%	40	72	0.81	0.25	0.058
	150%	40	75	0.81	0.23	0.054
	160%	40	78	0.81	0.21	0.050
	170%	40	81	0.81	0.20	0.046
	180%	40	84	0.81	0.18	0.042
	190%	40	87	[illegible]	[illegible]	[illegible]
	200%	40	90	0.81	0.17	0.039

AFW2	% Change in Client Load	# of clients per consultant	c	ρ	Lq	1-Wq(0)
Efficiency Domain Method	-30%	40	23	0.74	0.35	0.123
	-20%	40	26	0.75	0.34	0.114
	-10%	40	28	0.78	0.56	0.154
	0%	40	30	0.81	0.86	0.199
	10%	40	33	0.81	0.78	0.180
	20%	40	35	0.84	1.14	0.224
	30%	40	38	0.83	1.02	0.202
	40%	40	40	0.85	1.42	0.244
	50%	40	43	0.85	1.25	0.221
	60%	40	45	0.87	1.70	0.261
	70%	40	47	0.88	2.26	0.304
	80%	40	50	0.88	1.97	0.275
	90%	40	52	0.89	2.58	0.316
	100%	40	55	0.89	2.24	0.288
	110%	40	57	0.90	2.88	0.327
	120%	40	60	0.89	2.51	0.298
	130%	40	62	0.90	3.18	0.336
	140%	40	65	0.90	2.77	0.307
	150%	40	67	0.91	3.46	0.344
	160%	40	69	0.92	4.31	0.383
	170%	40	72	0.91	3.73	0.351
	180%	40	74	0.92	4.61	0.388
	190%	40	77	[illegible]	[illegible]	[illegible]
	200%	40	[illegible]	0.93	4.90	0.393

- Constant Server Utilization
- Flex Factor

Air University: The Intellectual and Leadership Center of the Air Force

Aim High...Fly - Fight - Win



AFW2 Staffing Yr1 Forecast Line Search

+116 clients (12% growth) forecasted from Step 1

The AFIT of Today is the Air Force of Tomorrow!

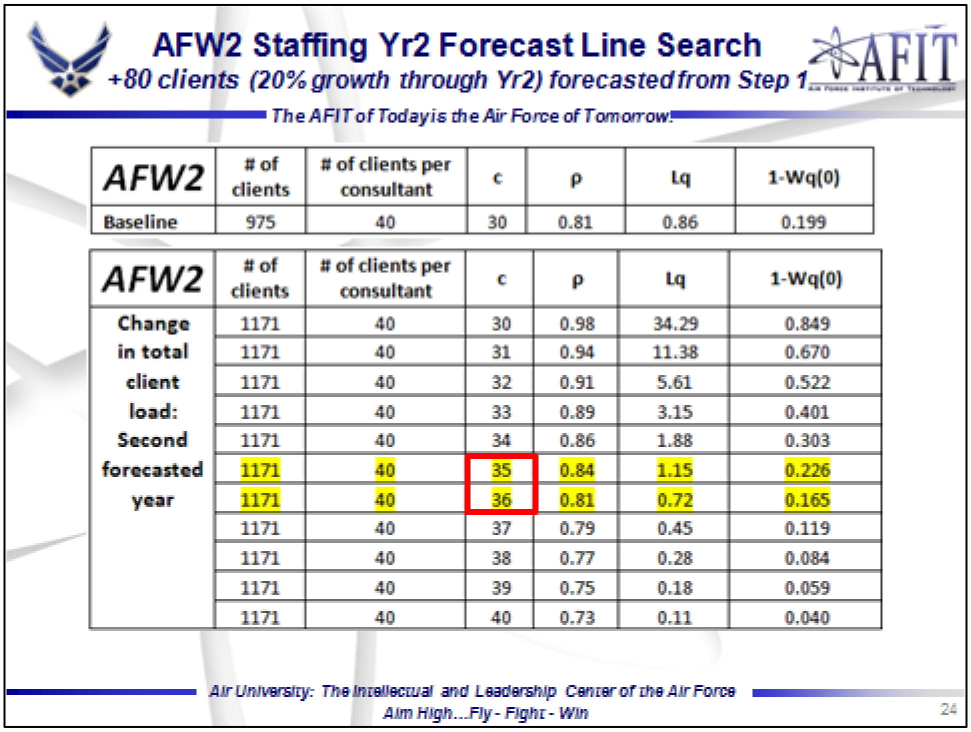
AFW2	# of clients	# of clients per consultant	c (# of consultants)	ρ - Server utilization rate	Lq - Mean expected # of people waiting	1-Wq(0): Probability of waiting for server
Current Ops	975	40	17	1.43	UNSTABLE	4.220

- Grossly under-staffed

AFW2	# of clients	# of clients per consultant	c	ρ	Lq	1-Wq(0)
Baseline	975	40	30	0.81	0.86	0.199

AFW2	# of clients	# of clients per consultant	c	ρ	Lq	1-Wq(0)
Change in total client load: First forecasted year	1091	40	28	0.97	31.77	0.844
	1091	40	29	0.94	10.45	0.661
	1091	40	30	0.91	5.10	0.510
	1091	40	31	0.88	2.83	0.387
	1091	40	32	0.85	1.67	0.289
	1091	40	33	0.83	1.01	0.213
	1091	40	34	0.80	0.62	0.154
	1091	40	35	0.78	0.39	0.109
	1091	40	36	0.76	0.24	0.076
	1091	40	37	0.74	0.15	0.052
	1091	40	38	0.72	0.09	0.035
	1091	40	39	0.70	0.05	0.023

Air University: The Intellectual and Leadership Center of the Air Force
Aim High…Fly - Fight - Win



AFW2 Staffing Yr2 Forecast Line Search

+80 clients (20% growth through Yr2) forecasted from Step 1

The AFIT of Today is the Air Force of Tomorrow!

AFW2	# of clients	# of clients per consultant	c	ρ	Lq	1-Wq(0)
Baseline	975	40	30	0.81	0.86	0.199

AFW2	# of clients	# of clients per consultant	c	ρ	Lq	1-Wq(0)
Change in total client load: Second forecasted year	1171	40	30	0.98	34.29	0.849
	1171	40	31	0.94	11.38	0.670
	1171	40	32	0.91	5.61	0.522
	1171	40	33	0.89	3.15	0.401
	1171	40	34	0.86	1.88	0.303
	1171	40	35	0.84	1.15	0.226
	1171	40	36	0.81	0.72	0.165
	1171	40	37	0.79	0.45	0.119
	1171	40	38	0.77	0.28	0.084
	1171	40	39	0.75	0.18	0.059
	1171	40	40	0.73	0.11	0.040

Air University: The Intellectual and Leadership Center of the Air Force
Aim High…Fly - Fight - Win

Step 2: AFW2 Q&E Domain

- Server Formula $c = \omega + \Delta \text{ ,where } \omega = \frac{\#\ of\ Clients}{40} \text{ and recall } \Delta > 0$
- Baseline $30 = \omega + \Delta, \quad where\ \omega = \frac{975}{40} = 24.375$
- Basic SQRT Formula $c = v\omega + \beta\sqrt{v\omega}$

$$v = \frac{(Expected\ \#wounded\ in\ forecasted\ year) + (\#\ Current\ Clients)}{(\#\ Current\ Clients)}$$

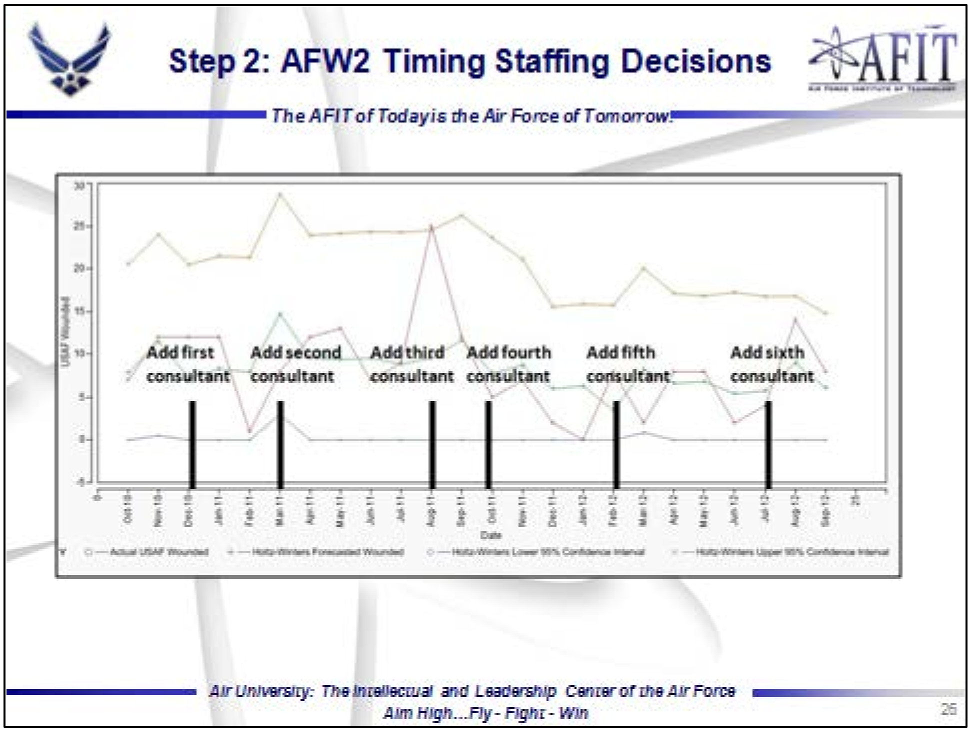
- 116 forecasted first year (1.12 x Current Load)

$$c \approx 1.12\left(\frac{975}{40}\right) + 1.2816\sqrt{1.12\left(\frac{975}{40}\right)} \approx 33.996 \sim 34\ consultants$$

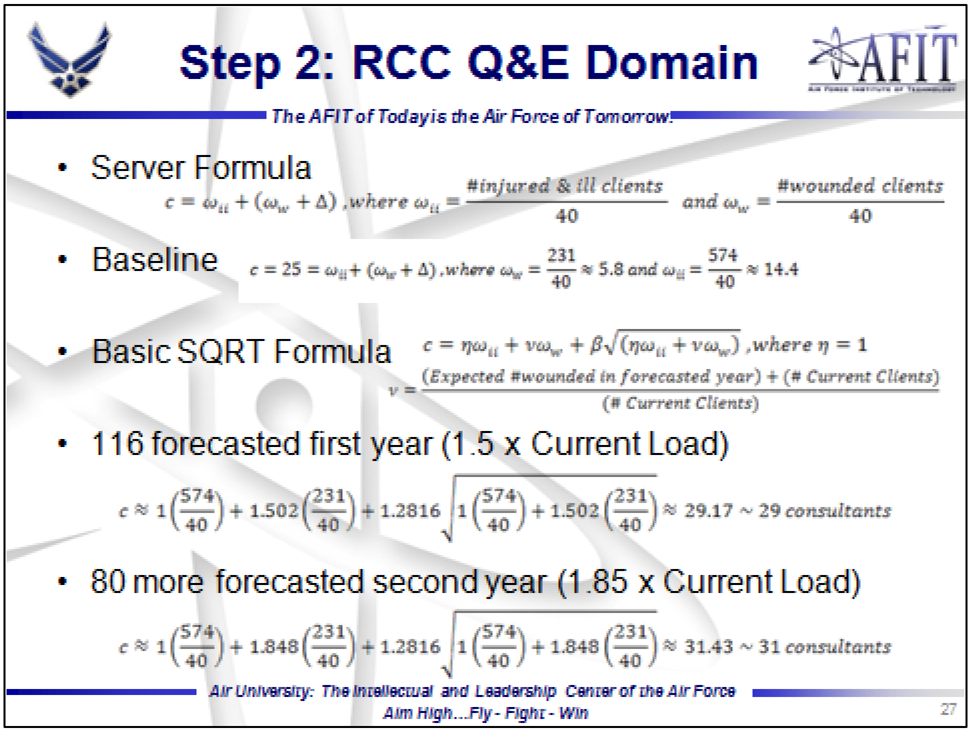
- 80 more forecasted second year (1.20 x Current Load)

$$c \approx 1.20\left(\frac{975}{40}\right) + 1.2816\sqrt{1.20\left(\frac{975}{40}\right)} \approx 36.181 \sim 36\ consultants$$

Source: Gross, Fundamentals of Queueing Theory



Step 2: AFW2 Timing Staffing Decisions



Step 2: RCC Q&E Domain

The AFIT of Today is the Air Force of Tomorrow!

- Server Formula

$$c = \omega_{ii} + (\omega_w + \Delta)\ ,where\ \omega_{ii} = \frac{\#injured\ \&\ ill\ clients}{40}\ \ and\ \omega_w = \frac{\#wounded\ clients}{40}$$

- Baseline $c = 25 = \omega_{ii} + (\omega_w + \Delta)\ ,where\ \omega_w = \frac{231}{40} \approx 5.8\ and\ \omega_{ii} = \frac{574}{40} \approx 14.4$

- Basic SQRT Formula $c = \eta\omega_{ii} + v\omega_w + \beta\sqrt{(\eta\omega_{ii} + v\omega_w)}\ ,where\ \eta = 1$

$$v = \frac{(Expected\ \#wounded\ in\ forecasted\ year) + (\#\ Current\ Clients)}{(\#\ Current\ Clients)}$$

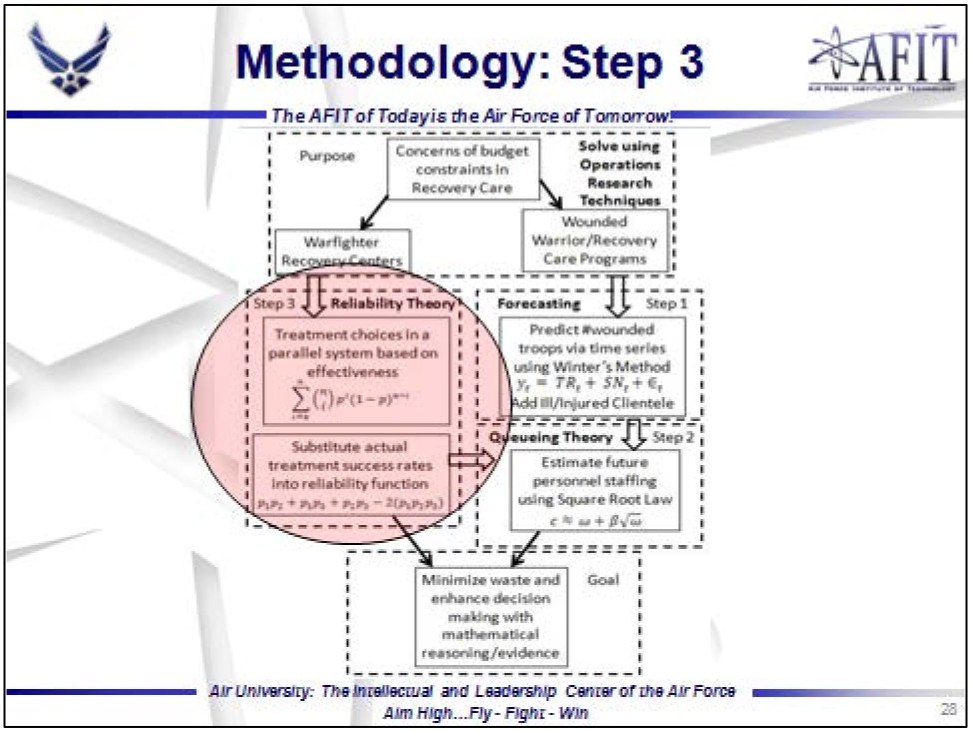
- 116 forecasted first year (1.5 x Current Load)

$$c \approx 1\left(\frac{574}{40}\right) + 1.502\left(\frac{231}{40}\right) + 1.2816\sqrt{1\left(\frac{574}{40}\right) + 1.502\left(\frac{231}{40}\right)} \approx 29.17 \sim 29\ consultants$$

- 80 more forecasted second year (1.85 x Current Load)

$$c \approx 1\left(\frac{574}{40}\right) + 1.848\left(\frac{231}{40}\right) + 1.2816\sqrt{1\left(\frac{574}{40}\right) + 1.848\left(\frac{231}{40}\right)} \approx 31.43 \sim 31\ consultants$$

Air University: The Intellectual and Leadership Center of the Air Force
Aim High…Fly - Fight - Win



Methodology: Step 3

The AFIT of Today is the Air Force of Tomorrow!

Purpose

Concerns of budget constraints in Recovery Care

Solve using Operations Research Techniques

Warfighter Recovery Centers

Wounded Warrior/Recovery Care Programs

Step 3 **Reliability Theory**

Treatment choices in a parallel system based on effectiveness

$$\sum_{i=k}^{n}\binom{n}{i}p^i(1-p)^{n-i}$$

Substitute actual treatment success rates into reliability function

$$p_1p_2 + p_1p_3 + p_2p_3 - 2(p_1p_2p_3)$$

Forecasting Step 1

Predict #wounded troops via time series using Winter's Method

$$y_t = TR_t + SN_t + \epsilon_t$$

Add Ill/Injured Clientele

Queueing Theory Step 2

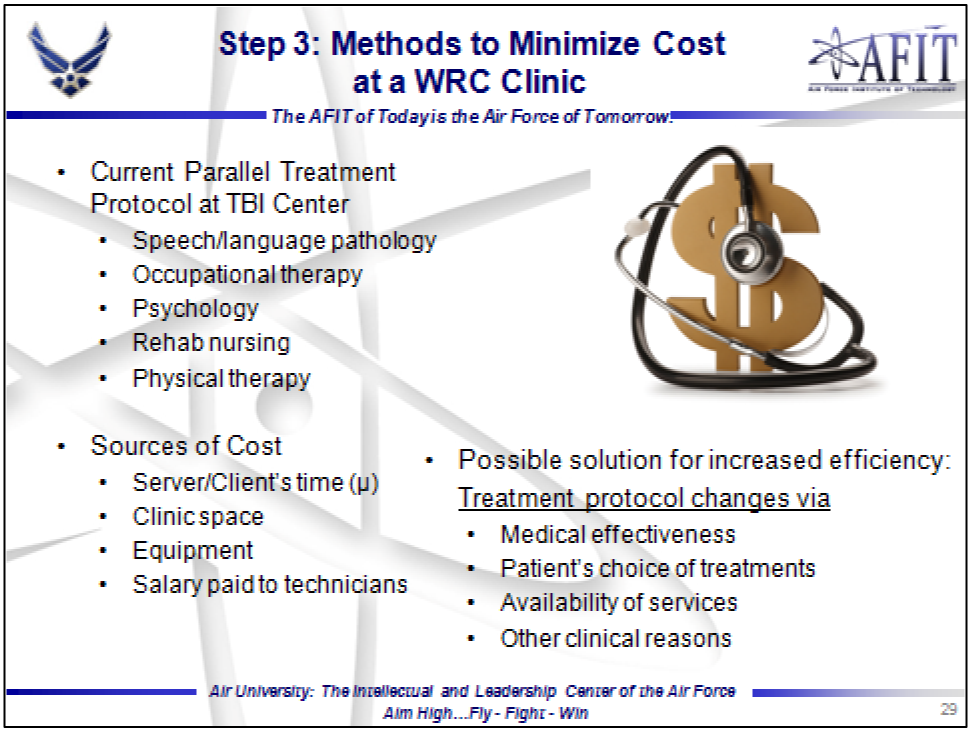
Estimate future personnel staffing using Square Root Law

$$c \approx \omega + \beta\sqrt{\omega}$$

Minimize waste and enhance decision making with mathematical reasoning/evidence

Goal

Air University: The Intellectual and Leadership Center of the Air Force
Aim High…Fly - Fight - Win

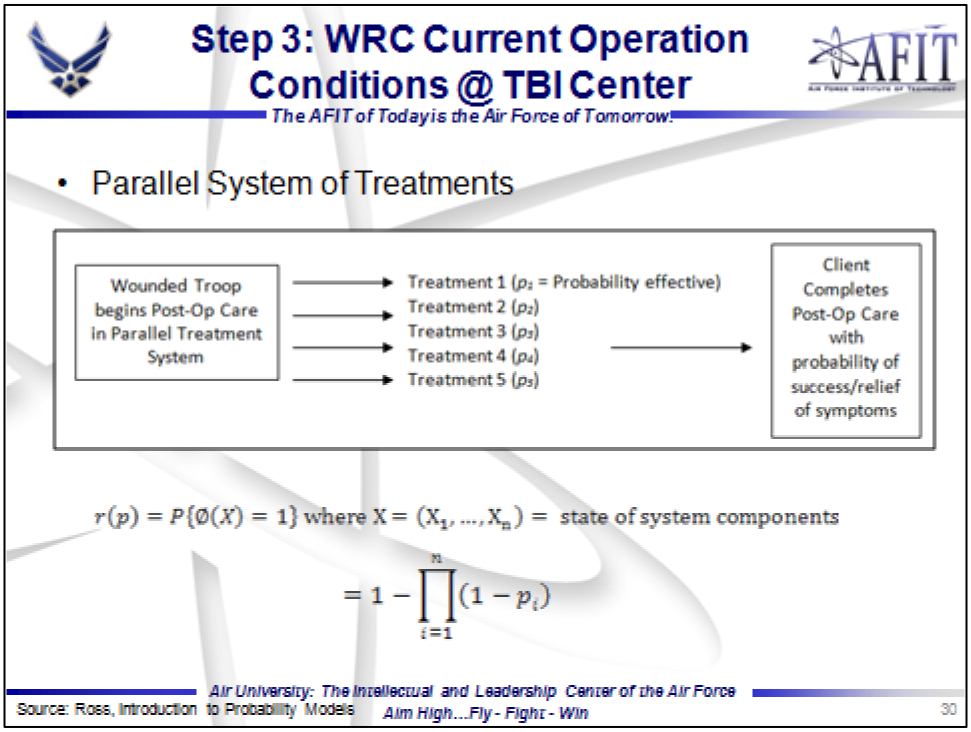


Step 3: Methods to Minimize Cost at a WRC Clinic

The AFIT of Today is the Air Force of Tomorrow!

- Current Parallel Treatment Protocol at TBI Center
    - Speech/language pathology
    - Occupational therapy
    - Psychology
    - Rehab nursing
    - Physical therapy
- Sources of Cost
    - Server/Client's time (μ)
    - Clinic space
    - Equipment
    - Salary paid to technicians
- Possible solution for increased efficiency: Treatment protocol changes via
    - Medical effectiveness
    - Patient's choice of treatments
    - Availability of services
    - Other clinical reasons

Air University: The Intellectual and Leadership Center of the Air Force
Aim High...Fly - Fight - Win



Step 3: WRC Current Operation Conditions @ TBI Center

The AFIT of Today is the Air Force of Tomorrow!

- Parallel System of Treatments

Wounded Troop begins Post-Op Care in Parallel Treatment System

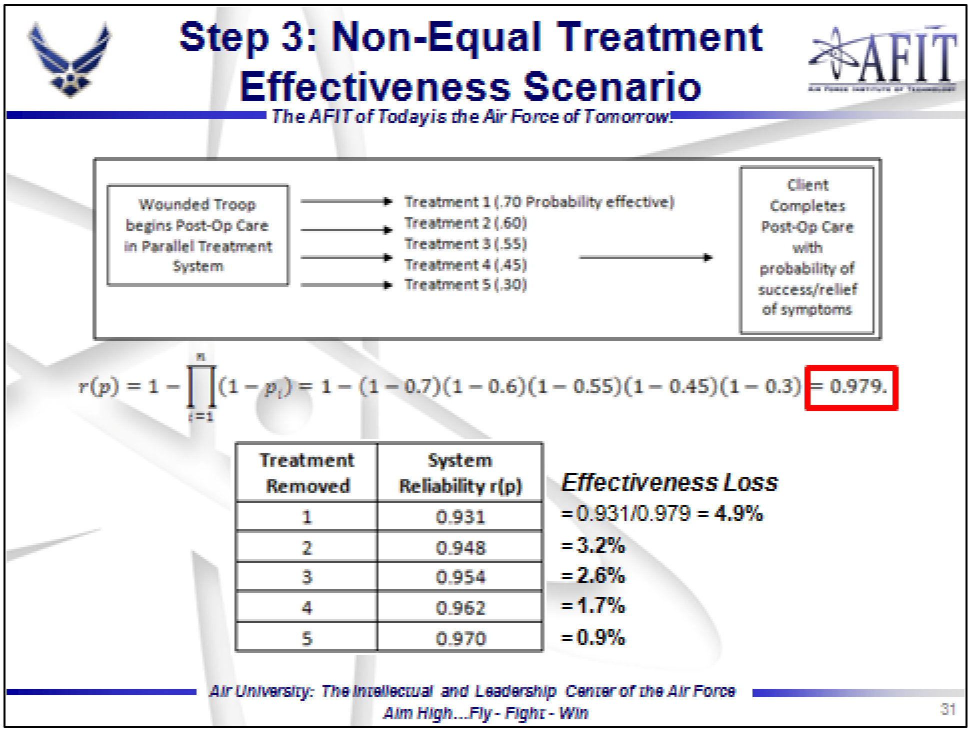
Treatment 1 (p_1 = Probability effective)
Treatment 2 (p_2)
Treatment 3 (p_3)
Treatment 4 (p_4)
Treatment 5 (p_5)

Client Completes Post-Op Care with probability of success/relief of symptoms

$$r(p) = P\{\emptyset(X) = 1\} \text{ where } X = (X_1, \ldots, X_n) = \text{ state of system components}$$

$$= 1 - \prod_{i=1}^{n}(1 - p_i)$$

Air University: The Intellectual and Leadership Center of the Air Force
Source: Ross, Introduction to Probability Models
Aim High...Fly - Fight - Win



Step 3: Non-Equal Treatment Effectiveness Scenario

The AFIT of Today is the Air Force of Tomorrow.

Wounded Troop begins Post-Op Care in Parallel Treatment System

Treatment 1 (.70 Probability effective)
Treatment 2 (.60)
Treatment 3 (.55)
Treatment 4 (.45)
Treatment 5 (.30)

Client Completes Post-Op Care with probability of success/relief of symptoms

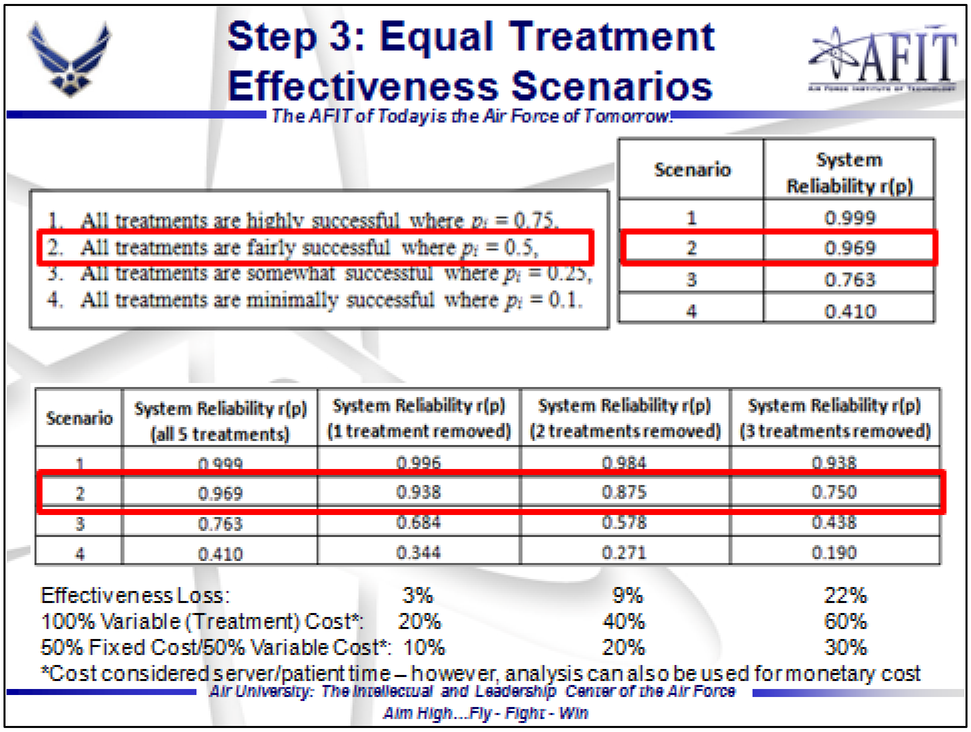
$$r(p) = 1 - \prod_{i=1}^{n}(1 - p_i) = 1 - (1 - 0.7)(1 - 0.6)(1 - 0.55)(1 - 0.45)(1 - 0.3) \boxed{= 0.979.}$$

Treatment Removed	System Reliability r(p)
1	0.931
2	0.948
3	0.954
4	0.962
5	0.970

Effectiveness Loss
= 0.931/0.979 = **4.9%**
= **3.2%**
= **2.6%**
= **1.7%**
= **0.9%**

Air University: The Intellectual and Leadership Center of the Air Force
Aim High…Fly - Fight - Win

31



Step 3: Equal Treatment Effectiveness Scenarios

The AFIT of Today is the Air Force of Tomorrow.

1. All treatments are highly successful where $p_i = 0.75$,
2. All treatments are fairly successful where $p_i = 0.5$,
3. All treatments are somewhat successful where $p_i = 0.25$,
4. All treatments are minimally successful where $p_i = 0.1$.

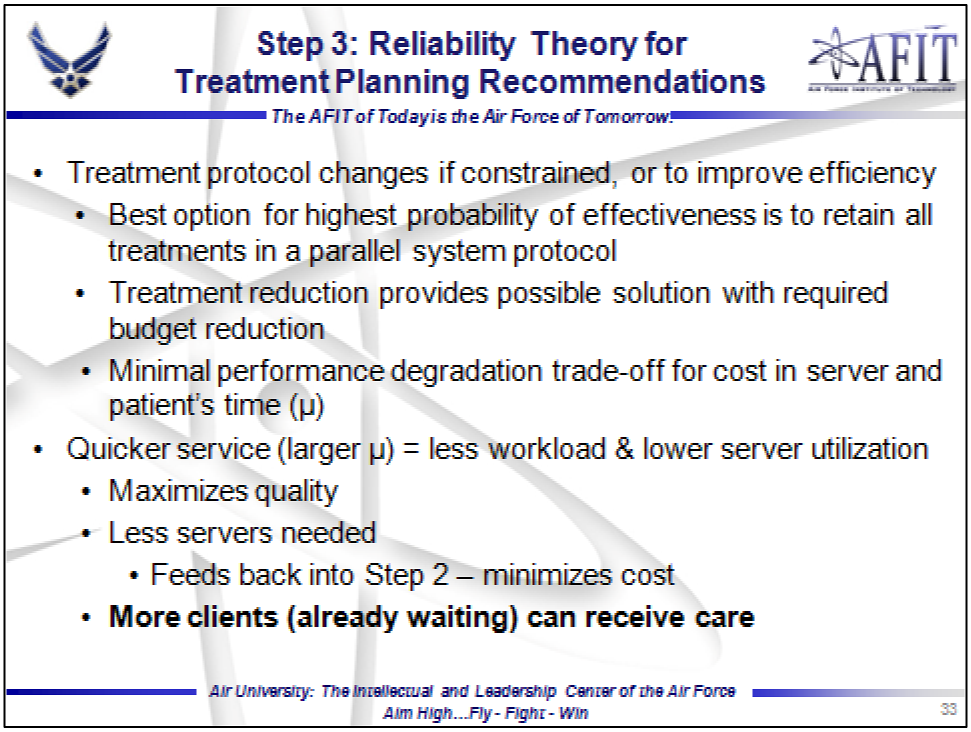
Scenario	System Reliability r(p)
1	0.999
2	0.969
3	0.763
4	0.410

Scenario	System Reliability r(p) (all 5 treatments)	System Reliability r(p) (1 treatment removed)	System Reliability r(p) (2 treatments removed)	System Reliability r(p) (3 treatments removed)
1	0.999	0.996	0.984	0.938
2	0.969	0.938	0.875	0.750
3	0.763	0.684	0.578	0.438
4	0.410	0.344	0.271	0.190

Effectiveness Loss:	3%	9%	22%
100% Variable (Treatment) Cost*:	20%	40%	60%
50% Fixed Cost/50% Variable Cost*:	10%	20%	30%

*Cost considered server/patient time – however, analysis can also be used for monetary cost

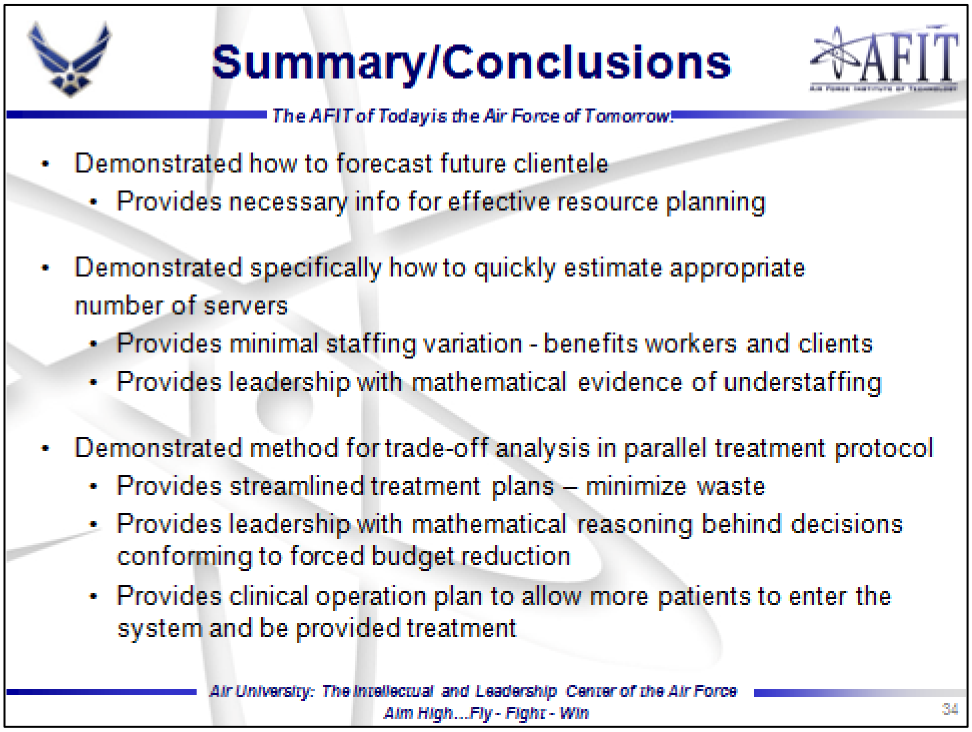
Air University: The Intellectual and Leadership Center of the Air Force
Aim High…Fly - Fight - Win



Step 3: Reliability Theory for Treatment Planning Recommendations

The AFIT of Today is the Air Force of Tomorrow!

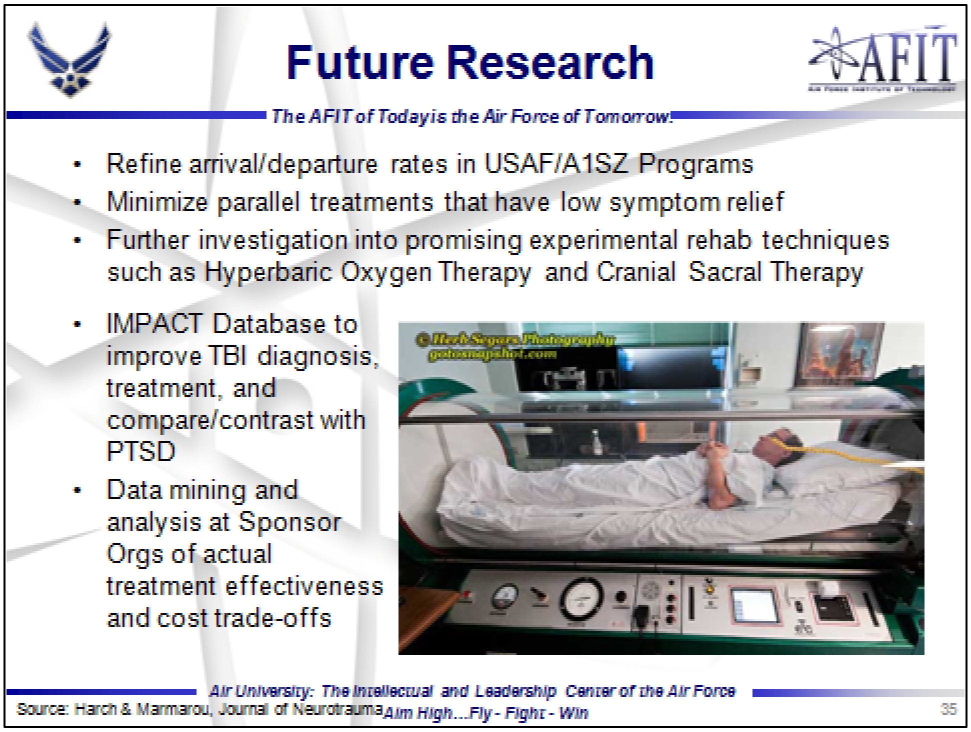
- Treatment protocol changes if constrained, or to improve efficiency
  - Best option for highest probability of effectiveness is to retain all treatments in a parallel system protocol
  - Treatment reduction provides possible solution with required budget reduction
  - Minimal performance degradation trade-off for cost in server and patient's time (μ)
- Quicker service (larger μ) = less workload & lower server utilization
  - Maximizes quality
  - Less servers needed
    - Feeds back into Step 2 – minimizes cost
  - **More clients (already waiting) can receive care**



Summary/Conclusions

The AFIT of Today is the Air Force of Tomorrow!

- Demonstrated how to forecast future clientele
  - Provides necessary info for effective resource planning
- Demonstrated specifically how to quickly estimate appropriate number of servers
  - Provides minimal staffing variation - benefits workers and clients
  - Provides leadership with mathematical evidence of understaffing
- Demonstrated method for trade-off analysis in parallel treatment protocol
  - Provides streamlined treatment plans – minimize waste
  - Provides leadership with mathematical reasoning behind decisions conforming to forced budget reduction
  - Provides clinical operation plan to allow more patients to enter the system and be provided treatment



Future Research

The AFIT of Today is the Air Force of Tomorrow!

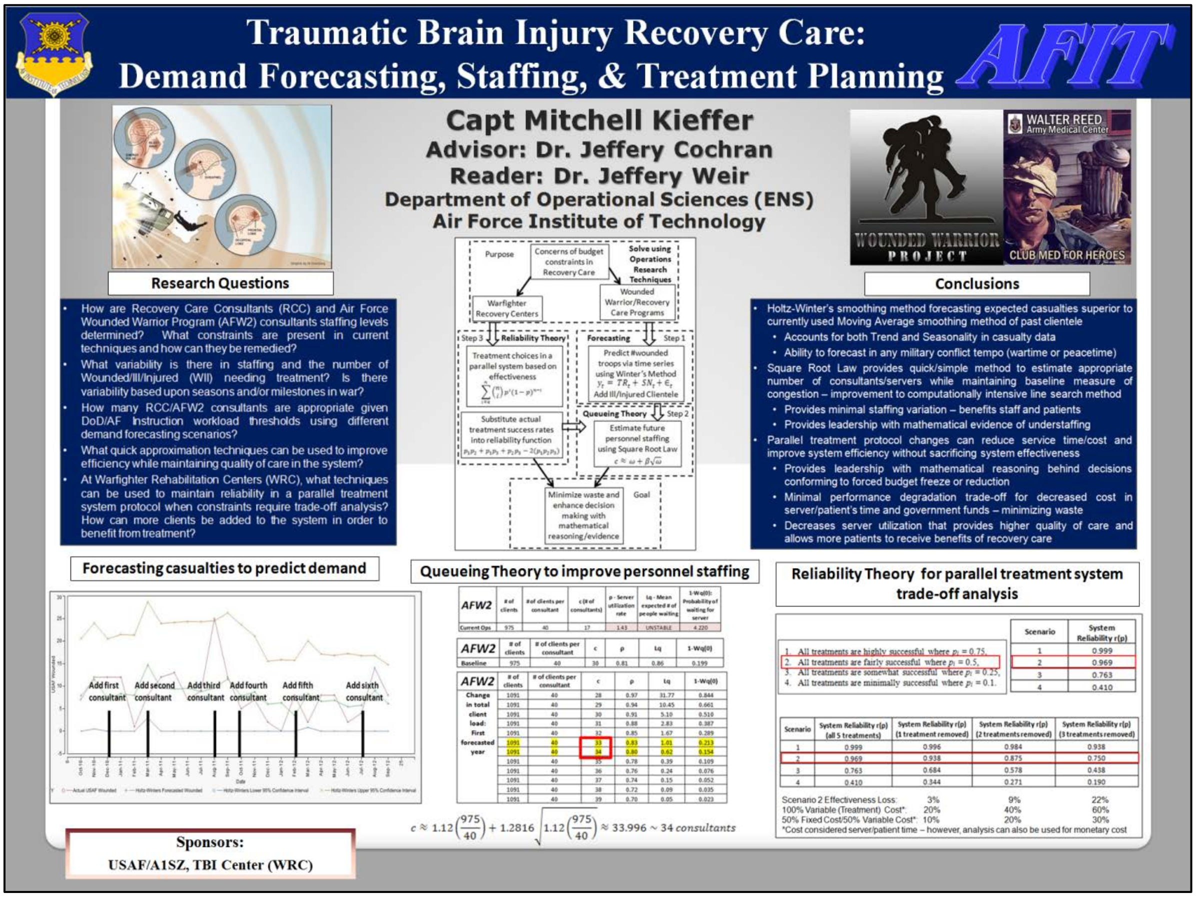
- Refine arrival/departure rates in USAF/A1SZ Programs
- Minimize parallel treatments that have low symptom relief
- Further investigation into promising experimental rehab techniques such as Hyperbaric Oxygen Therapy and Cranial Sacral Therapy
- IMPACT Database to improve TBI diagnosis, treatment, and compare/contrast with PTSD
- Data mining and analysis at Sponsor Orgs of actual treatment effectiveness and cost trade-offs

Air University: The Intellectual and Leadership Center of the Air Force

Aim High…Fly - Fight - Win

Source: Harch & Marmarou, Journal of Neurotrauma

Quad Chart



Traumatic Brain Injury Recovery Care: Demand Forecasting, Staffing, & Treatment Planning

Capt Mitchell Kieffer
Advisor: Dr. Jeffery Cochran
Reader: Dr. Jeffery Weir
Department of Operational Sciences (ENS)
Air Force Institute of Technology

Research Questions

- How are Recovery Care Consultants (RCC) and Air Force Wounded Warrior Program (AFW2) consultants staffing levels determined? What constraints are present in current techniques and how can they be remedied?
- What variability is there in staffing and the number of Wounded/Ill/Injured (WII) needing treatment? Is there variability based upon seasons and/or milestones in war?
- How many RCC/AFW2 consultants are appropriate given DoD/AF Instruction workload thresholds using different demand forecasting scenarios?
- What quick approximation techniques can be used to improve efficiency while maintaining quality of care in the system?
- At Warfighter Rehabilitation Centers (WRC), what techniques can be used to maintain reliability in a parallel treatment system protocol when constraints require trade-off analysis? How can more clients be added to the system in order to benefit from treatment?

Purpose

Concerns of budget constraints in Recovery Care

Solve using Operations Research Techniques

Warfighter Recovery Centers

Wounded Warrior/Recovery Care Programs

Step 3 **Reliability Theory**

Treatment choices in a parallel system based on effectiveness

$\sum_{i=k}^{n}\binom{n}{i}p^i(1-p)^{n-i}$

Substitute actual treatment success rates into reliability function

$p_1p_2 + p_1p_3 + p_2p_3 - 2(p_1p_2p_3)$

Forecasting Step 1

Predict #wounded troops via time series using Winter's Method

$y_t = TR_t + SN_t + \epsilon_t$

Add Ill/Injured Clientele

Queueing Theory Step 2

Estimate future personnel staffing using Square Root Law

$c \approx \omega + \beta\sqrt{\omega}$

Minimize waste and enhance decision making with mathematical reasoning/evidence

Goal

Conclusions

- Holtz-Winter's smoothing method forecasting expected casualties superior to currently used Moving Average smoothing method of past clientele
  - Accounts for both Trend and Seasonality in casualty data
  - Ability to forecast in any military conflict tempo (wartime or peacetime)
- Square Root Law provides quick/simple method to estimate appropriate number of consultants/servers while maintaining baseline measure of congestion – improvement to computationally intensive line search method
  - Provides minimal staffing variation – benefits staff and patients
  - Provides leadership with mathematical evidence of understaffing
- Parallel treatment protocol changes can reduce service time/cost and improve system efficiency without sacrificing system effectiveness
  - Provides leadership with mathematical reasoning behind decisions conforming to forced budget freeze or reduction
  - Minimal performance degradation trade-off for decreased cost in server/patient's time and government funds – minimizing waste
  - Decreases server utilization that provides higher quality of care and allows more patients to receive benefits of recovery care

Forecasting casualties to predict demand

Queueing Theory to improve personnel staffing

AFW2	# of clients	# of clients per consultant	c (# of consultants)	ρ - Server utilization rate	Lq - Mean expected # of people waiting	1-Wq(0): Probability of waiting for server
Current Ops	975	40	17	1.43	UNSTABLE	4.220

AFW2	# of clients	# of clients per consultant	c	ρ	Lq	1-Wq(0)
Baseline	975	40	30	0.81	0.86	0.199

AFW2	# of clients	# of clients per consultant	c	ρ	Lq	1-Wq(0)
Change in total client load: First forecasted year	1091	40	28	0.97	31.77	0.844
	1091	40	29	0.94	10.45	0.661
	1091	40	30	0.91	5.10	0.510
	1091	40	31	0.88	2.83	0.387
	1091	40	32	0.85	1.67	0.289
	1091	40	33	0.83	1.01	0.213
	1091	40	34	0.80	0.62	0.154
	1091	40	35	0.78	0.39	0.109
	1091	40	36	0.76	0.24	0.076
	1091	40	37	0.74	0.15	0.052
	1091	40	38	0.72	0.09	0.035
	1091	40	39	0.70	0.05	0.023

$$c \approx 1.12\left(\frac{975}{40}\right) + 1.2816\sqrt{1.12\left(\frac{975}{40}\right)} \approx 33.996 \sim 34\ consultants$$

Reliability Theory for parallel treatment system trade-off analysis

1. All treatments are highly successful where $p_i = 0.75$,
2. All treatments are fairly successful where $p_i = 0.5$,
3. All treatments are somewhat successful where $p_i = 0.25$,
4. All treatments are minimally successful where $p_i = 0.1$.

Scenario	System Reliability r(p)
1	0.999
2	0.969
3	0.763
4	0.410

Scenario	System Reliability r(p) (all 5 treatments)	System Reliability r(p) (1 treatment removed)	System Reliability r(p) (2 treatments removed)	System Reliability r(p) (3 treatments removed)
1	0.999	0.996	0.984	0.938
2	0.969	0.938	0.875	0.750
3	0.763	0.684	0.578	0.438
4	0.410	0.344	0.271	0.190

Scenario 2 Effectiveness Loss:	3%	9%	22%
100% Variable (Treatment) Cost*:	20%	40%	60%
50% Fixed Cost/50% Variable Cost*:	10%	20%	30%

*Cost considered server/patient time – however, analysis can also be used for monetary cost

Sponsors:
USAF/A1SZ, TBI Center (WRC)

Bibliography

Artalejo, J. R. (2002). A note on the optimality of the N- and D- policies for the M/G/1 queue. *Operations Research Letters , 30* (6), 375-376.

Bales (M.D.), J. (2013, January 28). Major, USAF, Air Force Wounded Warrior Adaptive Sports Program. (M. Kieffer, Interviewer)

Belasco, A. (2011). *Specialist in US Defence Policy and Budget.* Congressional Research Service.

Blood, C. G., Zouris, J. M., & Rotblatt, D. (1997). *Using the Ground Forces Casualty Forecasting System (FORECAS) to Project Casualty Sustainment.* San Diego: Naval Health Research Center.

Bowerman, B. L., O'Connell, R. T., & Koehler, A. (2005). *Forecasting, Time Series, and Regression: an applied approach* (4 ed.). Thomson Brooks/Cole.

Brenner, L. (2011). Traumatic Brain Injury and PTSD: Diagnosis and Treatment Issues. *3rd Annual Conference on Complexities and Challenges of PTSD and TBI.* Boston: Medscape LLC.

Broyles, J. R., Cochran, J. K., & Montgomery, D. C. (2011). A Markov decision process to dynamically match hospital inpatient staffing to demand. *IIE Transactions on Healthcare Systems Engineering , 1*, 116-130.

Carlson, K., Kehle, S., Meis, L., Greer, N., MacDonald, R., Rutks, I., et al. (2009). *The Assessment and Treatment of Individuals with History of Traumatic Brain Injury and Post-Traumatic Stress Disorder: A Systematic Review of the Evidence.* Minneapolis: Department of Veterans Affiars, Health Services Research and Development.

Cochran, J. K., & Bharti, A. (2006). Stochastic bed balancing of an obstetrics hospital. *Health Care Management Science , 9*, 31-45.

Cochran, J. K., & Roche, K. (2008). A queueing-based decision support methodology to estimate hospital inpatient bed demand. *Journal of the Operational Research Society , 59*, 1471-1482.

Darrell D. Jones, Lt Gen, USAF, Director, Manpower, Personnel and Services. (2012, June 21). AFI 34-1101: Air Force Warrior and Survivor Care. United States: AF/Manpower and Personnel (AF/A1SZ).

Demmons, B. (2011, Feb 23). Lieutenant Colonel, USAF/Manpower and Personnel, Warrior and Survivor Care. *Air Force Warrior and Survivor Care* .

DePalma, R. G., Burris, D. G., Champion, H. R., & Hodgson, M. J. (2005). *The New England Journal of Medicine* , 352:1335-42.

Federgruen, A., & So, K. C. (1991). Optimality of threshold policies in single-server queueing systems with server vacations. *Advances in Applied Probability , 23*, 388-405.

Gail H. McGinn, Deputy Under Secretary of Defense (Plans). (2009, Dec 1). DoDI 1300.24: Recovery Coordination Program. USA.

Gorunescu, F., McClean, S. I., & Millard, P. H. (2002). Using a queuing model to help plan bed allocation in a department of geriatric medicine. *Health Care Management Science , 5*, 307-312.

Gross, D., Shortle, J. F., Thompson, J. M., & Harris, C. M. (2008). *Fundamentals of Queueing Theory* (4th ed.). Hoboken, NJ: John Wiley & Sons, Inc.

Harch, P., Andrews, S., Fogarty, E., Amen, D., Pezzullo, J., Lucarini, J., et al. (2012). A Phase I Study of Low-Pressure Hyperbaric Oxygen Therapy for Blast-Induced Post-Concussion Syndrome and Post-Traumatic Stress Disorder. *Journal of Neurotrauma , 29* (1), 168-185.

Lillo, R. E., & Martin, M. (2000). On optimal exhaustive policies for the M/G/1 queue. *Operations Research Letters , 27*, 39-46.

Lindsay, M. A. (2012, Nov 15). Nurse Practitioner & Deputy Chief, Traumatic Brain Injury Service. (M. R. Kieffer, Interviewer)

Littig, S. J., & Isken, M. W. (2007). Short term hospital occupancy prediction. *Health Care Management Science* (10), 47-66.

Livingston, I. S., & O'Hanlon, M. (2012). *Afghanistan Index.* Brookings Institution.

Long, J. B. (2010). *Brain Vulnerability to Repeated Blast Overpressure and Polytrauma.* Lakewood, WA: The Geneva Foundation.

Marmarou, A., Lu, J., Butcher, I., McHugh, G., Mushkudiani, N., Murray, G., et al. (2007). IMPACT Database of Traumatic Brain Injury: Design and Description. *Journal of Neurotrauma , 24* (2), 239-250.

Mayo Clinic Staff. (2012, October 12). *Mayo Clinic*. Retrieved 11 21, 2012, from Traumatic Brain Injury: http://www.mayoclinic.com/health/traumatic-brain-injury/DS00552/DSECTION=treatments-and-drugs

Millard, P., & McClean, S. (1994). Modelling Hospital Resource Use. *Royal Society of Medicine Press* .

Morgan, R. (2012, August 8). Recovery Care Coordinator, AFMC 88 MDG/SGHM. (M. R. Kieffer, Interviewer)

O'Hankes, M. E., & Livingston, I. (2011). *Iraq Index.* Brookings Institution.

Okamura, H., Dohi, T., & Osaki, S. (2000). Optimal policies for a controlled queueing system with removable server under a random vacation circumstance. *Computers and Mathematics with Applications , 39*, 215-227.

Operation Enduring Freedom. (n.d.). Retrieved August 7, 2012, from http://icasualties.org/OEF/index.aspx

Owens, B. D., Kragh, J. F., Wenke, J. C., Macaitis, J., Wade, C. E., & Holcomb, J. B. (2008). Combat wounds in Operation Iraqi Freedom and Operation Enduring Freedom. *Journal of Trauma* , 64, 295-299.

Painton, L., & Campbell, J. (1995). Genetic algorithms in optimization of system reliability. *IEEE Transactions on Reliability , 44* (2), 172-178.

Ross, S. M. (2007). *Introduction to Probability Models* (9th ed.). Berkeley, CA: Elsevier Inc.

Runy, L. A. (2005). The cost of hospital care. *Hospitals and health Networks , 79* (12), 32.

Spencer, L. (2013). *General, USAF, Vice Chief of Staff.* Memorandum: Fiscal Year 2013 (FY13) Sequestration Guidance.

Taber, K. H., & Hurley, R. A. (2010, Winter). *PTSD Research Quarterly , 21* (1).

Tadj, L. (2003). A quorum queueing system under T-policy. *Journal of the Operational Research Society , 54*, 466-471.

Takagi, H. (1991). *Queueing Analysis: a Foundation of Performance Evaluation* (Vol. 1). Amsterdman, Netherlands: Elsevier Science.

Townes, T. (2012, 11 3). Headquarters, USAF/Manpower and Personnel, Warrior & Survivor Care, Pentagon. (M. Kieffer, Interviewer)

Traumatic Brain Injury. (2011, Aug 10). Retrieved Nov 15, 2012, from US Army Medical Department, Brooke Army Medical Center: http://www.bamc.amedd.army.mil/departments/orthopaedic/tbi/

Viana, A., & Rais, A. (2010). Operations Research in Healthcare: a survey. *International Transactions in Operational Research* , 1-31.

Wang, P. (1996). Markovian queueing models with periodic review. *Computers and Operations Research , 23*, 741-754.

Whitt, W. (2007). What you should know about queueing models to set staffing requirements in service systems. *Naval Research Logistics , 54* (5), 476-484.

Wojcik, B. E., Stein, C. R., Orosco, J., Bagg, K. A., & Humphrey, R. J. (2010, 7). Creation of an Expanded Barell Matrix to Identify Traumatic Brain Injuries of U.S. Military Members. *The Journal of Defense Modeling and Simulation: Applications, Methodology, Technology* , 157-166.

Yadin, M., & Naor, P. (1967). On queueing systems with variable service capacities. *Naval Research Logistics Quarterly* (14), 43-54.

Zoroya, G. (2009, March 5). 360,000 veterans may have brain injuries. *USA Today* .

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 074-0188

1. REPORT DATE *(DD-MM-YYYY)*
21 Mar 2013

2. REPORT TYPE
Master's Thesis

3. DATES COVERED *(From – To)*
31 Aug 2011 – 21 Mar 2013

4. TITLE AND SUBTITLE
Traumatic Brain Injury Recovery Care: Demand Forecasting, Staffing, and Treatment Planning

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Kieffer, Mitchell R., Captain, USAF

5d. PROJECT NUMBER
N/A

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAMES(S) AND ADDRESS(S)
Air Force Institute of Technology
Graduate School of Engineering and Management (AFIT/EN)
2950 Hobson Way
WPAFB OH 45433

8. PERFORMING ORGANIZATION REPORT NUMBER
AFIT-ENS-13-M-08

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Attn: Dr. James Bales, Air Force Wounded Warrior Program
301 Fisher Street
Keesler AFB, MS 39534
Phone: 591-5511, james.bales@us.af.mil

10. SPONSOR/MONITOR'S ACRONYM(S)
AFW2

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.

13. SUPPLEMENTARY NOTES
This material is declared a work of the U.S. Government and is not subject to copyright protection in the United States.

14. ABSTRACT

Improvised Explosive Device attacks have skyrocketed since the start of the War on Terror. Many troops wounded by these tactics receive long-lasting unseen wounds including Traumatic Brain Injury (TBI). TBI sufferers are treated along with other casualties. This has created an increasing, and varying, demand for ongoing post operative recovery care for troops returning from deployments. Diagnosis and treatment for TBI wounded troops is costly. This thesis is motivated by the recognition that budgets are constrained yet quality of care should not be compromised. Additive Holtz–Winters smoothing is used to forecast overall patient care demand, a regression based on queueing theory determines care consultant staffing levels, and reliability theory quantifies the idea of reducing cost by reducing parallel treatment planning. The scope is the Warfighter Rehabilitation Centers and AF Warrior and Survivor Care with data from SMEs, the Brookings Institution, and icasualties.org. This thesis provides a step-by-step methodology and analyzes the actual situation that leadership encountered from 2010-2012.

15. SUBJECT TERMS
Traumatic Brain Injury, Wounded Warrior, Rehabilitation, Forecasting, Queueing Theory, Reliability Theory

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
					Dr. Jeffery K. Cochran, ENS
a. REPORT	b. ABSTRACT	c. THIS PAGE			19b. TELEPHONE NUMBER *(Include area code)*
U	U	U	UU	113	785-3636 x4521, (Jeffery.Cochran@afit.edu)

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std. Z39-18