Journal homepage: www.jmisci.com
Contact: editor@jmisci.com

Journal of Military and Information Science
Institute of Science

JMIS

A Study on Casualty Profile Using Logistics Regression

Sezgin Özcan*

*War Colleges Command, Army War College, Yenilevent-34330, İstanbul, Turkey. Tel: +90 212 398-0100, e-mail: szgozcan@gmail.com

Abstract- The environment faced by today's servicemen is characterized by continual deployments to combat zones, where troops are exposed to the risks of the battlefield. Casualty, whether to combatants or noncombatants, is an unavoidable reality of war. Although the primary goal of combat is to defeat the enemy, keeping casualties down is important as well. Low numbers of injured or killed soldiers not only maintain the ranks of service members, but also have an incredible effect on morale. The purpose of this study is to create a profile of U.S. Army troops killed or injured due to hostile incidents in Afghanistan and Iraq between 2003 and 2011. The analysis of study my help decision makers to see profile that is most vulnerable to casualty. The first part of our study analyzes the descriptive statistical results and the second part contains the results of multivariate analysis of the casualty status of servicemen of the U.S. Army. As a conclusion for our multivariate model, an actual-duty person who is female, married, serving in the reserve forces, serving in a combat troop, between pay grades E1–E3, serving in Iraq, serving the first deployment is the serviceman with most potential to get injured or killed in the U.S. Army.

Keywords- U.S. Army, The Iraq War, The Afghanistan War, Casualty, Hostile Incident, Logistic Regression

1. Introduction

U.S. troops have been involved in two major wars since 2001 and over 35,000 servicemen of the Army have been either injured or killed due to hostile incidents in Iraq and Afghanistan. The deaths and injuries of U.S. troops in Iraq and Afghanistan are well publicized, but casualty totals alone do not indicate the risk for a given serviceman. A study by Buzzel and Preston reveals that the risk of death in Iraq shows considerable variability. It is highest in the Marine Corps, lowest in the Air Force, higher among enlisted troops than officers, much higher for men than women, and declines sharply with age. Hispanics have a higher death rate than non-Hispanics, and blacks have unusually low mortality in Iraq. (Buzzel and Preston, 2007). There are few publicized studies concerning the hostility casualties of U.S. troops in the two major wars the U.S. has been involved in since 2003. Both studies, The differential impact of mortality of American troops in the Iraq War: The non-metropolitan dimension by Curtis and Payne and Mortaility of American Troops in the Iraq War by Buzzel and Preston include only deaths that occurred between 2003 and 2007 and do not differentiate causes of death. These deaths include both combat and non-combat related and covers only the incident that happened the Iraq War, whereas, in our study, we focus on combat-related casualties only, including death and injury, and cover incidents that happened in either Iraq or Afghanistan between 2003 and 2011. Both studies cover the Marine Corps, Air Force, Army, and Navy. In our study, we focus on the combat-related casualties of Army servicemen. The purpose of this study is to create a profile of U.S. troops killed or injured due to hostile incidents in Afghanistan and Iraq. To assess the risk to a serviceman who is exposed to a hostile incident, a multivariate analysis of factors causing these casualties will produce a better understanding of this least desired outcome of war. Our study answers the following research questions.

- Does the number of deployments of servicemen have an effect on casualty status?

- Does experience (time in service) have an effect on casualty status?
- Is there a significant difference in casualty rates among the combat, combat service, and combat-service support branches in the Army?
- Is there a significant difference in casualty rates among the regular, guard, and reserve forces in the Army?
- Is there a significant difference in casualty rates among the ranks in the Army?
- Do demographics like age, gender, and marital status have any effect on casualty status?

2. Descriptive Statistics

The file used in this study was obtained from the Defense Manpower Data Center (DMDC). It was built from active-duty personnel extract files, covering the period from 2003 to 2011. To avoid the official limitations of using SSNs(Social Security Number), we received a file arranged according to an identification number for each individual rather than by SSN. The file contained 48,312 records of servicemen either injured or killed and 98,812 records of servicemen who served in the 2003–2011 period without injury. In our data, we used only army servicemen killed or injured in hostile action, of whom there were 35,698.

2.1. Data Description by Number of Casualties

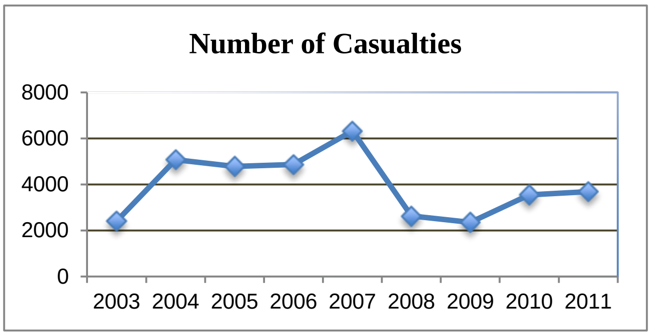

Fig. 1. Number of Casualties

Figure 1 shows the total number of hostile casualties that occurred between 2003 and 2011 in Iraq and Afghanistan. The number of hostile casualties among U.S. troops more than doubled from 2003 to 2004. In 2003, the number was 2,411 and by 2004, this number reached 5,079. In the following three years, the number of casualties did not change considerably. In 2007, the number of casualties increased by 1,461, reaching 6,324. This increase can be partly explained by the increased number of troops deployed to Afghanistan and Iraq, and is the highest number of casualties in 2003–2011. The number of casualties dropped significantly in 2008 to 2,631 and the following year reached its lowest value, 2,356. From 2009, the casualty number went up to around 3,500 for the following two years. This change can be explained by the change in focus of U.S. troops, from Iraq to Afghanistan.

2.2 Data Description by Incident Country

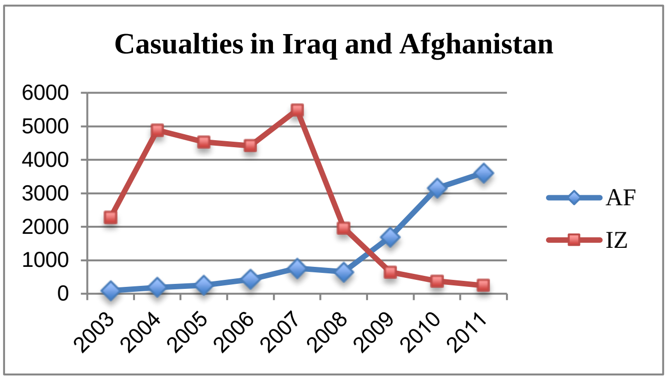

Fig. 2. Casualties in Iraq and Afghanistan

Figure 2 shows the number of casualties in Iraq and Afghanistan. There was a significant change in casualty numbers in Iraq in 2008, with a 65% drop in hostile casualties. The number decreased until 2011 when the U.S. left Iraq officially. The picture in Afghanistan is a little different. As the number of troops incremented, the number of casualties increased as well, specifically, after 2008.

2.3. Data Description by Pay Grade

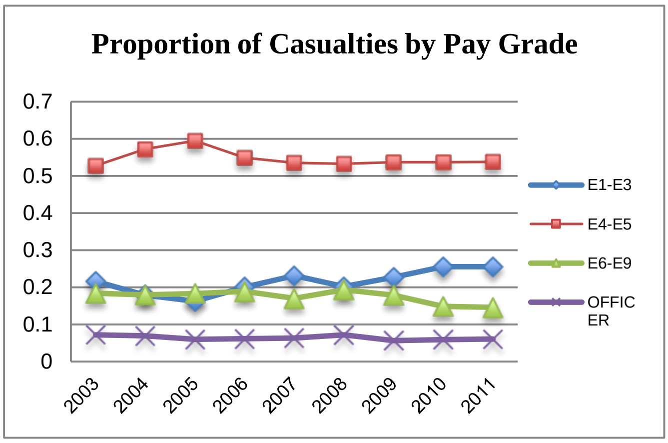

Fig. 3. Number of Casualties

Figure 3 shows casualty rates by pay grade in the U.S. Army between 2003 and 2011. Of all casualties, pay grades E1 (Private), E2 (Private), and E3 (Private First Class) (shown by E13, make up 21%, E4 (Specialist or Corporal) and E5 (Sergeant) (shown by E45) made up 55%, E6 (Staff Sergeant), E7 (Sergeant First Class), E8 (First Sergeant) and E9 (Sergeant Major) (shown by E69) made up 18%, and all officers makes up 6%. For the officers shown by O, the casualty ratio changed between 6% and 7% between 2003 and 2011. It reached its highest number in 2007, in which 401 officers were killed or injured. After 2008, there was a steady increase in both the number and ratio of casualties.

2.4 Data Description by Number of Combat (C), Combat Support (CS), and Combat Service Support (CSS) Troop Casualties

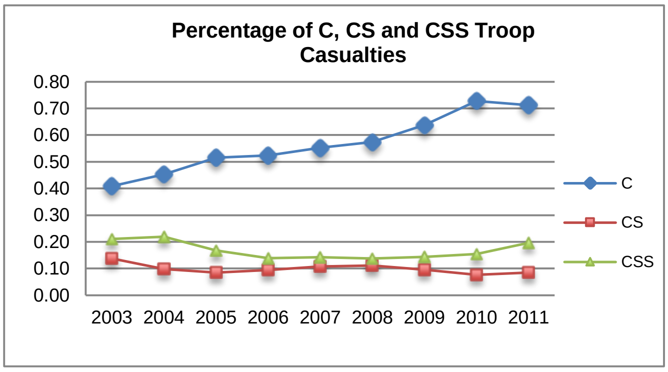

Fig. 4. Percentage of C, CS and CSS Troop Casualties

Figure 4 presents the casualty numbers of C, CS, and CSS troops serving in Iraq and Afghanistan between 2003 and 2011. For the CSS troops, the casualty percentage changed between 21% and 14%. It reached its highest value in 2004, with 1,115 troops killed or injured. For CS troops, the casualty percentage ranged between 8% and 14%. It reached its highest value in 2007, with 680 casualties. For the C troops, the casualty percentage changed between 41% and 71%. Its highest value was in 2007, with 3,492 casualties.

2.5. Data Description by Average Years in the Army

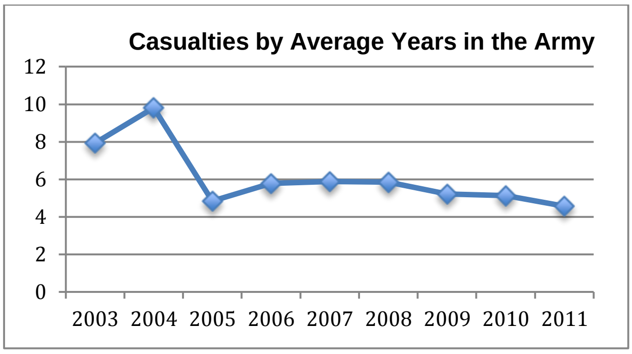

Fig. 5. Casualties by Average Years in the Army

Figure 5 presents the time in Army service before members were killed or injured. In the first years of the wars, years spent in the Army were above eight, but in 2005, the number dropped to 4.85 years. Between 2005 and 2008, the time in service before injury increased slightly to 5.85 years, but after that year, the number declined steadily and reached its lowest value, 4.57 years, in 2011.

2.6. Data Description by Organization Component Code

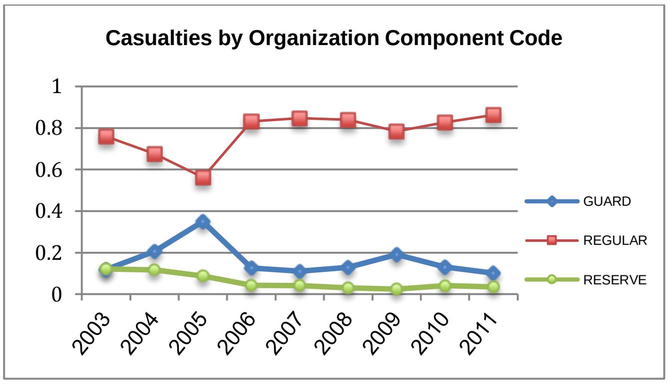

Fig. 6. Casualties by Organization Component Code

Figure 6 presents hostile casualty numbers and percentages of regular, reserve, and guard forces that served in Iraq and Afghanistan between 2003 and 2011. In 2003, 76% of casualties were regular forces and the ratio of guard forces and reserve forces was about the same. The casualty ratio of reserve forces decreased steadily up to 3.5% in 2011. For the guard forces, there was a big jump in casualties between 2003 and 2005, from 11.8% to 35%. The average casualty ratio of regular forces was 77.6% between 2003 and 2011. There had been a sharp decline until 2005, dropping to 56%, but in the same period there was a sharp increase in casualties among guard forces. A close look at the regular and guard forces reveals a symmetrical pattern. Whenever there is a change in the ratio of one force, there is an equal and opposite change in the other force.

2.7. Data Description by Average Number of Deployments

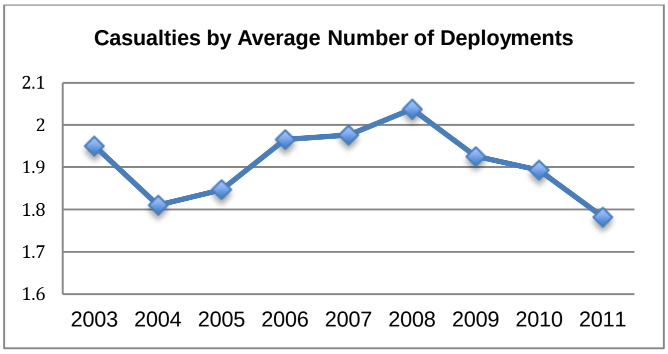

Fig. 7. Casualties by Average Number of Deployments

Figure 7 shows the average number of deployments of U.S. troops before servicemen were injured or killed. The overall average number was 1.91 and fluctuated between 1.78 and 2.03 from 2003 to 2011. In 2004, the number dropped to 1.81 and after that year climbed constantly until 2008, reaching its highest value, 2.03 deployments. Following that year, the number of deployments decreased steadily until 2011, reaching its lowest value, 1.78 deployments. The declining trend after 2008 can be partly explained by an overall change in the U.S. Army.

2.8. Data Description by Average Age

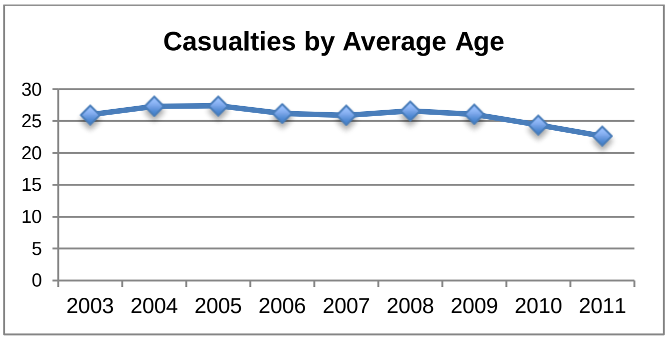

Fig. 8. Casualties by Average Age

Figure 8 presents the average age of U.S. troops injured or killed in Iraq or Afghanistan. The average age was 25.83 years between 2003 and 2011. We observe no significant change until 2010, when the average casualty age fluctuated between 26 and 27, but in 2010, the average casualty age dropped to 24.4. In the following year, the average age decreased to 22.6 years. In this figure, we see a pattern similar to that observed in deployment numbers and years spent in the Army. All three variables decreased after 2008.

2.9 Data Description by Gender

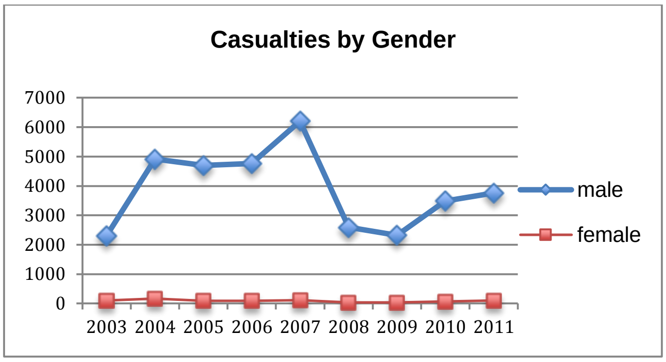

Fig. 9. Casualties by Gender

Figure 9 presents the gender of U.S. troops injured or killed in Iraq or Afghanistan between 2003 and 2011. Ninety-eight percent of the casualties were male and 2% were female. This discrepancy may be explained by the fact that female service members were not allowed to serve as combat troops, which kept them away from exposure to hostile attacks.

2.10. Data Description by Marital Status

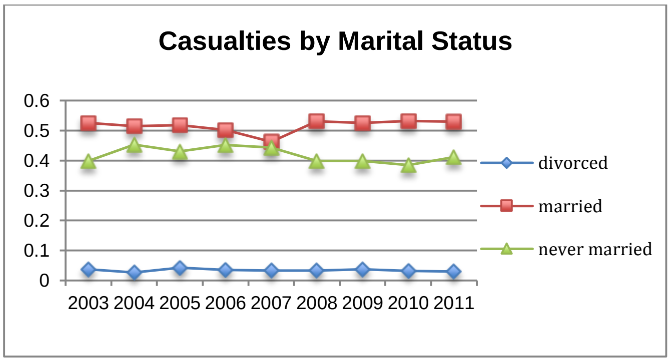

Fig. 10. Casualties by Marital Status

Figure 10 presents the marital status of U.S. troops injured or killed in Iraq and Afghanistan between 2003 and 2011. There was not a significant observation in the ratio of divorced service members. The ratio of casualty changed between 3% and 4%. Married service members made the highest proportion with a ratio of over 50%. But this can be related with the proportion of married service members within the army. The casualty rate moved between 46% and 53.2%.

3. Analytical Method

Regression methods are an integral part of any data analysis concerned with describing the relationships between a response variable and explanatory variables (Smith and Campbell, 1980; Wei et al., 2006). Logistic regression is a well-known statistical technique for modeling data with binary outcomes (Hosmer and Lemeshow, 2000; Menard, 2002). We wish to analyze the reasons of causality status (killed/injured or not). Therefore, logistic regression is preferred, because the outcome of casualty is binary. The dependent variable for every observation (i) is defined as Y_i, which is coded 1 if the servicemen is injured or killed, and 0 otherwise.

The theoretical model is:

$$\log(p_i / (1-p_i)) = b_0 + b_1 X_1 + \ldots + b_n X_n \qquad (1)$$

where,

$\log(p_i / (1-p_i))$ = Log of odds ratio for individual i

p_i = Probability injured or killed

b_0 = Intercept

b = Estimated coefficient (change in log odds for a unit change in X s)

X = Values of explanatory variables

The coefficients in the model represent the change in the log odds for a unit change in an X covariate. The X s capture the various demographic characteristics for the individuals, such as marital status, age, occupation component code, time in service, and rate of deployment. In logistic regression, the log-odds are generally assumed to be a linear function of various covariates.

The odds are defined as the probability that an individual with a particular set of characteristics was injured or killed in a hostile action, divided by the probability that he was not (Paisant, 2008). The odds can be any number between zero and infinity. Odds of one mean that a serviceman with a set of characteristics is equally likely to get injured or killed. Odds greater than one mean that such a serviceman is more likely to get injured or killed, while odds less than one mean the serviceman is less likely to get killed or injured (Fricker and Buttrey, 2008).

In our study, we analyzed two different models. In the first model, which we took as a base model, we have the variables without interactions, and in the second model, which is the alternative model; we have the interactions derived from the stepAIC function in R (Yamaoka, et al., 1978), which performs stepwise model selection by Akaike Information Criterion (AIC). (Venables and Ripley, 2002). In order to choose the best model, we performed pseudo-R^2 and Hosmer-Lomeshow tests.

Pseudo-R^2 test (Veall and Zimmermann, 1996) basically measures the percent of the variation in the dependent variable depending on changes in explanatory variables. The pseudo-R^2 is the logistic regression analog to the R^2 in linear regressions. It measures the proportion of deviance accounted for by the regression (Nagelkerke, 1991). Hosmer-Lomeshow test is a statistical test for goodness of fit for logistic regression models. The test assesses whether or not the observed event rates match expected event rates in subgroups of the model population. Hosmer and Lemeshow (2000) recommend partitioning the observations into 10 equal sized groups according to their predicted probabilities. Then,

$$G^2_{HL} = \sum_{j=1}^{10} \frac{\left(O_j - E_j\right)^2}{E_j\left(1 - E_j / n_j\right)} \sim \chi^2_8 \qquad (2)$$

where,

$E_j = \sum_i y_{ij}$ = Expected number of cases in the j^{th} group

n_j = Number of observartion in the j^{th} group

$O_j = \sum_i y_{ij}$ = Observed number of cases in the j^{th} group (Hosmer, 2000)

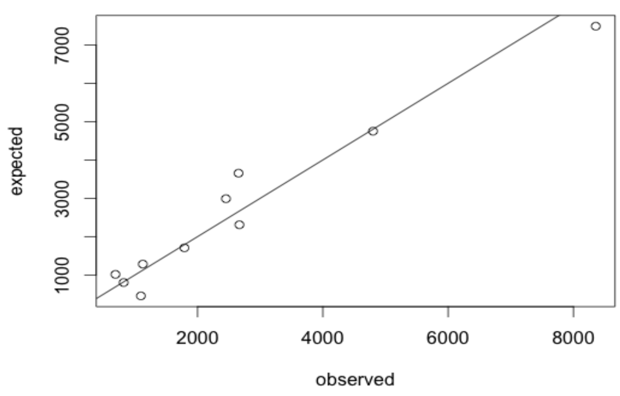

Fig. 11. Hosmer-Lomeshow test result for the base model

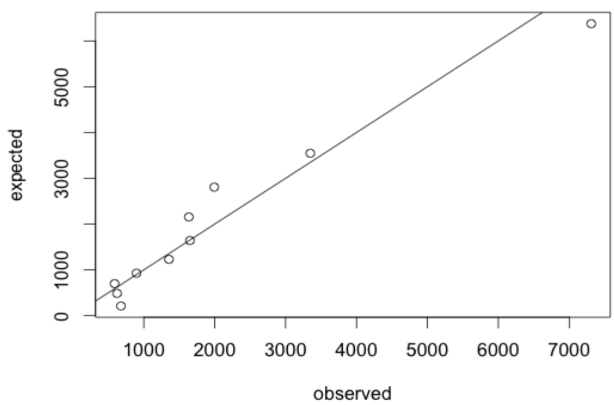

Fig. 12. Hosmer-Lomeshow test result for the alternative model

The pseudo-R^2 value for the base model is 0.17 and is 0.22 for the alternative model. Evaluating the two models, we see that pseudo-R^2 value is higher for the alternative model. The results of Hosmer-Lomeshow tests are presented in Figure 11 for the base model and Figure 12 for the alternative model and we see that there is no significant difference between two figures. The p values for the both models are zero and chi-square values are 1941.7 for the base model and 2702.9 for the alternative one. Since a large p-value shows a good fit, neither of our models fit very well by this measure. Even though the pseudo-R^2 square value is higher for the model with interactions is better, we preferred to analyze the base model because the results of the base model are easier to interpret.

3.1. Dependent Variable

A binary variable was used to define a serviceman who was injured or killed. If the serviceman was killed or injured in a hostile action, this variable was given the value of 1 and the serviceman was classified as “killed or injured.” If the serviceman was not killed or injured, the variable was given the value of 0. In this study, no distinction was made between killed and injured. Both these cases were classified in the same category.

3.2. Explanatory Variables

3.2.1. Pay Grade

Pay grades E1, E2, and E3 are taken as a group and presented by E13. Pay grades E4 and E5 are taken as a group and indicated by E45. Pay grades E6, E7, E8 and E9 are taken as a group and shown by E69, and all officers are taken as a group and indicated by Officer. The E45 pay-grade group casualty status makes the 55% of all casualties, followed by the E13 pay-grade group with 21%. With those statistics, servicemen serving in these two pay grades are more likely to get injured or killed. Pay grades E69 are expected to be less involved in hostile actions than other pay-grade groups because they are more likely assigned to staff or headquarter positions. So fewer casualties for servicemen of pay grades E69 are expected.

3.1.2. Gender

This variable presents the sex of the serviceman. Male servicemen constitute the majority of the army and they are expected to be more involved in hostile actions. Females are not allowed to serve in combat troops, which decreases exposure to hostile attacks. This makes males more exposed to hostile incidents.

3.2.3. Age

This variable gives the age of the serviceman at the time of the incident. The general expectation is that a junior serviceman is more likely to get killed or injured than the senior serviceman, because juniors are assigned to posts involving more danger than are seniors.

3.2.4. Marital Status

This variable gives the marital status of the serviceman at the time of the incident. Marital status was categorized into the levels of married, never married, and divorced.

3.2.5. Injury Country

This variable explains where the incident happened. In this study, only hostile incidents in Iraq and Afghanistan were taken into account.

3.2.6. Organization Component Code

The U.S. Army consists of regular, reserve and guard forces; the organization component code was categorized into these three variables. Regular forces make up the majority of the army, and it is expected that regular forces are more likely to be killed or injured. Analyzing the univariate statistics in Figure 6, we see that 27% of regular forces, 25.9% of guard forces, and 23.2% of reserve forces were injured during the two wars, which is consistent with our anticipation.

3.2.7. Time in Service

This variable explains the total years served before a serviceman was injured or killed. Time in service was calculated with the difference between the death or injury date and enlistment date. It is expected that as serviceman gains more experience and training, he is better prepared for hostile actions and less likely to get killed or injured. But a contrary assumption would be that in the early years of service, a serviceman is assigned to posts involving more danger.

3.2.8. Military Occupation Specialties (MOS)

This variable explains the military occupation specialties of the servicemen. To simplify, infantry, armor, field artillery, combat-engineer and air-defense artillery were grouped as combat troops. The chemical corps, general engineering, military-intelligence corps, military-police corps, signal corps and Army aviation were grouped as combat-support troops. The remaining military occupation specialties were grouped as combat-service support troops. The general anticipation is that those serving as combat troops are more exposed to danger and more likely to get killed or injured.

3.2.9. Deployment Count

This variable explains the number of deployment a serviceman had at time of incident. The Deployment Count variable was categorized in five groups. The servicemen were categorized to group 1 through group 4 with respect to number of deployments and group 5 includes the servicemen with 5 or more deployments. The common-sense expectation is that as the deployment numbers go up, the serviceman gains more experience and orientation to the battle environment and is less likely to get killed or injured.

4. Analysis Results

The effect of each independent variable was determined by trying all one-term deletions from the base model using the dropterm function. (Venables, 2002). Based on the results of this function, we decide whether a variable is statistically significant to analyze or not.

Table 1. Dropterm function results for the base model

	Df	Deviance	AIC	LRT	Pr(Chi)
none		105485	105523		
Gender	1	105499	105535	13.3	0.000272 ***
Injury Country	1	105551	105587	66.2	4.132e-16 ***
Marital_ Status	2	105610	105644	124.8	< 2.2e-16 ***
Deployment Count	4	105809	105839	323.6	< 2.2e-16 ***
Pay_Grade	3	106874	106906	1389.1	< 2.2e-16 ***
Dependent Quantity	1	106944	106980	1458.9	< 2.2e-16 ***
Age	1	107746	107782	2260.9	< 2.2e-16 ***
Org_Comp Code	2	108256	108290	2770.2	< 2.2e-16 ***
MOS	2	114532	114566	9046.5	< 2.2e-16 ***
Time_In_ Service	1	115306	115342	9820.6	< 2.2e-16 ***

Table 1 shows the results of the dropterm function our model. The effects of the variables are presented in ascending order. We see that the Gender was the least effective while Time_In_Service was the most effective among all variables included in the model.

Table 2 shows the results of summary function for the base model. This model demonstrates the effects of injury country, marital status, gender, pay grade, organization component code, dependent-quantity code, time in service, number of deployments, and MOS.

Table 2. Summary function results for the base model

Term	Estimate	Std. Error	z value	Pr(>|z|)
Intercept	2.0586	0.08824	23.327	< 2e-16 ***
Marital_Status Code_M	0.2005	0.04172	4.807	1.53e-06 ***
Marital_Status Code_N	0.0087	0.04367	0.201	0.8406
Age	-0.0897	0.00200	-44.744	< 2e-16 ***
Gender_ M	-0.1820	0.04918	-3.701	0.000215 ***
Pay_Grade_E45	-0.1275	0.02160	-5.903	3.58e-09 ***
Pay_Grade_E69	-1.1026	0.03481	-31.669	< 2e-16 ***
Pay_Grade Officers	-0.4880	0.04100	-11.903	< 2e-16 ***
Org_Comp Code_R	-1.1560	0.02529	-45.704	< 2e-16 ***
Org_Comp Code_V	0.2075	0.03792	5.473	4.43e-08 ***
Dependent Quantity_Code	0.0203	0.00052	38.803	< 2e-16 ***
Time_Inservice	0.2305	0.00256	90.022	< 2e-16 ***
Deployment Count_2	-0.1590	0.01879	-8.460	< 2e-16 ***
Deployment Count_3	-0.4157	0.02473	-16.810	< 2e-16 ***
Deployment Count_4	-0.3791	0.03828	-9.905	< 2e-16 ***
Deployment Count_5	-0.1038	0.05489	-1.892	0.058432
MOS_CS	-1.4797	0.01952	-65.091	< 2e-16 ***
MOS_CSS	-1.5419	0.02273	-78.993	< 2e-16 ***
Injury_Country _Iraq	-0.1363	0.01671	-8.156	3.46e-16 ***

* significant at 10 %; ** significant at 5 %; *** significant at 1 %

4.1. Pay Grade

In our model, the base case for pay grade is E1–E3. In Figure 3, we see that the majority of casualties occurred in the E4–E5 pay grades, with 55% followed by the E13 pay-grade group with 21%. Our initial descriptive study is not very consistent with one of the oldest observations in the social sciences, that lower-ranking servicemen are exposed to a greater risk of death than higher-ranking individuals (Buzzel and Preston, 2007). The study by Buzzel and Preston also shows that in the Army, privates first class have a death risk three times greater than the combined categories of major, colonel, and general; and enlisted men have a 38% higher mortality than officers. In the Marine Corps, lance corporals have a death risk 4.18 times greater than that of majors, colonels, and generals, and the single highest mortality group in any service consist of lance corporals, whose death risk is 3.1 times that of all troops in Iraq.

The base case for pay grade is E1–E3. After adjusting to other casualties, in our multivariate model we see that all other pay grades have lower odds ratio. The odds ratio values are 0.88 for E4–E5, 0.66 for Officers and 0.33 for E6–E9 pay grades. With our new results, we may infer that all the pay grades have lower odds ratios of getting killed or injured than pay grades E1–E3. But that difference is not as large as has been mentioned in previous studies.

4.2. Marital Status

The base case for marital status in our model is divorced servicemen. In descriptive statistics, we see that 27.7% of married, 28.7% of never married, and 22% of divorced servicemen were either injured or killed. Our model results show that a married serviceman is more likely to get injured or killed than both divorced and never married serviceman. The odds ratios for divorced and never married servicemen are very close that we can say that there is no significant difference between those two. In our model, we may conclude married servicemen are more likely to get injured or killed.

4.3. Gender

The base case for gender in the model is female. In Figure 9, we see that 98% of casualties were males. In both wars, 29.2% of male and 18.6% of female servicemen were injured or killed. With these statistics, the casualty of a male serviceman is more likely to occur, but in our model we perceive what is contrary to general expectation. A male serviceman has a lower odds ratio than an otherwise female serviceman. We may interpret that a male serviceman is less likely to get killed or injured due to hostile incident. The reasons for this result might be the physical capacity, better training and higher orientation to the conditions of the battlefield.

4.4. Injury Country

In our model, the base case for country of injury is Afghanistan. The number of casualties in Iraq was more than twice the number that occurred in Afghanistan. Dividing the casualty numbers by the total number of troops served, we get the ratio of 29.4% casualty percentage for troops who served in Iraq and 27.1% for troops who served in Afghanistan. In multivariate analysis, a serviceman serving in Iraq has an odds ratio of 0.87, which is very close to 1. We may infer that there is not much difference in the casualty status between the two countries.

4.5. Organization Component Code

The base case for organization component code in our model is guard forces. In the descriptive-statistics part, we see that 29.2% of regular forces, 28.3% of guard forces and 25.5% of reserve forces were either injured or killed. These ratios looks logical because regular forces are expected to be more exposed to hostile attacks than the other two groups. In our model, we see that regular forces are less likely to get killed or injured than other forces. The reserve forces have the highest odds ratio, which means they are predicted by the model more susceptible to death or injury. The level of training the forces get might explain this contradiction. Regular forces have more experience in the field and are more accustomed to the conditions of the battlefield, while the other two forces might have less experience.

4.6. Military Organization Specialties

The base case for MOS in our model is combat troops, which make the majority in the sample data. In our model, we see that the odds ratios for combat support and combat-service support troops indicate that they are less likely to get injured or killed than combat troops. This result looks convincing, because CS and CSS troops do not engage the enemy as C troops do.

4.7. Military Organization Specialties

The base case for Deployment Count in our model is the first deployment. Interpreting the coefficients in Table 2, we see that servicemen in their first deployment are more likely to get killed or injured. Servicemen with two deployments have odds ratio of 0.85 and servicemen with three and four deployments got nearly the same odds ratios, which are about 0.66. Even though the odds ratio for five or more deployments is 0.90, the *p* value is 0.06, which is statistically insignificant. We may interpret that as the number of deployments increases, the servicemen are less likely to killed or injured. The reason for this result might be that as the number of deployment increases, the serviceman gains more experience and is better oriented to the battlefield. Another factor is the fact that the pay grade of the servicemen increases as the number of deployments increases. Our pay grade analysis suggests that servicemen with higher pay grades are less likely to involve in hostile incidents.

4.8. Age

Fig. 13. Probability of Casualty of MOS vs. Age

Figure 13 shows the probability of casualty of MOS with respect to age. We got Figure 13 with the help of the predict function in R. Predict is a generic function for predictions from the results of various model fitting functions. The function invokes particular *methods*, which depend on the class of the first argument. (Chambers and Hastie, 1992) The first argument in our study was a male never married serviceman who was 28 years old and serving in Afghanistan. His pay grade was E45, and was serving as regular force in CSS troops. He had been serving in the Army for eight years and was on his second deployment. This particular serviceman was replicated for a number of times and was fitted in our model to see the effect of time in service on MOS. Figure 13 shows the probability of casualty of MOS with respect to the time they served. We see that the probability of casualty decreases for all troops as the age of the servicemen grows. C troops have the highest probability of casualty in all ages. CS and CSS troops have about the same probability of being killed or injured. As the age grows, we see that the gap between C and other two troops decreases. This might be explained by the fact that as the age grows C troops serviceman might be assigned to posts, which are less susceptible to hostile incidents.

5. Conclusion

The deaths and injuries of U.S. troops in Iraq and Afghanistan are well publicized, but casualty totals alone do not indicate the risk for a given serviceman. Decision makers need to see profile that is most vulnerable to casualty and take precautions to minimize that risk. To assess the risk to a serviceman who is exposed to a hostile incident, a multivariate analysis of factors causing these casualties produces a better understanding of this least desired outcome of war.

Our multivariate model suggests that pay grades E1 through E3 are more likely to be involved in hostile incidents than other pay-grade groups, and this exposure to danger decreases as the pay grades increase. The findings for gender are not parallel to popular ideas. It is expected that males will be more exposed to hostile incidents than females. But even though the difference is not large, our results show that males are less likely to get killed or injured. In terms of the effects of marital status, our study shows that married servicemen are more likely to be involved in hostile incidents. In our model, we found that regular forces have a lower risk of engaging in hostile incidents than guard and reserve forces, which is contrary to general expectation. The results for MOS were as expected. Combat troop are more likely to be killed or injured than other troops.

References

Buzzel,E. & Preston, S. (2007). Mortaility of American Troops in the Iraq War. Population and Development Review.

Chambers, J. M. & Hastie, T. J.(1992). *Statistical Models in S*. Wadsworth & Brooks/Cole.

Curtis, J. & Payne, F. (2010). The differential impact of mortality of American troops in the Iraq War: The non-metropolitan dimension. Demographic Research.

Fricker, R. D., & Buttrey, S. (2008). Assessing the Effects of Individual Augmentation (IA) on Active Component Navy Enlisted and Officer Retention. Monterey, CA.

Hosmer, D. W. & Lemeshow, S. (2000). Applied Logistic Regression. Wiley.

Hosmer, D. W., & Lemeshow, S. (2000). Model-Building Strategies and Methods for Logistic Regression. Applied Logistic Regression, Second Edition, 91-142.

Menard, S. (2002). Applied logistic regression analysis (No. 106). Sage.

Nagelkerke, N. J. (1991). A Note On A General Definition Of The Coefficient Of Determination. Biometrica.

Paisant, M. A. (2008). The effects of Individual Augmentation (IA) on Navy junior officer retention. Naval Postgraduate School Monterey Ca.

Smith, G., & Campbell, F. (1980). A critique of some ridge regression methods. Journal of the American Statistical Association, 75(369), 74-81.

Yamaoka, K., Nakagawa, T., & Uno, T. (1978). Application of Akaike's information criterion (AIC) in the evaluation of linear pharmacokinetic equations.Journal of pharmacokinetics and biopharmaceutics, 6(2), 165-175.

Veall, M. R., & Zimmermann, K. F. (1996). Pseudo-R2 Measures For Some Common Limited Dependent Variable Models. Journal of Economic Surveys, 10(3), 241-259.

Venables, W. N. & Ripley, B. (2002). Modern Applied Statistics with S. Springer.

Wei, Y., Pere, A., Koenker, R., & He, X. (2006). Quantile regression methods for reference growth charts. Statistics in medicine, 25(8), 1369-1382.