Forecasting Wounded-In-Action Casualty Rates for Ground Combat Operations

NAVAL HEALTH RESEARCH CENTER

Edwin D'Souza[1,2]
James Zouris[2]
Trevor Elkins[1,2]
Vern Wing[1,2]
LCDR Andrew Olson[2]

[1]Leidos, San Diego, CA;
[2]Naval Health Research Center, San Diego, CA

Report No. 20-32 was supported by Navy Bureau of Medicine and Surgery under work unit no. N1214. The views expressed in this article are those of the authors and do not necessarily reflect the official policy or position of the Department of the Navy, Department of Defense, nor the U.S. Government.

Executive Summary

The methodology to calculate wounded-in-action (WIA) casualty rates for ground combat operations was updated to implement recommendations from the Defense Health Agency Casualty Rate Development Work Group. The new model will be implemented in the Casualty Rate Estimation Tool (CREstT), which resides within the Department-of-Defense-accredited Medical Planners' Toolkit. Data were collected from various sources for major combat and contingency operations, including World War II, Korea, Vietnam, Iraq, and Afghanistan. Building upon Trevor Dupuy's extensive research on casualty rate estimation, we explored the effects of factors such as duration, population at risk or force size, terrain, climate, enemy capability, and tactical operation on WIA rates. We developed several regression and ensemble models, and selected the Ridge Regression model to estimate ground combat WIA rates. The Ridge Regression model outperformed the current model in CREstT (where a baseline-derived WIA rate is modified using adjustment factors) as well as Dupuy's model of estimating casualties. This work will refine WIA casualty projections in CREstT and improve resource estimates for medical planners.

Acknowledgments

We are indebted to Robert Syverston and James Young for their feedback over the course of the study.

We thank Carrie Brown from the Naval Health Research Center for editorial review.

Table of Contents

Background

The United States Department of Defense (DoD) medical planners need accurate forecasts for the number of wounded-in-action (WIA) casualties to reliably project ground combat operations' personnel and materiel resources. The Casualty Rate Estimation Tool (CREstT) within the Medical Planners' Toolkit (MPTk) is currently the only DoD-accredited tool that medical planners use to estimate combat and non-combat injuries and illnesses during the operational plan development.

In October 2015, as part of the accreditation process for MPTk, the Defense Health Agency Casualty Rate Development (CaRD) Work Group presented the following recommendations to the Strategic Analysis Working Group (SAWG) for future enhancements to CREstT:

- Expand the number of data points in the baseline WIA rate calculation.
- Develop data to replace the Battle Intensity selections with Mission Sets incorporating Tactical Operations and Tactical Actions—in accordance with guidance the medical planner is likely to receive from the Joint Staff J-3 (Operations) and J-5 (Strategic Plans and Policy) communities.
- Modify the WIA adjustment factor for Region to account for Enemy Capability within a combatant command region.
- Replace the concept of Troop Type with Battlespace or Operational Environment to ensure consistent terminology between warfighters and medical planners.

These recommendations are captured in Appendix A and the current version of CREstT (MPTk Version 1.3) has already implemented the last recommendation to replace Troop Type with Operational Environment.

The baseline WIA rates in CREstT are based on the user-selected Operational Environment (number of troops located in the Close Area, Support Area [Forward], and Support Area [Rear]) and Battle Intensity (Intense, Heavy, Moderate, Light, Peace Operations, and None). Battle Intensity was based on the overall evaluation of mission risk, weather and terrain, percentage of troops engaged, and the strength of the enemy forces.

The baseline WIA rates currently in CREstT were derived using seven data points:

- Operation New Dawn (Iraq after September 2010) and Operation Iraqi Freedom (OIF) 2008 provided Peace Operation intensity rates
- Operation Enduring Freedom (OEF) in Afghanistan 2010 provided Light intensity rates
- Vietnam and Korea provided Moderate intensity rates
- The Second Battle of Fallujah (OIF, November 2004) provided Heavy intensity rates
- The Battle of Hue (Tet Offensive in Vietnam) provided Intense rates

CREstT modifies the baseline WIA rates using user-input adjustment factors to calculate the WIA rate for the operation. WIA rate is expressed as the number of WIA casualties per 1,000 population at risk (PAR) or force size per day. The adjustment factors are Region, Terrain, Climate, and PAR. All the adjustment factors, including Operational Environment, are scalars that are applied multiplicatively to the baseline WIA rate (Zouris, D'Souza, Honderick, Tolbert, & Wing, 2013; Blood, Zouris, & Rotblatt, 1997). The methods used to estimate an operation's

WIA casualty rate are documented in the MPTk Methodology Manual (Naval Health Research Center, 2018).

The primary objective of this study was to implement the CaRD Work Group's recommendations to improve CREstT's WIA casualty rate forecasts for ground combat operations. The Navy Bureau of Medicine and Surgery sponsored this study in March 2017

Approach

Naval Health Research Center (NHRC) analysts collected 291 data points from numerous sources spanning major combat operations in World War II, Korea, Vietnam, Iraq, and Afghanistan, as well as contingency operations, including Grenada, Somalia, and Panama. This satisfied the CaRD Work Group's recommendation to expand the number of raw data points for estimating baseline WIA rates.

Various data attributes were collected for the 291 operations, including WIA rate, PAR, Duration, Climate, Terrain, Posture, Enemy Capability and Tactical Operation. Collecting Enemy Capability and Tactical Operation data for each operation ensured that the new model would satisfy the CaRD Work Group's remaining recommendations (to include enemy capability, and to replace battle intensity with mission sets in the WIA rate calculation).

Instead of using the current approach of modifying a baseline-derived WIA rate with adjustment factors, we developed various regression and ensemble models (Hastie, Tibshirani, & Friedman, 2009) to forecast the WIA rate. Ensemble models are prediction models that combine the strength of a collection of simpler base models. The final model provided the best predictions on out-of-sample data.

Limitations and Assumptions

- Operational Environment data were not available for the 291 data points. Empirically-derived scalar values for the three operational environments (Close Area, Support Area (Forward), and Support Area (Rear)) were applied multiplicatively to the WIA rate computed by the regression model. These Operational Environment scalar values were derived for each Enemy Capability: the WIA rates for each Operational Environment differ among the major combat operations.
- The Mission Sets concept proposed by the CaRD Work Group to replace Battle Intensity includes both Tactical Operation and Tactical Action data. Tactical Operation values were limited to Defensive, Offensive, Stability, Sustainment, Counterinsurgency, and Peace Operations. No combat mission data were available for Peace Operations. Tactical Action data were limited to the Korean War operations, with only a subset of the tactical actions (depicted in Appendix 1) available. Ten tactical actions were documented: Amphibious Assault; US Attack, Heavy Resistance; US Attack, Light Resistance; US Counterattack, Heavy Resistance; US Counterattack, Light Resistance; Airdrop Attack; Assault River Crossing; Defense, enemy attack or counterattack; Mopping Up; and Patrolling. Tactical actions only apply to Offensive and Defensive tactical operations, and scalars were derived for each tactical operation and tactical action combination. These scalars were applied multiplicatively to the WIA rate computed by the regression model.

- The PAR or force size ranges from 800 to 45,000. If the user inputs a PAR less than 800, the regression model uses the PAR lower limit (800) to compute the WIA rate. Likewise, if the user inputs a PAR greater than 45,000, the regression model uses the PAR upper limit (45,000).
- The duration of an operation ranges from 1 to 180 days. If the user inputs a longer duration greater, the regression model defaults to the upper limit of 180 days to compute the WIA rate.

Methodology

Overview

The methodology has five stages that proceed sequentially: data collection, exploratory data analysis, model development, model selection, and model comparison (see Figure 1).

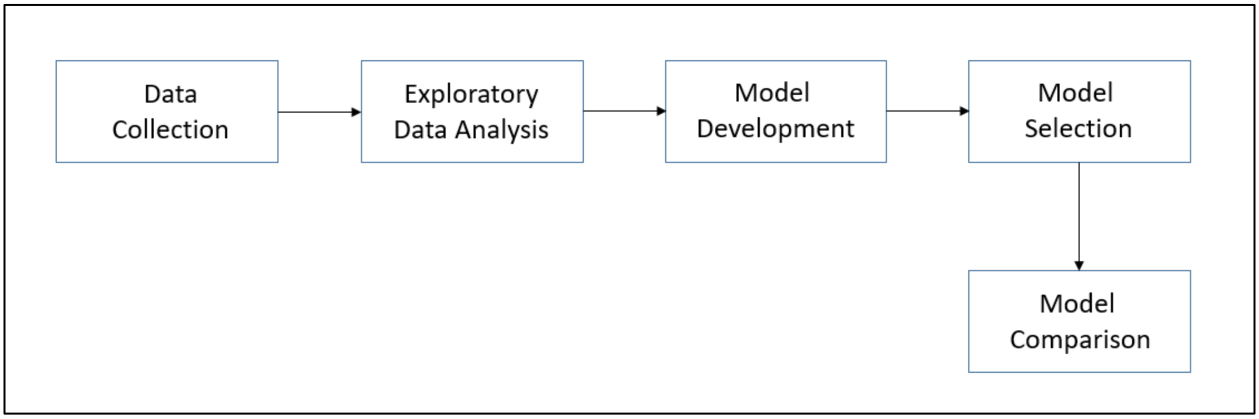

Figure 1. Methodology overview.

In the data collection stage, combat operational data were collected from several sources and cleaned. This was followed by exploratory data analysis, where the effects of factors such as terrain, climate, duration, PAR, enemy capability, posture, and tactical operation on WIA rates were explored. Using machine learning algorithms, several regression and ensemble models (Hastie, Tibshirani, & Friedman, 2009) were developed. The best model was selected using 10-fold cross validation; the Root Mean Squared Error (RMSE) on out-of-sample data was used as the metric for selection. Finally, the selected model's performance was compared with CREstT's current version (MPTk Version 1.3) and Trevor Dupuy's casualty estimation methodology (Dupuy, 1990).

The following five sections describe the methodology in more detail.

Data Collection

As recommended by the CaRD Work Group, several sources were examined to add more raw data points to forecast WIA rates. Only operations from World War II on were considered relevant to model WIA rates for future ground combat operations. Combat warfare today is vastly different from combat warfare a century ago in terms of weapon lethality, tactics, technology, casualty evacuation, medical care, and personal protective equipment (Holcomb, Stansbury, Champion, Wade, & Bellamy, 2006).

The data were cleaned and verified for accuracy before analysis. The data cleaning process included standardizing, recoding, and grouping the data into similar categories.

Table 1 summarizes the sources used to collect post-World War II data. The duration for each operation listed denotes the period for which we had data, not the actual operation period. Only operations with battalion-sized units or larger were recorded—data were not available for operations with units smaller than a battalion. Beebe and De Bakey (1952) provided data for World War II, Reister (1973) for the Korean War, Thayer (1975) for Vietnam, and Dupuy, Brinkerhoff, Bader, and Johnson (1985) for assorted contingency operations, such as those in Grenada, Somalia, and Panama. Detailed Vietnam operational situation reports were obtained from the Virtual Vietnam Archive (1965–1970). The Defense Casualty Information Processing System (DCIPS) and the Defense Manpower Data Center's (DMDC) Contingency Tracking System Deployment file provided data for combat operations in Iraq and Afghanistan.

Table 1
Combat Operations Post-World War II Data Sources for Available Time Periods

Military combat operation	Data source/reference
World War II (1940–1945)	Beebe & De Bakey (1952)
Korean War (1950–1953)	Reister (1973)
Vietnam War (1965–1970)	Thayer (1975), The Virtual Vietnam Archive (1965–1970)
Iraq War (2003–2007)	DCIPS, DMDC Deployment File
Afghanistan War (2004–2013)	DCIPS, DMDC Deployment File
Assorted contingency operations	Dupuy, Brinkerhoff, Bader, & Johnson (1985)

Data were collected for 291 combat operations: 40 from World War II, 44 from the Korean War, 124 from Vietnam, 39 from Iraq, 28 from Afghanistan, and 16 from assorted contingency operations.

The opposition force for each data point was assigned an Enemy Capability. Operations in World War II and Korea were classified as Near Peer (n = 88), those in Vietnam as Hybrid (n = 130), parts of the conflict in Iraq as Failed State (n = 16), and operations in Afghanistan, parts of Iraq, Grenada, Somalia, and Panama as Asymmetric (n = 57).

The Tactical Operation attribute had six levels: Defensive, Offensive, Stability, Sustainment, Counterinsurgency, and Peace Operations. These values were based on the United State Marine Corps doctrinal guidance for combat operations (Marine Corps Doctrinal Publication, 2017).

The Posture binary variable, consisting of either Attacker or Defender, was included in the dataset based on an analysis of the U.S. Army Concepts Analysis Agency (CAA) Database of Battles (1990). The CAA database had 1,320 data points: 660 attackers and 660 defenders. For battles occurring between 1600 and 1973, attackers generally suffered lower casualty rates compared with defenders (Figure 2).

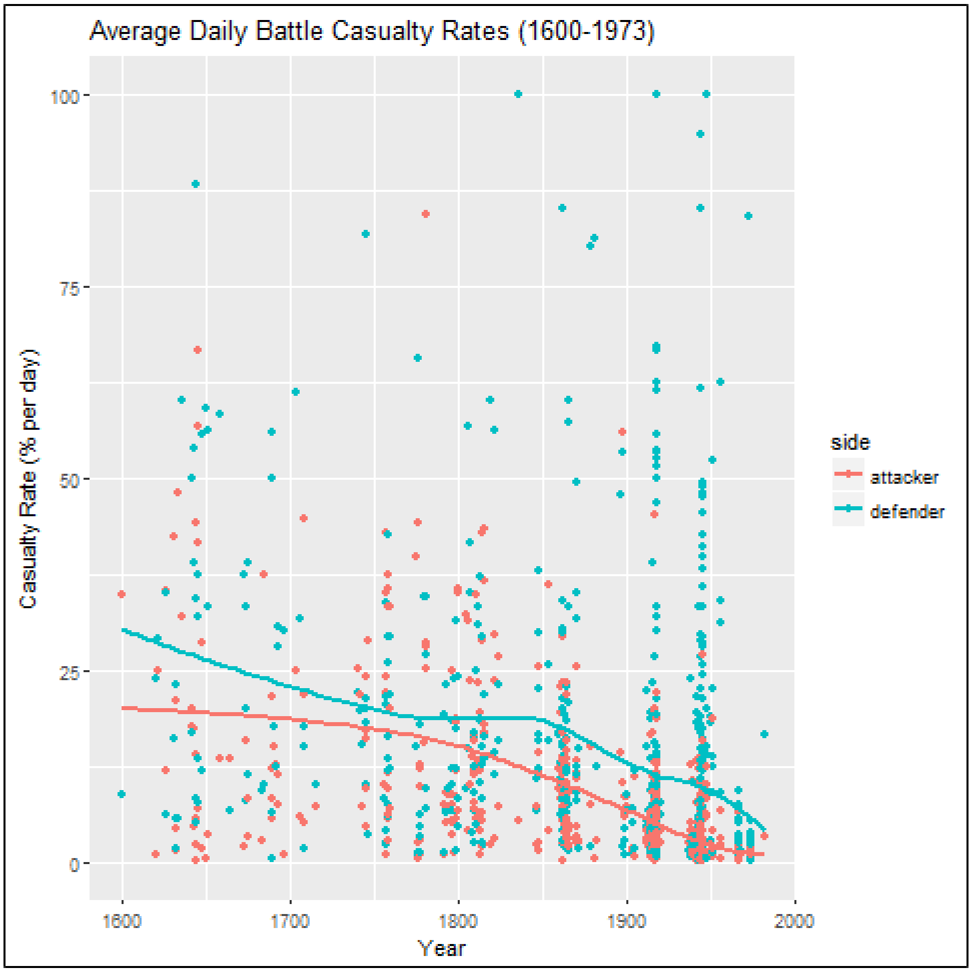

Figure 2. Casualty rates over time for operations in the U.S. Army CAA Database of Battles (1990).

Table 2 summarizes the attributes in the analysis dataset and lists the value ranges for numeric variables and discrete values for the categorical variables

Table 2

Data Attributes for 291 Combat Operations

Data attribute	**Type (values)**
PAR	Numeric (800–45,000)
Duration (days)	Numeric (1–180)
Service	Categorical (Army, USMC, Joint, other)
Terrain	Categorical (urban, rolling, flat, swamp, rugged)
Climate	Categorical (temperate, hot, cold)
Region	Categorical (SOUTHCOM, EUCOM, AFRICOM, CENTCOM, PACOM)
Posture	Categorical (attacker, defender)
Enemy Capability	Categorical (near peer, hybrid, failed state, asymmetric)
Tactical Operation	Categorical (defensive, offensive, stability, sustainment, counterinsurgency, peace operations)

Exploratory Data Analysis

Exploratory data analysis (Tukey, 1977) effectively summarized the data, and provided useful insights into the relationships among the data attributes.

Figure 3 plots histograms of the baseline WIA rate and the logarithm of the baseline WIA rate side-by-side. The baseline WIA rate is positive skewed, while the logarithm of the baseline WIA rate is bell-shaped and reasonably normal.

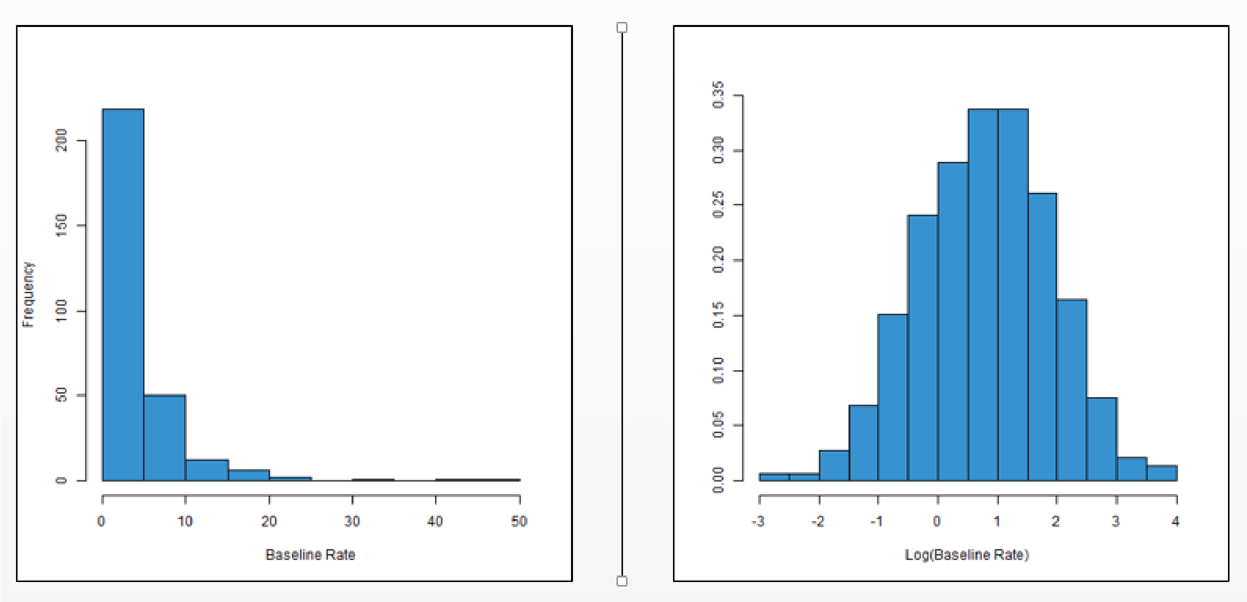

Figure 3. Histograms of the baseline WIA rate (left) and the logarithm of the baseline WIA rate (right).

We modeled the more normally distributed logarithm of the WIA rate (rather than the skewed baseline WIA rate) as the dependent variable in the regression and ensemble models.

Figure 4 displays a scatter plot of the WIA rate logarithm against the PAR logarithm juxtaposed with a scatter plot of the WIA rate logarithm against the Duration logarithm.

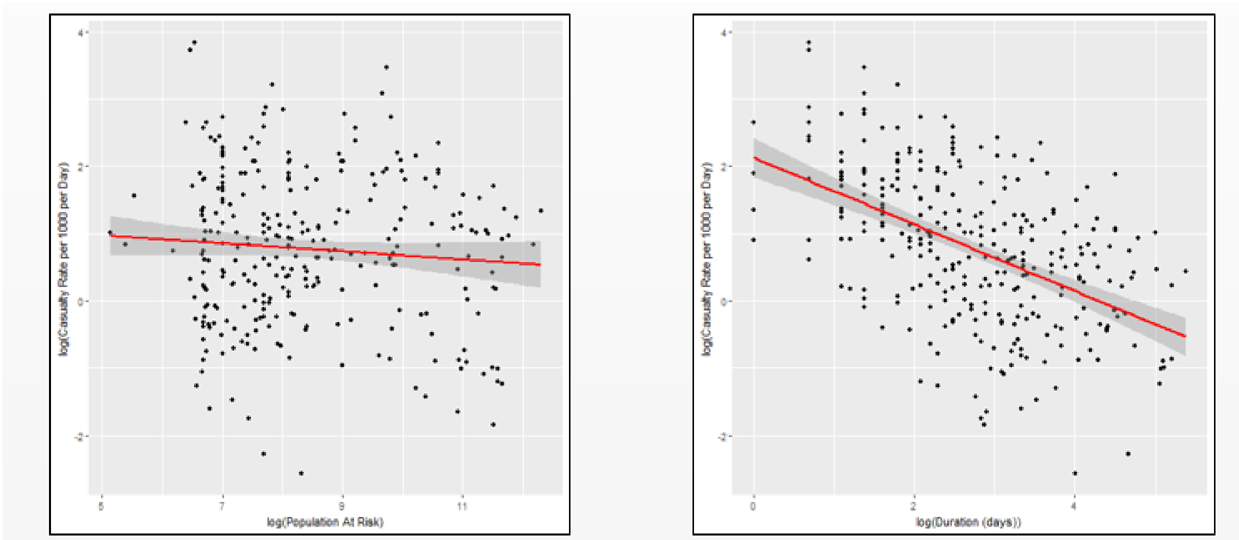

Figure 4. Scatter plot of the WIA rate logarithm against the PAR logarithm (left) and scatter plot of WIA rate logarithm against the Duration logarithm (right).

PAR has a very slight negative correlation with the WIA rate (Pearson's correlation coefficient, r = -0.08), indicating that the WIA rate decreases as the PAR increases. The Duration in days has a fairly moderate negative correlation (r = -0.5) with the WIA rate, indicating that the WIA rate decreases as an operation's duration increases. The very weak negative correlation of PAR with the WIA rate is explained by Trevor Dupuy (1990), who documents a very slight decrease in WIA rates as PAR increases over 1,000. The median PAR in our analysis dataset was 3,000—most of the PAR values were above 700.

Figure 5 shows the box plots of Service by the logarithm of the WIA rate and the box plot of Terrain by the logarithm of the WIA rate, side by side. A box plot shows the median, along with the lower and upper quartiles within the box, and displays outliers as dots beyond the box's whiskers.

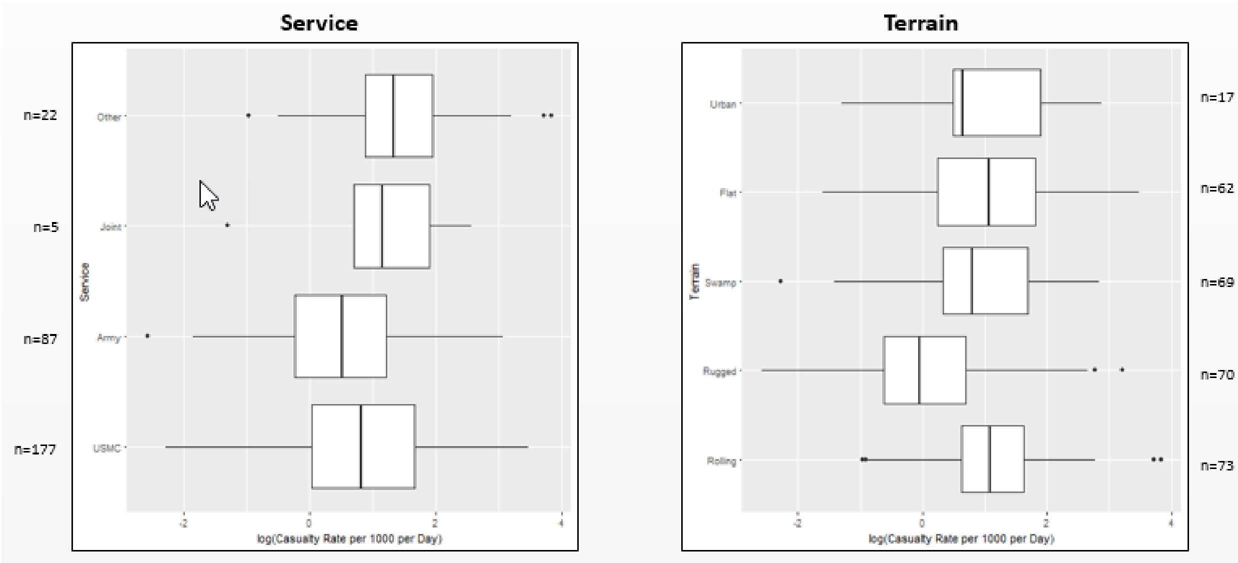

Figure 5. Box plots of Service (left) and Terrain by WIA rate logarithm (right).

For the box plot of Service by WIA rate, the United States Army and Marine Corps (USMC) rates have a lot of overlap, with the median rates for the USMC slightly higher than the Army. The box plot of Terrain by WIA rate shows a lot of overlap among Urban, Flat, Swamp, and Rolling, with Rugged terrain having the lowest rates.

Figure 6 shows the juxtaposed box plots of Climate and Region by the logarithm of the WIA rate.

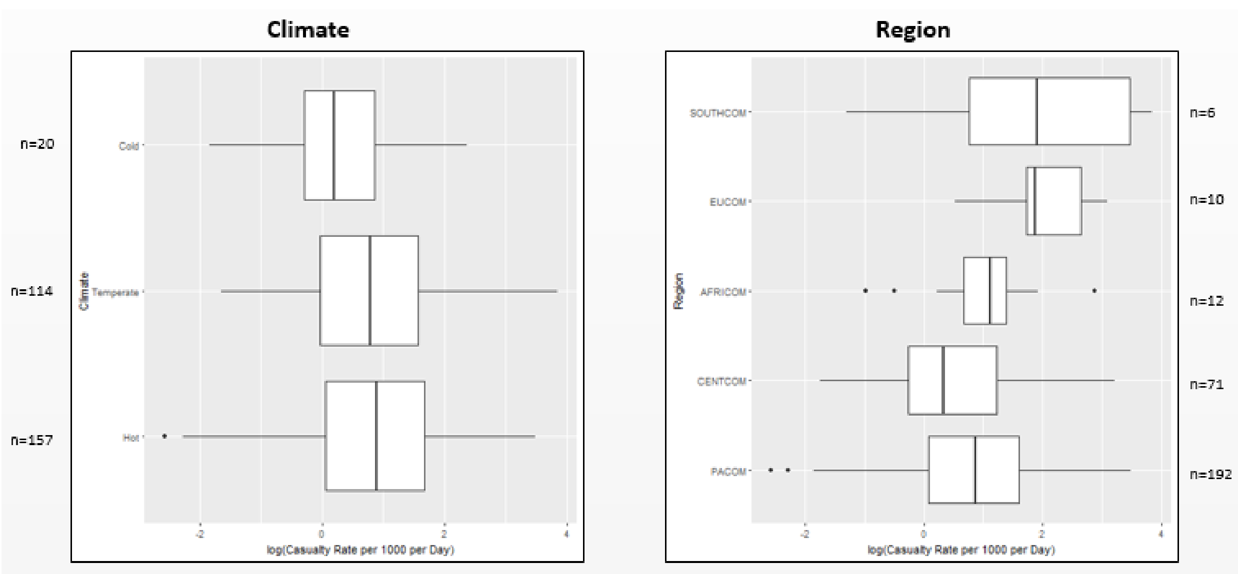

Figure 6. Box plots of Climate (left) and Region (right) by WIA rate logarithm.

The rates for Temperate and Hot climates are very similar, with a lower distribution of rates for Cold climates. The SOUTHCOM (n = 6), EUCOM (n = 10), and AFRICOM (n = 12) regions have fairly small counts. As expected, the PACOM region (consisting of World War II, Vietnam, and Korean War operations) has a higher median rate compared with CENTCOM (consisting primarily of Iraq and Afghanistan conflict operations).

Figure 7 shows the juxtaposed box plots of Tactical Operation and Posture by the logarithm of WIA rate.

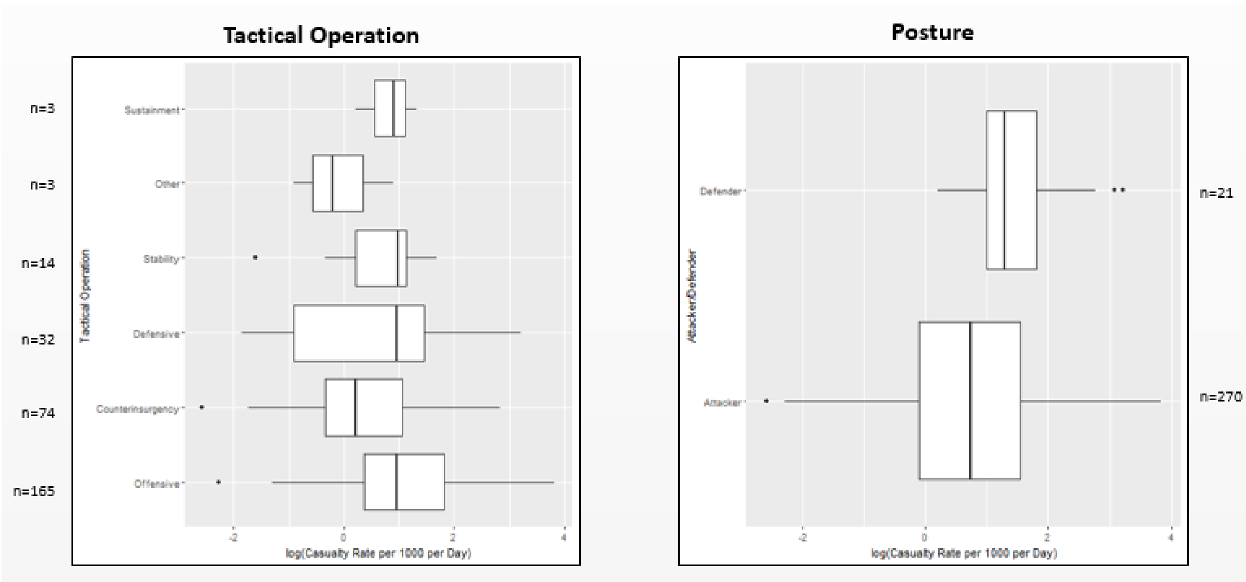

Figure 7. Box plot of Tactical Operation (left) Posture (right) by WIA rate logarithm (right).

The distribution of WIA rates has overlap among the Tactical Operation categories. Rates for Counterinsurgency operations are generally lower than rates for Offensive operations, while Stability and Sustainment operation rates are similar. The box plot of Posture by the logarithm of WIA rate shows that Attackers generally have lower rates than Defenders.

Figure 8 shows a box plot of Enemy Capability by the logarithm of WIA rate. All four enemy capability rates overlap. Operations against Near Peer enemies have the highest median rates, followed by operations against Hybrid, Failed State, and Asymmetric enemies.

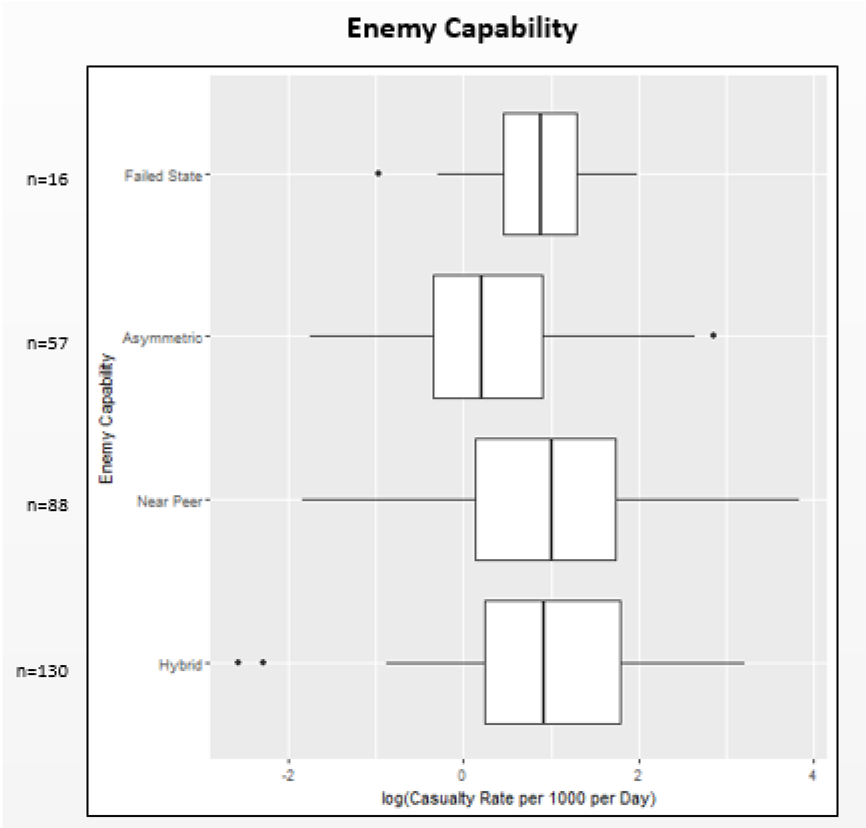

Figure 8. Box plot of Enemy Capability by WIA rate logarithm.

Model Development

Model development starts by randomly splitting the analysis dataset of 291 observations into 2: a training/validation set, and a test set. Seventy percent of the data is allocated to the training/validation set, and the remaining 30% to the test set. The training/validation set was used to develop various regression and ensemble models (Hastie, Tibshirani, & Friedman, 2009), and to select the best model. The held-out test set, on which none of the models were trained, was used to compare the performance of the selected model with the CREstT's current version (MPTk Version 1.3) and with Dupuy's casualty estimation methodology (Dupuy, 1990).

Seven models were initially constructed: multiple regression, ridge regression, classification and regression trees (CART), random forests, support vector machines (SVM), gradient boosting machines (GBM), and neural networks. The multiple regression, SVM, and neural network models were subsequently abandoned because they performed worse than the rest of the models on out-of-sample data. Only the remaining four models—Ridge Regression, CART, Random Forests, and GBM—were considered for final model selection. The WIA rate logarithm was the outcome or response variable, and the predictors or independent variables consisted of PAR, Duration, Service, Terrain, Climate, Region, Posture, Enemy Capability, and Tactical Operation.

The Ridge Regression model (Hastie, Tibshirani, & Friedman, 2009) is an enhancement to the multiple regression linear model. It is a shrinkage method that regularizes or shrinks the coefficient estimates towards zero. The shrinkage of the coefficients prevents model overfitting by reducing the estimates' variance. Ridge Regression is an effective technique for modeling data that suffers from multicollinearity (where the independent variables are correlated). Unlike ensemble models like Random Forest and GBM, the Ridge Regression model is interpretable.

The CART model (Brieman, Friedman, Olshen, & Stone, 1984) is either a classification or regression tree-based model, where the tree is constructed by splitting on predictors at decision points, starting with the most significant predictor at the top of the tree (Figure 9). Tree-based models capture nonlinear relationships and do not assume a linear relationship between the response variable and predictors. Complex interactions among the predictors are captured, and the plotted classification or regression tree is interpretable.

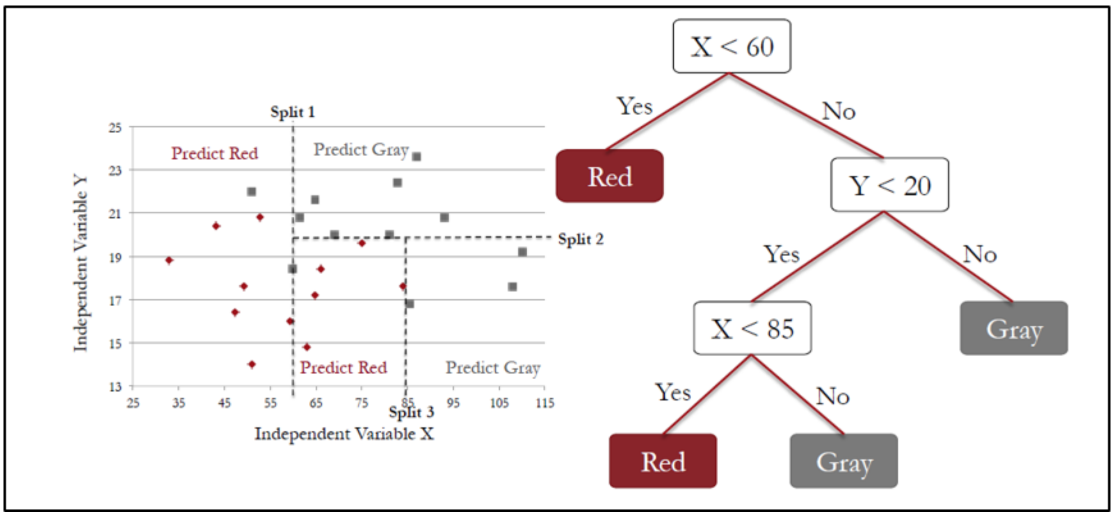

Figure 9. Classification and Regression Tree (CART) model representation. Credit: Analytics Edge course, edX.

The Random Forest (Hastie, Tibshirani, & Friedman, 2009) is an ensemble model that works by building a large number of classification or regression trees. To make predictions for a new observation, each tree "votes" on the outcome and the outcome with the majority vote is selected (classification), or the average of the continuous response from each tree is used (regression). Each tree is built from a bagged or bootstrapped aggregated sample of the data, where observations are randomly selected with replacement. At each decision split point, only a random subset of the variables can be selected. A Random Forest model typically improves upon CART's prediction accuracy, but the model is less interpretable than CART.

GBM (Friedman, 2001) is a popular machine learning method for regression and classification problems, and produces a prediction model in the form of an ensemble of weak prediction models. The intuition behind gradient boosting is to repetitively leverage the pattern in model residuals, and strengthen a model with weak predictions to make it better. When the residuals do not have any pattern that can be modeled, modeling of the residuals stops to prevent overfitting.

Model Selection

The four models—Ridge Regression, CART, Random Forest, and GBM—were developed using the training/validation set, and the best model selected based on the lowest RMSE using 10-fold, cross-validation (CV; James, Witten, Hastie, & Tibshirani, 2015). RMSE is the average forecasting error for a specific combat operation—the lower the RMSE, the better the model.

In 10-fold CV, the training/validation set is randomly divided into 10 groups, or folds, of approximately equal size. The first fold is treated as a validation set, and the model is fit on the remaining nine folds. The $RMSE_1$ is then computed on the observations in the held-out, or validation fold. This procedure is repeated 10 times; each time, a different group of observations is treated as a validation set. This process results in 10 estimates of the validation error: $RMSE_1$, $RMSE_2$, …, $RMSE_{10}$. The 10-fold CV estimate is then computed by averaging these values.

We ensured that the four models used identical folds for training and estimating the validation folds' RMSE over the 10 CV iterations.

Figure 10 is a dot plot that shows the mean RMSE of the WIA rate logarithm and the associated 95% confidence interval (CI) for the four models. The Ridge Regression model outperformed the rest based on the lowest mean RMSE and 95% CI.

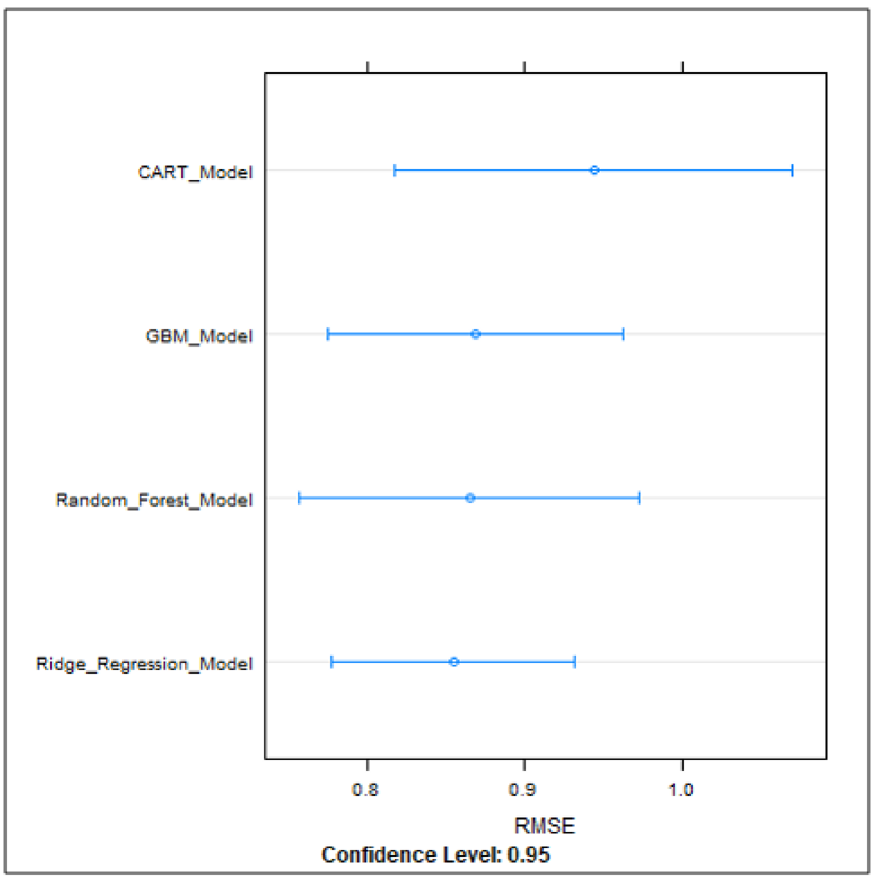

Figure 10. Dot plot of model performance based on the RMSE. The lower the RMSE, the better the model. CART = Classification and Regression Tree model; GBM = Gradient Boosting Machine model.

Model Comparison

The performance of the selected Ridge Regression model was compared with the current version of CREstT (MPTk Version 1.3) and with Dupuy's casualty estimation methodology (Dupuy, 1990). The RMSE of the WIA rate logarithm on the held-out test set (n = 87) was used as the comparison performance metric.

Figure 11 summarizes the Ridge Regression model performance.

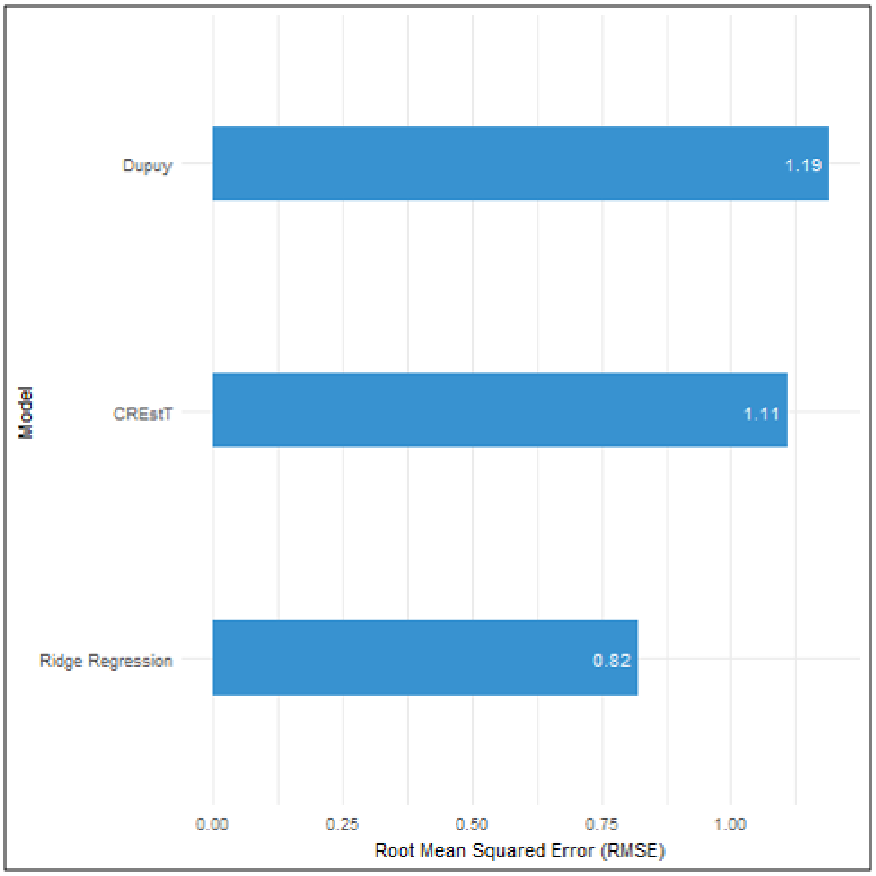

Figure 11. Model performance comparison using RMSE on held-out test set (n = 87). The lower the RMSE, the better the model performance. Dupuy is Trevor Dupuy's casualty estimation model (Dupuy, 1990).

The Ridge Regression model had the lowest test set RMSE of 0.82, and outperformed the current version of CREstT (RMSE = 1.11, and Dupuy's model (RMSE = 1.19).

Proposed Model

A Ridge Regression model is used to estimate the mean WIA rate for ground combat operations. The final WIA rate is calculated by applying empirically based adjustments to the mean WIA rate.

Equation 1 is the Ridge Regression model with seven predictors: logarithm of PAR, logarithm of Duration, Tactical Operation (TO), Posture, Climate, Terrain, and Enemy Capability (EC). The WIA rate logarithm is the model's response variable. Service and Region predictors were not significant, and therefore, excluded from the model.

$$\boldsymbol{ln(WIARate) = \hat{\beta}_{Intercept} + \hat{\beta}_{ln(PAR)} * ln(x_{PAR}) + \hat{\beta}_{ln(Duration)} * ln(x_{Duration}) + \hat{\beta}_{TO} * x_{TO} + \hat{\beta}_{Posture} * x_{Posture} + \hat{\beta}_{Climate} * x_{Climate} + \hat{\beta}_{Terrain} * x_{Terrain} + \hat{\beta}_{EC} * x_{EC}} \quad (1)$$

$$\boldsymbol{MeanWIARate = e^{(\ln(WIARate))} * \hat{\beta}_{TA} * \hat{\beta}_{Operational\ Environment}} \quad (2)$$

Tactical Action and Operational Environment scalar values were empirically derived: Tactical Action from casualties sustained in various actions during Korean War (Reister, 1973), and Operational Environment from casualties in each environment during the Korean War, OIF and OEF. Equation 2 computes the mean WIA rate (*MeanWIARate*) by applying these scalars multiplicatively to the exponentiated WIA rate from Equation 1.

The user will also be allowed to select a Combat Advantage factor. Combat Advantage is a categorical variable with three levels: Neutral, US Advantage, and Enemy Advantage. The final WIA rate (*FinalWIARate*) is calculated based on the Combat Advantage and Enemy Capability user selections (Table 3).

Table 3
Final WIA Rate Calculation Based on User-Selected Combat Advantage and Enemy Capability

Combat advantage	**Enemy capability**	**Final WIA rate *(FinalWIARate)***
Neutral	Any	*MeanWIARate*
US Advantage	Near Peer	50th percentile of Exponential distribution with mean = *MeanWIARate*
US Advantage	Any except Near Peer	50th percentile of Gamma distribution with mean = *MeanWIARate*
Enemy Advantage	Near Peer	80th percentile of Exponential distribution with mean = *MeanWIARate*
Enemy Advantage	Any except Near Peer	80th percentile of Exponential distribution with mean = *MeanWIARate*

The *FinalWIARate* is the *MeanWIARate* from Equation 2 when the Combat Advantage is Neutral. When the Enemy Capability is Near Peer and the Combat Advantage is either US Advantage or Enemy Advantage, the *FinalWIARate* is the 50th percentile of an exponential distribution (with a mean equal to *MeanWIARate*) or the 80th percentile of an exponential distribution (with a mean equal to *MeanWIARate*) respectively. However, when the Enemy Capability is any selection except Near Peer and the Combat Advantage is US Advantage or Enemy Advantage, the *FinalWIARate* is the 50th percentile of a Gamma distribution (with a mean equal to *MeanWIARate*) or the 80th percentile of a Gamma distribution (with a mean equal to *MeanWIARate*) respectively.

Casualties may be stochastically simulated in CREstT using a Poisson distribution with a mean equal to *FinalWIARate*.

Model Parameters

Table 4 displays the parameter estimates or coefficients of the proposed Ridge Regression model. The *R* software package *glmnet* (Friedman, Hastie & Tibshirani, 2010) was used to develop and test the model. The *glmnet* software deliberately does not provide standard errors for the model coefficients because standard errors are not meaningful for biased estimates that arise from penalized estimation methods like Ridge Regression (Hastie, Tibshirani, & Friedman, 2009).

Table 4
Ridge Regression Model Parameter Estimates

Variable	**Parameter**	**Estimate ($\hat{\beta}$)**
Intercept	Intercept	4.0986
ln(PAR)	ln(PAR)	-0.1577
ln(Duration)	ln(Duration)	-0.4630
Tactical Operation	Offensive (baseline)	0
	Stability	-0.3832
	Sustainment	-0.4818
	Counterinsurgency	-0.5727
	Defensive	-1.0436
	Peace Operations	-3.0000
Posture	Attacker (baseline)	0
	Defender	1.1864
Climate	Temperate (baseline)	0
	Hot	-0.0072
	Cold	-0.3233
Terrain	Urban (baseline)	0
	Rolling	-0.0060
	Flat	-0.1289
	Swamp	-0.3504

Variable	Parameter	Estimate ($\widehat{\beta}$)
	Rugged	-0.4060
Enemy Capability	Near Peer (baseline)	0
	Hybrid	-0.2095
	Failed State	-0.7422
	Asymmetric	-0.7422

The parameter estimate for Peace Operations (Tactical Operation) in Table 4 was derived using Subject Matter Expert (SME) judgment. Likewise, the parameter estimate for *Failed State* (Enemy Capability) was SME-adjusted to be the same as for Asymmetric enemies: the original estimate for Failed State was not reliable because of the small sample size (n = 16).

Table 5 displays the empirically derived scalars for the Tactical Action adjustment factor. Note that the Tactical Action scalar values are only associated with Offensive and Defensive tactical operations.

Table 5
Tactical Action Scalar Values.

Tactical operation	Tactical action	Estimate ($\widehat{\beta}$)
Defensive	US Counterattack, heavy resistance	1.651
	Defense, enemy attack or counterattack	1.463
	US Counterattack, light resistance	0.979
	Patrolling	0.603
Offensive	Amphibious Assault	2.582
	US Attack, heavy resistance	2.278
	Airdrop Attack	1.912
	Assault River Crossing	1.610
	Defense, enemy attack or counterattack	1.609
	US Attack, light resistance	0.942
	Mopping Up	0.887
	Patrolling	0.474

Table 6 displays the empirically derived scalars for the Operational Environment adjustment factor by Enemy Capability. These Operational Environment scalar values were calculated for

Near Peer, Failed State, and Asymmetric enemies using WIA-rate ratios estimated from the Korean War, OIF, and OEF respectively. The Operational Environment scalars for Hybrid enemies were estimated using SME judgment.

Table 6
Operational Environment Scalar Values by Enemy Capability

	Enemy capability			
Operational environment	**Near Peer estimate ($\hat{\beta}$)**	**Hybrid estimate ($\hat{\beta}$)**	**Failed State estimate ($\hat{\beta}$)**	**Asymmetric estimate ($\hat{\beta}$)**
Close area	1.000	1.000	1.000	1.000
Support area (forward)	0.188	0.209	0.257	0.459
Support area (rear)	0.070	0.090	0.135	0.310

The effects of the seven predictors in the Ridge Regression model and the two adjustment factors—Tactical Action and Operational Environment—on the forecasted WIA rate are summarized in Table 7.

Table 7
Predictor and Adjustment Factor Effects on WIA Rate in the Ridge Regression Model.

Predictor/adjustment factor	**Effect on forecast WIA**
PAR	As PAR increases, WIA rate decreases
Duration (days)	As Duration increases, WIA rate decreases
Climate	WIA rate decreases in the following sequence:
	Temperate → Hot → Cold
Terrain	WIA rate decreases in the following sequence:
	Urban → Rolling → Flat → Swamp → Rugged
Tactical operation	WIA rate decreases in the following sequence:
	Defensive → Offensive → Stability → Sustainment → Counterinsurgency → Peace Operations
Posture	Attackers have lower rates than Defenders
Enemy capability	WIA rate decreases in the following sequence:
	Near Peer → Hybrid → Failed State = Asymmetric
Tactical action	WIA rate decreases in the following sequence:
	Tactical Operation = Offensive:
	Amphibious Assault → US Attack, Heavy Resistance → Airdrop Attack → Assault River Crossing → Defense, enemy attack or counterattack → US Attack, Light Resistance → Mopping Up → Patrolling
	Tactical Operation = Defensive:
	US Counterattack, Heavy Resistance → Defense, enemy attack or counterattack → US Counterattack, Light Resistance → Patrolling
Operational environment	WIA rate decreases in the following sequence:
	Close Area → Support Area (Forward) → Support Area (Rear)

Conclusion

To implement the CaRD Work Group recommendations , we proposed a Ridge Regression model to estimate the WIA rate for ground combat operations, and demonstrated that this model generates more accurate estimates than both the current version of CREstT (MPTk Version 1.3) and Trevor Dupuy's methodology (Dupuy, 1990). We expanded the number of data points underpinning the baseline WIA rate from seven to 291, spanning combat operations in World War II, Korea, Vietnam, Iraq, and Afghanistan. The proposed model successfully addresses each of the CaRD Work Group's recommendations.

We developed several regression and ensemble models and, after discarding the poorly performing ones, were left with four models to evaluate: Ridge Regression, CART, Random Forest, and Gradient Boosting Machine. The Ridge Regression model was selected based on the lowest RMSE using 10-fold cross-validation on the training/validation set. The Ridge Regression model was also shown to outperform the current version of CREstT (MPTk Version 1.3) and Trevor Dupuy's methodology (Dupuy, 1990) using RMSE as the comparison metric on the out-of-sample test set.

The Ridge Regression model uses seven predictors to forecast a ground combat operation's WIA rate: PAR or force size, Duration, Tactical Operation, Posture, Climate, Terrain, and Enemy Capability. Notably, the Region factor currently used in CREstT was not significant, and was excluded from the model.

Tactical Actions data that the CaRD Work Group recommended for inclusion in operations' Mission Sets (Appendix A) were not available for all 291 data points. We were able to find a limited set of tactical actions only for Korean War operations (Reister, 1973), and derived a set of tactical action scalars to be applied multiplicatively to the model-estimated WIA rate. Likewise, Operational Environment scalars were derived for each Enemy Capability and applied multiplicatively to the model-estimated WIA rate.

The model parameter for operations against a *Failed State* enemy (n = 16) could not be reliably estimated because of insufficient data. Using SME judgment, the parameter for Failed State in the model was set equal to the parameter for Asymmetric enemies.

Model parameters are limited by the available data; our model was trained on combat operational data primarily from World War II, Korea, Vietnam, Iraq, and Afghanistan. Though we had recent operational data from Iraq and Afghanistan against Asymmetric and Failed State enemies, the earliest data we had for operations against Near-Peer enemies were from the Korean War over 6 decades ago, and the earliest data we had for battles against Hybrid enemies were from the Vietnam War almost 5 decades ago. Changes in combat warfare—weapon lethality, tactics, technology, casualty evacuation, medical care, and personal protective equipment—affect the ground combat WIA rate (Holcomb, Stansbury, Champion, Wade, & Bellamy, 2006). As combat warfare evolves, we should continue to iteratively add combat operations to our dataset, and dynamically update our regression model. Future efforts that provide additional insight into casualty rates for specific missions should be included in the model when available.

This study proposed a Ridge Regression model to estimate WIA casualty rates for ground combat operations that will be implemented in MPTk Version 1.4/1.5. This work will refine WIA casualty projections in CREstT, and improve resource estimates for medical planners.

References

Beebe, G., & De Bakey M. (1952). *Battle Casualties – Incidence, Mortality, and Logistic Considerations*. Springfield, IL: Charles C. Thomas.

Blood, C. G., Zouris, J. M., & Rotblatt, D. (1997). *Incorporating adversary-specific adjustments into the Ground Forces Casualty Forecasting System (FORECAS) to project casualty sustainment* (Report No. 97-39). San Diego, CA: Naval Health Research Center.

Brieman, L., Friedman, J. H., Olshen, R. A., & Stone, C. J. (1984). *Classification and Regression Trees*. Boca Raton, FL: Chapman & Hall/CRC.

Dupuy, T. N. (1990). *Attrition: Forecasting Battle Casualties and Equipment Losses in Modern War*. Falls Church, VA: Nova Publications.

Dupuy, T. N., Brinkerhoff, J. R., Bader, B., & Johnson, C. C. (1985). *Casualty Estimates for Contingencies* (Report No. 118, Vol I-II). Fairfax, VA: Data Memory Systems, Inc.

Hastie, T., Tibshirani, R., & Friedman, J. H. (2009). *The elements of statistical learning: Data mining, inference, and prediction*. Second Edition. New York, NY: Springer.

Holcomb, J. B., Stansbury, L. G., Champion, H. R., Wade, C., & Bellamy, R.F. (2006). Understanding combat casualty care statistics. *Journal of Trauma, 60*(2), 397–401.

Friedman, J. H. (2001). *Greedy function approximation: A gradient boosting machine*. The Annals of Statistics. Vol 29, No. 5, 1189–1232.

Friedman, J. H., Hastie, T., & Tibshirani, R. (2010). Regularization paths for generalized linear models via coordinate descent. *Journal of Statistical Software*, *331*(1), 1–22. Retrieved from http://www.jstatsoft.org/v33/i01/

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2015). *An introduction to statistical learning with applications in R*. New York, NY: Springer.

Marine Corps Doctrinal Publication (MCDP) 1-0 (2017). *Marine Corps operations*. Quantico, VA: Marine Corps Combat Development Command.

Naval Health Research Center. (2018). *Medical Planners' Toolkit methodology manual* (Version 1.3). San Diego, CA: Author.

Reister, F. (1973). *Battle Casualties and Medical Statistics: U.S. Army Experience in the Korean War*. Washington, DC: Department of the Army, Office of the Surgeon General.

Thayer, T. (1975). *A Systems Analysis View of the Vietnam War*. Volume 4: Allied Ground and Naval Operations. Washington, DC: OASD (SA) RP, Southeast Asia Intelligence Division, Pentagon.

The Virtual Vietnam Archive. (1965–1970). Lubbock, TX: Texas Tech University. Vietnam operational situation reports retrieved from https://www.vietnam.ttu.edu/virtualarchive

Tukey, J. W. (1977). *Exploratory data analysis*. Reading, MA: Addison-Wesley.

U.S. Army Concepts Analysis Agency (CAA) Database of Battles (1990). Retrieved from https://github.com/jrnold/CDB90

Zouris, J., D'Souza, E., Honderick, P., Tolbert, B., & Wing, V. (2013). *Development of a Methodology for Estimating Casualty Occurrences and the Types of Illnesses and*

Injuries for the Range of Military Operations (Report No. 13-06). San Diego, CA: Naval Health Research Center.

Appendix A
CaRD Work Group Recommendations for WIA Rate Enhancements in CREstT