**Supplementary Figure 1.** Example classification tree used in random forest (RF) models to assess entanglement and vessel strike injury severity. Each tree is unique in construction, based on a bootstrap sample taken from all known-outcome cases. In our study, we force each tree to be built with equal numbers of ‘*dead.decline*’ vs ‘*recovered*’ health classes, as we are equally-interested in accurate prediction of each class. This example tree starts with a sample of 10 injury cases (5 of each known outcome class). Out of *n* variables, are randomly chosen as potential splitters of data in the parent (top) node. The variable that best separates cases into *dead.decline* vs *recovered* classes (in this case, was gear constricting? Y/N) is chosen for the initial split. Data continue to be split based on successively chosen variables until all cases in terminal nodes (=squares) are of a single class. This tree represents one of 1,000 random bootstrap trees used in construction of RF models. Each tree is unique because it is based on a different initial bootstrap set of data and contains different variables.



**Supplementary Figure 2.** Ensemble of bootstrap trees (= a random forest) used as models for entanglement and vessel strike risk. Model cross-validation is accomplished by introducing out-of-bag (OOB) samples (known outcome cases excluded from individual tree construction) to each tree. Predictions for OOB samples are dependent upon the variable characteristics, which determine which terminal node (squares) the case is assigned to. All cases are eventually used as OOB through the bootstrap process and a sufficient number of trees, which allows for cross-validated error rates of each model to be calculated and summarized as confusion matrices.



**Supplementary Figure 3.** Novel data (unknown outcome cases) not used in the construction of the random forest are introduced to the existing ensemble of trees and predictions of health status are made for each case, depending on the values of variables in each narrative and which terminal node the case falls into. The probability of ‘*Dead.Decline*’ or ‘*Recovery*’ represents the fraction of forest trees that ‘vote’ for each health status for each case introduced to the forest. In the example below, if the forest consisted of 3 trees, the health status of the unknown outcome case is predicted as ‘*Dead.Decline*’, with a probability of 2/3.

