

Purpose: Binary outcomes (toxicity=1, no-toxicity=0) are frequently used as inputs to fit parametric models. However, in reality, outcomes span the range in-between these extremes; binary values are used for lack of more precise quantification. We propose a method to estimate non-binary outcomes, which can then be used to more accurately fit conventional parametric models (e.g., Lyman-probit, relative-seriality).

Methods and Materials: The non-binary outcomes were estimated using 4 machine learning algorithms. The algorithms used variables selected from the available data to best-fit binary outcomes. Based on the fitting, each algorithm outputted binary estimates of whether or not the patient suffered toxicity. By fitting each algorithm to randomly selected subsets of the patient data, multiple binary toxicity estimates were generated for each patient. Averaging these estimates for each algorithm produced a non-binary estimate. Furthermore, averaging these non-binary estimates over all algorithms reduced prediction bias. This final algorithm-averaged non-binary patient estimate was then used as input to conventional parametric algorithms. This method is demonstrated in the context of radiotherapy-induced pneumonitis.

Results: One-hundred estimates from each algorithm x 4 algorithms were averaged to produce a non-binary toxicity estimate for each patient. Both the original binary toxicity outcome and the non-binary estimate were fitted to the parametric Lyman-probit and relative-seriality models. Both parametric models had large confidence interval limits when fitted to the original binary outcomes. The fits improved when using the non-binary estimates. The improvement was substantially better for the Lyman-probit model, but only marginally so for the relative-seriality model. Thus, the Lyman-probit model is better suited to modeling radiation-induced pneumonitis, a conclusion that would not have been possible from the original binary outcomes.

Conclusions: Non-binary estimates from machine learning algorithms can be used to not only generate better fits to conventional parametric models, but also to deduce which model is methodologically better suited.