Online resource 3 – classification and regression tree analysis.

Title

IDH mutation status and role of WHO grade and mitotic index in overall survival in grade II-III diffuse gliomas

Acta Neuropathologica

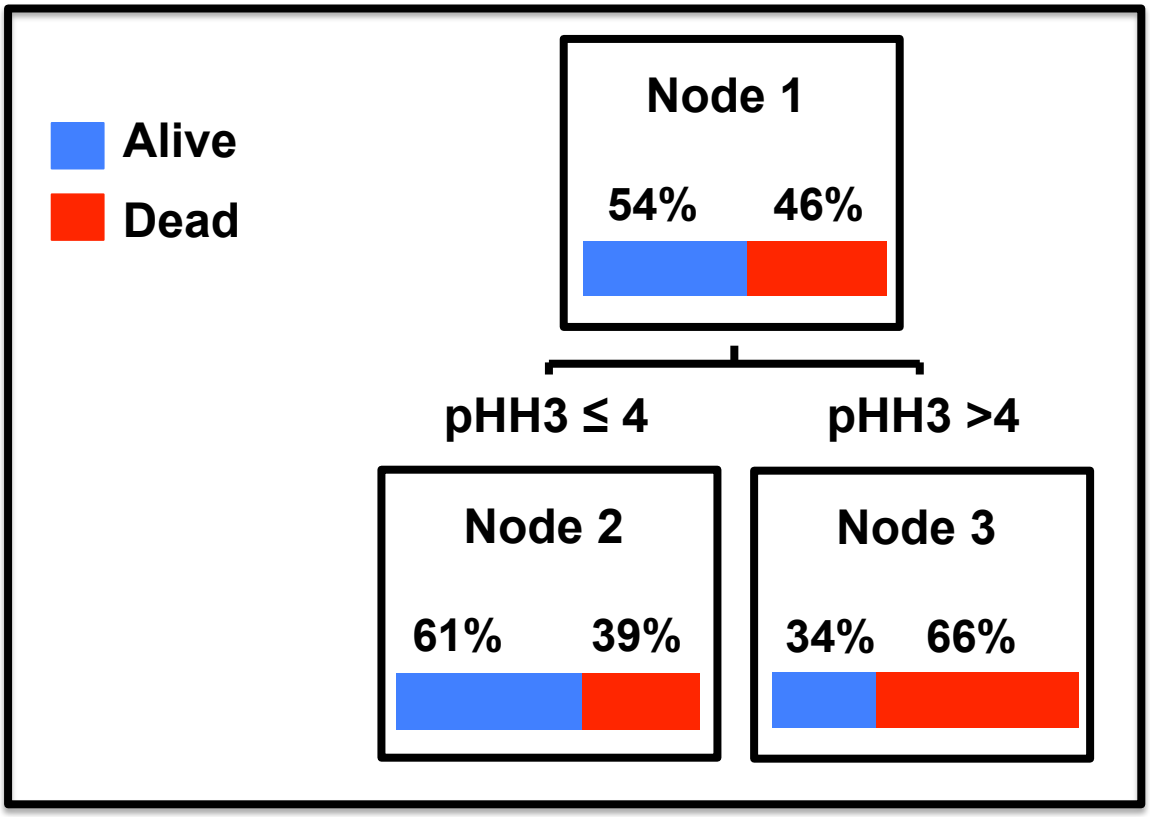
Authors: Adriana Olar, Khalida M. Wani, Kristin Diefes, Lindsey E. Heathcock, Hinke F. van Thuijl, Mark R. Gilbert, Terri S. Armstrong, Erik P. Sulman, Daniel P. Cahill, Elizabeth Vera-Bolanos, Ying Yuan, Jaap C. Reijneveld, Bauke Ylstra, Pieter Wesseling, Kenneth D. Aldape

Corresponding authors:
Adriana Olar, MD
The University of Texas MD Anderson Cancer Center
Dept. of Pathology, G1.3510
1515 Holcombe Blvd.
Houston, TX, 77030, USA
Email: adriana_olar@yahoo.com

Kenneth Aldape, MD
Princess Margaret Cancer Centre and Ontario Cancer Institute
610 University Ave
Toronto, ON M5G 2M9, Canada
Email: ken.aldape@uhnresearch.ca

Classification and regression tree (CART) analysis

In order to select the most appropriate cut-off for mitotic index we used the CART data-mining algorithm. This is a data-driven method of analysis that identifies criteria that best separate the data. The output is a tree structure in which each node splits based on data-defined criteria [1]. The analysis was performed in Statistica, v. 11, StatSoft, Inc., Tulsa, OK, USA using the data mining feature, specifically the general classification and regression tree option. Multiple models were run using outcome (alive vs. dead) as the dependent variable and the continuous mitotic index as the predictor variable. The best model chosen had the following parameters: Gini splitting rule, prune on variance stopping rule, and a case limit per node of 10. Node 1 had the lowest resubstitution cost in this model and suggested 4.5 as the best mitotic index splitting value. Because in real practice, mitotic index has to be a natural number, based on this suggested value and on the algorithm's splitting rules (the value of 4 had to be included on the left) we chose mitotic index subgroups (0 to 4 mitotic figures and greater than 4 mitotic figures) that best subclassify outcome in WHO II-III diffuse glioma. This is represented in the figure tree below.



Reference:

1. Breiman L, Friedman J, Olshen RA, Stone JC (1983) Classification and Regression Trees. Wadsworth Statistics/ Probability Series, 1st edn. Chapman and Hall/CRC, FL (USA).