Supplementary Online Content

Bertsimas D, Margonis GA, Sujichantararat S, et al. Using artificial intelligence to find the optimal margin width in hepatectomy for colorectal cancer liver metastases. *JAMA Surg*. Published online June 1, 2022. doi:10.1001/jamasurg.2022.1819

This supplementary material has been provided by the authors to give readers additional information about their work.

eMethods

Definitions: Patients with CRLM and resected concurrent extrahepatic disease were included to avoid selection biases, reflect real world practices, and allow for comparisons to contemporary studies from other institutions. Concurrent extrahepatic disease was defined as extrahepatic disease known at the time of hepatic resection and either resected at the same time or within 6 months of hepatectomy. Patients with incompletely resected extrahepatic disease (R2 resections) were excluded. Margin width values were presented as discrete variables (integer numbers). R1 resection was defined as tumor cells at the resection margin (0 mm). A narrow margin group was defined as either margin clearance less than 1 mm or involved margins.[1] Intraoperative ablation was not used to control intraoperatively positive margins. In cases of multiple CRLMs, the closest resection width was recorded as the margin distance.

First step- Counterfactual Estimator: The RF was trained using 11 clinically relevant patient-, tumor- and treatment-related predictors of prognosis that were known at the time of surgery.[2,3] Patient-related predictors were age and gender. Tumor-related predictors were the site and the T and N stage of the primary tumor, as well as the size, distribution, and number of liver metastases. The presence of completely resected extrahepatic disease was another tumor-related predictor included in the RF model. Margin width and the use of intra-operative ablation were the treatment-related predictors. This created a matrix of probabilities; each column reflects the different values of margin width while each row includes the predicted probability of death for a single patient if he/she was treated with each margin. The matrix of probabilities is also called rewards matrix as reward estimation is an approach from the causal inference literature that seeks to utilize the observed data to make inferences about the unobserved outcomes.

The counterfactual Random Forests (RF) model was selected as our predictive model because it has demonstrated superior predictive ability due to its capacity to identify complex interactions and nonlinearities of predictor effect while minimizing overfitting.[4-6] Unlike variable selection methods such as Cox regression analysis, interactions among predictors do not need to be explicitly specified to be utilized by RF as no implicit assumptions about the underlying relationships between the predictor variables and prognosis are made.[7] In fact, any interactions between prognostic factors serve to increase the importance of the individual interacting factors, making them more likely to be given high importance relative to other factors.

Estimating interactions between predictors with the SHAP model: These values reflect the additional combined impact of two predictors after accounting for the effect of each individual variable on the model's output. For example, if the model output is risk of death within a certain time interval, the presence of a positive interaction value between two prognostic factors suggests that the co-occurrence of these two factors confers a higher risk of death than would be expected by adding the risks conferred by each factor individually. We assessed whether tumor number interacts with surgical margin.

External validation: It is critical for a fair evaluation that we do not evaluate the quality of the policy using rewards from our existing reward estimator trained on the training set. This is to avoid any information from the training set leaking through to the out-of-sample evaluation. Instead, we estimated a new set of rewards using only the external validation set and evaluated the policy against these rewards.

eResults

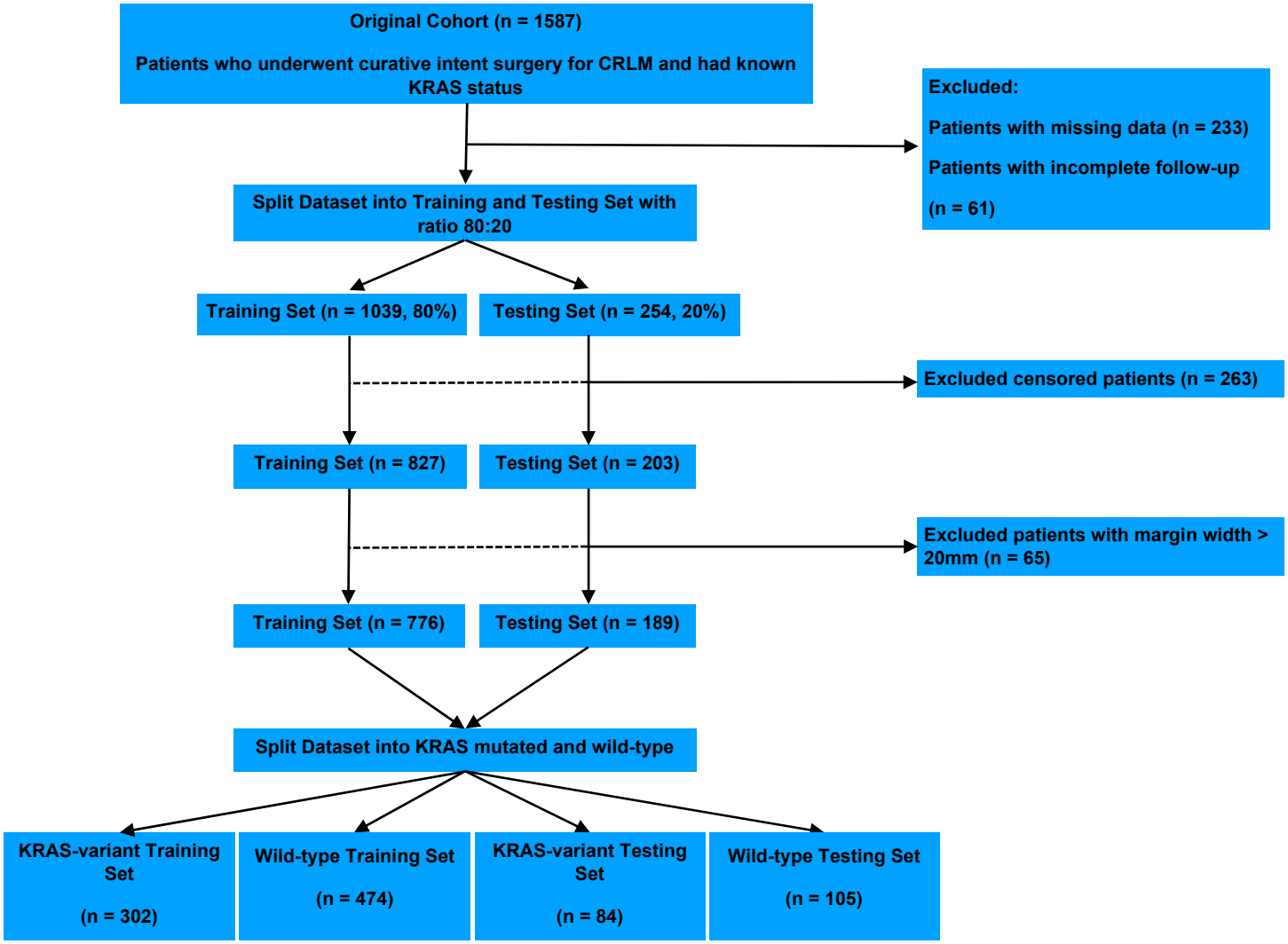
SHAP: On visual inspection of eFigure 2A, for tumor size greater than 3 cm (dashed line), right tumor side (red dots) is consistently associated with higher risk of death (higher SHAP value) compared to left side (blue dots). In contrast, on visual inspection of eFigure 2B, the presence of metastatic lymph nodes (red dots) does not change the risk of death in patients with tumors larger than 3 cm (blue and red dots are randomly scattered). Thus, tumor side and not lymph node status is a prognostic factor in patients with tumors larger than 3 cm. This confirms the selection of prognostic factors in the OPT.

SHAP interaction analysis: The strongest positive interaction (SHAP interaction values > 0, conferring worse prognosis) was between high number of tumors (red) and positive/narrow margins up to around 1 mm (black dashed line). Interestingly, this pattern is reversed for margins greater than 3 mm (red dashed line; SHAP interaction value = 0) and peaks at around 6 mm, with a negative interaction between high number of tumors and margins (yellow dashed line; SHAP

interaction values < 0, conferring better prognosis). The negative interaction suggests that patients with high number of tumors and wide margins (especially wider than 6 mm) fare better than would be expected from the additive prognostic effect of the two variables.

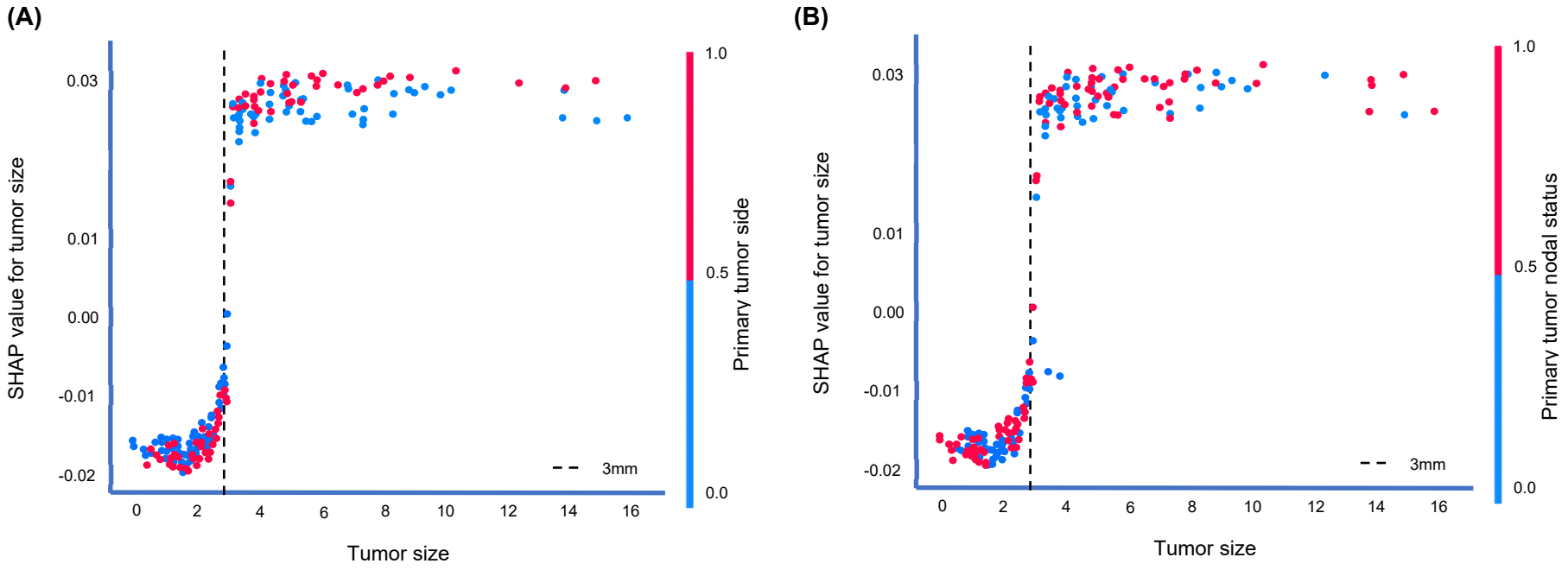
eReferences

1. Sadot E, Groot Koerkamp B, Leal JN, et al. Resection margin and survival in 2368 patients undergoing hepatic resection for metastatic colorectal cancer: surgical technique or biologic surrogate? *Ann Surg.* 2015;262(3):476-485; discussion 483-475.
2. Margonis GA, Buettner S, Andreatos N, et al. Prognostic Factors Change Over Time After Hepatectomy for Colorectal Liver Metastases: A Multi-institutional, International Analysis of 1099 Patients. *Ann Surg.* 2019;269(6):1129-1137.
3. Margonis GA, Sasaki K, Gholami S, et al. Genetic And Morphological Evaluation (GAME) score for patients with colorectal liver metastases. *Br J Surg.* 2018;105(9):1210-1220.
4. Breiman LJMl. Random forests. 2001;45(1):5-32.
5. Hastie T, Tibshirani R, Friedman J. The elements of statistical learning. Springer series in statistics. In: ∴ Springer; 2001.
6. Wongvibulsin S, Wu KC, Zeger SL. Clinical risk prediction with random forests for survival, longitudinal, and multivariate (RF-SLAM) data analysis. *BMC Med Res Methodol.* 2019;20(1):1.
7. Lunetta KL, Hayward LB, Segal J, Van Eerdewegh P. Screening large-scale association study data: exploiting interactions using random forests. *BMC Genet.* 2004;5:32.



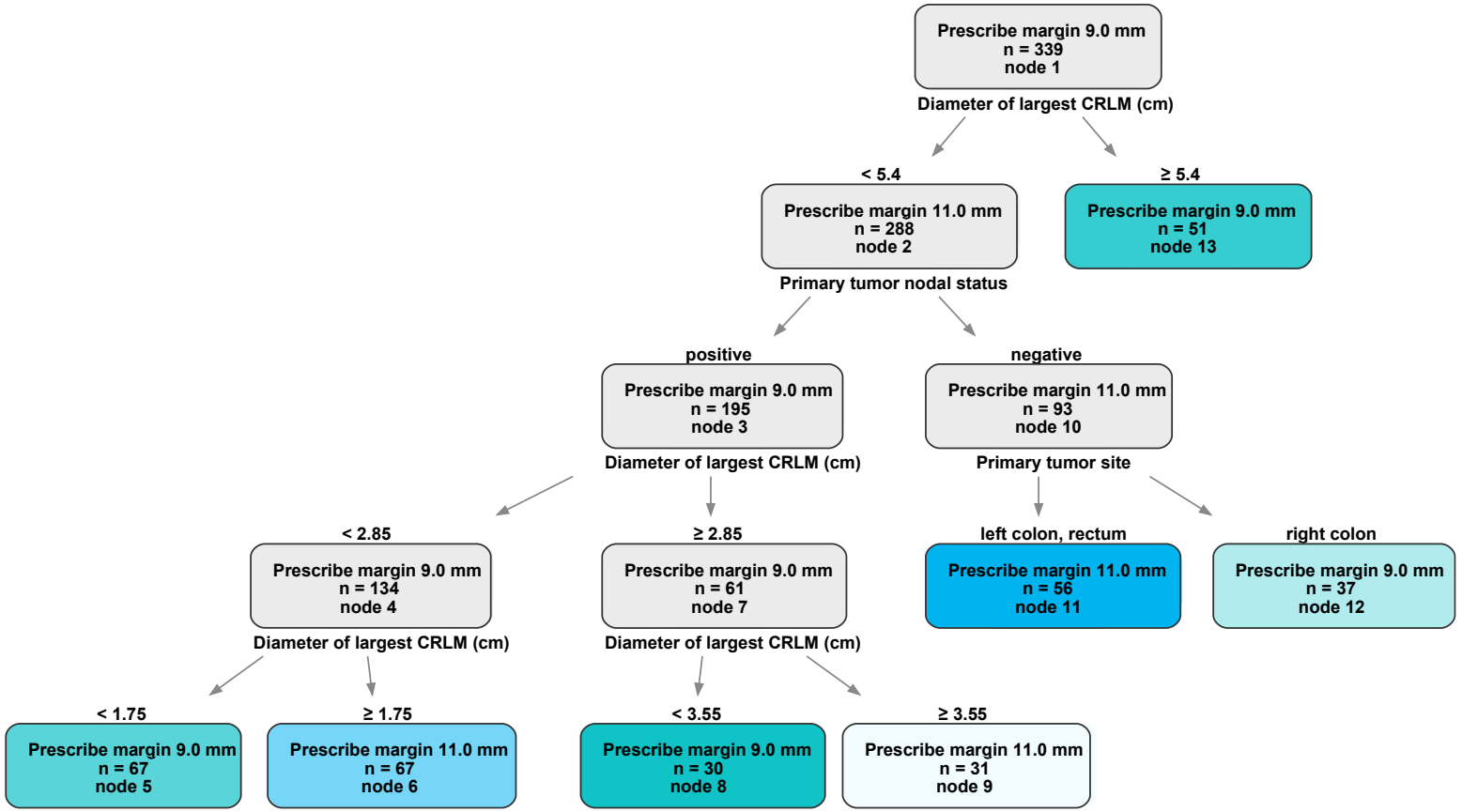
eFigure 1. Internal Cohort Selection

Patients with margins wider than 20 mm were excluded because no study has reported on an optimal margin width > 20 mm.



eFigure 2. SHAP Dependence Plot of Tumor Size vs its SHAP Value in the RF Model

The interaction effect with primary tumor side (A) and primary tumor nodal status (B) is tested.



eFigure 3. OPT for Liver-Specific Recurrence in Patients with KRAS-Variant Tumors