**Table S5** Chi-squared Automatic Interaction Detection (CHAID) method to grow a decision tree. To minimize cohort selection bias and prevent model overfitting, we also employed split-sample validation, which randomly assigned roughly half of the cases to a new training set, and half of the cases to a new test set.

|  |
| --- |
| **Classification** |
| Sample | Observed | Predicted |
| CCC | EC | HGSC | LGSC | MC | Percent Correct |
| Training | CCC | 121 | 4 | 1 | 0 | 9 | 89.6% |
| EC | 9 | 104 | 13 | 0 | 14 | 74.3% |
| HGSC | 2 | 3 | 484 | 0 | 9 | 97.2% |
| LGSC | 0 | 0 | 26 | 0 | 0 | 0.0% |
| MC | 0 | 0 | 1 | 0 | 71 | 98.6% |
| Overall Percentage | 15.2% | 12.7% | 60.3% | 0.0% | 11.8% | 89.6% |
| Test | CCC | 140 | 0 | 2 | 0 | 10 | 92.1% |
| EC | 14 | 102 | 14 | 0 | 20 | 68.0% |
| HGSC | 1 | 6 | 473 | 0 | 11 | 96.3% |
| LGSC | 0 | 0 | 29 | 0 | 1 | 0.0% |
| MC | 5 | 1 | 0 | 0 | 62 | 91.2% |
| Overall Percentage | 18.0% | 12.2% | 58.1% | 0.0% | 11.7% | 87.2% |

HGSC-high-grade serous carcinoma, EC-endometrioid carcinoma, CCC-clear cell carcinoma, MC-mucinous carcinoma, LGSC-low-grade serous carcinoma