**Prognostic Factors Across Poorly Differentiated Neuroendocrine Neoplasms: a Pooled Analysis**

**SUPPLEMENTARY MATERIALS.**

**Supplementary Methods**

*Statistical Analysis*

Data were analyzed by descriptive statistics. Clinical and pathologic characteristics of patients were stratified according to tumor site: colorectal, gastroesophageal, lung, pancreas, gallbladder/biliary and ileum–cecum–duodenum. Differences in frequencies were assessed as appropriate with the χ2 or the Fisher exact test. Ki-67 cut-off values that best identify subjects who died early (within 3 years from surgery) were evaluated with a receiver operating characteristic (ROC) curve. The area under the curve (AUC) was calculated to determine the diagnostic value of the test. Optimal cut-off points for Ki-67 values were determined using the Youden index, which calculates the cut-off point based on the sum of sensitivity and specificity maximization.

*TP53, KRAS, BRAF* and *RB1* molecular markers showed a high number of missing values. Variables with less than 30% missing values were imputed with the MissForest algorithm of the randomForestSRC package.1 *TP53, KRAS, BRAF* and *RB1* molecular risk factors were evaluated only in the subgroup of 169 patients for whom the molecular analyses were available. OS curves were calculated using the Kaplan–Meier method. The log-rank test assessed the survival difference between patient groups. Univariable and multivariable Cox proportional regression analyses assessed the association between clinicopathological characteristics and overall survival (OS). Hazard ratios (HRs) are presented with respective 95% CIs.

Random survival forests (RSF)2,3 is an extension of Random Forest techniques to survival settings, allowing efficient non-parametric analysis of time to event data. A forest of survival trees was grown using a log-rank splitting rule to select the optimal candidate variables. For each node *q=√p,* the candidate variables were randomly selected from all *p* variables in the survival tree to maximize the survival difference between the child nodes. The function used 63.3% samples to grow the forest, while the remaining 36.8% of observations, called the out-of-bag4 sample, was used as a hold-out test set for each tree. Specifically, the function used 267 samples, while the remaining 155 samples (opaque binary blob - OBB samples) were used as a hold-out test set for each. The out-of-bag prediction error estimate was 33.2%. A random log-rank split rule was used for continuous variables, which were randomly selected from split = 10 split point values. Variable importance (VIMP) was used to estimate the importance of each variable. VIMP for x is the prediction error for the original ensemble subtracted from the prediction error for the new ensemble obtained using randomizing x assignments.3 Positive values indicate variables with predictive ability, whereas zero or negative values identify non-predictive variables.3,5

The performance of prediction models was evaluated after bootstrapped cross-validation performed 100-times by discrimination ability and calibration with AUC and Brier scores, respectively. The multivariable Cox Model and RSF showed a high-quality prediction performance at 1, 2 and 3 years. Moreover, RSF performed better in discrimination ability (AUC 1 year 76 vs 77; 2 years 85 vs 87.7; 3 years 87.2 vs 89.1) with similar calibration (Brier score 1 year 0.199 vs 0.195; 2 years 0.114 vs 0.106; 3 years 0.074 vs 0.073).

Data analysis was performed using the R environment for statistical computing and graphics (R Foundation, Vienna, Austria – Version 3.6.2). All tests were two-sided, and *p*-values <0.05 were considered statistically significant.

**References**

1. Ishwaran H, Kogalur UB. Random Forests for Survival, Regression, and Classification (RF-SRC), R package version 2.6. 1. (2018). https://cran.r-project.org/web/packages/randomForestSRC/randomForestSRC.pdf
2. Ishwaran H, Kogalur UB. Random Survival Forests for R. R News. (2007). https://arxiv.org/pdf/0811.1645.pdf
3. Ishwaran H, Kogalur UB, Blackstone EH et al. Random survival forests. Ann Appl Stat 2008;2:841–860.
4. Breiman L. Out-of-bag estimation. (1996). https://www.stat.berkeley.edu/~breiman/OOBestimation.pdf
5. Ishwaran H, Kogalur UB, Gorodeski E et al. High-dimensional variable selection for survival data. J Am Stat Assoc 2010;105(489):205-217.

**Supplementary Figure 1: Overall survival of 422 poorly differentiated neuroendocrine carcinomas. Source: original.**



**Supplementary Figure 2: Overall survival of 422 poorly differentiated neuroendocrine carcinoma according to morphology. Source: original.**



**Supplementary Figure 3: Variable dependence of predicted survival of the forest at 1, 2 and 3 years on Ki-67 (A), morphology (B), site (C) and stage (D).** Each predicted point is dependent on the full combination of all risk factors, not only on the feature displayed in the dependence plot. Symbols with red circles indicate censored cases, and blue triangles indicate death events. Source: original.





**Supplementary Figure 4: Overall survival according to Ki-67 55% cut-off in stage IV treated patients.**

**Supplementary Figure 5: Overall survival according to Ki-67 55% cut-off in Stage IV patients who received platinum-based chemotherapies as first-line treatment.**



**Supplementary Table 1. Univariable\* and multivariable\* analysis of overall survival of 169 patients with molecular data**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Variable** | **Univariate** | **p-value** | **Multivariable model 2, HR (95% CI)** | **p-value** | **Multivariable model 3, HR (95% CI)** | **p-value** |
| Sex (male vs female) | 1.02 (0.71–1.47) | 0.93 | – | – | – |   |
| Age (increase of 10 years) | 0.99 (0.85–1.17) | 0.95 | – | – | – |   |
| Site: |  |  |  |  |  |  |
| Colorectal | 1.00 |  | 1.00 | – | – |  |
| Gallbladder/Biliary | 0.86 (0.21–3.57) | 0.83 | 0.54 (0.05–5.35) | 0.60 | – |  |
| Gastroesophageal | 0.76 (0.47–1.21) | 0.25 | 0.72 (0.41–1.25) | 0.25 | – |  |
| Lung LCNEC | 0.41 (0.24–0.69) | **0.001** | 0.74 (0.38–1.45) | 0.38 | – |  |
| Pancreas | 0.44 (0.22–0.90) | **0.02** | 0.67 (0.30–1.47) | 0.31 | – |  |
| Stage (III–IV vs I–II) | 1.81 (1.05–3.10) | **0.03** | 1.71 (0.92–3.17) | 0.09 | 1.65 (0.94–2.90) | 0.08 |
| Morphology (pure vs combined) | 1.88 (1.22–2.91) | **0.004** | 1.74 (1.07–2.83) | **0.03** | 1.70 (1.07–2.70) | **0.02** |
| Ki-67 (≥55 vs <55) | 6.53 (3.88–11.00) | **<0.0001** | 6.19 (3.10–12.35) | **<0.0001** | 6.87 (3.87–12.18) | **<0.0001** |
| SSTR2A (2–3 vs 0–1) | 0.79 (0.55–1.14) | 0.21 | – | – | – |  |
| p53 IHC (present vs absent) | 1.12 (0.75–1.66) | 0.58 | – | – | – |  |
| rb1 IHC (present vs absent) | 0.63 (0.44–0.91) | **0.01** | 1.55 (0.95–2.51) | 0.08 | 1.27 (0.84–1.93) | 0.26 |
| *TP53* (mutation vs wild-type)  | 2.61 (1.70–4.00) | **<0.0001** | 1.54 (0.91–2.61) | 0.11 | – |  |
| *KRAS* (mutation vs wild-type) | 1.27 (0.79–2.02) | 0.32 | – | – | – |  |
| *RB1* (mutation vs wild-type) | 1.18 (0.60–2.28) | 0.65 | – | – | – |  |
| *BRAF* (mutation vs wild-type) | 1.89 (0.69–5.21) | 0.22 | – | – | – |  |
| \*Stratified for Center; Multivariable model 2 includes all factors associated (P≤0.05); §Multivariable model 3 retains only variables showing an association (P≤0.10) with overall survival.  |  |  |  |  |  |  |

**Supplementary Table 2.** Number of patients with NEC according to stage, diagnosis and therapy

|  |  |  |
| --- | --- | --- |
| **Patients, n (%)** | **Diagnosis** | **Therapy** |
| **Stage I** |
| 2 (0.8) | LCNEC | Not Performed |
| 1 (0.4) | LCNEC | Other |
| 13 (5.4) | LCNEC | Chemotherapy (platinum) |
| 3 (1.3) | Combined LCNEC | Not Performed |
| 0 (0.0) | Combined LCNEC | Other |
| 6 (2.5) | Combined LCNEC | Chemotherapy (platinum) |
| 0 (0.0) | MINEN | Not Performed |
| 0 (0.0) | MINEN | Other |
| 0 (0.0) | MINEN | Chemotherapy (platinum) |
| 1 (0.4) | NEC | Not Performed |
| 0 (0.0) | NEC | Other |
| 0 (0.0) | NEC | Chemotherapy (platinum) |
| **Stage II** |
| 5 (2.1) | LCNEC | Not Performed |
| 0 (0.0) | LCNEC | Other |
| 9 (3.8) | LCNEC | Chemotherapy (platinum) |
| 2 (0.8) | Combined LCNEC | Not Performed |
| 0 (0.0) | Combined LCNEC | Other |
| 1 (0.4) | Combined LCNEC | Chemotherapy (platinum) |
| 2 (0.8) | MINEN | Not Performed |
| 0 (0.0) | MINEN | Other |
| 0 (0.0) | MINEN | Chemotherapy (platinum) |
| 0 (0.0) | NEC | Not Performed |
| 0 (0.0) | NEC | Other |
| 8 (3.3) | NEC | Chemotherapy (platinum) |
| **Stage III** |
| 1 (0.4) | LCNEC | Not Performed |
| 0 (0.0) | LCNEC | Other |
| 9 (3.8) | LCNEC | Chemotherapy (platinum) |
| 3 (1.3) | Combined LCNEC | Not Performed |
| 0 (0.0) | Combined LCNEC | Other |
| 3 (1.3) | Combined LCNEC | Chemotherapy (platinum) |
| 22 (9.2) | MINEN | Not Performed |
| 6 (2.5) | MINEN | Other |
| 13 (5.4) | MINEN | Chemotherapy (platinum) |
| 5 (2.1) | NEC | Not Performed |
| 9 (3.8) | NEC | Other |
| 30 (12.5) | NEC | Chemotherapy (platinum) |
| **Stage IV** |
| 0 (0.0) | LCNEC | Not Performed |
| 0 (0.0) | LCNEC | Other |
| 6 (2.5) | LCNEC | Chemotherapy (platinum) |
| 0 (0.0) | Combined LCNEC | Not Performed |
| 0 (0.0) | Combined LCNEC | Other |
| 1 (0.4) | Combined LCNEC | Chemotherapy (platinum) |
| 4 (1.7) | MINEN | Not Performed |
| 2 (0.8) | MINEN | Other |
| 1 (0.4) | MINEN | Chemotherapy (platinum) |
| 3 (1.3) | NEC | Not Performed |
| 7 (2.9) | NEC | Other |
| 62 (25.8) | NEC | Chemotherapy (platinum) |

**Supplementary Table 3. Tumor site of stage IV NEC according to therapy**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Therapy** | **Total, n (%)** | **Colorectal, n (%)** | **Gallbladder/biliary, n (%)** | **Gastroesophageal, n (%)** | **Ileum–cecum–duodenum, n (%)** | **Lung LCNEC, n (%)** | **Pancreas, n (%)** |
| Total | 86 (100) | 37 (100) | 1 (100) | 15 (100) | 8 (100) | 7 (100) | 18 (100) |
| Chemotherapy (platinum) | 70 (81.4) | 30 (81.1) | 1 (100) | 11 (73.3) | 7 (87.5) | 7 (100) | 14 (77.8) |
| Other | 9 (10.5) | 2 (5.4) | 0 (0.0) | 2 (13.3) | 1 (12.5) | 0 (0.0) | 4 (22.2) |
| Not performed | 7 (8.1) | 5 (13.5) | 0 (0.0) | 2 (13.3) | 0 (0.0) | 0 (0.0) | 0 (0.0) |