**Supplementary Tables**

**Table S1.** Numbers of features remained after each selection step for CT-based signature building.

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Feature selection steps | Extracted from tumor region | | | | | Extracted from lymph node region | | | | |
| Shape  features | Histogram  features | GLCM  features | GLRLM  features | DL  features | Shape  features | Histogram  features | GLCM  features | GLRLM  features | DL  features |
| Before selection | 4 | 51 | 36 | 42 | 136 | 4 | 51 | 36 | 42 | 136 |
| Reproducibility measurement | 4 | 30 | 22 | 26 | 95 | 4 | 39 | 32 | 41 | 131 |
| Univariate analysis | 0 | 4 | 5 | 6 | 48 | 4 | 18 | 24 | 35 | 100 |
| Feature grouping | 0 | 1 | 3 | 4 | 12 | 1 | 3 | 8 | 12 | 24 |
| LASSO | 0 | 0 | 0 | 0 | 2 | 0 | 0 | 0 | 1 | 2 |

Abbreviations: GLCM, gray-level co-occurrence matrix; GLRLM, gray-level run-length matrix; DL, deep learning; LASSO, least absolute shrinkage and selection operator.

**Table S2.** Numbers of features remained after each selection step for PET-based signature building.

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Feature selection steps | Extracted from tumor region | | | | | Extracted from lymph node region | | | | |
| Shape  features | Histogram  features | GLCM  features | GLRLM  features | DL  features | Shape  features | Histogram  features | GLCM  features | GLRLM  features | DL  features |
| Before selection | 4 | 51 | 36 | 42 | 136 | 4 | 51 | 36 | 42 | 136 |
| Reproducibility measurement | 2 | 40 | 27 | 21 | 116 | 4 | 32 | 16 | 32 | 119 |
| Univariate analysis | 0 | 1 | 0 | 0 | 70 | 4 | 26 | 15 | 22 | 80 |
| Feature grouping | 0 | 1 | 0 | 0 | 26 | 0 | 6 | 6 | 6 | 33 |
| LASSO | 0 | 1 | 0 | 0 | 7 | 0 | 1 | 0 | 1 | 3 |

Abbreviations: GLCM, gray-level co-occurrence matrix; GLRLM, gray-level run-length matrix; DL, deep learning; LASSO, least absolute shrinkage and selection operator.

**Table S3**. Results of radiomics feature selection and Rad-score building.

|  |  |  |
| --- | --- | --- |
| Rad-score | Feature | Coefficient |
| CT-based | CT\_T\_DL2\_H\_Energy | 0.2670 |
|  | CT\_LN\_DL2\_H\_Range | 0.1576 |
|  | CT\_LN\_XL\_GLRLM\_Mean | 0.1122 |
|  | CT\_LN\_DL1\_H\_Energy | 0.0921 |
|  | CT\_T\_DL4\_H\_StandardEntropy | -0.0088 |
| PET-based | PET\_LN\_DL7\_H\_StandardEntropy | -0.3197 |
|  | PET\_T\_X\_H\_Variance | -0.2778 |
|  | PET\_T\_DL1\_H\_StandardEntropy | -0.1920 |
|  | PET\_T\_DL8\_H\_StandardEntropy | -0.1309 |
|  | PET\_LN\_DL8\_H\_Uniformity | 0.0955 |
|  | PET\_LN\_XH\_H\_Range | 0.0913 |
|  | PET\_LN\_DL2\_H\_Range | 0.0573 |
|  | PET\_T\_DL3\_H\_Energy | 0.0482 |
|  | PET\_T\_DL2\_H\_StandardEntropy | -0.0382 |
|  | PET\_T\_DL7\_H\_Entropy | 0.0115 |
|  | PET\_LN\_XL\_GLRLM\_Mean | 0.0097 |
|  | PET\_T\_DL4\_H\_Skewness | 0.0063 |
|  | PET\_T\_DL5\_H\_StandardEntropy | -0.0005 |

Abbreviations: LN, lymph node region; T, tumor region; X, original PET/CT images; XL, original image filtered directionally with low-pass filter along x and y directions; XH, original image filtered directionally with high-pass filter along x and y directions; DLx, xth feature map; H, histogram feature; GLRLM, GLRLM feature.

**Table S4**. Comparison of signatures combining features of different categories.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Signature | Number of involved features | | C-index (95% CI) | |
| DL features | HC features | Training set | Test set |
| CT-based signature | 3 | 2 | 0.738(0.690-0.796) | 0.707(0.635-0.779) |
| CT-based signature with only DL features | 5 | 0 | 0.724(0.676-0.772) | 0.704(0.633-0.775) |
| CT-based signature with only HC features | 0 | 4 | 0.662(0.609-0.716) | 0.634(0.555-0.712) |
| PET-based signature | 10 | 3 | 0.730(0.683-0.776) | 0.683(0.610-0.755) |
| PET-based signature with only DL features | 8 | 0 | 0.693(0.644-0.743) | 0.676(0.603-0.749) |
| PET-based signature with only HC features | 0 | 10 | 0.665(0.613-0.718) | 0.655(0.581-0.729) |

Abbreviations: DL, deep learning; HC, hand-crafted; CI, confidence interval.

**Table S5**. Results of univariate analysis.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Characteristics | Training set | | Test set | |
|  | HR (95% CI) | *P* value | HR (95% CI) | *P* value |
| Age (> 45 vs. ≤ 45y) | 0.93 (0.64-1.36) | 0.721 | 0.90 (0.48-1.69) | 0.750 |
| Gender (female vs. male) | 0.66 (0.41-1.09) | 0.103 | 0.96 (0.56-1.66) | 0.887 |
| Family history of cancer (yes vs. no) | 0.69 (0.45-1.08) | 0.102 | 0.79 (0.42-1.52) | 0.484 |
| Smoking (yes vs. no) | 1.28 (0.87-1.88) | 0.212 | 1.43 (0.83-2.47) | 0.193 |
| Drinking (yes vs. no) | 0.79 (0.42-1.47) | 0.457 | 1.09 (0.51-2.32) | 0.818 |
| HGB (≥ 131 vs. < 131 g/L) | 0.86 (0.53-1.40) | 0.540 | 1.02 (0.52-1.99) | 0.950 |
| ALB (≥ 35 vs. < 35 g/L) | 0.43 (0.16-1.16) | 0.085 | 1.40 (0.78-2.53) | 0.258 |
| CRP (> 8.2 vs. ≤ 8.2 mg/L) | 1.07 (0.64-1.77) | 0.796 | 1.15 (0.56-2.36) | 0.706 |
| LDH (> 245 vs. ≤ 245 U/L) | 1.15 (0.63-2.09) | 0.651 | 1.40 (0.56-3.51) | 0.477 |
| Pre-DNA (> 4000 vs. ≤ 4000 copies/ml) | 1.99 (1.32-2.99) | 0.0007 | 2.50 (1.35-4.66) | 0.0027 |
| T category (T3-4 vs. T1-2) | 0.88 (0.54-1.44) | 0.615 | 1.25 (0.45-3.47) | 0.666 |
| N category (N2-3 vs. N0-1) | 1.96 (1.34-2.88) | 0.0005 | 1.58 (0.92-2.73) | 0.096 |
| Overall stage (IVA vs. III) | 2.47 (1.69-3.61) | < 0.0001 | 2.39 (1.37-4.14) | 0.0014 |

Abbreviations: HR, hazard ratio; CI, confidence interval; HGB, hemoglobin; ALB, albumin; CRP, C-reaction protein; LDH, lactate dehydrogenase; Pre-DNA, pre-treatment plasma Epstein-Barr Virus DNA.

**Table S6**. Multivariate analysis results for disease-free survival in training set.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Endpoint | Variables | HR (95% CI) | Coefficient | *P* value |
| DFS | CT-based signature  (per 1 increase) | 2.99 (1.84-4.86) | 1.095 | < 0.001 |
|  | PET-based signature  (per 1 increase) | 2.32 (1.55-3.46) | 0.840 | < 0.001 |

Abbreviations: DFS, disease-free survival; CT, computed tomography; PET, positron emission tomography; HR, hazard ratio; CI, confidence interval; LDH, lactate dehydrogenase; HGB, hemoglobin; ALB, albumin; CRP, C-reaction protein; Pre-DNA, pre-treatment plasma Epstein-Barr Virus DNA;

*P*-values were calculated by multivariate Cox proportional hazard model with stepwise backward elimination including the following factors: age (> 45 vs. ≤ 45 y), gender (female vs. male); smoking (yes vs. no), drinking (yes vs. no), family history of cancer (yes vs. no), LDH (> 245 vs. ≤ 245 U/L), HGB (≥ 131 vs. < 131 g/L), ALB (≥ 35 vs. < 35 g/L), CRP (> 8.2 vs. ≤ 8.2 mg/L), pre-DNA (> 4000 vs. ≤ 4000 copies/ml), T category (T3-4 vs. T1-2), N category (N2-3 vs. N0-1), overall stage (IVA vs. III), CT-based signature and PET-based signature.

**Table S7**. Baseline information of nomogram-defined high-risk and low-risk groups.

|  |  |  |  |
| --- | --- | --- | --- |
| Characteristics | High-risk group  (n = 206) | Low-risk group  (n = 501) | *P* values a |
|  | No. (%) | No. (%) |  |
| Age (y) |  |  | 0.119 |
| Median (range) | 43 (12-76) | 45 (9-76) |  |
| Gender |  |  | 0.298 |
| Female | 45 (21.8) | 128 (25.5) |  |
| Male | 161 (78.2) | 373 (74.5) |  |
| Smoking |  |  | 0.031 |
| Yes | 85 (41.3) | 164 (32.7) |  |
| No | 121 (58.7) | 337 (67.3) |  |
| Drinking |  |  | 0.843 |
| Yes | 26 (12.6) | 66 (32.8) |  |
| No | 180 (87.4) | 435 (67.2) |  |
| WHO pathology type | |  | 0.380 |
| I | 2 (0.9) | 2 (0.4) |  |
| II-III | 204 (99.1) | 499 (99.6) |  |
| Family history of cancer | |  | 0.073 |
| Yes | 50 (24.3) | 155 (30.9) |  |
| No | 156 (75.7) | 345 (69.1) |  |
| LDH (U/L) |  |  | 0.001 |
| Median (range) | 186 (116-658) | 174 (100-626) |  |
| HGB (g/L) |  |  | 0.383 |
| Median (range) | 143 (95-171) | 146 (79-178) |  |
| ALB (g/L) |  |  | 0.493 |
| Median (range) | 43.8 (31-52) | 44.4 (25-54) |  |
| CRP (mg/L) |  |  | < 0.001 |
| Median (range) | 3.2 (0.3-126.6) | 1.8 (0-127.2) |  |
| Pre-DNA (copies/ml) b | |  | < 0.001 |
| Median (range) | 33000 (0-68700000) | 1760 (0-9240000) |  |
| T category c |  |  | < 0.001 |
| T1 | 5 (2.4) | 29 (5.8) |  |
| T2 | 26 (12.6) | 37 (7.4) |  |
| T3 | 92 (44.7) | 346 (69.0) |  |
| T4 | 83 (40.3) | 89 (17.8) |  |
| N category c |  |  | < 0.001 |
| N0 | 7 (3.4) | 63 (12.6) |  |
| N1 | 52 (25.2) | 265 (53.0) |  |
| N2 | 74 (36.0) | 119 (23.6) |  |
| N3 | 73 (35.4) | 54 (10.8) |  |
| Overall stage c |  |  | < 0.001 |
| III | 70 (34.0) | 365 (72.8) |  |
| IVA | 136 (66.0) | 136 (27.2) |  |
| Treatment |  |  | < 0.001 |
| IC+CCRT | 173 (84.0) | 296 (59.0) |  |
| CCRT alone | 33 (16.0) | 205 (41.0) |  |

Abbreviations: WHO, world health organization; LDH, lactate dehydrogenase; HGB, hemoglobin; ALB, albumin; CRP, C-reaction protein; Pre-DNA, pre-treatment plasma Epstein-Barr Virus DNA; IC, induction chemotherapy; CCRT, concurrent chemoradiotherapy; AC, adjuvant chemotherapy.

a *P* values were calculated by Chi-square test for categorical variables and non-parametric test for continuous variables.

b 23 patients lost this data.

c According to the 8th edition of the International Union against Cancer/American Joint Committee on Cancer (UICC/AJCC) staging manual.

**Table S8**. Survival outcomes between radiomics nomogram-defined high and low risk groups.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Endpoints | Training set | | *P* values | Test set | | *P* values |
|  | High-risk group | Low-risk group |  | High-risk group | Low-risk group |  |
| 1-year DFS | 85.8% | 97.3% | < 0.0001 | 84.5% | 98.2% | < 0.0001 |
| 3-year DFS | 54.6% | 91.0% | 64.6% | 89.7% |
| 5-year DFS | 46.7% | 88.6% | 57.4% | 85.6% |
| 1-year OS | 95.5% | 100% | < 0.0001 | 97.2% | 100% | < 0.0001 |
| 3-year OS | 82.6% | 97.0% | 84.5% | 97.0% |
| 5-year OS | 64.0% | 93.4% | 70.9% | 94.3% |
| 1-year DMFS | 91.7% | 98.5% | < 0.0001 | 90.1% | 98.2% | < 0.0001 |
| 3-year DMFS | 68.9% | 96.1% | 78.5% | 94.0% |
| 5-year DMFS | 62.3% | 94.8% | 72.3% | 91.7% |
| 1-year LRRFS | 94.6% | 98.8% | < 0.0001 | 92.9% | 100% | 0.004 |
| 3-year LRRFS | 79.5% | 95.5% | 81.8% | 93.3% |
| 5-year LRRFS | 75.8% | 93.9% | 76.6% | 89.7% |

Abbreviations: DFS, disease-free survival; OS, overall survival; DMFS, distant metastasis-free survival; LRRFS, locoregional relapse-free survival.

**Table S9**. Results of multivariate analysis for the whole cohort.

|  |  |  |  |
| --- | --- | --- | --- |
| Endpoints | Variable | HR (95% CI) | *P* values b |
| DFS | EBV DNA (> 4000 vs. ≤ 4000 copies/ml) a | 1.704 (1.196-2.428) | 0.003 |
|  | Overall stage (IVA vs. III) | 2.097 (1.501-2.930) | < 0.001 |
|  | Treatment (IC+CCRT vs. CCRT) | 0.793 (0.553-1.135) | 0.204 |
| OS | Gender (female vs. male) | 0.518 (0.293-0.914) | 0.023 |
|  | EBV DNA (> 4000 vs. ≤ 4000 copies/ml) a | 1.581 (1.020-2.452) | 0.041 |
|  | Overall stage (IVA vs. III) | 2.554 (1.674-3.896) | < 0.001 |
|  | Treatment (IC+CCRT vs. CCRT) | 0.774 (0.489-1.224) | 0.273 |
| DMFS | EBV DNA (> 4000 vs. ≤ 4000 copies/ml) a | 2.031 (1.254-3.290) | 0.004 |
|  | N category (N2-3 vs. N0-1) | 1.636 (1.043-2.568) | 0.032 |
|  | Overall stage (IVA vs. III) | 2.054 (1.329-3.177) | 0.001 |
|  | Treatment (IC+CCRT vs. CCRT) | 0.694 (0.437-1.101) | 0.121 |
| LRRFS | Smoking (yes vs. no) | 1.650 (1.047-2.599) | 0.031 |
|  | EBV DNA (> 4000 vs. ≤ 4000 copies/ml) a | 1.718 (1.049-2.812) | 0.031 |
|  | T category (T3-4 vs. T1-2) | 2.744 (1.158-6.504) | 0.022 |
|  | N category (N2-3 vs. N0-1) | 1.759 (1.083-2.857) | 0.022 |
|  | Treatment (IC+CCRT vs. CCRT) | 0.758 (0.453-1.268) | 0.292 |

Abbreviations: DFS, disease-free survival; OS, overall survival; DMFS, distant metastasis-free survival; LRRFS, locoregional relapse-free survival; HR, hazard ratio; CI, confidence interval; EBV, Epstein-Barr virus; IC, induction chemotherapy; CCRT, concurrent chemoradiotherapy; LDH, lactate dehydrogenase; HGB, hemoglobin; ALB, albumin; CRP, C-reaction protein;

a 21 patients lost the data on EBV DNA and were not included.

b *P* values were calculated by adjusted Cox proportional hazard model with backward elimination including the following factors: age (> 45y vs. ≤ 45y), gender (female vs. male), smoking (yes vs. no), drinking (yes vs. no), family history of cancer (yes vs. no), LDH (> 245 vs. ≤ 245 U/L), HGB (≥ 131 vs. < 131 g/L), ALB (≥ 35 vs. < 35 g/L), CRP (> 8.2 vs. ≤ 8.2 mg/L), EBV DNA (> 4000 vs. ≤ 4000 copies/ml), T category (T3-4 vs. T1-2), N category (N2-3 vs. N0-1), overall stage (IVA vs. III) and treatment (IC+CCRT vs. CCRT alone).

**Table S10**. Results of multivariate analysis within the high-risk group.

|  |  |  |  |
| --- | --- | --- | --- |
| Endpoints | Variable | HR (95% CI) | *P* values a |
| DFS | Treatment (IC+CCRT vs. CCRT) | 0.472 (0.291-0.766) | 0.002 |
| OS | Treatment (IC+CCRT vs. CCRT) | 0.314 (0.181-0.545) | < 0.001 |
| DMFS | Treatment (IC+CCRT vs. CCRT) | 0.333 (0.193-0.576) | < 0.001 |
| LRRFS | Treatment (IC+CCRT vs. CCRT) | 0.665 (0.288-1.534) | 0.338 |

Abbreviations: DFS, disease-free survival; OS, overall survival; DMFS, distant metastasis-free survival; LRRFS, locoregional relapse-free survival; HR, hazard ratio; CI, confidence interval; EBV, Epstein-Barr virus; IC, induction chemotherapy; CCRT, concurrent chemoradiotherapy; LDH, lactate dehydrogenase; HGB, hemoglobin; ALB, albumin; CRP, C-reaction protein;

a *P* values were calculated by adjusted Cox proportional hazard model with backward elimination including the following factors: age (> 45y vs. ≤ 45y), gender (female vs. male), smoking (yes vs. no), drinking (yes vs. no), family history of cancer (yes vs. no), LDH (> 245 vs. ≤ 245 U/L), HGB (≥ 131 vs. < 131 g/L), ALB (≥ 35 vs. < 35 g/L), CRP (> 8.2 vs. ≤ 8.2 mg/L), EBV DNA (> 4000 vs. ≤ 4000 copies/ml), T category (T3-4 vs. T1-2), N category (N2-3 vs. N0-1), overall stage (IVA vs. III) and treatment (IC+CCRT vs. CCRT alone).

**Table S11**. Results of multivariate analysis within the low-risk group.

|  |  |  |  |
| --- | --- | --- | --- |
| Endpoints | Variable | HR (95% CI) | *P* values a |
| DFS | Overall stage (IVA vs. III) | 1.843 (1.080-3.145) | 0.025 |
|  | Treatment (IC+CCRT vs. CCRT) | 0.683 (0.397-1.174) | 0.168 |
| OS | Overall stage (IVA vs. III) | 2.227 (1.079-4.593) | 0.03 |
|  | Treatment (IC+CCRT vs. CCRT) | 1.007 (0.464-2.186) | 0.986 |
| DMFS | Treatment (IC+CCRT vs. CCRT) | 0.783 (0.350-1.750) | 0.551 |
| LRRFS | Smoking (yes vs. no) | 2.022 (1.021-4.002) | 0.043 |
|  | Treatment (IC+CCRT vs. CCRT) | 0.670 (0.338-1.329) | 0.252 |

Abbreviations: DFS, disease-free survival; OS, overall survival; DMFS, distant metastasis-free survival; LRRFS, locoregional relapse-free survival; HR, hazard ratio; CI, confidence interval; EBV, Epstein-Barr virus; IC, induction chemotherapy; CCRT, concurrent chemoradiotherapy; LDH, lactate dehydrogenase; HGB, hemoglobin; ALB, albumin; CRP, C-reaction protein;

a *P* values were calculated by adjusted Cox proportional hazard model with backward elimination including the following factors: age (> 45y vs. ≤ 45y), gender (female vs. male), smoking (yes vs. no), drinking (yes vs. no), family history of cancer (yes vs. no), LDH (> 245 vs. ≤ 245 U/L), HGB (≥ 131 vs. < 131 g/L), ALB (≥ 35 vs. < 35 g/L), CRP (> 8.2 vs. ≤ 8.2 mg/L), EBV DNA (> 4000 vs. ≤ 4000 copies/ml), T category (T3-4 vs. T1-2), N category (N2-3 vs. N0-1), overall stage (IVA vs. III) and treatment (IC+CCRT vs. CCRT alone).

**Table S12**. Results of multivariate analysis within the low-risk group defined by nomogram A.

|  |  |  |  |
| --- | --- | --- | --- |
| Endpoints | Variable | HR (95% CI) | *P* values a |
| DFS | Drinking (yes vs. no) | 3.423 (1.073-10.913) | 0.038 |
|  | EBV DNA (> 4000 vs. ≤ 4000 copies/ml) | 3.249 (1.903-5.547) | < 0.001 |
|  | Overall stage (IVA vs. III) | 12.078 (1.607-90.79) | 0.015 |
|  | Treatment (IC+CCRT vs. CCRT) | 1.137 (0.653-1.979) | 0.651 |
| OS | ALB (≥ 35 vs. < 35 g/L) | 0.13 (0.031-0.552) | 0.006 |
|  | EBV DNA (> 4000 vs. ≤ 4000 copies/ml) | 2.892 (1.387-6.029) | 0.005 |
|  | Overall stage (IVA vs. III) | 20.283 (2.524-163.009) | 0.005 |
|  | Treatment (IC+CCRT vs. CCRT) | 1.139 (0.536-2.422) | 0.734 |
| DMFS | ALB (≥ 35 vs. < 35 g/L) | 0.149 (0.035-0.636) | 0.01 |
|  | EBV DNA (> 4000 vs. ≤ 4000 copies/ml) | 3.429 (1.602-7.341) | 0.002 |
|  | N category (N2-3 vs. N0-1) | 1.998 (1.027-3.889) | 0.042 |
|  | Overall stage (IVA vs. III) | 18.612 (2.298-150.717) | 0.006 |
|  | Treatment (IC+CCRT vs. CCRT) | 1.477 (0.674-3.237) | 0.33 |
| LRRFS | EBV DNA (> 4000 vs. ≤ 4000 copies/ml) | 2.497 (1.249-4.993) | 0.01 |
|  | Treatment (IC+CCRT vs. CCRT) | 0.858 (0.42-1.751) | 0.674 |

Abbreviations: DFS, disease-free survival; OS, overall survival; DMFS, distant metastasis-free survival; LRRFS, locoregional relapse-free survival; HR, hazard ratio; CI, confidence interval; EBV, Epstein-Barr virus; IC, induction chemotherapy; CCRT, concurrent chemoradiotherapy; LDH, lactate dehydrogenase; HGB, hemoglobin; ALB, albumin; CRP, C-reaction protein;

a *P* values were calculated by adjusted Cox proportional hazard model with backward elimination including the following factors: age (> 45y vs. ≤ 45y), gender (female vs. male), smoking (yes vs. no), drinking (yes vs. no), family history of cancer (yes vs. no), LDH (> 245 vs. ≤ 245 U/L), HGB (≥ 131 vs. < 131 g/L), ALB (≥ 35 vs. < 35 g/L), CRP (> 8.2 vs. ≤ 8.2 mg/L), EBV DNA (> 4000 vs. ≤ 4000 copies/ml), T category (T3-4 vs. T1-2), N category (N2-3 vs. N0-1), overall stage (IVA vs. III) and treatment (IC+CCRT vs. CCRT alone).

**Table S13**. Results of multivariate analysis within the high-risk group defined by nomogram A.

|  |  |  |  |
| --- | --- | --- | --- |
| Endpoints | Variable | HR (95% CI) | *P* values a |
| DFS | Smoking (yes vs. no) | 1.696 (1.117-2.575) | 0.013 |
|  | Family history of cancer (yes vs. no) | 0.529 (0.307-0.911) | 0.022 |
|  | Treatment (IC+CCRT vs. CCRT) | 0.898 (0.524-1.538) | 0.695 |
| OS | Gender (female vs. male) | 0.436 (0.197-0.962) | 0.04 |
|  | Treatment (IC+CCRT vs. CCRT) | 1.379 (0.650-2.924) | 0.402 |
| DMFS | T category (T3-4 vs. T1-2) | 1.905 (1.001-3.624) | 0.05 |
|  | Treatment (IC+CCRT vs. CCRT) | 0.863 (0.423-1.758) | 0.685 |
| LRRFS | Smoking (yes vs. no) | 1.972 (1.044-3.726) | 0.036 |
|  | N category (N2-3 vs. N0-1) | 2.101 (1.034-4.268) | 0.04 |
|  | Treatment (IC+CCRT vs. CCRT) | 0.875 (0.396-1.932) | 0.740 |

Abbreviations: DFS, disease-free survival; OS, overall survival; DMFS, distant metastasis-free survival; LRRFS, locoregional relapse-free survival; HR, hazard ratio; CI, confidence interval; EBV, Epstein-Barr virus; IC, induction chemotherapy; CCRT, concurrent chemoradiotherapy; LDH, lactate dehydrogenase; HGB, hemoglobin; ALB, albumin; CRP, C-reaction protein;

a *P* values were calculated by adjusted Cox proportional hazard model with backward elimination including the following factors: age (> 45y vs. ≤ 45y), gender (female vs. male), smoking (yes vs. no), drinking (yes vs. no), family history of cancer (yes vs. no), LDH (> 245 vs. ≤ 245 U/L), HGB (≥ 131 vs. < 131 g/L), ALB (≥ 35 vs. < 35 g/L), CRP (> 8.2 vs. ≤ 8.2 mg/L), EBV DNA (> 4000 vs. ≤ 4000 copies/ml), T category (T3-4 vs. T1-2), N category (N2-3 vs. N0-1), overall stage (IVA vs. III) and treatment (IC+CCRT vs. CCRT alone).

**Table S14**. Results of multivariate analysis within the low-risk group defined by nomogram B.

|  |  |  |  |
| --- | --- | --- | --- |
| Endpoints | Variable | HR (95% CI) | *P* values a |
| DFS | EBV DNA (> 4000 vs. ≤ 4000 copies/ml) | 1.647 (1.006-2.699) | 0.047 |
|  | Overall stage (IVA vs. III) | 2.098 (1.306-3.371) | 0.002 |
|  | Treatment (IC+CCRT vs. CCRT) | 1.335 (0.730-2.440) | 0.348 |
| OS | Overall stage (IVA vs. III) | 2.660 (1.512-4.679) | 0.001 |
|  | Treatment (IC+CCRT vs. CCRT) | 1.645 (0.781-3.468) | 0.191 |
| DMFS | N category (N2-3 vs. N0-1) | 2.428 (1.214-4.856) | 0.012 |
|  | Overall stage (IVA vs. III) | 2.253 (1.156-4.392) | 0.017 |
|  | Treatment (IC+CCRT vs. CCRT) | 1.421 (0.590-3.423) | 0.434 |
| LRRFS | N category (N2-3 vs. N0-1) | 2.054 (1.004-4.200) | 0.049 |
|  | Treatment (IC+CCRT vs. CCRT) | 1.248 (0.536-2.909) | 0.607 |

Abbreviations: DFS, disease-free survival; OS, overall survival; DMFS, distant metastasis-free survival; LRRFS, locoregional relapse-free survival; HR, hazard ratio; CI, confidence interval; EBV, Epstein-Barr virus; IC, induction chemotherapy; CCRT, concurrent chemoradiotherapy; LDH, lactate dehydrogenase; HGB, hemoglobin; ALB, albumin; CRP, C-reaction protein;

a *P* values were calculated by adjusted Cox proportional hazard model with backward elimination including the following factors: age (> 45y vs. ≤ 45y), gender (female vs. male), smoking (yes vs. no), drinking (yes vs. no), family history of cancer (yes vs. no), LDH (> 245 vs. ≤ 245 U/L), HGB (≥ 131 vs. < 131 g/L), ALB (≥ 35 vs. < 35 g/L), CRP (> 8.2 vs. ≤ 8.2 mg/L), EBV DNA (> 4000 vs. ≤ 4000 copies/ml), T category (T3-4 vs. T1-2), N category (N2-3 vs. N0-1), overall stage (IVA vs. III) and treatment (IC+CCRT vs. CCRT alone).

**Table S15**. Results of multivariate analysis within the high-risk group defined by nomogram B.

|  |  |  |  |
| --- | --- | --- | --- |
| Endpoints | Variable | HR (95% CI) | *P* values a |
| DFS | Smoking (yes vs. no) | 1.927 (1.230-3.019) | 0.004 |
|  | Family history of cancer (yes vs. no) | 0.548 (0.315-0.952) | 0.033 |
|  | EBV DNA (> 4000 vs. ≤ 4000 copies/ml) | 2.218 (1.320-3.726) | 0.003 |
|  | Overall stage (IVA vs. III) | 2.082 (1.300-3.336) | 0.002 |
|  | Treatment (IC+CCRT vs. CCRT) | 1.037 (0.609-1.766) | 0.893 |
| OS | Overall stage (IVA vs. III) | 2.528 (1.298-4.923) | 0.006 |
|  | Treatment (IC+CCRT vs. CCRT) | 1.301 (0.596-2.840) | 0.509 |
| DMFS | EBV DNA (> 4000 vs. ≤ 4000 copies/ml) | 2.203 (1.147-4.228) | 0.018 |
|  | Overall stage (IVA vs. III) | 1.997 (1.114-3.579) | 0.02 |
|  | Treatment (IC+CCRT vs. CCRT) | 1.527 (0.760-3.066) | 0.235 |
| LRRFS | Smoking (yes vs. no) | 3.435 (1.783-6.618) | < 0.001 |
|  | EBV DNA (> 4000 vs. ≤ 4000 copies/ml) | 3.686 (1.690-8.042) | 0.001 |
|  | Treatment (IC+CCRT vs. CCRT) | 0.686 (0.345-1.365) | 0.283 |

Abbreviations: DFS, disease-free survival; OS, overall survival; DMFS, distant metastasis-free survival; LRRFS, locoregional relapse-free survival; HR, hazard ratio; CI, confidence interval; EBV, Epstein-Barr virus; IC, induction chemotherapy; CCRT, concurrent chemoradiotherapy; LDH, lactate dehydrogenase; HGB, hemoglobin; ALB, albumin; CRP, C-reaction protein;

a *P* values were calculated by adjusted Cox proportional hazard model with backward elimination including the following factors: age (> 45y vs. ≤ 45y), gender (female vs. male), smoking (yes vs. no), drinking (yes vs. no), family history of cancer (yes vs. no), LDH (> 245 vs. ≤ 245 U/L), HGB (≥ 131 vs. < 131 g/L), ALB (≥ 35 vs. < 35 g/L), CRP (> 8.2 vs. ≤ 8.2 mg/L), EBV DNA (> 4000 vs. ≤ 4000 copies/ml), T category (T3-4 vs. T1-2), N category (N2-3 vs. N0-1), overall stage (IVA vs. III) and treatment (IC+CCRT vs. CCRT alone).