

**Appendix 1 The tumor-node-metastasis (TNM) re-stage criteria based on the 8th edition of the American Joint Committee on Cancer (AJCC) Cancer Staging Manual**

*T stage*

**T3**

Any of the following characteristics:

- (I) Size: >5 cm but <7 cm;
- (II) Local invasion: direct invasion of chest wall (including superior sulcus tumors), parietal pleura (PL3), phrenic nerve, or parietal pericardium.

**T4**

Any of the following characteristics:

- (I) Size >7 cm;
- (II) Airway location: invasion of the carina or trachea;
- (III) Local invasion: diaphragm, mediastinum, heart, great vessels, recurrent laryngeal nerve, esophagus or vertebral body.

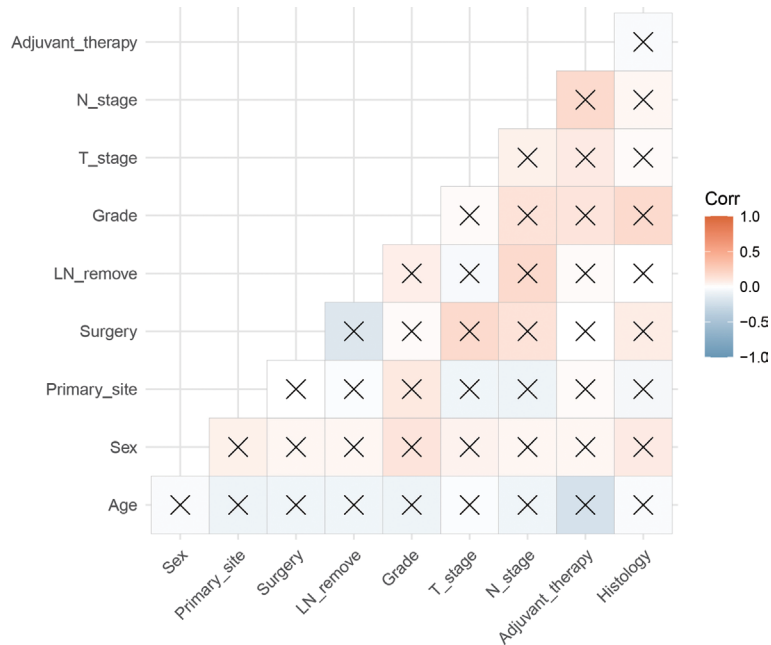
*N stage (not corrected; same as the 7th edition AJCC)*

**N0**

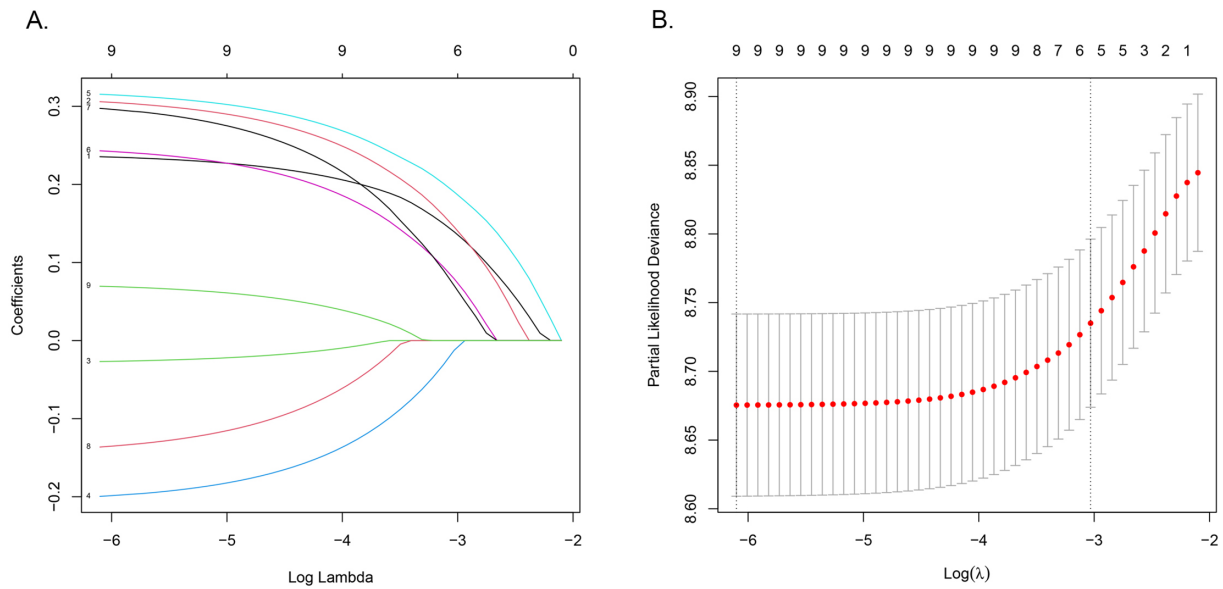
No regional lymph nodes involvement.

**N1**

Involvement of ipsilateral peribronchial and/or ipsilateral hilar lymph nodes (includes direct extension to intrapulmonary nodes).



**Figure S1** Variable correlation analysis. The correlation analysis plot demonstrates that there is no strong correlation between any of the variables included in this study. The symbol “X” represents a lack of significant correlation. LN, lymph node; Corr, correlation.



**Figure S2** LASSO regression detects multicollinearity of variables. (A) The coefficient distribution plot; (B) the error parameter plot. LASSO regression results suggest that all nine variables we selected are important, and the final model does not suffer from multicollinearity issues. LASSO, least absolute shrinkage and selection operator.