Comments to the Authors,

Dr. Li and colleagues provided a complementary system to TNM cancer staging system which is based on lymph node ratio (LNR). Two independent dataset with large sample size (N= 4262 and 154) were collected to test the performance of LNR stage. The authors found LNR classification should be considered as a complementary system to TNM staging. The study was performed rigorously and the findings are interesting. I only have several quite slight concerns.

**Major Compulsory Revisions**

1, In the validation stage, why only 154 samples were collected while the test dataset is as high as 4362? Why not provided power estimation for the validation stage? It would be better to tell the readers the sample in validate dataset is power enough for the validation.

2, Please provide the correlation between LNR with all the matric in the present study as the supplementary, such as T, N, M, TNM, age and so on for two datasets.

**Minor Essential Revisions**

1, Line number should be provided in the revised version or else it is difficult to mention where is the problems of the manuscript.

2, How to determined 10,000 iterations not 100,000 iterations or others?

3, in the table 2, the author mentioned The multivariate analysis was adjusted using the same Cox regression model at Table 2, However, in the footnotes of table 2, no any confounder were mentioned to be adjusted.

4, I am not interesting why TRM stage was not validated in FUDAN dataset? sample size is not enough?

5, Table 1 is not the basic characteristic description of public dataset. Maybe it is not qualified to be the main table in the manuscript without any analysis or conclusion. It would be transfer to supplementary.

6, It seems the authors didn’t consider multiple test correction and the P<0.001 was applied for the minimum P-value threshold, why not provided accurate P-value?

7, How to explain the C-INDEX for N, R, TNM, TRM models are quite similar with each other in Table 4?

8, In the abstract, the values or regions of R0 should be mentioned since the value for other categories have been list.

9, in the manuscript, (such as page 6) the authors used R classification frequently, However, I think should be “R categorical classification” right?