Comments to the Authors,

This manuscript conducted a case-control association study to evaluate the association of four SNPs rs4035540, rs5762746, rs2236141 and rs2236142, located in CHEK2 and toxicity, survival characteristics in a Chinese lung cancer population. The authors found rs4035540 was significantly associated with overall survival and gastrointestinal toxicity, respectively. The findings are interesting, However, I have several major concerns.

2, All the association which are claimed as significant association actually very weak, the P-value ranges from 0.01 to 0.05. Such kinds of P-value would turn to non-significant with multiple test correction. That means the possibility of the finding to be false positive is very high. How to deal with this problem should be mentioned by the authors to guarantee the conclusion is solid.

2, The authors claimed “it may be a new biomarker to predict the prognosis to non-smoking NSCLC female patients in Chinese population”, However, the prediction sensitivity, specificity, accuracy or ROC were not presented at all? I think it is over-claimed or the author should provide the evidence.

3, The authors design this study to demonstrate the prognostic prediction value of CHEK2. I recommend the authors to describe this clinic question more comprehensive, such as how many genes were potential to be as the prediction biomarker till now, such as XRCC1, so that the readers can be clear about this field with more background knowledge. And after the author finished CHEK2, could you compare the performance between CHEK2 and other reported relevant genes?

4,  In the last section of the result, the authors conducted haplotype analysis. However, What’s the motivation to do it? What’s conclusion? The haplotype construction were conducted in case or control samples or the whole samples? Any association analysis between haplotype and disease OS or toxicity？

5, Actually, the detection of the SNPs were quite easy and low cost, while the samples and patients clinical information is more valuable, more SNPs should be detected at the same time to provide more comprehensive and valuable discoveries. Any specific reason for the author only conducted 4 SNPs in the current project?