Comments to the Authors,

In the present study, Dapeng Hao and his colleagues investigated the correlation between cancer risk and tumor genomic mutation rate and provided an important evidence of the predominant contribution of somatic mutation to cancer risk. This study is a critical and prompt response to the recent paper published on the journal of *Science* few month ago. The authors collected huge number of genome-wide or exon-sequence data from 5,542 cancer samples throughout 41 different cancer types. The authors found the revised Armitage-Doll model can interpret the relationship between mutation counts and cancer risk with high accuracy. The study was performed rigorously and the findings are very interesting. In general, I'd recommend publication if the authors can address the following concerns.

**Maor Essential Revisions**

1, What would happen if the number of the stem cell divisions was adjusted in the models?

**Minor Essential Revisions**

1， the inclusion criteria and exclusion criteria were not stated. The diagnosis critieria of RA were not mentioned

2， No disease controls were provided.

3， Whether the most important confounder was analyzed.

There is increasing awareness that sharing high-dimensional genomics data is essential for the cancer genomics field to make sustainable translational contributions. This way, safe long-term data storage is ensured, independent researchers can reproduce and thereby validate analyses, perform powerful meta-analysis or use publicly available data to help interpret new data.

1, In the section of “Robustness Analysis”, the distribution of Pearson, Spearman and corresponding P-value in the 10,000 iterations should be provided in the supplementary.

2, The influence of the variation in the estimation of lifetime risk for each cancer to the correlation between mutation ratio and cancer risk should be evaluated.

3, A detailed definition and selection of consensus mutations as well as a complete list of consensus mutations should be provided as the supplementary and the influence of the variation of the number of the mutation to the conclusion should be validated.

4, In the supplementary section, authors has provided the mutation counts per Mb and lifetime risk for each cancer, however, the explicit data which was used to establish the Figure 1 should be showed as the form of table.

5, In the analysis, the authors actually established a log-log relationship between mutation counts and lifetime risk. However, in the title, the author used “mutation rate”. Is there any difference between these two terms, especially when considering difference division speed for different cell type?

6, It seems some importantly related works were not mentioned in the background section, such as \*\* to introduced \*\* and to introduced \*\*.

**Discretionary Revisions**

1, the role of the factorial of N in the formula of logarithm of cancer lifetime risk (Armitage-Doll model) might be need some introduction.

Tables longer than one page should be provided as an Excel or similar file type

Figures should be well prepared with high-resolution and the optimum font size is 8pt.