Reviewer's report:

In this study, the authors explored the cancer risks among patients with type 2 diabetes mellitus (T2DM) through a national population-based cohort study that included diabetic patients and the general population in Taiwan by using standardized incidence ratios (SIRs) and population attributable fractions (PAFs). The authors demonstrate that increased risks of cancer are observed in liver, colorectal, oral, pancreatic, and kidney cancers in men and in liver, colorectal, breast, pancreatic, endometrium, bladder, and kidney cancers in women. The finding indicates that unusual risks of cancer are associated with T2DM. This study is the largest study to examine the SIRs and PAFs of diabetes on site-specific cancer incidence for the Taiwanese population. It was valuable to show the risks of site-specific cancer among men and women with DM in Taiwan. However, a few concerns should be addressed:

Major Compulsory Revisions

1. Study subjects were selected by ICD-9-CM code 250 and A-code A181. Depending on these criteria, the study subjects included type 1 and type 2 DM patients. Owing to the different mechanism of type I and type II DM patients, it would be nice if only to screen specific type 2 DM patients in this study.

2. The residential area information should not use the variable in insured dataset directly because this area variable did not present where the man lives, it presented the area where his company was.

3. Metformin is the most commonly prescribed drug for type II diabetes. Previous studies show that metformin is associated with decreased cancer risk. As I know, the prescription information was included in NHI database, please group study subject by different therapy in order to eliminate the effect of DM treatment.

4. The results showed some higher risks of site-specific cancers and some lower risks of other site-specific cancers. Please add and discuss the clinical implication of these results in this study.

5. In discussion, page 12, section 2, the authors explained the possible situation for the inconsistency results analyzed from the Registry of Catastrophic Illness database “because catastrophic illness status is not compulsory”. But, in clinical practice these days, the cancer patients will get the catastrophic application form
when he or she was diagnosed with catastrophic disease. And nurses or administrators will help them to apply the catastrophic status to free from copayment to NHI. Most of cancer patients get the catastrophic card. Also, the Taiwan Cancer Registry used the Catastrophic Illness database to double check the cancer patients. The result from Catastrophic Illness database was authentic.

Minor Essential Revisions

1. Authors used “Chinese” and “Taiwanese” to describe study subjects. All of the insured in NHIRD (National Health Insurance Research Database) were Taiwan people. If “Chinese” and “Taiwanese” used together, that will confuse audience. Please correct all the “Chinese” to “patients in Taiwan” or “population in Taiwan”.

2. Due to Human subject protection statement, all of the studies use NHI database must receive approval from institutional review board (IRB), please provide clarifying information.

3. In methodology, Page 6, section 2, the authors identify lots of site-specific cancers but without showing ICD-9-CM code. Please show the ICD-9-CM codes that used to identify the outcome cancer you observed in the manuscript or in Table 2.

4. Please check carefully for grammar, syntax, and usage. Some examples stated below:
   a. Page 14, this study was supported ….., this sentence was not complete, please delete.
   b. In Table 1, column 2 showed “general population” and column 4 showed “no diabetes”, please be consistent.
   c. In Table 1, column 1, age “ 90#” please be consistent with insurance premium “ #19200”, change # to # .
   d. In Table 2, annotation (b) and (c) were not note in the table.
   e. In Table 2, column 3 and 5, there was an annotation “#” next to the incidence density, but without explain below the table.
   f. The number of study subject in Figure 1 and Page 7 was inconsistent. Page 7, line 9, with type 2 diabetes aged <20 years (N=19986). In Figure 1, the number was 19098.
   g. The number of subjects were eligible on the left side in Figure 1 was wrong, the 478,910 should be 478,022.
   h. The number of subjects without diabetes were eligible on the right side in Figure 1 was wrong, the 16,535,274 should be 9,478,609, please correct it.
   i. In Figure 2, the bar charts of “pancreas in woman” and “bladder in woman” were absent. Please provide these 2 bars.

5. Finally, in Figure 1, the included participants were 474686. It was not consistent with the number of patients with type2 diabetes in 1997-1998 minus subjects that was excluded or without information for gender and residential area.
That might due to the overlapping of those excluding criteria, please clarify.

**Level of interest:** An article whose findings are important to those with closely related research interests

**Quality of written English:** Acceptable

**Statistical review:** Yes, and I have assessed the statistics in my report.

**Declaration of competing interests:**

I declare that I have no competing interests.