Author's response to reviews

Title: Incidence rates of tuberculosis in chronic hepatitis C infected patients with or without interferon based therapy: a population-based cohort study in Taiwan

Authors:

Shang-Yi Lin (amoe616@gmail.com)
Tun-Chieh Chen (kmtthidchen@gmail.com)
Po-Liang Lu (idpaul@gmail.com)
Chun-Yu Lin (infectionman@gmail.com)
Wei-Ru Lin (lwru@mail2000.com.tw)
Yi-Hsin Yang (yihsya@kmu.edu.tw)
Yen-Hsu Chen (infchen@gmail.com)

Version: 3
Date: 17 October 2014

Author's response to reviews: see over
Re: MS: 4705416191378548 Incidence rates of tuberculosis in chronic hepatitis C infected patients with or without interferon based therapy: a population-based cohort study in Taiwan

Dear Editor,

A point-by-point response to the peer reviews of our original manuscript, #MS: 4705416191378548, is presented in this reply letter according to the reviewers’ comments and suggestion. All authors have read and agreed to the re-submitted version of the manuscript. The reviewers’ queries and suggestions are addressed item by item as follows, highlighted in Bold format:

Reviewer 1, Ulla Beijer

The results section begins with "Baseline characteristics of the study population". This entire first paragraph should instead belong to method section (including Figure 1).

***

We rewrote the first paragraph and added the comparisons between the two cohorts. Thanks for reviewer’s comments!
P13. line 241-250

Table 1 should not show both characteristics from the baseline and the results / outcome of TB. The main part of Table 1 also belong to the method section.

***

We deleted “outcome of TB” in table 1 and added it to the results section
P14, line 257-260
Table 4 belongs to literature part in the Background. This does not prevent that the literature are also discussed in the discussion, but without a reference to the table.

***

Thanks for reviewer’s kind comments. We revised this paragraph and deleted these statements about case reports listed in table 4. P16 line 288-292
Reviewer 2, Andrew Edmonds

We do appreciate the reviewer’s comments and suggestion to improve the manuscript. Many thanks.

Major Compulsory Revisions
(which the author must respond to before a decision on publication can be reached)

1. Line 19: According to the title, the paper is about TB incidence rates in HCV-infected patients with or without IBT. Shouldn’t the running head be reflect the fact that you’re not just looking at patients *ON* IBT?

***

We modified the running head as “Tuberculosis in HCV infected patients with or without interferon-based therapy”.

P 1, line 19

2. Line 40-41: Do you mean in patients in general, or amongst HCV-infected patients? Clarify. Also, isn’t it debated whether IBT increases/decreases the rate of TB, not only whether it increases the rate?

***

We revised the sentence as “It is debated whether interferon-based therapy (IBT) would affect the incidence of active tuberculosis (TB) among hepatitis C (HCV) infected patients.”

P3, line 40-41

3. Lines 45-51: The methodology is not clear. Did you take a random sample of the 1,000,000 and find 8,321 HCV-infected people in that sample? If so, how large was the random sample? Or, did you just find 8,321 in the sample of 1,000,000?

***

This study used data from the Taiwan National Health Insurance Research Database (NHIRD). In 1995, universal health insurance was initiated in Taiwan, and this covers 99% of the population of 23 million people because of mandatory, universal enrollment. In 1999, the Bureau of National Health Insurance began to release all decoded patient data in electronic form under the NHIRD project. A
systematic, random sampling method was used to build this representative database of 1,000,000 NHI enrollees by biostatistic experts. There were no statistically significant differences in age, sex, or healthcare costs between the sample group and all the enrollees. This data set spans from January 2000 through December 2009 and includes all claims data for these 1,000,000 individuals.[1] Various extracted datasets are available to researchers, and hundreds of published papers have used the NHIRD as the basis for their studies. [2]

1. National Health Insurance Research Database, Taiwan. Available at http://nhird.nhri.org.tw/date_cohort.htm#2

We modified the sentence as “This 9-year cohort study was based on the Longitudinal Health Insurance Database 2000 (LHID 2000) consisting of 1,000,000 beneficiaries randomly selected from all Taiwan National Health Insurance enrollees in 2000 (>23.7 million).”
P 3, line 46-48

4. Line 49: Can you clarify whether it is the case that the 621 subjects received IBT at time of HCV diagnosis, and whether active TB was excluded in these 621? Also, how did you get 2460? Is it four matched subjects for each IBT subject? If so, why did you not match 2,484?

***
a. Yes, active TB cases were excluded in the IBT treated cohort (n=621). Figure 1 is the study flow chart for the enrollment of participants.
b. The control subjects were matched at a ratio of 1:4 with the non-treated cohort in terms of age, sex and the year and month of the index visit. Although we could not find 100% matched subjects within the non-treated cohort (n=7587), a total of 95.98% of the control subjects (n=2460) were matched.

P10, line 184-190

5. Line 57: Suggest using incidence rate ratios and 95% confidence intervals rather than just p-values. This would need to be done in the main paper as well, obviously.

***

We added the results of Adjusted Hazard Ratio and 95% Confidence Interval in 1-year follow-up and long-term follow-up in the abstract and the Results section. And we added Table 2 for clear presentation as well. P4 line 62-64
P14, line 262-267
Table 2. Results of hazard ratios in IBT-treated cohorts in 1-year and long-term follow-up

<table>
<thead>
<tr>
<th>Cohort</th>
<th>TB case n. (%)</th>
<th>Crude HR (95% CI)</th>
<th>Adjusted HR (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1-year follow-up</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Control (n=621)</td>
<td>3 (0.5)</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>IBT-treated (n=2640)</td>
<td>5 (0.2)</td>
<td>2.36(0.57-9.89)</td>
<td>2.81(0.61-12.98)</td>
</tr>
<tr>
<td>Long-term follow-up</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Control (n=621)</td>
<td>3 (0.5)</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>IBT-treated (n=2640)</td>
<td>12 (0.5)</td>
<td>0.99(0.28-3.50)</td>
<td>1.02(0.28-3.78)</td>
</tr>
</tbody>
</table>


***
We deleted the term “modified”, thanks for reviewer’s kind comments!

P4, line 64
P15, line 272

7. Line 61: Don’t you mean “rate” rather than “risk”?

***
We revised “risk” to “rate”.

P4, line 67

8. Line 62: You’re not taking about TB infection… you’re talking about TB disease, right? I think you don’t have to say reactivation.

***
Thanks for reviewer’s kind comments, we deleted these words “(infection or reactivation)”

P4, line 67

9. Line 64: Why active TB? If the concern is that IBT causes TB incidence, why would you be concerned if someone already has active TB? Is it because IBT would exacerbate the TB disease already ongoing? If so, I would explain this more in the main paper.
***

To avoid misunderstanding, these statements were deleted according to reviewer’s comments.
P4, line 67, 68

10. Line 63-65: This conclusion seems contradictory to your finding (IBT doesn’t cause TB disease, but patients should be evaluated for latent TB before taking IBT) – why? Something else needs to be explained in the abstract if you keep this conclusion there, and if you make this recommendation in the main paper, it needs to be explained.

***

Thanks for the reviewer’s kind comments. We deleted these statements in both abstract and main text.
P4, line 67, 68
P17, line 304-309

11. Line 133: OK, now I see what you did. Your sample of one million is from a larger database of 24 million. Please clarify this in the abstract.

***

We clarified this in the abstract.

We modified the sentence as “This 9-year cohort study was based on the Longitudinal Health Insurance Database 2000 (LHID 2000) consisting of 1,000,000 beneficiaries randomly selected from all Taiwan National Health Insurance enrollees in 2000 (>23.7 million).”
P 3, line 46-48

12. Lines 133-139: It would be helpful to provide more details about how you selected the one million people for your dataset. Did each of the 24 million people who had at least one day in the system have an equal probability of being selected for the sample population? I suppose so. If this is the case, then people with a longer duration of enrollment would have a higher probability of having HCV, IBT, and TB…”

***
The data analyzed in this study were retrieved from the National Health Insurance Research Database (NHIRD), a database established and maintained by the National Health Research Institutes (NHRI). Before releasing medical claims data to the NHRI, the Bureau of National Health Insurance (BNHI) scrambles the identification codes of each patient. The National Health Insurance (NHI) program in Taiwan has been operating since 1995. The program covers approximately 99% of Taiwan’s 23.74 million people and was contracted by 97% of hospitals and clinics by the end of 2009.[1] The NHI medical claims database, including an outpatient service, inpatient care, Chinese medicine, dental care, childbirth, physical therapy, preventive health care, home care, and rehabilitation for chronic mental illness. Therefore, the NHIRD is one of the largest and most complete nationwide population-based datasets in Taiwan and there are no statistically significant differences in age, sex, and average insured payroll-related amount between the sample group and all enrollees.

We used a systemic sampling of the patient data of 1 million people from all insured beneficiaries that was released by the NHRI as the Longitudinal Health Insurance Database (LHID 2000). The NHRI reported no significant variations in age and sex between the LHID and all insurants.[2,3] The high accuracy and validity of diagnoses in the NHIRD have been described in previous studies. [4-9] International Classification of Diseases Ninth Revision Clinical Modification (ICD-9-CM) codes were used for diagnoses.

This data set spans from January 2000 through December 2009 and includes all claims data for these 1,000,000 individuals and it offers a good opportunity to design this population-based-cohort study.


10. Line 151: How can this be a case-control design when you selected “controls” not based on their disease status, but rather based on their exposure status?

***

Thanks for reviewer’s kind comments. This database dose not record the exposure status, this study was based on the disease status. We revised this sentence “…population-based and record-based case-control study design…” to “…retrospective cohort study design…”.

P8, line 152

11. Line 153: So, it wasn’t a random sample of the 24 million. It was a random sample of the subset of the 24 that was over 20? (21 and older?)

***

This database (LHID 2000) contains one million randomly selected subjects from the Taiwan National Health Insurance Research Database (NHIRD), which was developed for research purposes. A systematic, random sampling method was used to build this representative database of 1,000,000 NHI enrollees.[1]

Current recommendations for treatment of persons with chronic HCV infection derive from data collected in the randomized registration trials. These trials have usually been restrictive in their exclusion criteria, such as teenage or children and thus have not reflected the general population who require therapy.[2]

Therefore, we excluded patients < 20 y/o in this study.

1. National Health Insurance Research Database, Taiwan. Available at http://nhird.nhri.org.tw/date_01.html

2. Ghany MG, Strader DB, Thomas DL, Seeff LB: Diagnosis, management, and treatment of

15. Lines 154-155: Why at least two service claims? And why for treatment only? Why did you not call people HCV-infected based on laboratory diagnostic criteria?

***

The Taiwan NHIRD did not contain direct laboratory results (such as biochemical data, viral genotype, viral load, histological characteristics). Therefore, we were unable to identify the HCV infected patients based on laboratory diagnostic criteria.

With approval from the NHRI, we were able to use the scrambled patient identification numbers to interlink files, including registry of medical facilities, details of inpatient orders, ambulatory care, and prescriptions. For only once service claim (Diagnoses are coded according to the International Classification of Diseases, Ninth Revision, Clinical Modification (ICD-9-CM).) would overestimate the diagnosis of HCV. Therefore, we used at least two service claims of ambulatory or inpatient treatment care to identify this group. (HCV infected patients)[1]


16. Line 159: I don’t think you can say that 12,547 subjects with HCV infection were identified if your definition is treated at least two times for HCV.

***

Initially, we screened the diagnosis of HCV from this database (LHID 2000) by ICD-9-CM codes of HCV in 070.41, 070.44, 070.51, 070.54, and V02.62. A total of 12,547 subjects with an HCV infection were identified. To avoid 1, the possibility of overestimation, 2, exclusion of missing data (loss of follow-up) in this database, we used at least two service claims of ambulatory or inpatient treatment care to identify this group. (HCV infected patients)[1]

17. Lines 160-164: So it was a random sample. But then you threw out some cases based on certain criteria? Why 8,312 when the abstract says 8,321? Why did the history of TB diagnosis have to come before the first HCV diagnosis; what if it was at the same time? Also, you’re not talking about diagnosis; you’re talking about treatment for HCV, right?

***

1. Yes, LIHD 2000 is a database by a systematic, random sampling method to build this representative database of 1,000,000 NHI enrollees from the population of 23 million people in Taiwan. Thereafter, we defined some diseases by certain criteria.
2. It is our typographical error. The corrected number is “8,312” not “8,321”. Many thanks for reviewer’s kind comment!
3. The primary purpose of this study was to observe the association between IBT and the rate of TB in HCV infected patients. The index date for patients with HCV infection was the date of their first medical visit. After excluding a diagnosis of TB code (ICD-9 code:010-018) before the HCV index date, so that we could analyze the association between TB and the presence of IBT in this HCV infected population.[1,2]


18. Lines 178-180: You’re talking about HCV treatment, not HCV diagnosis. Please clarify. Also, why exclude these 77? Give a justification. Need to add “for” after “those patients who had received IBT,” and add a space before “(n=621).”

***

a. This study was to observe the rate of TB in HCV infected patients (with or without IBT). In order to identify both cohorts treated and not treated with IBT, we added another figure for better understanding this flow chart (figure 1)
b. Because NHI in Taiwan reimbursed CHC patients 4–6 months of interferon or 6 months of pegylated interferon-based treatment during 2000–2008, and most patients who did not achieve early virologic response after 3 months of treatment discontinued IBT [1-3], we selected patients receiving >= 2 months of IBT into our analyses(IBT treated cohort) and patients who had received
IBT for a period of <2 months were excluded. (n=77)

c. We revised the sentence as comments, P10, line 179,180


19. Line 182: I thought it was ambulatory or inpatient?

***
We revised “…the index ambulatory care visit” to “…the index date.”
P10, line 183

20. Lines 186-187: Did they have to match exactly by age? How close did they have to match? Also, what was the index ambulatory (inpatient?) care visit for a “control”? Same as for the HCV treatment group?

***

a. There were significant difference in ago and gender (P=0.0025 and P<0.0001, respectively) before comparing both cohorts. Age, sex are risk factors for TB.[1]
For the control group, we used a simple random sampling method and selected 4 insured HCV infected patients without treatment for every HCV infected patients with IBT during the same period. Patients and controls were matched for age, sex and the year and month of the index visit.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Interferon-based therapy</th>
<th></th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Yes</td>
<td>No</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>621</td>
<td>2460</td>
<td></td>
</tr>
<tr>
<td>Sex, n(%)</td>
<td>Male</td>
<td>Female</td>
<td></td>
</tr>
<tr>
<td></td>
<td>351(56.5)</td>
<td>270(43.5)</td>
<td>0.849</td>
</tr>
<tr>
<td></td>
<td>1380(56.1)</td>
<td>1080(43.9)</td>
<td></td>
</tr>
</tbody>
</table>
### Age, n(%) Table

<table>
<thead>
<tr>
<th>Age Range</th>
<th>n</th>
<th>(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>20 – 29</td>
<td>22</td>
<td>(3.5)</td>
</tr>
<tr>
<td>30 – 39</td>
<td>54</td>
<td>(8.7)</td>
</tr>
<tr>
<td>40 – 49</td>
<td>123</td>
<td>(19.8)</td>
</tr>
<tr>
<td>50 – 59</td>
<td>231</td>
<td>(37.2)</td>
</tr>
<tr>
<td>60 – 69</td>
<td>154</td>
<td>(24.8)</td>
</tr>
<tr>
<td>70</td>
<td>37</td>
<td>(6.0)</td>
</tr>
</tbody>
</table>

Adapted from table 1

---

b. We revised “...the index ambulatory care visit” to “...the index date.”

P10, line 188


21. Lines 183-189: This is a confusing. So you took a random sample of one million from 24 million, identified HCV cases from the one million (even though “HCV diagnosed” really means treated at least twice for HCV), split these cases into IBT and no IBT. There were 621 in the IBT group, and 7587 in the no IBT group. But rather than just looking at TB incidence in the no IBT group and comparing this to the TB incidence in the IBT group, you matched those in the IBT group to those from the no IBT group based on age, sex, year/month (1:4 ratio) and then looked at incidence in the two groups? Is this correct? Why not just compare incidence in the IBT and no IBT groups, particularly if the groups don’t differ by age, gender, and calendar time, and forget the matching altogether? How can you call it “randomly selected” if you’re matching by those three characteristics? If you randomly select 2,460 from 7,587, how would they match on the characteristics?

***

There were significant difference in age and gender (P=0.0025 and P<0.0001, respectively) before comparing both cohorts. Age, sex are risk factors for TB.

[1-4] We used a simple random sampling method and selected 4 insured HCV infected patients without treatment for every HCV infected patients with IBT during the same period. For each controls were randomly selected from the HCV cohort, matched on gender, age (same year of birth), and calendar time by means of incidence density sampling.[5]

1. Havlir DV, Getahun H, Sanne I, Nunn P: Opportunities and challenges for HIV care in


22. Lines 193-196: There wasn’t at least one year of follow-up if TB was diagnosed in the first year, right? What about competing risks (death, etc.)? There were more reasons for withdrawal than loss to follow-up, assuredly.

***

The interferon-therapy group would be associated with higher risk of malignancies compared with the control group. During the long-term follow period, competing mortality should be adjusted because of the high mortality for these patients with malignancies. We do competing risk analysis (including death), as following. There was no significant difference in both models.

<table>
<thead>
<tr>
<th>Model</th>
<th>1-year follow-up Adjusted HR (95% CI)</th>
<th>Long-term follow-up Adjusted HR (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cox proportional hazards model</td>
<td>2.81(0.61-12.98)</td>
<td>1.02(0.28-3.78)</td>
</tr>
<tr>
<td>Competing risk model</td>
<td>2.81(0.61-12.98)</td>
<td>1.02(0.28-3.78)</td>
</tr>
</tbody>
</table>

23. Lines 199-204: Why does treatment with TB figure into your definition of TB incidence? This is the same problem I’m having with needing to be treated for HCV in order to be called diagnosed with HCV. If you are looking at TB incidence, why does it matter if the person was treated for TB or how long the were treated for? Isn’t the most important thing whether they were diagnosed clinically or microbiologically?
The Taiwan NHIRD did not contain direct laboratory results. Because of the limitation of this database, there was no standard procedure for diagnosis of MTB infection (acid-fast stain, culturing and histopathology). Instead, we diagnosed MTB infection if a patient had taken anti-MTB medication for more than 3 months. This definition of TB is consistent with prior literature.[1-4]


24. Lines 207-220: It would be helpful to say here what you do with these factors. For example, what’s the point of clarifying the etiology and severity of liver disease? How did you decide on what factors to include?

To determine the impact of IBT on the risk of active TB, it is important to take into consideration the influences of known risk factors.[1,2] Therefore, we identified these diseases by ICD-9 codes then adjusted them. References are needed to substantiate these statements.

P.12, line 214

Because IBT is used with caution in patients with cirrhosis and is contraindicated in patients with decompensated liver disease [3], we clarified the etiology and severity of liver disease by ICD-9 codes to adjusted these situations by multivariate-adjusted Cox proportional hazards mode

1. Pai M. Diagnosis of latent tuberculosis infection (tuberculosis screening) in HIV-negative adults. In: Elinor L Baron, editor. UpToDate. Waltham, MA: UpToDate; 2014 [cited 2014 Jul 31]

25. Lines 223-237: Now there’s a time-to-event analysis? Where was this in the abstract?

***

a. Yes, the methodology of time-to-event analysis in this study is Kaplan-Meier analysis and the proportional hazards model.

b. Thanks for reviewer’s kind comments! We added the sentence “The Cox proportional hazards models were used to estimate the hazard ratio (HR) and associated confidence interval (CI) of rate of active TB.” to the abstract.

P3. line 53

26. Line 228: I thought the definition of TB was not detection (by culture/microscopy) but rather an ICD-9 code plus two anti-TB medications for more than 90 days. What does “were coded by the date of their last visit” mean? Do you just mean that those lost to follow-up were censored on the date of last visit?

***

a. The Taiwan NHIRD did not contain direct laboratory results. Because of the limitation of this database, there was no standard procedure for diagnosis of MTB infection (acid-fast stain, culturing and histopathology). Instead, we diagnosed MTB infection if a patient had taken anti-MTB medication for more than 3 months.

b. Yes, for patients lost to follow-up were coded by the date of their last visit, creating “censored” data (right-censored data)

We also do competing risk analysis, as following table.

<table>
<thead>
<tr>
<th></th>
<th>1-year follow-up</th>
<th>Long-term follow-up</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Adjusted HR (95% CI)</td>
<td>Adjusted HR (95% CI)</td>
</tr>
<tr>
<td>Cox proportional hazards model</td>
<td>2.81(0.61-12.98)</td>
<td>1.02(0.28-3.78)</td>
</tr>
<tr>
<td>Competing risk model</td>
<td>2.81(0.61-12.98)</td>
<td>1.02(0.28-3.78)</td>
</tr>
</tbody>
</table>

27. Line 234: Do you mean “in the proportional hazard model, we adjusted for…”? How were these covariates coded? This would be good to specify in the preceding paragraph. Wasn’t IBT also included in the model? Did you estimate hazard ratios
and 95% confidence intervals? If so, specify this.

***

a. These covariates were identified by ICD-9 codes (in the preceding paragraph) P11,12 line 207-223

b. Because these covariates were described in the preceding paragraph, we deleted the sentence “In this study, the proportional hazards we adjusted for included age, sex, HIV, COPD, connective tissue disease, silicosis, DM, end stage renal disease, malignancy, and the etiology and severity of liver disease.” to clarify this paragraph. P13, line 235

c. IBT was included in the model and hazard ratio, 95% confidence intervals and p value were calculated in table 2,3,4 P27,28,29

28: Lines 241-251: I would recommend just giving all of these numbers here, rather than also in the Methods section. It is hard to reconcile the numbers in so many different places… why does it matter that there were 8,312 after exclusions in the Methods section (is this the same as the 8,321 in the abstract?) when that number is not here? The reader, like I did, will try to match the number he/she read earlier, and not find it… so just give the important numbers here and in the figure, and leave the rest out. Why does it say 7,586 here but 7,587 in the Methods? Edit carefully! Also, why say that 699 were enrolled for analysis when the actual number after an additional exclusion was 621? This whole paragraph seems like a repeat of the Methods section. Why not talk a little bit about the how the two groups compared in Table 1? Also, I would not call them “matched controls” – “control” implies disease status, not exposure status.

***

a. We modified the Methods section in abstract as reviewer’s comments P3, line 46-56

b. “7,587” is correct number, not “7,586”. Thanks for reviewer’s comments!

c. 699 HCV patients who ever received IBT. Among them, 621 patients (88.84%) were treated for a minimum of 2 months, were defined “IBT treated cohort”

d. We clarified the paragraph of Baseline characteristics of the study population in the Results section according to reviewer’s comments
   - To simply this section
   - To clarify this section
   - Talk about the comparison in both cohorts
P13, line 241-250
e. We revised the term of “matched controls” to “control cohort”.
P13, line 244


***
“±” means standard deviation and we add “(± SD)” to this sentence and abstract.
P14, line 254
P4, line 58

30. Line 261: I don’t think you defined “long-term” follow-up in the Methods section, nor did you say how these p-values were calculated. What do they represent? Again, why not give effect measures and 95% confidence intervals? Now that I look at Table 1, these p-values seem just to compare proportions in the different groups who got TB? Are they indeed not a comparison of the incidence rates, even though you attribute the p-values to the incidences? This all needs to be clarified.

***
a. We added the definition of “long-term” follow-up: 9 years.
P13, line 237
b. We deleted the item of “outcome of TB” in Table 1 and added the statements about the “outcome of TB” in the results section
P14, line 257-260
P27, Table 1
c. We added Table 2. Results of hazard ratios in IBT-treated cohorts in 1-year and long-term follow-up

<table>
<thead>
<tr>
<th>Cohort</th>
<th>TB case n. (%)</th>
<th>Crude HR (95% CI)</th>
<th>Adjusted HR (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1-year follow-up</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Control (n=621)</td>
<td>3 (0.5)</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>IBT-treated (n=2640)</td>
<td>5 (0.2)</td>
<td>2.36(0.57-9.89)</td>
<td>2.81(0.61-12.98)</td>
</tr>
<tr>
<td></td>
<td>Long-term follow-up</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Control (n=621)</td>
<td>3 (0.5)</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>IBT-treated (n=2640)</td>
<td>12 (0.5)</td>
<td>0.99(0.28-3.50)</td>
<td>1.02(0.28-3.78)</td>
</tr>
</tbody>
</table>

31. Lines 262-263: The figure is labeled “cumulative incidence” but this is not what it
is… if you used 1 minus Kaplan Meier to approximate cumulative incidence, say this. Cumulative incidence implies that you took into account competing risks, which you did not.

***

We revised the figure 2 labeled as “1 minus Kaplan Meier to approximate cumulative incidence…”
P32 Figure 2.

32. Lines 264-266: Why do you keep giving the definition of the IBT treated cohort (more than two months IBT)? Once was enough. The number who developed TB should be moved up to the part with the proportions and the p-values (line 261)… as written, it seems like this is something new, whereas in actuality it is more of the same. Why talk about the incidence of TB in a group that you excluded (n=77) even if it was for an unjustified reason?

***

a. Because NHI in Taiwan reimbursed CHC patients 4–6 months of interferon or 6 months of pegylated interferon-based treatment during 2000–2008, and most patients who did not achieve early virologic response after 3 months of treatment discontinued IBT [1-3], we selected patients receiving >= 2 months of IBT into our analyses (IBT treated cohort) and patients who had received IBT for a period of <2 months were excluded. (n=77) Because we presumed that some patients in this group (IBT <2 months) may withdrawal IBT due to the development of active TB, this group were also included for evaluation the association with active TB. However, among the HCV patients who had ever received IBT < 2 months (n=77), no TB developed during the study period.

b. This sentence of “Among the IBT treated cohort (received IBT ≥ 2 months), 3 patients developed TB during the study period.” was deleted. P14,line265

c. We presumed that some patients in this group (IBT <2 months) may withdrawal IBT due to the development of active TB. However, no active TB developed during the therapy period.

d. This sentence of “Among the HCV patients who had ever received IBT < 2 months (n= 77), no TB developed during the therapy period.” was deleted to clarify the statements in the Results section.

Thanks for reviewer’s comment


33. Lines 269-271: Why not report the adjusted hazard ratios and 95% confidence intervals for your primary exposure (IBT)? Again, what is “modified”?

***

a. We reported the adjusted hazard ratios and 95% CI in this paragraph P15, line 274-276

b. We deleted the term “modified”, thanks for reviewer’s kind comments! P15,line 272

34. Line 274: Now you’re not calling it a case-control study. I would make the necessary correction.

***

Thanks for reviewer’s kind comments! We revised all terms of “case-control study” to “population-based cohort study”, which is compatible with this design of study. P15, line 279

35. Line 274: Hard to say it’s not associated when you don’t state the hazard ratio here or in the Results. Looking at the table, there is an indication that IBT is associated with TB in the short-term model, although the measure is imprecise. One wonders what would happen if you didn’t use such a stringent definition for the outcome (TB)?

***

a. We added the hazard ratio in the Results section P14 ,line 264-269 and Table 2

b. Although the AHR in IBT cohort is 2.81 in 1-year follow-up, it is not statistically significant.

c. We defined active TB by: at least one outpatient visit or one hospital
admission during the follow-up period with ICD-9-CM codes of TB (010-018) plus the prescription of more than two anti-TB medications for more than 90 days during the study period. The TB guidelines in Taiwan recommend the use of chest X-ray followed by sputum smear and culture examination for the diagnosis of TB.[1] Laboratory facilities able to perform sputum smear or culture examination of Mycobacterium tuberculosis are widely accessible throughout Taiwan.[2]

It is possible that patients with other diseases (e.g., lung cancer, nontuberculous mycobacterial infection, or latent TB infection) were misdiagnosed with active TB and put on anti-tuberculosis medications initially. To avoid this misclassification of outcome, we screened the NHI records of patients who were classified as active TB cases by our study definition. If the ICD-9 codes of TB (010-018) in these patients were replaced by those of nontuberculous mycobacterial infection (031), lung cancer (162), or positive tuberculin skin test (795.5) during subsequent follow-up with discontinuation of anti-tuberculosis medications, the patients would be reclassified as non-TB.

1. Luh KT, editor. Taiwan Guidelines on TB Diagnosis & Treatment. Taipei: Centers for Disease Control, Department of Health, Executive Yuan, Taiwan; 2008.

36. Line 278: Again, you’re not comparing incidence rates, you’re comparing the proportions of the different groups that developed TB.

***

We added Table 2. Results of hazard ratios in IBT-treated cohorts in 1-year and long-term follow-up

<table>
<thead>
<tr>
<th>Cohort</th>
<th>TB case n. (%)</th>
<th>Crude HR (95% CI)</th>
<th>Adjusted HR (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>1-year follow-up</td>
<td></td>
</tr>
<tr>
<td>Control (n=621)</td>
<td>3 (0.5)</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>IBT-treated (n=2640)</td>
<td>5 (0.2)</td>
<td>2.36 (0.57-9.89)</td>
<td>2.81 (0.61-12.98)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Long-term follow-up</td>
<td></td>
</tr>
<tr>
<td>Control (n=621)</td>
<td>3 (0.5)</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>
37. Line 287: It’s not just reactivation; in addition to latent TB developing into active TB, it could be primary infection leading to active TB.

***
We revised “TB reactivation” to “active TB”
P15, line 289

38: Lines 288-295: What does HDV have to do with your study? Consider not including this report, and if you do, define before using abbreviation. Do you mean one case report, or one case? Or, do you mean that there is one case per case report? This paragraph describing isolated case reports does not seem to be placed in the context of your study… we don’t even know the details of the TB cases in your study population. Are there no population-level data whatsoever on TB incidence post-IBT? I don’t find these case reports to be informative, and would their removal from the text (and also remove Table 4).

***
a. We deleted the paragraph describing isolated case reports, including Table 4.
b. We cited one HCV clinical trial in Taiwan (multicenter, randomized trial) reported that 308 treatment-naive HCV-1–infected patients receiving IBT resulted in one case with TB reactivation at week 32 of IBT. [1]
P15,16 line 289-292


39. Line 303: If you’re going to talk about a hepatic encephalopathy finding, at least give the hazard ratio in the Results section (or here).

***
We added the results of hazard ration and 95% CI in the Results section.
P15, line 274-276

40. Line 305: Did you look at dose-response? I don’t see anything about this.
We deleted this sentence of “Among cases with active TB, there was no obvious cumulative dose-response in the IBT and active TB cases during hepatitis treatment in our study.” Thanks for reviewer’s kind comments!
P16 line 301

41. Lines 315-329: How does your study support any of these recommendations? For example… you looked in a HCV/TB endemic area and appeared to find nothing of consequence, so why would you say that close monitoring is required in Turkey, Eastern Europe, etc.?

***
Thanks for reviewer’s kind comments. We deleted these statements in abstract and main paper.
P4, line 67,68
P17, line 307-309

42. Lines 330-334: So you didn’t have any HCV or TB lab results? Or did you just not have viral genotype, viral load, and CD4. No wonder you had to use those exposure/outcome definitions. Even so, they were strict – can you discuss your choice of definitions? I don’t understand “a substantial proportion of patients exposed to corticosteroids would be underestimated.”

a. Yes, this is the limitation of this database. Taiwan NHIRD did not contain direct laboratory results
b. Because of the limitation of this database, we defined the diagnosis of HCV and TB according to previous studies [1-5]
c. We revised “substantial proportion…” to “a proportion…”
P18, line 321


43. Line 351: What informs this recommendation?

***

Thanks for reviewer’s kind comments. We deleted these statements in abstract and main paper.
P4, line 67,68
P17, line 307-309

44. In general, I am concerned about the selection of the “control” population (why?), the way incidences were compared (and the analysis overall), the exposure and outcome definitions…

***

Thanks for reviewer’s very kind comments.

a. This database does not record the exposure status, this study was based on the disease status. We revised all terms of “case-control study” to “population-based cohort study”, which is compatible with this design of study. These eligible controls were selected randomly and matched at a ratio of 1:4 with the non-treated cohort in terms of age, sex and the year and month of the index visit. Although we could not find 100% matched subjects within the non-treated cohort (n=7587), a total of 95.98% of the control subjects (n=2460) were matched. Thereafter, we compared the incidence of active TB in both cohorts (IBT-treated and non-treated cohorts) in this database.

b. We added the results of Adjusted Hazard Ratio and 95% Confidence Interval in 1-year follow-up and long-term follow-up in the abstract and the Results section.(Add a Table 2)

45. Figure 2: How can this be both for one-year and long-term follow-up (as the legend says)… there are only two curves (treated and untreated) in the figure; is this
figure supposed to represent either one-year or long-term (rather than both)?

***
This figure presented the cumulative incidences in both cohorts during “1-year” follow-up. We revised the figure legends of figure 2.
P. 31, figure 2

Minor Essential Revisions
(such as missing labels on figures or the wrong use of a term which the author can be trusted to correct)

1. Line 46: I recommend writing out the number instead of depending on the reader to make the calculation of 106.

***
We revised “10^6” to “1,000,000”.
P3, line 47

2. Line 46: Say “of” not “from.” And, say approximately 1,000,000 (as there were not exactly 1,000,000).

***
We used this database (LHID2000) containing one million randomly selected subjects from the Taiwan National Health Insurance Research Database (NHIRD), which was developed for research purposes. A systematic, random sampling method was used to build this representative database of 1,000,000 NHI enrollees. Therefore, we did not use the term “approximately 1,000,000”. We revised the sentence as “This 9-year cohort study was based on the Longitudinal Health Insurance Database 2000 (LHID 2000) consisting of 1,000,000 beneficiaries randomly selected from all Taiwan National Health Insurance enrollees in 2000 ( >23.7 million).”
P3, line 46-48

3. Line 55: Need space before “per” (typo).
***
We add a space before “per”
P4, line 60

4. Line 59: Say “a risk factor” not “the risk factor.”

***
We revised “the risk factor” as “a risk factor”
P3, line 65

5. Line 61: Why not use the abbreviations IBT and TB?

***
We used abbreviations of IBT and TB instead of “Interferon-based therapy” and “tuberculosis”.
P3, line 67

6. Line 63: Say “is” not “was.”

***
This sentence has been deleted as recommended.

7. Line 64 and 65: Why not use the abbreviation TB? Don’t give the abbreviation if you’re only going to use it sometimes.

***
We used abbreviations of TB instead of “tuberculosis” in this main article.

7. Lines 78-79: References are needed to substantiate these statements.

***
We added the reference in this sentence.
P5, line 81

9. Line 81: Why “therefore”? It would make sense that effective anti-HCV therapies are needed because HCV can lead to cancer and cirrhosis, not because many people are chronically and newly infected. Change the order of your sentences.
We changed the order of first and second sentences according to reviewer’s recommendation.
P5, line 78-81

10. Line 87: In general, I would recommend including references when you make biological / “factual” statements like these.

We cited the reference in this statement.
P5, line 87

11. Lines 90-91: In general, always spell out words before using abbreviations (e.g., TB, BCG).

Thanks for reviewer’s kind comments!
We spell out these words in P 5, line 90,91

12. Line 94: And then, use the abbreviation (e.g. TB) rather than spelling out the word in full.

We use the abbreviation “TB” instead of “tuberculosis” later.

13. Line 98: Now you’re defining BCG? Why not do this the first time you mention it? Also, what is complex? Or is this a typo?

a. We revised “bacillus Calmette-Guerin” to “BCG”
b. It is a typo and we revised “M. avium intracellular ecomplex” to “M. avium intracellulare complex”
P6, line 98,99

14. Line 100: Is “Th1type” a word?

We revised “Th1type” to ”Th1 type””

***
We revised “infected mice” to “TB infected mice”

16. Line 107: Say “are considered.”

***
We revised “were considered…” to “are considered…”

17. Line 109: I think you can give the TB abbreviation sooner and also use the HCV abbreviation here.

***
We use TB and HCV abbreviations instead of the following tuberculosis and hepatitis C virus words.

18. Line 111: Do you mean “a difficulty” rather than “in difficulty”?

***
We revised “in difficulty” to ”a difficulty”.

19. Line 114: Do you mean “in different geographic areas”?

***
Yes, we revised “…geographic variations” to “…different geographic areas”

20. Lines 119-122: I think you mean to say that this study used the Longitudinal Health Insurance Database 2000, rather than assembling the database for the purpose
of this study. Also, say “a risk factor” rather than “the risk factor,” insert a comma after “population,” and make sure to put a period at the end of the sentence.

***

a. We revised “assembled” to “used” in P7, line 120
b. We revised “the risk factor” to “a risk factor” in P7, line 122
c. We inserted a comma after “…population” in P7, line 121
d. We put a period “during January 2000 to December 2009” at the end of the sentence in P7, line 123

21. Line 130: NHI isn’t just a registry, correct? Isn’t it a program or system?

***

We revised “registry” to “program”
P7, line 131

22. Line 178: Need space after “excluded.”

***

We inserted a space after “excluded”.
P10, line 179

23. Line 184: Need space after “cohort.”

***

We inserted a space after “cohort”.
P10, line 185

24. Line 189: Need space after “excluded.”

***

We inserted a space after “excluded”.
P10, line 190

25. Line 193, 196, 204, 275, 293, 294, 311, 325, etc.: Why not use “TB” abbreviation? Edit carefully!

***
Thanks for reviewer’s kind comments
We use TB abbreviation instead of the following “tuberculosis”

26. Line 204: Need space after “90.”

***
We inserted a space after “90”.
P11, line 205

27. Line 209: Need space after “silicosis.”

***
We inserted a space after “silicosis”.
P11, line 210

28. Line 211: Need space after “ICD-9.”

***
We inserted a space after “ICD-9”.
P12, line 212

29. Line 212: Need space after “malignancy” and before “because.”

***
We inserted a space after “malignancy” and before “because.”
P12, line 214

30. Line 217: Need space after “cirrhosis.”

***
We inserted a space after “cirrhosis”.
P12, line 220

31. Line 224. Need space after “(USA).” Distribution should be plural./

***
We inserted a space after “cirrhosis”.
We revised “distribution” to “distributions”
P12, line 227

32. Line 227: Now endpoint is two words? Be consistent.

***
We revised “end point” to “endpoint”

P13, line 230

33. Lines 258-259: Say “the cumulative incidences were” and add a space after 0.151.

***
We revised “the cumulative incidence was…” to “the cumulative incidences were…”
We inserted a space after “0.151”.

P14, line 260,261

34. Line 283: What does anti-TNF therapy have to do with this study?

***
We deleted this statement about anti-TNF.

P15, line 288

35: Line 291: Why give the abbreviation for HAART?

***
This paragraph had been revised and these statements about case reports were deleted as reviewer’s comments.

P16, line 289-292

36. Line 309: You already gave the abbreviation for HCV.

***
We revised “Hepatitis C virus” to “HCV”

P16, line 304

37. Lines 315-322: You can’t start a complete sentence with 1) or 2) or 3) or 4).
38. Line 317: “May be” should be two words, not one.
39: Line 350: Use IBT abbreviation and insert a space after “active.”

***

We rewrote the sentences according to the reviewer’s comments.

40. Table 1: Make sure spaces are present where they should be. Why is “n(%)” not next to the “Outcome” title, like is for the other variables?

***

We modified the table 1 according to both reviewers’ comments.
We deleted the item of ”outcome”
P27, table 1