PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Factors associated with long-term prescription of benzodiazepine: a retrospective cohort study using a health insurance database in Japan
AUTHORS	Takano, Ayumi; Ono, Sachiko; Yamana, Hayato; Matsui, Hiroki; Matsumoto, Toshihiko; Yasunaga, Hideo; Kawakami, Norito

VERSION 1 – REVIEW

REVIEWER	Donovan Maust Assistant Professor University of Michigan USA
REVIEW RETURNED	01-Mar-2019

GENERAL COMMENTS	This is a retrospective cohort study of n=88,001 patients who were started on a bzd and went on to use for ≥8 months. The authors present relatively straightforward analysis of a common problem that isn't well understand relative to how common bzd prescribing is. I have several overall suggestions for their analysis: 1. They note in the introduction how the definition of "long-term" is all over the place in the literature. It would significantly strengthen their analysis if they would vary their definition of "long-term" and demonstrate whether their findings are stable (i.e., what if long-term is 120 days or ≥6 months or ≥9 months, to use some other common definitions). Not that they have to do the whole range, but at least some alternatives. 2. Why not include size of the initial prescription? Recent publication by Gerlach et al. (JAMA IM 2018) as well as Simon et al. that is already cited both show this as a predictor of chronic use. A few more specific points: Introduction: - the evidence linking benzos with dementia is a bit murkier than suggested here Methods: - you say "BZD-related" but is that anything other than the "Z-drugs"—might just be clear here that it is "BZD and the related Z-drugs" or something - why do you exclude clobazam but include clonazepam? Is it because in Japan clonazepam is not used as an anti epileptic (even though classified by WHO that way per your table)? - are there people who could have received a 90d supply and *not* be counted as long-term b/c they didn't have a prescription every month? - What if someone received a prescription on Jan 1, then Jan 29,

	none in February, then March 1, etc. Is this person considered long-term? In other words, does your definition (one claim per month) account for how many days supplied and the period of coverage? - "worker or dependent": so would a dependent include both spouse and children of the person through whom coverage is obtained? Does this really add to the model? I would think the combination of age and gender would be more important. What theoretical reason would you have to think age v. Dependent—apart from age and gender—matters? - why not include insomnia as a diagnosis? - why not include any measure of medical comorbidity? At least some sort of overall measure. - is physician specialty derived from the first bzd prescription? - "if a patient was prescribed multiple BZDs . . . we selected the BZD with the longest half-life." For all of these characteristics of the prescribed medication, would it make more sense to use the medication with the most days prescribed? Selected the first one seems a bit arbitrary—or convince the reader why that makes the most sense. - given the large sample size, the authors should consider lowering the threshold for significance. - also, Table 1 (comparing differences b/t groups) would potentially be more meaningful if you used a measure of standardized differences between groups. For example, the age differs by less than 6 mos but is statistically significant simply by virtue of the sample size. Results: - in Table 1, it appears that the psychiatric diagnoses column add up to 100%—but this doesn't make sense, as patients can clearly have more than one diagnosis. Or is this somehow the single specific diagnosis linked to the prescription? Would be better for the statistical test to be for each row.

REVIEWER	Jacob Simmering Department of Internal Medicine Carver College of Medicine University of Iowa
REVIEW RETURNED	20-Mar-2019

GENERAL COMMENTS	Takano et al present an interesting analysis quantifying the rates of long-term benzodiazepines (BZD) in Japan among people newly started on BZD. Additionally, the analysis describes factors associated with long-term BZD use. Given the risk of dependence, withdrawal and lack of evidence for long-term efficacy of BZD in anxiety and other mood disorders, this is an important analysis. They found that despite guidelines and other systems to discourage long term use of BZD that approximately 9% of new users were still taking BZD after 8 months and increased age, cancer, mood disorder, prescription by a psychiatrist, and high initial dose were associated with increased odds of still taking BZD at 8 months. The paper has many strengths. Restricting the analysis to new users (at least 6 months BZD-free enrollment) and non-elderly adults avoids the limitations faced by other studies. Use of the large claims database provides a very large sample with generally sufficiently long follow-up (only 14% were lost due to exiting the data before 9 months). Limitations chiefly are those faced by all studies using claims data. They are not able to measure the severity of the mood disorder or construct a history of mental health for the enrollees.

	However, on balance, I find this to be a compelling analysis. Major Comments: Based on the results in Table 2, it seems that people with mood/psych disorders or treatment directed by a psychiatrist are generally the people who take BZD after 8 months. A common use of BZD in the US is in people with specific transient fears (e.g., fear of flying) to manage the associated anxiety. These people would take the BZD on the day of the flight but would not have any indication for continued treatment after the flight is over. This is in contrast to someone with an anxiety disorder that would have a more persistent need for treatment. How frequent is this type of prescribing in Japan? The results in Table 2 (increased risk of remaining on BZD at 8 months with seeing a psychiatrist, having a psych diagnosis and having regular dosing) seems to suggest that transient versus chronic anxiety/disorder may drive the majority of the findings. Given that the half-life was associated with long-term use in the unadjusted analysis and the distinction between chronic versus transient anxiety, I would be interested in seeing a sensitivity analysis focusing on the effect of half-life on long-term use among those patients who have an RX from a psychiatrist or a specific psych diagnosis versus and, separately, those who do not. If the underlying disease type is correlated with medication half-life (which seems like a reasonable assumption), the relationship between medication half-life and long-term use would be confounded unless the analysis was stratified. I suspect those with transient anxiety (as in the fear of flying) will be prescribed shorter half-life BZDs while those with more chronic anxiety will tend to have longer half-life BZDs. It is possible that those taking longer-acting BZDs without a psych diagnosis have greater odds of remaining on BZDs for longer than 8 months than those on shorter half-life BZDs for the same treatment indication. Additionally, I would like to see expanded discussion on the possibility of different durations of need resulting from transient or situational anxiety and longer term/chronic anxiety in the discussion. Minor Comments: Line 75 says the data runs to November 2014 while lines 82-83 suggest the study period was April 1st 2012 to December 31st 2015. Why is the November 2014 date cited in Line 75? Line 218 mentions a limitation of loss-to-follow-up of about 14% of the cohort and suggests there may be a bias caused by this due to the healthy-worker-bias. A possible sensitivity analysis could consider the baseline characteristics of those lost to those remaining in data at 9 months (where they more likely to have a psych diagnosis?) or to consider a survival analysis (outcome = days until off BZD) where those lost prior to 8 months are censored at the time they exited the data and those who remain on BZD during the full data span are censored at the end of the study period. Line 219 mentions not including the elderly as a limitation of this study. I am not sure that is true given that the interest of this study is in describing long-term BZD use in the non-elderly adults. Typographical: Line 39: “any generation” is a weird phrasing in English. “any age

	group" would be more standard. Line 56/Line 57: "female was a risk factor" and "male was a risk factor" – add the word sex or gender to read "female gender was a risk factor" and "male gender was a risk factor." Line 59: "we subjected new BZDs users" I believe you mean "we followed new BZDs users" Line 60: Provide a reference for the two previous studies Line 88: "at least one of available oral BZD…" should probably read "at least one of the available BZD…" Line 171: "Also, previous studies using prevalent user design…" I am not sure how "also" transitions this sentence from the prior sentence. I believe I understand the intention (the prior sentence states the 9% figure is lower than that seen in prior publications; however, those publications suffer from prevalent user bias and may over-estimate long-term BZD user) but the sentence needs slight revision. Line 174: "Alternatively…" does not make sense here as a transition. Line 213: "patient' severity" should be "the patient's severity" Line 215: "claimed data" should be "claim data" Line 217: "quitted their job" should be "quit their job"

VERSION 1 – AUTHOR RESPONSE

Reviewer: 1
Reviewer Name: Donovan Maust
Institution and Country: Assistant Professor, University of Michigan, USA

This is a retrospective cohort study of n=88,001 patients who were started on a bzd and went on to use for ≥8 months. The authors present relatively straightforward analysis of a common problem that isn't well understand relative to how common bzd prescribing is. I have several overall suggestions for their analysis:

Thank you for your comments. We added sensitivity analyses and modified some variables based on your suggestions. The new methods, different from previous results, demonstrated a risk factor (half-life of BZD). We carefully revised our manuscript and highlighted the revised text in the manuscript. Point-by -point responses are provided below.

	Comment	Response
1	They note in the introduction how the definition of “long-term” is all over the place in the literature. It would significantly strengthen their analysis if they would vary their definition of “long-term” and demonstrate whether their findings are stable (i.e., what if long-term is 120 days or ≥6 months or ≥9 months, to use some other common definitions). Not that they have to do the whole range, but at least some alternatives.	We added analysis that considered differences in the definition of “long-term” to consecutive BZD prescription for 6 months and 12 months. The findings were almost the same as the results when using a definition of 8 months. We revised the methods (P9, L141-144), results (P14, L199-203), and discussion (P17, L263-265) and added supplementary tables (Supplementary Table 3 and 4).
2	Why not include size of the initial prescription? Recent publication by Gerlach et al. (JAMA IM 2018) as well as Simon et al. that is already cited both show this as a predictor of chronic use.	We modified the variables of “Administration instructions” to describe the size of the initial prescription based on the studies by Gerlach and Simon. This variable was based on BZDs with the longest prescription days in the first prescription. We divided the variable into four categories: “as needed”, “regular prescription: 1 week (1-7 days)”, “regular prescription: 2 weeks (8-14 days)”, and “regular prescription: more than 2 weeks (≥ 15 days)”. The results indicated that an initial prescription of a larger size was associated with long-term prescription compared to "as needed". We revised the methods (P8, L124-126), results (P13, L189-191), and Table 1 and 2.
A f ew more specific points:		
Intr oduction:		
3	- the evidence linking benzos with dementia is a bit murkier than suggested here	BZD is considered as a risk factor of dementia in previous studies, but the studies were mostly
		conducted among elderly people. We deleted this text because this study focused on the risks of BZD in non-elderly people, and as such, the information was not particularly relevant here. (P4, L39)
Met hods:		

4	- you say “BZD-related” but is that anything other than the “Z-drugs”—might just be clear here that it is “BZD and the related Z-drugs” or something	We revised “BZD-related” to “Z-drug” throughout the manuscript.
5	- why do you exclude clobazam but include clonazepam? Is it because in Japan clonazepam is not used as an anti-epileptic (even though classified by WHO that way per your table)?	We reconsidered whether anti-epileptic drugs (not only clobazam, but also clonazepam) should be excluded from this study because these drugs are likely to be prescribed for a long-term. Accordingly, the number of included BDZ was changed to 31 (P7, L97-102). We also revised Supplementary Table 1.
6	- are there people who could have received a 90d supply and *not* be counted as long-term b/c they didn’t have a prescription every month?	Yes, there were people who received a 90-day supply because two kinds of BZDs can be prescribed up to 90 days (diazepam, and nitrazepam) and six kinds of BZDs have no limitation (tofisopam, eszopiclone, flutazolam, flutoprazepam, mexazolam, and rilmazafone) in Japan. These people were not counted as long-term and the percentage for long-term prescription may be underestimated. This is one of the limitations and we added this point to the study limitations (P18, L279-281). Additionally, we added the maximum number of prescription days of each BZD in Supplementary Table 1.
7	- What if someone received a prescription on Jan 1, then Jan 29, none in February, then March 1, etc. Is this person considered long-term? In other words, does your definition (one claim per month) account for how many days supplied and the period of coverage?	Someone who received a prescription on Jan 1, then Jan 29, none in February, then March 1 was not considered long-term in this study. We considered that most people were prescribed BZD for less than 30 days because 23 of 31 BZDs (74.2%) are limited to a 30-day prescription in Japan and are prescribed in small amounts at the first prescription to observe how the drugs work. As presented in Table 1, most participants were prescribed BZDs for 14 days or less. Also, we confirmed that 99.3% of patients were given a prescription for less than 30 days. We added a justification for the definition of long-term (P7-8, L107-110) and results of prescription days in the results (P10, L160-162).
8	- “worker or dependent”: so would a dependent	Yes, the dependent of workers (employed person)

	include both spouse and children of the person through whom coverage is obtained? Does this really add to the model? I would think the combination of age and gender would be more important. What theoretical reason would you have to think age v. Dependent—apart from age and gender—matters?	included family members who were covered by the worker's insurance, including spouse, children aged over 18 years old, and parents under 65 years depending on their situation. We included "worker (employed) or dependent" because we believe occupational status affects psychiatric conditions and has an influence on long-term use of BZD. Also, there were male dependents and female workers.
9	- why not include insomnia as a diagnosis?	It was difficult to identify insomnia using one category of ICD-10 because insomnia symptoms were included as diagnosis criteria in the F category (e.g., depression). Therefore, we did not include insomnia as a diagnosis.
10	- why not include any measure of medical comorbidity? At least some sort of overall measure.	We did not include variables of medical comorbidity because medical comorbidities were not clearly associated with long-term prescription in previous studies. We did include the diagnosis of cancer as one of the major diseases related to psychological distress.
11	- is physician specialty derived from the first bzd prescription?	Yes, the variable of medical specialty was derived from the first BZD prescription. We added an explanation (P8, L121-122).
12	- "if a patient was prescribed multiple BZDs . . . we selected the BZD with the longest half-life." For all of these characteristics of the prescribed medication, would it make more sense to use the medication with the most days prescribed? Selected the first one seems a bit arbitrary—or convince the reader why that makes the most sense.	We reconsidered our methods to assess the characteristics of BZD of the first prescription and used BZDs with the longest prescription days for "type of BZD", "administration instructions", and "half-life of BZD". The findings were almost identical to previous results, but the result for half-life of BZD was different. Patients who were prescribed a BZD with a medium half-life were significantly more likely to be prescribed for a long-term. We revised the methods (P9, L128-130), results (P13, L189-191) discussion (P16, L237-245), and Table 1 and 2.
13	- given the large sample size, the authors should consider lowering the threshold for significance.	We used a general threshold for significance ($p < 0.05$) in accordance with methods from previous studies.

	- also, Table 1 (comparing differences b/t groups) would potentially be more meaningful if you used a measure of standardized differences between groups. For example, the age differs by less than 6	It is our understanding that standardized difference is suitable for comparison of baseline characteristics between two groups (for example, difference in age between two treatment groups). However, Table 1
	mos but is statistically significant simply by virtue of the sample size.	included the results of the univariate comparison of outcomes. Although our main analysis is the multivariable logistic regression, we believe it is important to include the univariate analyses as well. Another reviewer suggested a comparison of background characteristics between those who were followed for 8 months and those who were censored. In that comparison, we followed your advice and used the standardized difference.
Re sults:		
14	- in Table 1, it appears that the psychiatric diagnoses column add up to 100%—but this doesn't make sense, as patients can clearly have more than one diagnosis. Or is this somehow the single specific diagnosis linked to the prescription? Would be better for the statistical test to be for each row.	We divided patients into four categories: only F3, only F4, both F3 and F4, and other psychiatric disorders (F0, 1, 5-9). If a patient had alcohol use disorder (F1) and depressive episode (F3), the person was categorized in other psychiatric disorders. We used this method to make mutually exclusive categories and there were only a few people with other psychiatric disorders among new BZD users. We added an explanation of this point in the methods (P8, L119-121).

Reviewer: 2
Reviewer Name: Jacob Simmering
Institution and Country: Department of Internal Medicine, Carver College of Medicine, University of Iowa

Takano et al present an interesting analysis quantifying the rates of long-term benzodiazepines (BZD) in Japan among people newly started on BZD. Additionally, the analysis describes factors associated with long-term BZD use. Given the risk of dependence, withdrawal and lack of evidence for long-term efficacy of BZD in anxiety and other mood disorders, this is an important analysis. They found that despite guidelines and other systems to discourage long term use of BZD that approximately 9% of new users were still taking BZD after 8 months and increased age, cancer, mood disorder, prescription by a psychiatrist, and high initial dose were associated with increased odds of still taking BZD at 8 months.

The paper has many strengths. Restricting the analysis to new users (at least 6 months BZD-free enrollment) and non-elderly adults avoids the limitations faced by other studies. Use of the large claims database provides a very large sample with generally sufficiently long follow-up (only 14% were lost due to exiting the data before 9 months). Limitations chiefly are those faced by all studies using claims data. They are not able to measure the severity of the mood disorder or construct a history of mental health for the enrollees. However, on balance, I find this to be a compelling analysis.

Thank you for your comments. We added a stratified analysis based on your suggestions. We carefully revised our manuscript and highlighted the revised text in the manuscript. Point-by-point responses are provided below.

	Comment	Response
Maj or Comments:		
1	Based on the results in Table 2, it seems that people with mood/psych disorders or treatment directed by a psychiatrist are generally the people who take BZD after 8 months. A common use of BZD in the US is in people with specific transient fears (e.g., fear of flying) to manage the associated anxiety. These people would take the BZD on the day of the flight but would not have any indication for continued treatment after the flight is over. This is in contrast to someone with an anxiety disorder that would have a more persistent need for treatment. How frequent is this type of prescribing in Japan? The results in Table 2 (increased risk of remaining on BZD at 8 months with seeing a psychiatrist, having a psych diagnosis and having regular dosing) seems to suggest that transient versus chronic anxiety/disorder may drive the majority of the findings.	We could not find data that described the frequency of prescriptions for people with transient anxiety. In Japan, BZDs are prescribed for both persistent anxiety (e.g., generalized anxiety disorder) and transient anxiety. It was difficult to identify whether a patients' condition was chronic or not at the time of their first prescription. However, people who had transient anxiety would likely be prescribed BZDs as needed. Therefore, we believe that the prescription status (as needed or regular prescription) reflects the condition. We revised the methods (P8, L124-126) and results (P13, L189-191).

2	Given that the half-life was associated with long-term use in the unadjusted analysis and the distinction between chronic versus transient anxiety, I would be interested in seeing a sensitivity analysis focusing on the effect of half-life on long-term use among those patients who have an RX from a psychiatrist or a specific psych diagnosis versus and, separately, those who do not. If the underlying disease type is correlated with medication half-life (which seems like a reasonable assumption), the relationship between medication half-life and long-term use would be confounded unless the analysis was stratified. I suspect those with transient anxiety (as in the fear of flying) will be prescribed shorter half-life BZDs while those with more chronic anxiety will tend to	We added an analysis stratifying participant by prescription status (as needed or regular prescription), which we believe represents the condition as stated above. As pointed out, risk factors of long-term prescription were different between groups. In patients who were prescribed as needed, sex, medical specialty, multiple BZDs prescription, and half-life were not associated with a long-term prescription. On the other hand, the results among patients with a regular prescription were the same as the results for all participants. We revised the methods (P9-10, L144-146) and results (P14, L203-207), and discussion (P17-18, L265-270) and added Supplementary Table 5.
	have longer half-life BZDs. It is possible that those taking longer-acting BZDs without a psych diagnosis have greater odds of remaining on BZDs for longer than 8 months than those on shorter half-life BZDs for the same treatment indication.	
3	Additionally, I would like to see expanded discussion on the possibility of different durations of need resulting from transient or situational anxiety and longer term/chronic anxiety in the discussion.	It was difficult to distinguish an anxiety condition using only the information from the claims data at the initial prescription. We thought further study will be needed using other data including detailed patient information because prescription status was associated with the long-term prescription. We added this point to the discussion (P17-18, L265-270).
Minor Comments:		
4	Line 75 says the data runs to November 2014 while lines 82-83 suggest the study period was April 1st 2012 to December 31st 2015. Why is the November 2014 date cited in Line 75?	This sentence is based on a reference #14 (JMDC Inc. https://www.jmdc.co.jp/pharma/database.html). In this JMDC website, the representativeness of their claim database at the time of November 2014 was explained.

5	Line 218 mentions a limitation of loss-to-follow-up of about 14% of the cohort and suggests there may be a bias caused by this due to the healthy-worker-bias. A possible sensitivity analysis could consider the baseline characteristics of those lost to those remaining in data at 9 months (where they more likely to have a psych diagnosis?) or to consider a survival analysis (outcome = days until off BZD) where those lost prior to 8 months are censored at the time they exited the data and those who remain on BZD during the full data span are censored at the end of the study period.	We added analysis to confirm characteristics of the participants between those who were censored from this study cohort and those who were followed up for 8 months. The censored patients (14.0%) were likely to be prescribed at psychiatry. We added this to the methods (P9, L139-141), results (P12, L178-181) and limitations (P18, L274-277), and Supplementary Table 2. Please note that we calculated standardized differences between groups because of the large sample size as recommended by Reviewer 1 (Supplementary Table 2).
6	Line 219 mentions not including the elderly as a limitation of this study. I am not sure that is true given that the interest of this study is in describing long-term BZD use in the non-elderly adults.	We agree with you. We deleted this limitation (P18, L277-279).
Typographical:		
7	Line 39: “any generation” is a weird phrasing in English. “any age group” would be more standard.	Thank you for your comments. We revised typos and inappropriate expressions in accordance with your suggestions. P4, L40
8	Line 56/Line 57: “female was a risk factor” and “male was a risk factor” – add the word sex or gender to read “female gender was a risk factor” and “male gender was a risk factor.”	P5, L58, L60
9	Line 59: “we subjected new BZDs users” I believe you mean “we followed new BZDs users”	P5, L61
10	Line 60: Provide a reference for the two previous studies	P5, L62
11	Line 88: “at least one of available oral BZD…” should probably read “at least one of the available BZD…”	P6-7, L90-91

12	Line 171: “Also, previous studies using prevalent user design…” I am not sure how “also” transitions this sentence from the prior sentence. I believe I understand the intention (the prior sentence states the 9% figure is lower than that seen in prior publications; however, those publications suffer from prevalent user bias and may over-estimate long-term BZD user) but the sentence needs slight revision.	P14, L219-221
13	Line 174: “Alternatively…” does not make sense here as a transition.	P15, L223
14	Line 213: “patient’ severity” should be “the patient’s severity”	P17, L261-262
15	Line 215: “claimed data” should be “claim data”	P18, L271
16	Line 217: “quitted their job” should be “quit their job”	P18, L277

VERSION 2 – REVIEW

REVIEWER	Donovan Maust University of Michigan
REVIEW RETURNED	12-Jun-2019

GENERAL COMMENTS	Thank you for responding to the suggestions.

REVIEWER	Jacob Simmering University of Iowa, USA
REVIEW RETURNED	17-Jun-2019

GENERAL COMMENTS	I find the analysis and revisions presented here to be suitable. My primary concern reading the associations measured were not so much risk factors for long-term use of BZD but rather associated with having conditions that require long-term BZD. The revised analysis and discussion adequately highlight this possibility.