Harvard Medical School

Division of Pharmacoepidemiology and Pharmacoeconomics

BRIGHAM AND WOMEN'S HOSPITAL

A Teaching Affiliate of Harvard Medical School

Elisabetta Patorno, M.D., Dr.P.H.
Assistant Professor of Medicine
epatorno@bwh.harvard.edu

Georg Röggla, M.D.
Research Editor, *British Medical Journal*

May 23, 2017

Dear Dr. Röggla,

We appreciate the opportunity to revise and resubmit our manuscript, "**Benzodiazepines and the Risk of All-cause Mortality in Adults: A Cohort Study**" for possible publication in the *British Medical Journal*.

The reviewers and the editors raised important points to be addressed. We considered each comment and made corresponding changes to the manuscript. We believe this process has resulted in a substantially improved version and are indebted to the reviewers for their helpful comments. The following pages provide our detailed response to the reviewers and the editors.

We look forward to your decision regarding this revised version.

With best regards,

Elisabetta Patorno and co-authors

12-Jan-2017

Dear Dr. Patorno

Manuscript ID BMJ.2016.036319 entitled "Benzodiazepines and the Risk of All-cause Mortality in Adults: A Cohort Study"

Thank you for sending us your paper. We sent it for external peer review and discussed it at our manuscript committee meeting. We recognise its potential importance and relevance to general medical readers, but I am afraid that we have not yet been able to reach a final decision on it because several important aspects of the work still need clarifying.

We hope very much that you will be willing and able to revise your paper as explained below in the report from the manuscript meeting, so that we will be in a better position to understand your study and decide whether the BMJ is the right journal for it. We are looking forward to reading the revised version and, we hope, reaching a decision.

Georg Roeggla
groggla@bmj.com

Report from The BMJ's manuscript committee meeting

These comments are an attempt to summarise the discussions at the manuscript meeting. They are not an exact transcript.

Manuscript meeting 12.01.2017

Elizabeth Loder (chair), Angela Wade (stats), Wim Weber, Jose Merino, Georg Roggla, Tiago Villanueva, Daoxin Yin, Amy Price.

Decision: Ask for Revision.

The committee was interested in your research. We thought this is relevant research question and the findings aren't in line with previous research. The following concerns were mentioned:

• The committee agreed with the reviewers concerns.

[Response] We have addressed each concern from the reviewers in the following pages.

• We think that this is an important topic with still lots of uncertainties. The study is very well analysed; in particular, the use of the high dimensional propensity score and careful exposure assessment. The authors further conduct several sensitivity analyses supporting their finding. The authors clearly show that the crude association between benzos and mortality is confounded and we believe that their results are as robust as it can get from observational research.

[Response] Thank you for the positive feedback.

• You may want to add a short discussion about other potential side effects of benzodiazepines for our clinical readership.

[Response] The Discussion section currently reports the following short discussion about potential side effects of benzodiazepines:

Discussion, page 12: "*BZDs confer their effects through their action on γ-amino-butyric acid (GABA) type A receptors in the central nervous system, which are molecular substrates for the regulation of vigilance, anxiety, muscle tension, epileptogenic activity and memory functions. Because of their psychotropic action, BZDs have been associated with hypnotic-related **side effects** such as daytime sleepiness, impairment of psychomotor and cognitive functioning, increased risk of motor vehicle collisions, and increased risk of falls and fractures, in particular among older patients and with possible greater risks for BZDs with longer half-life. BZDs have also been associated with increased risk for development of dependence and abuse. However, this risk is not as substantial as with older sedatives and other recognized drugs of abuse, and overdose with a BZD rarely causes severe cardiovascular or respiratory depression and death.*"

We are happy to further expand this section if more discussion is preferable.

• Please clarify that your paper refers to initiation rather than long-term use of Benzos.

[Response] We have revised the text as follows:

Abstract, Conclusions, page 3: "*This large population-based cohort study suggests either no increase or at most a minor increase in **all-cause** mortality risk associated with **BZD initiation**.*"

Conclusions, page 14: "*Results from this large cohort study **based on an intention-to-treat approach** suggest either no increase or a small increase in the risk of **all-cause** mortality associated with **BZD initiation**.*"

"What this study adds", page 15: "*In this study, which included over 1,250,000 BZD initiators and used multiple approaches in the study design and analysis to minimize the potential for residual confounding, we found either no increase or at most a minor increase in **all-cause** mortality risk associated with **BZD initiation**.*"

• You only focus on relatively young patients. What about old age patients?

[Response] Though representing a small portion of our study population, we assessed over 92,000 patients 65-year-old and older who initiated BZD treatment. This sub-cohort of elderly patients is more than 10 times larger than the largest previously published study assessing the association between BZD treatment and all-cause mortality in older adults (Jaussent I, Ancelin ML, Berr C, Pérès K, Scali J, Besset A, Ritchie K, Dauvilliers Y. Hypnotics and mortality in an elderly general population: a 12-year prospective study. BMC Med. 2013; 11:212).

Consistent with most previously published studies (Jaussent I et al. BMC Med. 2013; Vinkers DJ et al. JAMA 2003; Hogan DB et al. Can J Clin Pharmacol 2003; Rumble R et al. J Am Geriatr Soc 1992), we did not find an increased risk in all-cause mortality associated with initiation of BZD treatment among older adults (adjusted HR = 0.89 [0.85-0.94]). These findings further support the fact that the psychotropic effects of BZDs are likely not responsible for the increased risk observed in previous investigations.

• Please consider citing Increasing Benzodiazepine Prescriptions and Overdose Mortality in the United States, 1996-2013. Am J Public Health. 2016;106:686-8

[Response] Thanks for the suggestion. We now cite this relevant paper in the 1st paragraph of the Introduction section.

• Sensitivity and specificity of linkage should be given.

[Response] The sensitivity and specificity of the Social Security Administration Death Master File for detecting mortality status, have been estimated to range between 87%-98% and 97%-100%, respectively (Hermansen SW et al. Am J Epidemiol. 2009; Schisterman EF and Whitcomb BW. Popul Health Metr. 2004).

If the outcome is defined with high specificity, then relative risk estimates will be unbiased, assuming that misclassification is non-differential with respect to exposure. We have clarified this aspect in the text:

Discussion, page 13: "*The linkage with the Social Security Administration Death Master File ensured that mortality was captured with high specificity in our dataset, thus minimizing possible bias in the relative risk estimates.*"

First, please revise your paper to respond to all of the comments by the reviewers. Their reports are available at the end of this letter, below. Please also respond to the additional comments by the committee.

In your response please provide, point by point, your replies to the comments made by the reviewers and the editors, explaining how you have dealt with them in the paper.

Comments from Reviewers

Reviewer: 1

Recommendation:

Comments:
I read this paper with interest. It is a well-designed and clearly described propensity score matched cohort study that explored the risk of death with benzodiazepine use using a claims database. The topic is of great interest due to the clinical relevance of the outcome and inconsistencies of previous findings. There are however some methodological concerns which can certainly be addressed by the authors as itemized below:

Abstract

• It is not clear what the need is for taking treatment barriers into account in the analysis; so this deserves further explanations.

[Response] In this study, we used a non-active comparator group in the primary analyses. Non-users can have quite different characteristics compared with patients who initiate a drug treatment, such as a lower burden of comorbidities, and thus a lower mortality risk, or conversely higher barriers to treatment and surveillance for comorbidities, and thus a higher mortality risk. Therefore, a pharmacoepidemiologic study that chooses to use a non-active comparator group should account for the potential bias deriving from both these scenarios through study design and/or analysis.

To clarify this aspect, we have updated the Abstract and the Methods:
Abstract, Page 3, Objectives: "*To evaluate the risk of all-cause mortality associated with benzodiazepine (BZD) initiation vs. non-initiation in adults, trying to address **potential** treatment barriers and confounding **related to the use of a non-active comparator group***"

Methods, Page 6: "*Non-users can have quite different characteristics compared with patients **who initiate** BZD treatment, such as a lower burden of comorbidities, and thus a lower mortality risk, or conversely higher barriers to treatment and surveillance for comorbidities, and thus a higher mortality risk; **both scenarios will need to be accounted for through proper study design and/or analysis choices**.*"

• Authors restrict the study to those patients with a recent medical visit. However, having a recent medical visit per se is a weak proxy of the health status of the patient

as visit can be due to urgent symptoms or can be just a follow-up visit. This should be acknowledged among the limitations.

[Response] The main reason for requiring a recent medical visit is not to use it as a proxy of the health status of the patient, but to reduce the likelihood of barriers to treatment and barriers to surveillance for comorbidities, since these patients have recent evidence of contact with the healthcare system. In addition, it provides a well-defined cohort entry date, minimizing the chances of immortal time bias.

We would also like to point out that the health status of patients that initiate a BZD can itself range from lower to higher levels of severity and urgency of symptoms and underlying motivation of use, e.g. from insomnia driven by mild stress to severe insomnia/anxiety due to a recent diagnosis of malignancy or to end-of-life conditions associated with severe pain. Thus, we think the variability in the baseline health status of non-users (who could have had a medical visit for a variety of reasons) is a strength rather than a limitation, as it allowed to have a large and diverse pool of potential non-users to choose from for BZD initiators. Specifically, each BZD initiators was 1:1 matched with the most similar non-user on the basis of over 300 baseline characteristics identified via investigator-specified and high-dimensional propensity-score. This maximized the chances of finding suitable comparisons for BZD initiators and thus reduced confounding.

- When index date is mentioned for the first time, a definition for it is not provided Introduction

[Response] Thanks for the comment. We now provide the definition of index date at the first occurrence of the term:

Abstract, Page 3, Participants: "*To address treatment barriers and confounding, patients were required to have ≥1 filling for any medication in the 90 days and 91-180 days before the index date (i.e., the date of drug initiation for BZD users and the date of the selected medical visit for BZD non-users) and the high-dimensional propensity score (hdPS) was estimated on the basis of > 300 covariates.*"

- Authors claims that wide use of BZD is due to large range of indications. I would rather say that wide use of these drugs is mostly due to very frequent and increasing over time treatment of anxiety and insomnia as all the other indications account for only minor part of BZD use. Authors report also some estimates about prevalence of BZD use; however, it should be specified the setting in which the BZD use has been estimated (community? also in-hospital use? Some indications are for in-hospital use only (anaesthesia)). In addition, it would be helpful reporting (if available) some estimates about long term use of BZD in elderly in USA as this is more likely to be associated with increased risk of death. Based on the above mentioned comments I suggest rephrasing the first paragraph of the introduction.

[Response] Thanks for the comment. We have re-phrased the first paragraph accordingly.

Background, Page 4: "*Benzodiazepines (BZDs) are one of the most commonly prescribed classes of psychotropic medications in developed countries. In 2008, approximately 5.2% of U.S. adults aged 18 to 80 years used BZDs **in an outpatient setting**, **with use increasing from 4.1% in 1996 to 5.6% in 2013**. Similarly, an estimated 8.4% of the population in British Columbia, Canada used a BZD in 2006, and 5.8% up to 16.3% used BZDs across several European countries in 2008. **BZD use appears to increase with age, with a higher proportion of any and long-term use among patients over 50 years**. Because of their established efficacy, BZDs are widely used in the treatment of anxiety and sleep disorders, **which together with mood disorders have been found to be the most common indications for BZD prescription by the U.S. Medical Expenditure Panel Survey in 2013**.*"

- Authors state that in general several studies investigated the association of specific causes of death and BZD, but very limited details have been provided (how BZD were used? e.g. chronic vs occasional? What types of patients, e.g. elderly people?). I suppose that the issues of BZD-related mortality have been mostly observed in elderly people. If that's true more emphasis on the analysis restricted to over 65 years should be given.

[Response] Thanks for the comment. We have now emphasized that most studies identifying a positive association between benzodiazepines and all-cause mortality were conducted in adult or young adult populations. Conversely most studies conducted in populations of older adults (i.e., over 65 years) have not identified an increased risk in mortality associated with the use of benzodiazepines, beyond an increase in falls and fractures related mortality.

Background, Page 4: "*Despite earlier mixed results regarding a possible association between BZDs and all-cause mortality, **and mostly no indication of an increased risk among patients aged 65 years or older beyond falls and fractures related mortality**, more recent evidence has reported a three-fold or higher risk increase among **adult populations** exposed to BZDs, **even for durations of use shorter than one month**. Moreover, several studies **conducted in adult or young adult populations**, have **not provided support for a specific effect of BZDs that could explain the increased risk in all-cause mortality, but have rather suggested associations with a wide range of causes of death**, including cardiovascular disease, cancer, respiratory disease, and suicide.*"

Methods
- Study population, exposure and outcome
- Was BZD treatment starting a censoring factor in non-initiators of BZD? In addition to intention to treat approach was also considered somehow per protocol approach?

[Response] BZD treatment starting was not a censoring factor among non-initiators of BZDs, in line with the intention-to-treat approach that we used in this investigation. In the context of a non-user comparator, a per-protocol approach would have led to a differential opportunity for BZD users and non-users to be censored (i.e., the likelihood of discontinuing BZD treatment among BZD initiators is much higher than the likelihood of starting BZD treatment among non-users) and, thus, to differential follow-up and potential informative censoring. In order to limit this possibility, we opted for an intention-to-treat approach. This approach or variations of this approach have also been used by the recent literature reporting an over three-fold increased risk of all-cause mortality among adult populations exposed to BZDs.

-        How can the authors be sure that persons who are BZD users are current users shortly before their time of death? Is it not possible that a patient dies months away from a single BZD prescription? This is an important point that should be clarified

[Response] Yes, this is the correct interpretation of an intention-to-treat approach. We have further clarified this approach in the text:

Methods, page 6: "*Follow-up began on the day following the index date for BZD initiators and non-initiators. Patients were followed in an "intention-to-treat" approach until the occurrence of death, nursing home admission, end of continuous health plan enrollment, end of the study period (December 31st, 2013), or end of the observation period, whichever came first. Consistent with an "intention-to-treat" approach, we disregarded treatment variations occurring during the follow-up for BZD initiators and non-initiators.*"

-        Was there any chance that BZDs were not captured using the study data source as being purchased privately by the citizens? This may be the case when using claims databases from some European countries where BZDs are charged directly to citizens and not covered by NHS

[Response] BZDs are prescription medications in the U.S. and they are comprehensively covered. Thus, BZD exposure is well captured in the study data source.

-        Was in-hospital and out of hospital death equally captured? Authors should provide information to rule out possible outcome misclassification

[Response] Yes, both in-hospital and out of hospital death were equally captured through linkage of the Optum Clinformatics database with the Social Security Administration Death Master File. Though outcome misclassification cannot be completely ruled out in observational studies, we do not expect this to be a concern, as the capture of mortality in healthcare utilization databases is expected to be fairly specific. In particular, the specificity of the Social Security Administration Death Master File for detecting mortality status, has been estimated to range between 97%-100% (Hermansen SW et al., Am J Epidemiol. 2009; Schisterman EF et al., Popul Health Metr. 2004). High specificity of the outcome ensures unbiased relative risk estimates.

We have clarified this aspect in the text:

Discussion, page 13: "*The linkage with the Social Security Administration Death Master File ensured that mortality was captured with high specificity in our dataset, thus minimizing possible bias in the relative risk estimates.*"

- • Patient characteristics
- Authors identified patient characteristics during the six months prior to ID. This approach may lead to covariate underestimation and probably increase the chance of residual confounding. Was not possible to choose a longer time window for covariate assessment? Can the authors comment on it?

[Response] A 6-month baseline period, rather than a longer baseline period, was chosen in order to measure and adjust for the factors that may have more heavily influenced and precipitated the prescription of a BZD, i.e., aspects of care occurring in reasonable proximity to BZD initiation. In the context of a longer time window for covariate assessment, these aspects may become diluted by other factors that occurred at a greater temporal distance from BZD initiation, and that may therefore be of minor or no relevance to BZD initiation.

Nevertheless, to address the reviewer's concern, we re-ran our primary analysis using a 12-month baseline period for covariate definition and found consistent results, i.e., HdPS-adjusted HR = 1.04 (1.00-1.09), which provides re-assurance that the chosen covariate assessment window is adequate.

- • Statistical analysis
- Authors use antidepressants as active comparator. However, indications of use, especially of non-SSRI antidepressants, are not really comparable to those of BZD (e.g. tricyclics for headache, neuropathic pain, etc). I suggest restricting the comparator to SSRIs only

[Response] We thank the reviewer for this suggestion. We now restrict the active comparator analyses to patients initiating SSRI antidepressants only. Results were consistent with previous analyses including all antidepressants:

| Analysis | BZD Initiators | | | | Antidepressant Initiators | | | | |
	No. patients	No. Events	Person-years	IR[1]	No. patients	No. Events	Person-years	IR[1]	HR (95% CI)
Unadjusted	1,063,845	6,008	465,228	12.9	1,161,408	4,050	506,776	8.0	1.62 (1.55-1.68)
1:1 PS-matched	901,535	3,567	393,814	9.1	901,535	3,459	394,351	8.8	1.03 (0.99-1.08)
1:1 HdPS-matched	879,192	3,780	384,276	9.8	879,192	3,375	384,578	8.8	1.12 (1.07-1.17)
Analysis	BZD Initiators				SSRI Antidepressant Initiators				
	No. patients	No. Events	Person-years	IR[1]	No. patients	No. Events	Person-years	IR[1]	HR (95% CI)
Unadjusted	1,063,845	6,008	465,228	12.9	649,400	2,276	282,646	8.1	1.61 (1.53-1.68)

1:1 PS-matched	563,088	2,076	245,380	8.5	563,088	2,096	245,252	8.5	0.99 (0.93-1.05)
1:1 HdPS-matched	549,022	2,213	239,380	9.2	549,022	2,032	239,261	8.5	1.09 (1.03-1.16)

BZD: Benzodiazepine, IR: incidence rate, HR: hazard ratio, CI: confidence interval, PS: propensity score, Hd: high-dimensional
[1] Per 1,000 person-years

We have updated the text and tables accordingly.

- Psychiatric patients are often treated with multiple psychotropic drugs. Was polypharmacy accounted for? It is likely that the number of concomitant psychotropic drugs used is at least indirectly a risk factor for mortality, particularly for older persons

[Response] Yes, we balanced the use of each individual class of psychotropic medications and the total number of concomitant medications at baseline via 1:1 propensity-score matching (see Table 1 and eTable 2)

- Was the effect of duration and dosage of BZD use accounted for in some way? The association between benzodiazepines and death is likely to be stronger for high dosage and prolonged exposure. In addition, potential dose-effect relationship would increase the plausibility of the association between BZD use and death.

[Response]

Regarding duration of use: As we used an intention-to-treat approach, which disregards patterns of treatment during follow-up, the effect of duration of BZD use was not investigated. In order to address the reviewer's concerns, we implemented an exploratory hybrid approach where we 1) restricted the study population to (a) BZD initiators that remained on treatment and had no censoring events until the end of increasingly extended time windows (30, 60, 90, and 180 days) after treatment initiation and to (b) BZD non-users that had no censoring events until the end of increasingly extended time windows (30, 60, 90, and 180 days) after cohort entry, and 2) started the follow-up only after the accumulation of this time for each group (i.e. at day 31, 61, 91, and 181 days after the index date). We 1:1 PS-matched patients on the basis of baseline information measured during the six months preceding and including the index date up to the 30th, 60th, 90th, and 180th day after the index date, depending on the specific analysis, and followed-up patients in an "intention-to-treat" approach until the occurrence of death, nursing home admission, end of continuous health plan enrollment, end of the study period (December 31st, 2013), or end of a 180-day observation period, whichever came first. The results of these analyses are reported below and are consistent with primary findings:

1:1 HdPS-matched analysis	HR (95% CI)
Primary analysis	**1.00 (0.96-1.04)**
30-day duration of use	0.90 (0.85-0.94)
60-day duration of use	0.90 (0.84-0.97)
90-day duration of use	0.85 (0.77-0.95)
180-day duration of use	0.97 (0.86-1.11)

We would prefer not to include these results in the manuscript for two reasons: (1) the results don't provide additional insight since they are consistent with the primary analysis, and more importantly (2) we feel the standard intention-to-treat approach used in the primary analysis is overall a more valid method since it respects the temporality of events.

Regarding dose: Stratified analyses by dose have not been included in the manuscript due to the absence of a valid and recognized conversion scale across benzodiazepine agents. Despite this, to address the reviewer's concerns, in exploratory analyses we converted the daily dose of the BZD initiated on the index date to the lorazepam-equivalent daily dose, on the basis of the comparative potency provided at www.uptodate.com (Bystritsky A, Stein MB, Hermann R. Pharmacotherapy for generalized anxiety disorder in adults). Results stratified by tertiles of lorazepam equivalent daily doses are provided below and are consistent with primary findings:

1:1 HdPS-matched analysis	HR (95% CI)
Primary analysis	**1.00 (0.96-1.04)**
<1.5 mg/day lorazepam equivalent	0.83 (0.78-0.88)
1.5-2.9 mg/day lorazepam equivalent	0.91 (0.86-0.96)
≥ 3 mg/day lorazepam equivalent	1.06 (1.00-1.12)

Because these results are based on approximate equal potencies relative to lorazepam 1mg orally, which are not recommended for conversion between agents (Bystritsky A, Stein MB, Hermann R. Pharmacotherapy for generalized anxiety disorder in adults. www.uptodate.com), we would suggest not to include the results of these exploratory analyses in the manuscript.

-        In Figure 1, the authors specify that the BZD users and non-users are time matched; this should be specified in the methods. In particular, how was the time matching carried out? On calendar date (date of index prescription?)

[Response] We identified temporally-aligned (or time-matched) BZD-non-users as patients that had a medical visit +/– 14 days of the treatment initiation date (i.e., the index prescription) for the corresponding BZD-user and fulfilled the same inclusion criteria as BZD new users. We have provided further clarification in the Methods section:

Methods, page 6: "*In order to select non-BZD-initiators with similar opportunity to be evaluated and treated by a physician as BZD initiators **and within a similar time window**, for each BZD-exposed subject we identified a random patient that had a medical visit +/– 14 days of the treatment initiation date for the corresponding BZD-user and fulfilled the same inclusion criteria as BZD new users, i.e., six months of continuous health plan enrollment prior to the selected medical visit and no use of any BZD in the six months prior to and including the date of the visit (i.e., the index date for non-BZD-initiators).*"

Results
-        Authors indicate that after PS matching all characteristics were well balanced and ASD was <0.1. However, ASD for SSRI use was equal to 0.11 and if all psychotropic

drugs would be lumped together the ASD would be even higher, so I suggest to take this into account in the final analyses

[Response] Although we did not anticipate there to be residual confounding due to the minor residual imbalance in SSRI-use, to address the reviewer's concern, we have re-run the adjusted Cox models including the covariate SSRI use in addition to the exposure variable. Results are reported below and are identical to primary findings:

1:1 HdPS-matched analysis	HR (95% CI)
Primary analysis	1.00 (0.96-1.04)
Analysis further accounting for SSRI use	1.00 (0.96-1.04)

- How do the authors explain that the hdPS matched analysis yield higher HR (1.00; 0.96-1.04) than PS matched analysis (HR: 0.89; 0.85-0.93) and closer to crude estimates (HR: 1.79; 1.73-1.85)

[Response] Although the difference between the PS-matched HR=0.9 and the HdPS-matched HR=1.00 is small and probably not meaningful, we have investigated the list of empirical covariates identified by the hdPS algorithm and noted the presence of several pregnancy-related codes, which were included among the 200 empirically identified confounders used in fully adjusted analyses. In the context of this study, normal pregnancy may have functioned as a proxy for fair health, which is unmeasured in administrative data, and may have contributed to better confounding control (confounding may bias in either direction).

- E-Table 2: can the authors explain why the use of BZDs appears to double between 2004 and 2005?

[Response] The calendar year 2004 only contributed 6 months of data as we required a 6-month baseline period.

- E-Table 3: It is surprising that the frequency of depression in BZD users is higher than that of anxiety; can the authors comment on the accuracy of the psychiatric diagnoses coding I the study data source?

[Response] The prevalence of depression was modestly higher than the prevalence of anxiety among BZD initiators in our study (16% vs. 14%). This is not surprising to us as depression, in addition to being a fairly prevalent condition in the general population, frequently co-exists with aspects related to anxiety or sleep disorders. It is not unexpected that among patients where depression represents the major condition, anxiety and sleep disorders may be under-coded in administrative datasets. All these conditions as well as the use of psychotropic medications were well balanced after propensity-score-matching.

Discussion

- The absence of risk of death in long-acting BZD users and in older persons should be further commented in the discussion.

[Response] We now further comment on this aspect in the Discussion section:

Discussion, page 12: "*Residual confounding rather than true effect modification may also explain the moderate discrepancy in the risk of all-cause mortality between patients under and over 65 years and between patients initiating short- and long-acting BZDs. **In addition, the lack of an increased risk in all-cause mortality among BZD initiators over 65 years and among initiators of long-acting agents further suggests that the hypnotic effects of BZDs are likely not responsible for the increased risk observed in previous investigations**.*"

- The burden of psychotropic drugs in elderly people may play an important role in increasing risk of death and deserve further analyses and discussion by the authors.

[Response] As previously noted, most studies identifying a positive association between benzodiazepines and all-cause mortality were among adult or young adult populations. Conversely most studies conducted in populations of older adults (i.e., over 65 years) have not identified an increased risk in all-cause mortality associated with the use of benzodiazepines, beyond an increase in falls and fractures related mortality (where the burden of psychotropic medications may indeed play a major role). We therefore believe that further analyses and discussion of specific causes of mortality that may be affected by the burden of psychotropic medications in older adults would be out of scope of the current study.

- In general, description of previous studies on BZD and all-cause or cause specific mortalities should be more detailed: was any dose and duration effect observed? Were patients' subgroups (e.g. elderly people) at increased risk of death? Was there any specific indication of use as strong risk factor of death?

[Response] We now provide more detailed information in the Background and the Discussion sections:

Background, page 4: "*Despite earlier mixed results regarding a possible association between BZDs and all-cause mortality, **and mostly no indication of an increased risk among patients aged 65 years or older beyond falls and fractures related mortality**, **more** recent evidence has reported a three-fold or higher risk of all-cause mortality among **adult populations** exposed to BZDs, **even for durations of use shorter than one month**. Moreover, several studies **conducted in adult or young adult populations**, have **not provided support for a specific effect of BZDs that could explain the increased risk in all-cause mortality, but have rather suggested associations with a wide range of causes of death**, including cardiovascular disease, cancer, respiratory disease, and suicide.*"

Discussion, page 12: "*Yet, a number of studies **performed in adult populations** have reported strong associations between BZDs and all-cause mortality**, even for short durations of treatment,** as well as associations with **a wide range of causes of deaths***, including cardiovascular disease and cancer, **for which the biological mechanism leading to an increased risk remains unclear**.*"

We could not identify literature supporting the existence a strong association between indications of use for BZDs and all-cause mortality.

-        In the conclusion authors first state that the study suggest that no or small increase in risk of death is associated with BZD. Later on, they add that residual confounding may explain such a small increase that seems contradicting previous statement. I suggest rephrasing the conclusion.

[Response] In the conclusions, we wanted to point out that *at least part* of the small increase in risk observed in selected analyses could be due to residual confounding, but we did not want to entirely rule out the existence of a small risk of all-cause mortality associated with BZDs. We have tried to clarify this aspect in the text:

Discussion, page 14: "*The direction and the extent of the attenuation in the point estimates with increasing levels of adjustment, suggest that residual confounding likely explains **at least part of the noted** small increase in mortality risk observed in selected analyses.*"

Additional Questions:
Please enter your name: Gianluca Trifirò

Job Title: Assistant Professor of Pharmacology

Institution: University of Messina

Reimbursement for attending a symposium?: No

A fee for speaking?: No

A fee for organising education?: No

Funds for research?: Yes

Funds for a member of staff?: No

Fees for consulting?: No

Have you in the past five years been employed by an organisation that may in any way gain or lose financially from the publication of this paper?: No

Do you hold any stocks or shares in an organisation that may in any way gain or lose financially from the publication of this paper?: No

If you have any competing interests <A HREF='http://www.bmj.com/about-bmj/resources-authors/forms-policies-and-checklists/declaration-competing-interests'target='_new'> (please see BMJ policy) please declare them here: As leader of an academic pharmacoepidemiology team the Department to whic I'm affiliated received from several pharmaceutical companies some unconditional grants for reserach projects not related to the topic of this paper

Reviewer: 2

Recommendation:

Comments:
 This a well powered observational prospective cohort study in a large American private health insurer's database. The researchers went at all lengths to assure sophisticated matching in study recruitment and extensive multivariate analysis of confounding. Selected patients from this health insured population were patients who visited a doctor, had no BZD prescription in the previous 180 days and at least one other medication prescription in each of two preceding periods of 90 days. The active group initiated a BZD during the visit and the control group did not. The study seems well designed, conducted and analysed.

[Response] We thank the reviewer for the positive feedback.

There is no information on the indication and dose of the initiation (high dose for anxiety, low dose for sleeping pill) and on the further fate of BZD use among those initiating. Did they use this for only a month, were there refills ? Who many patients stopped after the initial prescription (as indicated by guidelines, and how many turned into chronic users, which should be avoided ?

[Response]

Regarding indication of use: Administrative datasets do not provide information regarding the specific indication of use associated with the initiation of a specific therapy, but they provide rich information regarding patient characteristics up to and including the date of drug initiation (including inpatient and outpatient diagnoses, medication use, and measures of healthcare utilization, e.g., number of medical visits, hospitalizations, etc.), which can be used to approximate indications of use and to measure important risk factors for mortality. All these patient characteristics (plus 200 empirically identified confounders, selected by the hdPS methodology) were accounted for and balanced across groups via propensity-score-matching.

15

Regarding dose: As previously noted, stratified analyses by dose have not been included in the manuscript due to the absence of a valid and recognized conversion scale across benzodiazepine agents. However, exploratory analyses stratified by tertiles of lorazepam-equivalent daily dose produced results consistent with the primary findings (see below), suggesting that indication of use and dose do not play a major role in the association between BZDs and all-cause mortality.

1:1 HdPS-matched analysis	HR (95% CI)
Primary analysis	1.00 (0.96-1.04)
<1.5 mg/day lorazepam equivalent	0.83 (0.78-0.88)
1.5-2.9 mg/day lorazepam equivalent	0.91 (0.86-0.96)
≥ 3 mg/day lorazepam equivalent	1.06 (1.00-1.12)

Because these results are based on approximate equal potencies relative to lorazepam 1mg orally, which are not recommended for conversion between agents (Bystritsky A, Stein MB, Hermann R. Pharmacotherapy for generalized anxiety disorder in adults. www.uptodate.com), we would suggest not to include these results in the manuscript.

Regarding treatment duration: As previously mentioned, in the context of a non-user comparator, we opted for an intention-to-treat approach rather than a per-protocol approach as the latter would have led to a differential opportunity for BZD users and non-users to be censored and potential informative censoring. In line with an intention-to-treat approach, treatment use and variations occurring during the follow-up for BZD initiators and non-initiators were disregarded.

In order to address comments from the 1st reviewer, we also ran an exploratory analysis where we 1) restricted the study population to (a) BZD initiators that remained on treatment and had no censoring events until the end of an increasingly extended time windows (30, 60, 90, and 180 days) after treatment initiation and to (b) BZD non-users that had no censoring events until the end of increasingly extended time windows (30, 60, 90, and 180 days) after cohort entry, and 2) started the follow-up only after the accumulation of this time for each group (i.e. at day 31, 61, 91, and 181 days after the index date).
This approach identified the following number of patients that remained on BZD treatment over time:

1:1 HdPS-matched analysis	No. BZD users
Primary analysis	1,252,988
30-day duration of use	1,070,357
60-day duration of use	392,775
90-day duration of use	129,365
180-day duration of use	40,338

These results show that approximately two thirds of BZD initiators stopped treatment within the first two months and a much smaller proportion of the population turned into chronic users.

We now provide more information in the Methods section:

Methods, page 6: "*We considered a six-month observation period (180 days) in the main analysis, and observation periods of 12 and 48 months of follow-up in sensitivity analyses.* **A 180-day observation period was chosen for the primary analysis as we empirically observed that in routine care approximately two thirds of BZD initiators stopped treatment within the first two months**."

The study spans a long period from 2004 to 2013 (10 years), and the nature of the population of this database might have changed during that period.

[Response] The reviewer raises an important point. We addressed the potential change in patient characteristics over time by propensity-score matching study participants within each calendar year. This ensured BZD initiators and non-initiators were matched by calendar time.

The question remains whether long term use of benzodiazepines as sleeping pills, even in low doses, is detrimental for QOL and has an impact on mortality.

[Response] As previously noted, the results from the exploratory analyses evaluating increasingly extended periods of BZD treatment on mortality risk were consistent with primary findings, i.e., they showed no increased risk of all-cause mortality associated with BZD initiation:

1:1 HdPS-matched analysis	HR (95% CI)
Primary analysis	**1.00 (0.96-1.04)**
30-day duration of use	0.90 (0.85-0.94)
60-day duration of use	0.90 (0.84-0.97)
90-day duration of use	0.85 (0.77-0.95)
180-day duration of use	0.97 (0.86-1.11)

For further details regarding these exploratory analyses, please refer to the response to the 1st reviewer on pages 8-9.

We would prefer not to include these results in the manuscript for two reasons: (1) the results don't provide additional insight since they are consistent with the primary analysis, and more importantly (2) we feel the standard intention-to-treat approach used in the primary analysis is overall a more valid method since it respects the temporality of events.

The authors should take care that their conclusions mention that this is about initiation of BZD (with no information of further exposure after initiation) and 6 months measurement of outcome (all cause mortality).

[Response] Thanks for the comment. We have updated our Conclusions to emphasize this study reports results from an intention-to-treat analysis. However, as we also report results for 12- and 48-month follow-up periods, we have not edited our Conclusions to suggest this study only evaluated results for a 6-month follow-up period.

Conclusions, page 14: "*Results from this large cohort study **based on an-intention-to-treat approach** suggest either no increase or a small increase in the risk of **all-cause** mortality associated with **BZD initiation**. If a detrimental effect with regard to all-cause mortality exists, it is likely to be much lower than previously stated and to have only modest clinical relevance, given its magnitude from both an absolute and relative perspective. The direction and the extent of the attenuation in the point estimates with increasing levels of adjustment, suggest that residual confounding likely explains at least part of the noted small increase in mortality risk observed in selected analyses.*"

Additional Questions:
Please enter your name: Vander Stichele

Job Title: Senior Research Coordinator

Institution: Department of Pharmacology, Ghent University

Reimbursement for attending a symposium?: No

A fee for speaking?: No

A fee for organising education?: No

Funds for research?: No

Funds for a member of staff?: No

Fees for consulting?: No

Have you in the past five years been employed by an organisation that may in any way gain or lose financially from the publication of this paper?: No

Do you hold any stocks or shares in an organisation that may in any way gain or lose financially from the publication of this paper?: No

If you have any competing interests <A HREF='http://www.bmj.com/about-bmj/resources-authors/forms-policies-and-checklists/declaration-competing-interests'target='_new'> (please see BMJ policy) please declare them here: NONE

Reviewer: 3

Recommendation:

Comments:
This study uses a cohort matching process to compare patients who initiated a prescription for benzodiazepines versus those who did not. Although neither designed with patient input nor inteded for patient dissemination, the study results are useful information for both patients and primary care physicians who may worry about the impact of benzodiazepines.

[Response] We thank the reviewer for the positive feedback.

I wish that the authors would better explain the idea of residual confounding, perhaps by example. The tables and flow charts are very helpful in understanding the study. Is a 95% confidence interval appropriate for so many variables? Should it be even more stringent? The discussion of why benzodiazepines might or might not contribute to mortality is useful to the results.

[Response] The use of a 95% confidence interval is a standard approach consistently used by most medical literature assessing one contrast of interest (in our case benzodiazepine users vs. non-users). Consistency in the choice of measures of association and confidence intervals facilitates comparisons across studies in terms of results and conclusions. The adjustment of confidence intervals may be sometimes chosen by select studies assessing multiple contrasts of interest or performing multiple testing over time (such as in a randomized controlled trial assessing multiple treatment arms or planning for multiple looks over time). This does not apply to our study.

My one concern about the study is the definition of an initiator. One prescription over a 90 day period within two weeks of a physician visit. My concern is that it combines a wide range of initiators. The person prescribed a low dose of alprazolam for a few weeks, versus someone taking a longer acting variant over a longer time are both initiators. Is it appropriate to consider this a discrete variable, or perhaps it is an issue for further study. And of course, we have no data about compliance or dosing behavior. My point being that initiators are not necessarily a homogeneous group for mortality comparison purposes.

[Response] The reviewer raises an important point, i.e., BZD initiators being a "heterogeneous" group with regard to mortality risk. To address this issue, we purposely considered and adjusted for a large number of patient characteristics (over 300 covariates) in order to capture both overt risk factors for mortality and variables that may be proxies of risk factors that were unmeasured or incompletely measured in our dataset. All these covariates were balanced after propensity-

score-matching, which means that the "patient heterogeneity" of BDZ users and non-users with regard to mortality risk was successfully accounted for in our analyses.

We have run further exploratory analyses evaluating the effect of BZD *dose and duration of use* on all-cause mortality. In brief, these analyses produced results consistent with our primary findings, i.e., they showed no increased risk of all-cause mortality associated with BZD initiation. For further details regarding these exploratory analyses, please refer to the response to the 1st reviewer on pages 8-9 and the response to the 2nd reviewer on page 15.

However, it is an interesting study, and the analyses appear well thought out and presented.

Additional Questions:
Please enter your name: Elaine Sieff

Job Title: Lay Reviewer

Institution: None

Reimbursement for attending a symposium?: No

A fee for speaking?: No

A fee for organising education?: No

Funds for research?: No

Funds for a member of staff?: No

Fees for consulting?: No

Have you in the past five years been employed by an organisation that may in any way gain or lose financially from the publication of this paper?: No

Do you hold any stocks or shares in an organisation that may in any way gain or lose financially from the publication of this paper?: No

If you have any competing interests <A HREF='http://www.bmj.com/about-bmj/resources-authors/forms-policies-and-checklists/declaration-competing-interests'target='_new'> (please see BMJ policy) please declare them here: