PEER REVIEW HISTORY

BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form (**http://bmjopen.bmj.com/site/about/resources/checklist.pdf**) and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below.

ARTICLE DETAILS

TITLE (PROVISIONAL)	SYMPTOMS OF DEPRESSION AND ALL-CAUSE MORTALITY IN FARMERS, A COHORT STUDY. THE HUNT STUDY, NORWAY.
AUTHORS	Letnes, Jon Magne; Torske, Magnhild; Hilt, Bjørn; Bjorngaard, Johan; Krokstad, Steinar

VERSION 1 - REVIEW

REVIEWER	Jonathan Stieglitz Institute for Advanced Study in Toulouse (France); University of New Mexico (USA)
REVIEW RETURNED	18-Dec-2015

GENERAL COMMENTS	GENERAL COMMENTS This paper explores associations between depression and all-cause mortality among employed rural Norwegians using a prospective design. The ability to link depression data to National Registry mortality data among a large sample is a strength of the study. In general, data are presented clearly and appropriately analyzed. However, the study contains many serious weaknesses including insufficient motivation, insufficient explanation of both data structure and rationale for data analysis, lack of additional analyses that would greatly strengthen the paper, and insufficient acknowledgment of study limitations. SPECIFIC COMMENTS ABSTRACT Objectives. This sub-section lacks a clear statement of the main study objective and hypothesis to be tested. It is vague to say "…to evaluate [the association between depression and mortality] in a socioeconomic setting". For example, is the main study objective to determine if higher socioeconomic position (SEP) protects against depression-induced mortality (as stated in the Conclusion, pg. 17 lines 44-49)? If so, this is not clear from the Abstract or Introduction, and would entail additional statistical analyses (e.g. inclusion of education as a potential confounder). STRENGTHS AND LIMITATIONS It is inaccurate and redundant to say that the study is conducted among "a total population in a geographic area". In reality the study is conducted in one (rural) county; lacking a major city (pg. 15, line 51), this one county cannot be considered to be population-representative. INTRODUCTION Discussion of why farmers may be expected to have improved health compared to non-farmers should be expanded to include diet (e.g. reduced consumption of processed foods) and psycho-social

factors (e.g. reduced inequality in rural farming communities, and thus reduced psycho-social stress from inequality). Relatedly, discussion of why farmers may be expected to have poorer health than non-farmers should acknowledge prior literature on depression in subsistence populations. For example, subsistence populations may be marginalized and exposed to greater uncertainties (especially environmental) compared to their market-integrated peers (Pike and Patil 2006; Stieglitz et al. 2015a; Stieglitz et al. 2015b).

Most importantly, the Introduction does not list any hypothesis or predictions that the study is designed to test. Are farmers predicted to have lower or higher all-cause mortality vs. the general working population? What are potential moderators or mediators underlying this association? Why should the association between depression and mortality be affected (positively or negatively) by occupational status in general and farming in particular? I realize that the journal imposes a 4,000 word limit for research articles, but the Introduction is simply too short, lacking sufficient motivation for hypothesis tests. And when stating hypotheses, the authors should be explicit about which variables are controls, and which are the primary predictors of interest. For example, in Tables 2-3 it's not clear to me why other variables (e.g. education, smoking, etc.) aren't included as controls. There are several places in the Introduction where words with connotations of valence (e.g. “negative work-related factors”, pg. 4 line 46; “negative impact of frequently changing legislations…”, pg. 4 line 51) are used inappropriately. Words such as negative or positive should be used to describe statistical associations rather than assigning valence to lifestyle features. This is not an exhaustive list, since similar occurrences appear throughout the manuscript (e.g. pg. 11 line 10, “more favourable pattern of alcohol consumption”).

MATERIALS AND METHODS
Were the 66,140 study participants representative of the county and Norwegian population? Can the authors compare socio-demographics of their sample to census data?

Depression is often associated with unemployment, and unemployed individuals are omitted from this study. Does this omission introduce potential bias and threaten generalizability? If so this should be stated in the study limitations. If a study goal is to only examine employed individuals, than this needs to be better justified, especially if the authors wish to claim that the sample is population-representative.

Figure 1 is helpful but the numbers in Figure 1 are not consistent with numbers in the text. N=65,232 in the first box of Figure 1, but the text states an N=66,140. There are other discrepancies within Figure 1 (e.g. 65,232-[9,602+16,245] ≠ 39,134) making it difficult to understand how the final sample for analysis was determined. Based on these discrepancies it is also difficult to determine the Q2 response rate. Also, in Figure 1, why not include an upstream box indicating those who were “Invited to participate”?

Mortality. It would be helpful to include descriptive information on Norwegian age-specific mortality or life expectancy. It is also unclear if data on cause-specific mortality are available; if so, it would be interesting if the authors analysed specific death causes separately. Lastly, I would like more detail on how the National Registry mortality database is assembled. How regularly are data collected?

How are these data actually gathered (e.g. who collects it, who codes it, are any data missing, etc.)?

Occupation. Why are drivers classified as manual workers, and not routine non-manual workers? More generally, I would like more detail on the "EGP scheme" used to classify by SEP, especially since many individuals self-report >1 occupation. I am assuming data on income and/or household wealth (more direct SEP measures) are unavailable, but is this correct?

Symptoms of depression. Items are scored on a scale ranging from 0-3 (pg. 8 line 28), but what does each level of the scale mean? The authors should report measures of internal consistency of the depression questionnaire. I would also like to see the individual items, and basic descriptives (mean and SD, segregated by sex) for each item, perhaps in a supplement. Also, was there any attempt to validate the depression score (Stieglitz et al. 2015a)?

Somatic long-lasting limiting illness. What is a “three-split scale”?

Smoking status. Are there no data on history of smoking (e.g. pack-years)?

Several typos should be corrected. For example, on pg. 7 line 14, “every Norwegian citizen have their own unique 11-digit personal identification number…” Also on pg. 7, line 26, “Classification of occupation in HUNT2 was based on a Norwegian occupational standard that is similar the social class…” This is not an exhaustive list of typos.

RESULTS
Table 1. It would be helpful to include tests of statistical significance comparing farmers to AOO for all study variables, or farmers to specific professional categories. Also, a minor omission: variables “somatic long-standing illness” and “education level” are missing “(%)” in column 1.

Table 3. In Model 2 – controlling for somatic long-standing illness – the effect of depression on mortality among farmers substantially weakens (to the point of no longer being marginally significant [based on the CI], although p-values are not reported), whereas there remains a significant effect of depression among professionals and manual workers. So what is the interpretation, that somatic illness mediates the association between depression and mortality (and if so why)? Also, this mediation effect seems at odds with the authors’ claim (on pg. 13 lines 11-13) that “Model 2 was in addition adjusted for somatic long-lasting limiting illness, and the same main patterns were seen here.” Moreover, in Table 3 why not also include an “All Other Occupations” category as in Table 2?

Sensitivity and additional analyses. Did the authors re-run analyses after omitting individuals reporting >1 occupation? Are results similar if depression is coded as a continuous variable (summed item score)? Did the authors try modelling effects of individual depression items on mortality, to see if somatic symptoms are more strongly linked to mortality than cognitive or emotional symptoms? The authors mention selection bias (pg. 16 line 1) due to health status; is it possible to re-run analyses among only those reporting somatic long-lasting illness (perhaps including smokers and those with a probable alcohol problem)? If similar patterns are found among this

	relatively “unhealthy” subsample, then it would suggest that self-selection of healthy farmers into the study is not the main factor underlying associations reported here. DISCUSSION The authors should outline potential mechanisms underlying the association between depression and mortality in this Norwegian sample. And outline future research necessary to examine these mechanisms. Also, the link between the Introduction and Discussion is quite weak, in part because the overarching hypothesis motivating this research is unclear. Study limitations. What, if any, bias is introduced by omitting from analyses subjects who emigrated (n=172)? The authors should also acknowledge as a limitation use of self-report data (rather than objective biomarkers or other clinical data) to assess baseline health status. REFERECES Pike IL, and Patil CL. 2006. Understanding women’s burdens: preliminary findings on psychosocial health among Datoga and Iraqw women of northern Tanzania. Culture, medicine and psychiatry 30(3):299-330. Stieglitz J, Schniter E, von Rueden C, Kaplan H, and Gurven M. 2015a. Functional Disability and Social Conflict Increase Risk of Depression in Older Adulthood Among Bolivian Forager-Farmers. The Journals of Gerontology Series B: Psychological Sciences and Social Sciences 70(6):948-956. Stieglitz J, Trumble B, Emery Thompson M, Blackwell A, Kaplan H, and Gurven M. 2015b. Depression as sickness behavior? A test of the host defense hypothesis in a high pathogen population. Brain Behavior and Immunity 49:130-139.

REVIEWER	Cédric Lemogne Associate Professor in Psychiatry Université Paris Descartes France
REVIEW RETURNED	01-Jan-2016

GENERAL COMMENTS	The authors are dealing with one particularly important, yet overlooked issue, namely the differential association between depression and mortality across different occupational categories. This is a topic worthy of investigation and with considerable clinical relevance. Despite this interesting research question, and the adequate sample to perform related analyses, the authors may want to consider the following issues, which in my opinion, may prevent the acceptance of the manuscript as it currently stands. According to the BMJ Open guidelines for reviewers, I will merely focus on methodological issues. 1) In the introduction, although the authors nicely made the case for studying all-cause mortality in farmers, they did not make the same for studying the association between depression and all-cause mortality per se in this category specifically. This point is especially important as the authors did not apply correction for multiple

statistical tests. Therefore, unless they clearly had a a priori hypothesis, one could attribute any unexpected result to chance.

Here, the question the reader might have is: why should one expect the association between depression and all-cause mortality be of different magnitude among farmers particularly? Recent findings suggest that psychological variables, including but not limited to depression, may relate to mortality-related outcomes to a different extent accross occupational categories.

Redmond N, Richman J, Gamboa CM et al. Perceived stress is associated with incident coronary heart disease and all-cause mortality in low- but not high-income participants in the Reasons for Geographic And Racial Differences in Stroke study. J Am Heart Assoc 2013;2:e000447.

Lazzarino AI, Hamer M, Stamatakis E, Steptoe A. Low socioeconomic status and psychological distress as synergistic predictors of mortality from stroke and coronary heart disease. Psychosom Med 2013;75:311 6.

Wiernik E, Pannier B, Czernichow S et al. Occupational status moderates the association between current perceived stress and high blood pressure: evidence from the IPC cohort study. Hypertension 2013;61:571-7.

2) As acknowledged by the authors, both clinical and sub-clinical depression has been associated with increased mortality. There is indeed no compelling evidence that this association is specific to major depression. Therefore, the use of a binary indicator of depression is questionable.

Indeed, using a cut-off when it comes to examine the association between a continuous measure and any health-related outcome may produce unstable results and/or reduce statistical power.

MacCallum R.C., Zhang, S., Preacher, K.J., & Rucker, D.D. (2002). On the practice of dichotomization of quantitative variables. Psychological Methods, 7(1), 19-40.

In the present article, using a cut-off of 8 means that an individual scoring 7 on the HADS-D is considered closer to an individual scoring zero than to an individual scoring 8, which could blurr significant associations. Although assuming a linear relationship is also a strong postulate, the authors may at least consider to run sensitivity analyses taking depression as a continous measure as well (perhaps rescaling it to the 25th - 75th percentile interval to provide meaningful HRs).

This is indeed what is proposed by the statistical guidelines of the American Psychosomatic Society.

3) Although I understand that statistical power may limit the use of more than 5 occupational categories, I was wondering whether the "professionals" category would be more heterogenous than the others as regards SES, merging, for instance, nurses and lawyers. Did the authors attempt to split at least this very broad category ?

4) Current smoking status was crudely categorized to yes / no. Were more detailled data about smoking status available ?

	5) The authors may want to perform further adjustment for mortality risk factors including alcohol misuse, smoking and BMI in order to examine to what extent these factors may mediate 1) the association between occupation and mortality and 2) the association between depression and mortality, particularly in farmers. 6) Since the authors focused their study on farmers, they may want to choose this category as the reference category in the analyses, so that they may compare it with all the other categories, and not only the highest. 7) The authors focus their article on the association between depression and mortality among one subgroup but this association is not statistically significant, even when adjusting only for age and sex. However the discussion seems to consider it to be significant. This is particularly problematic in the context of multiple tests (at leat one per occupational category), which could have warranted a corrected threshold (e.g. P<0.01 according to a Bonferroni correction). Likewise, the authors write that "Our study indicates that the depression mortality association is present in all SEP occupational groups." As far as I understand table 3, this association is significantly in only two occupational groups. Therefore, unless I misunderstood the results section and tables, it seems that the discussion is not entirely supported by the results. Please note that I am all for publishing negative results from well-conducted studies, as this one, but I am bit confused by the discussion as it stands. 8) The authors may want to formally test the interaction between occupational category and depression.

VERSION 1 – AUTHOR RESPONSE

Reviewer #1:

1) In the introduction, although the authors nicely made the case for studying all-cause mortality in farmers, they did not make the same for studying the association between depression and all-cause mortality per se in this category specifically. This point is especially important as the authors did not apply correction for multiple statistical tests. Therefore, unless they clearly had a a priori hypothesis, one could attribute any unexpected result to chance.
Here, the question the reader might have is: why should one expect the association between depression and all-cause mortality be of different magnitude among farmers particularly? Recent findings suggest that psychological variables, including but not limited to depression, may relate to mortality-related outcomes to a different extent accross occupational categories.

Authors' response: The reviewer is quite right. The part of the introduction regarding the motivation for studying the association between depression and mortality is not properly justified as it stands. We have therefore made an extensive change to this part of the manuscript.

2) As acknowledged by the authors, both clinical and sub-clinical depression has been associated with increased mortality. There is indeed no compelling evidence that this association is specific to major depression. Therefore, the use of a binary indicator of depression is questionable. Indeed, using a cut-off when it comes to examine the association between a continuous measure and any health-related outcome may produce unstable results and/or reduce statistical power.
In the present article, using a cut-off of 8 means that an individual scoring 7 on the HADS-D is

considered closer to an individual scoring zero than to an individual scoring 8, which could blurr significant associations. Although assuming a linear relationship is also a strong postulate, the authors may at least consider to run sensitivity analyses taking depression as a continous measure as well (perhaps rescaling it to the 25th - 75th percentile interval to provide meaningful HRs). This is indeed what is proposed by the statistical guidelines of the American Psychosomatic Society.

Authors' response (to both reviewers): The measurement of symptoms of depression and classification of caseness of depression is, as the reviewers state, an important issue. We tried splitting HADS depression scores into four categories based on increasing symptom pressure (HADS depression scores of 0-7, 8-11, 12-14 and 15-21), as others have tried earlier. The problem we then met was that some of the analysed occupational groups end up with very low numbers of subjects and very few events - thus affecting power. For example, 18 farmers had a total symptom depression score in the highest of the four-category HADS classification, and only four of them died during the ~20 years of follow-up. In the self-employed workers group only five subjects were in the category of highest depression symptoms pressure, and none of them died.
Doing an analysis with symptoms of depression as a continuous variable might be a good idea. However, from our perspective, this offers a more theoretical and less practical approach to the question. Clinicians and health care providers often seek to classify patients in a dichotomous fashion; "healthy" or "ill". The interpretation of data with symptoms of depression as a continuous variable would hence be more difficult, and it is an important reason as to why we favour a dichotomisation.
We have provided a section on this in the discussion chapter.

3) Although I understand that statistical power may limit the use of more than 5 occupational categories, I was wondering whether the "professionals" category would be more heterogenous than the others as regards SES, merging, for instance, nurses and lawyers. Did the authors attempt to split at least this very broad category?

Authors' response: Indeed, splitting the occupational categories would be interesting. During the build-up of the analyses, much time was spent on both finding the most correct design of the revised EGP scheme and finding a tolerable division of the occupational groups. We tried division into both nine and seven occupational categories, but as mentioned, statistical power was a problem when taking into consideration the low numbers of deaths among participants with symptoms of depression.

4) Current smoking status was crudely categorized to yes / no. Were more detailed data about smoking status available?

Authors' response: Detailed data on smoking status was available in the form of number of cigarettes per day among those who smoked. This has been added in the revised table 1.

5) The authors may want to perform further adjustment for mortality risk factors including alcohol misuse, smoking and BMI in order to examine to what extent these factors may mediate 1) the association between occupation and mortality and 2) the association between depression and mortality, particularly in farmers.

Authors' response (to both reviewers):
We used Directed Acyclic Graphs (DAGs) to evaluate possible confounders and mediators. We considered age and sex to be confounders. We were, however, uncertain whether somatic long-lasting illness was a confounder or a mediator. Because of this uncertainty, we chose to adjust for it in a separate model. We deliberately did not adjust for variables we considered to be mediators, such as the ones mentioned – alcohol abuse, smoking, BMI etc.
Education was deliberately omitted from the analyses, as there already is an educational gradient in

the occupational hierarchy.
Effect decomposition – adjusting for a mediator in an attempt to distinguish direct and indirect effects on the outcome of an exposure – is commonly done in social epidemiology. However, according to Rothman, Greenland and Lash (2008), (Chapters 12 and 26 in Modern Epidemiology, 3rd edition, edited by Rothman, Greenland and Lash), adjusting for an intermediate may actually introduce confounding where there was none originally. They further argue that this approach is “not reliably valid” when used as it is commonly done in social epidemiology, and they caution against “considering the ratio between controlled direct and total effects as the “proportion explained” by the specified intermediate” (p. 546). We realize that opinions regarding adjustment for mediators may vary, but we have chosen to follow the recommendations by Rothman et al., and we have consequently not adjusted for any variables we considered to be mediators.

Reference: Glymour MM, Greenland S. Causal Diagrams. In: Rothman KJ, Greenland S, Lash TL, eds. Modern epidemiology. 3rd Ed. Philadelphia: Lippincott Williams & Wilkins, 2008: 183-212

6) Since the authors focused their study on farmers, they may want to choose this category as the reference category in the analyses, so that they may compare it with all the other categories, and not only the highest.

Authors’ response: Using farmers as a reference category is a good idea, and could help to visualise our results in regard to farmers in a clearer fashion. Although, using farmers as a reference category could make the comparison between farmers and the other occupations more easy, we find it better to use the highest (or alternatively the lowest) ranked category as the reference when comparing several occupational groups in a socioeconomic hierarchy. This makes for a more easy comparison of the categories.

7) The authors focus their article on the association between depression and mortality among one subgroup but this association is not statistically significant, even when adjusting only for age and sex. However the discussion seems to consider it to be significant. This is particularly problematic in the context of multiple tests (at leat one per occupational category), which could have warranted a corrected threshold (e.g. P<0.01 according to a Bonferroni correction). Likewise, the authors write that "Our study indicates that the depression mortality association is present in all SEP occupational groups." As far as I understand table 3, this association is significantly in only two occupational groups. Therefore, unless I misunderstood the results section and tables, it seems that the discussion is not entirely supported by the results. Please note that I am all for publishing negative results from well-conducted studies, as this one, but I am bit confused by the discussion as it stands.

Authors’ response: The use of p-values has been the topic of much debate in epidemiology and statistics over the years. Because P-values confound effect size and sample size, and because they are often misinterpreted, we did not provide them in Table 2 or Table 3. We provided 95% confidence intervals as an estimate of the uncertainty of the point estimate due to random error, according to Rothman, Greenland and Lash (2008) (chapter 10 in Modern Epidemiology, 3rd edition, edited by Rothman, Greenland and Lash). If one uses 95% confidence intervals to determine whether or not there is a statistically significant effect – that is, assess whether the confidence interval includes 1, or whether confidence intervals overlap - it is essentially the same as using an (arbitrary) cut-off p-value of p=0.05, with the same problems as mentioned previously. All effect estimates are >1, indicating that there is an association between symptoms of anxiety or depression and mortality, even though the confidence intervals in several groups include 1, because our data material was not large enough. Thus the size of the effect estimates has been our main focus, not statistical significance. However, given the many opinions regarding the use of p-values, we understand your confusion, and in the revised document we have stated more clearly that our main findings are not statistically significant.

Reference: Rothman KJ, Greenland S, Lash TL. Precision and Statistics in Epidemiologic Studies. In: Rothman KJ, Greenland S, Lash TL, eds. Modern epidemiology. 3rd Ed. Philadelphia: Lippincott Williams & Wilkins, 2008: 148-168

8) The authors may want to formally test the interaction between occupational category and depression.

Authors' response: We have tested the interaction between occupational category and symptoms of depression, and reported the findings in the results chapter.

Reviewer #2:

ABSTRACT

Objectives. This sub-section lacks a clear statement of the main study objective and hypothesis to be tested. It is vague to say "...to evaluate [the association between depression and mortality] in a socioeconomic setting". For example, is the main study objective to determine if higher socioeconomic position (SEP) protects against depression-induced mortality (as stated in the Conclusion, pg. 17 lines 44-49)? If so, this is not clear from the Abstract or Introduction, and would entail additional statistical analyses (e.g. inclusion of education as a potential confounder).

Authors' response: The reviewer is quite right, the main objective was not to determine if higher SEP protects against depression-induced mortality. However, we found the strong effect in the professionals category interesting. This however, should not have been included in the "conclusions and final comments" section and has therefore been omitted in the revised edition. This action should clear up the discrepancy noted by the reviewer.

STRENGTHS AND LIMITATIONS

It is inaccurate and redundant to say that the study is conducted among "a total population in a geographic area". In reality the study is conducted in one (rural) county; lacking a major city (pg. 15, line 51), this one county cannot be considered to be population-representative.

Authors' response: Although Nord-Trøndelag County does not have any major cities, demographical data are fairly similar to the national average, as we argue in the revised manuscript.

INTRODUCTION

Discussion of why farmers may be expected to have improved health compared to non- farmers should be expanded to include diet (e.g. reduced consumption of processed foods) and psycho-social factors (e.g. reduced inequality in rural farming communities, and thus reduced psycho-social stress from inequality).

Relatedly, discussion of why farmers may be expected to have poorer health than non-farmers should acknowledge prior literature on depression in subsistence populations. For example, subsistence populations may be marginalized and exposed to greater uncertainties (especially environmental) compared to their market-integrated peers (Pike and Patil 2006; Stieglitz et al. 2015a; Stieglitz et al. 2015b).

Authors' response: We certainly agree that farmers are marginalised and exposed to uncertainty, as we discuss in the introduction. Although, if we understand correctly, we do not see them as a subsistence population. Their life style in a developed country of today, like Norway, and in the span of the last 20 years, is not consistent with subsistance.

Most importantly, the Introduction does not list any hypothesis or predictions that the study is designed to test. Are farmers predicted to have lower or higher all-cause mortality vs. the general

working population? What are potential moderators or mediators underlying this association? Why should the association between depression and mortality be affected (positively or negatively) by occupational status in general and farming in particular? I realize that the journal imposes a 4,000 word limit for research articles, but the Introduction is simply too short, lacking sufficient motivation for hypothesis tests.

Authors' response: We have made extensive changes to the introduction chapter in line with comments from both authors, including specifying why one could expect depression-associated mortality in farmers to be different to other occupational groups, and a further justification of our hypotheses.

And when stating hypotheses, the authors should be explicit about which variables are controls, and which are the primary predictors of interest. For example, in Tables 2-3 it's not clear to me why other variables (e.g. education, smoking, etc.) aren't included as controls.

Authors' response: Explanations regarding adjustment for confounders and mediators have been addressed in an answer to reviewer #1 under authors' response to question number 5.

There are several places in the Introduction where words with connotations of valence (e.g. "negative work-related factors", pg. 4 line 46; "negative impact of frequently changing legislations...", pg. 4 line 51) are used inappropriately. Words such as negative or positive should be used to describe statistical associations rather than assigning valence to lifestyle features. This is not an exhaustive list, since similar occurrences appear throughout the manuscript (e.g. pg. 11 line 10, "more favourable pattern of alcohol consumption").

Authors' response: We have made changes regarding these comments.

MATERIALS AND METHODS Were the 66,140 study participants representative of the county and Norwegian population? Can the authors compare socio-demographics of their sample to census data?

Authors' response: Changes have been made regarding this comment in the "Study population and background" section.

Depression is often associated with unemployment, and unemployed individuals are omitted from this study. Does this omission introduce potential bias and threaten generalizability? If so this should be stated in the study limitations. If a study goal is to only examine employed individuals, than this needs to be better justified, especially if the authors wish to claim that the sample is population-representative.

Authors' response: HUNT participants who had been in paid employment but were unemployed when they participated in HUNT2 are included in our analyses. We excluded HUNT2 participants who have never been in paid employment and/or had not reported an occupational category. We appreciate the comment and have rephrased and expanded the "Occupation" section to explain the exclusion criteria more clearly.

Figure 1 is helpful but the numbers in Figure 1 are not consistent with numbers in the text. N=65,232 in the first box of Figure 1, but the text states an N=66,140. There are other discrepancies within Figure 1 (e.g. 65,232-[9,602+16,245] ≠ 39,134) making it difficult to understand how the final sample for analysis was determined. Based on these discrepancies it is also difficult to determine the Q2 response rate. Also, in Figure 1, why not include an upstream box indicating those who were "Invited to participate"?

Authors' response: The presented numbers might, as the reviewer states, be confusing. The discrepancy arises from different numbers being presented in the method article and in our material available for the selection of included participants. The method article was written in 2013, and the data material is under continuous quality assurance, which results in differences in the presented HUNT2 baseline population (65,232 vs. 66,140). Discrepancies can also be found in how participation in HUNT2 Q1 has been defined, e.g. finishing just one question, or finishing the whole survey, and these two mentioned factors combined are among the most important in accounting for the discrepancy. These misunderstandings and errors have been corrected in the revised edition, and the figure has been corrected and the upstream box ("invited to participate") has been added, as rightfully asked for by the reviewer.

Mortality. It would be helpful to include descriptive information on Norwegian age-specific mortality or life expectancy.

Authors' response: We have provided data on Norwegian life expectancy for the study start year 1995 and the last year of follow-up, 2014.

It is also unclear if data on cause-specific mortality are available; if so, it would be interesting if the authors analysed specific death causes separately. Lastly, I would like more detail on how the National Registry mortality database is assembled. How regularly are data collected? How are these data actually gathered (e.g. who collects it, who codes it, are any data missing, etc.)?

Authors' response: We did not obtain data on cause-specific mortality, and this is now further specified in the revised edition. We have also included further information on how mortality data was assembled.

Occupation. Why are drivers classified as manual workers, and not routine non-manual workers? More generally, I would like more detail on the "EGP scheme" used to classify by SEP, especially since many individuals self-report >1 occupation. I am assuming data on income and/or household wealth (more direct SEP measures) are unavailable, but is this correct?

Authors' response: One could certainly argue that drivers could be classified in the routine non-manual workers category, however, drivers were originally placed in the manual workers category in the original EGP scheme. Further we would also assume that this varied occupational group overall would have more in common with the manual workers category, especially as this group probably would be dominated by transport drivers with a substantial part of their work having the characteristics of manual work. In general, the use and the revising of the EGP scheme were thoroughly discussed among the authors. Our solution may not be the best one, but at the same time it certainly is difficult to know what is the most advantageous solution. One of the main reasons for using the EGP as a SEP gradient was to help pick the highest SEP occupation as the main one for those with several occupations, and at the same time being able to compare farmers with other socioeconomic groups. Using occupational groups as a marker for socioeconomic position was therefore the most natural solution. As the reviewer assumes, data on income and / or household wealth was not available.

Symptoms of depression. Items are scored on a scale ranging from 0-3 (pg. 8 line 28), but what does each level of the scale mean?

Authors' response: The levels of the scale differ between each question, but as described, an increasing score (from 0-3) reflects increasing symptom pressure. This has been reformulated in the revised manuscript.

The authors should report measures of internal consistency of the depression questionnaire.

Authors' response: A study performed by Mykletun (2001) has addressed this theme and showed that internal consistency regarding HADS in the HUNT2 is satisfactory. Information on this has now been included in the methods section under "Symptoms of depression", along with the full reference.

I would also like to see the individual items, and basic descriptives (mean and SD, segregated by sex) for each item, perhaps in a supplement.

Authors' response: We have provided a table as supplementary material of HADSD item scores with mean and SD segregated by sex and combined.

Also, was there any attempt to validate the depression score (Stieglitz et al. 2015a)?

Authors' response: No attempts were made to externally validate depression scores in the HUNT2 Study in the manner done in the study to which the reviewer refers. However, internal consistency has, as previously mentioned, been reported as satisfactory.

Somatic long-lasting limiting illness. What is a "three-split scale"?

Authors' response: We have described this using other wording in the revised version.

Smoking status. Are there no data on history of smoking (e.g. pack-years)?
Authors' response: Cigarettes per day among the different occupational groups is now presented in table 1.

RESULTS
Table 1. It would be helpful to include tests of statistical significance comparing farmers to AOO for all study variables, or farmers to specific professional categories. Also, a minor omission: variables "somatic long-standing illness" and "education level" are missing "(%)" in column 1.

Authors' response: These statistical tests were originally performed and have now been included in the revised manuscript. All the general characteristics were statistically significantly different between farmers and the all other occupations (AOO) group, except for numbers of cigarettes smoked per day among active smokers. The missing percentage marks have been included in Table 1.

Table 3. In Model 2 – controlling for somatic long-standing illness – the effect of depression on mortality among farmers substantially weakens (to the point of no longer being marginally significant [based on the CI], although p-values are not reported), whereas there remains a significant effect of depression among professionals and manual workers. So what is the interpretation, that somatic illness mediates the association between depression and mortality (and if so why)? Also, this mediation effect seems at odds with the authors' claim (on pg. 13 lines 11-13) that "Model 2 was in addition adjusted for somatic long-lasting limiting illness, and the same main patterns were seen here."

Authors' response: We have clarified our interpretation in the revised edition.

Moreover, in Table 3 why not also include an "All Other Occupations" category as in Table 2?

Authors' response: This has been included in the revised edition.

Sensitivity and additional analyses. Did the authors re-run analyses after omitting individuals reporting

>1 occupation?

Authors' response: We have performed sensitivity analyses which are now reported in the results chapter.

Are results similar if depression is coded as a continuous variable (summed item score)?

Authors' response: The same question has been answered to reviewer #1 under authors' response to question number 2.

Did the authors try modelling effects of individual depression items on mortality, to see if somatic symptoms are more strongly linked to mortality than cognitive or emotional symptoms?

Authors' response: We did not try modelling effects of individual depression items to see if somatic symptoms are more strongly linked to mortality. Further, the HADS (Hospital Anxiety and Depression Scale) was invented for hospital use, and for that reason, somatic depression symptoms are omitted from this depression-screening tool.

The authors mention selection bias (pg. 16 line 1) due to health status; is it possible to re-run analyses among only those reporting somatic long-lasting illness (perhaps including smokers and those with a probable alcohol problem)? If similar patterns are found among this relatively "unhealthy" subsample, then it would suggest that self- selection of healthy farmers into the study is not the main factor underlying associations reported here.

Authors' response: Analyses on only those reporting somatic LLI gives very uncertain estimates owing to low numbers in some of the occupational categories, including farmers. We have therefore refrained from presenting such an analysis.

DISCUSSION

The authors should outline potential mechanisms underlying the association between depression and mortality in this Norwegian sample. And outline future research necessary to examine these mechanisms.

Authors' response: We have made changes regarding this comment in the revised manuscript.

Also, the link between the Introduction and Discussion is quite weak, in part because the overarching hypothesis motivating this research is unclear.

Authors' response: We have improved this link by making changes to both the introduction and discussion.

Study limitations. What, if any, bias is introduced by omitting from analyses subjects who emigrated (n=172)?

Authors' response: The important subject of possible selection bias from emigration has been addressed in the revised edition. We expect the potential bias from emigrated subjects to be negligible, owing to the very low percentage (0.4%).

The authors should also acknowledge as a limitation use of self-report data (rather than objective biomarkers or other clinical data) to assess baseline health status.

Authors' response: We have mentioned this limitation in the revised manuscript.

Other comments from the authors regarding the revised manuscript:
In the revision we have changed a reference to a recently published paper as well as deleted some references to fit the manuscript demands from BMJ Open, as some new references were added to the introduction and discussion chapters.

VERSION 2 – REVIEW

REVIEWER	Jonathan Stieglitz Institute for Advanced Study in Toulouse, France
REVIEW RETURNED	19-Feb-2016

GENERAL COMMENTS	The revised paper has improved substantially, but there are still several important weaknesses that must be addressed. The Introduction still lacks clarity. The authors end the first paragraph by saying, “With the earlier mentioned changes in the farming occupation, one could expect an increasing mortality risk in farmers, approaching the level of the general population – a hypothesis that requires further investigation.” But earlier the authors mention why mortality among farmers should be lower, not higher than the general population. Only in the next paragraph do they acknowledge why mortality might be higher among farmers. The Introduction still does not sufficiently motivate the study. The authors state, “There is a well-established link between depression and mortality which makes the combination of low mortality rates and high levels of depression among farmers particularly interesting.” But the combination of low mortality rate and high levels of depression is very common globally, and across many socioeconomic classes. So if not specific to farmers then what makes this link interesting in the Norwegian farming context? Another example of insufficient motivation is when the authors state “The mentioned high rate of suicides among farmers could further indicate that depression in this occupational group is particularly threatening”. But threatening in what sense, reducing economic productivity? Can one estimate among Norwegian farmers disability adjusted life years (DALYs) or other measures of foregone economic productivity or social costs related to depression, to get a sense of the public health burden? In Methods the authors state “Average income, prevalence of higher education, and smoking prevalence are somewhat lower than the national average.” This appears to contradict claims throughout the paper that the sample is, in fact, population-representative. If income and education are lower - and more common among depressed (as is often found), then depression prevalence should be higher in this region vs. the general population. I would refrain from referring to the sample as population-representative. In Results, regarding further analysis of those reporting multiple occupations, I would like to know among farmers reporting multiple occupations (22% of farmers), for what percentage was farming the higher SEP occupation? Is depression, mortality, or the link between depression and mortality higher for these individuals vs. other farmers (who report farming as the lower SEP occupation)? This

could serve as a robusticity check of the main hypothesis of interest. The authors even state in the Discussion that “An alternative and probably better approach might have been to classify study participants based on their main occupation, but information on the main occupation was not available to us. This may have led to a biased classification of workers as farmers, which would subsequently affect our estimates. This is especially relevant since one out of five farmers in our population stated that they had more than one occupation.”

Why the authors don’t include a model 3 in Table 3 – including other available data shown in Table 1 (e.g. smoking, alcoholism, education, etc.) – is a complete mystery to me. Inclusion of these variables might help identify potential mechanisms underlying the link between depression and mortality, which would contribute to study novelty. There is no theoretical or practical justification for only including three control variables (age, sex, self-reported somatic illness). In an expanded model 3 I would also like to see potential interactions between occupation and these other variables currently omitted from table 3. The authors even mention in the Discussion that “Lifestyle factors such as the low prevalence of possible alcohol abuse and smoking may contribute to the low mortality of farmers.” If this can be directly tested with existing data, and if it falls under the general study objective (which I think it does) then why not test it?

Figure 1 would be improved if the authors included an additional box after “study population” to include the number of participants who died, the number lost to follow-up (e.g. due to emigration), and the number who were censored.

Several study limitations are not mentioned. Depression symptoms are only assessed for the previous week, which is a very short temporal interval for symptom assessment. For this reason other validated and reliable depression instruments often use much longer temporal intervals (e.g. past month, past six months, past year). This affects categorization of who is vs. isn’t depressed, despite the authors’ claim in the Strengths and Limitations bullet points that the study offers “a low risk of misclassification”. Another study limitation not currently mentioned is the use of self-reported illness data, rather than objective biomarkers or clinical measures of illness. This limitation could, in principle affect other self-reported measures used here; for example if you self-report a long-standing illness then you might report "nearly always" or "quite often" being worn out from work even if work is not physically demanding. Have the authors tried to verify the accuracy of these and other self-reports using available data? Also, the authors state that a 70.2% response rate is “high” (lines 4-5, pg. 19) when, in fact, many social scientists would disagree with this claim. This limitation should be acknowledged too.

In the Discussion the authors state “Although our results cannot be directly generalised to all countries, they are probably valid for countries with a similar socioeconomic environment and agricultural evolution.” Please provide examples.

Lastly, there are several redundancies throughout the text that should be eliminated. For example, in the Abstract the Objectives are: “To explore all-cause mortality and the association between symptoms of depression and all-cause mortality in farmers, and to evaluate this in a socioeconomic setting using a prospective cohort design.” Delete “and to evaluate this in a socioeconomic setting”.

	Also, in the Strength and Limitations bullet #3 the word depression appears twice consecutively.

REVIEWER	Cédric Lemogne Université Paris Descartes Paris France
REVIEW RETURNED	24-Feb-2016

GENERAL COMMENTS	The authors have improved their manuscript in many ways. However, I am still concerned about three issues: - the lack of analyses adjusted for behavioral risk factors; in my opinion, such analyses could provide meaningful results. - the lack of sensitivity analyses using a continuous measure of depression; in my opinion, dichotomization may result in both Type I and Type II error. Even if clinicians are more comfortable with ill vs healthy categorization, such approach may not adequately capture the association between depressive symptoms and mortality - the interpretation of the main results; in my opinion, confidence interval including 1 may signal a high probability to observe these results under the assumption of a null hypothesis I do appreciate that the authors responded point-by-point to these comments and I realize that I could be wrong. Perhaps a statistical reviewer could have a look on the authors' responses regarding these three points?

VERSION 2 – AUTHOR RESPONSE

Reviewer #1: Jonathan Stieglitz
The Introduction still lacks clarity. The authors end the first paragraph by saying, "With the earlier mentioned changes in the farming occupation, one could expect an increasing mortality risk in farmers, approaching the level of the general population - a hypothesis that requires further investigation." But earlier the authors mention why mortality among farmers should be lower, not higher than the general population. Only in the next paragraph do they acknowledge why mortality might be higher among farmers.
The Introduction still does not sufficiently motivate the study. The authors state, "There is a well-established link between depression and mortality which makes the combination of low mortality rates and high levels of depression among farmers particularly interesting." But the combination of low mortality rate and high levels of depression is very common globally, and across many socioeconomic classes. So if not specific to farmers then what makes this link interesting in the Norwegian farming context?

Authors' response: We have made changes to the introduction chapter addressing these comments.

Another example of insufficient motivation is when the authors state "The mentioned high rate of suicides among farmers could further indicate that depression in this occupational group is particularly threatening". But threatening in what sense, reducing economic productivity? Can one estimate among Norwegian farmers disability adjusted life years (DALYs) or other measures of foregone economic productivity or social costs related to depression, to get a sense of the public health burden?

Authors' response: We have tried to clear up these statements in the text, as we meant "threatening" in regard to mortality risk. We agree that disability adjusted life years (DALYs) or other measures can

probably give a sense of the public health burden, but that was not the scope of this paper.

In Methods the authors state "Average income, prevalence of higher education, and smoking prevalence are somewhat lower than the national average." This appears to contradict claims throughout the paper that the sample is, in fact, population-representative. If income and education are lower - and more common among depressed (as is often found), then depression prevalence should be higher in this region vs. the general population. I would refrain from referring to the sample as population-representative.

Authors' response: We cannot see that we have claimed that the material is representative of the Norwegian population. Although, we did refer to Nord-Trøndelag County as a "miniature Norway" in regard to geographical characteristics, which we have now specified.

In Results, regarding further analysis of those reporting multiple occupations, I would like to know among farmers reporting multiple occupations (22% of farmers), for what percentage was farming the higher SEP occupation?

Authors' response: Firstly, to clear things up, 22% of those coded as farmers in our study reported multiple occupations (now 23,5% after changes to the study population, see comments further down). Among participants coded with a higher SEP occupation instead of being coded as farmers, 792 were farmers (4.8%). Farming was the highest SEP occupation among 83% of the participants who reported farming as at least one of their occupations. We have added this information in the methods section.

Is depression, mortality, or the link between depression and mortality higher for these individuals vs. other farmers (who report farming as the lower SEP occupation)? This could serve as a robusticity check of the main hypothesis of interest. The authors even state in the Discussion that "An alternative and probably better approach might have been to classify study participants based on their main occupation, but information on the main occupation was not available to us. This may have led to a biased classification of workers as farmers, which would subsequently affect our estimates. This is especially relevant since one out of five farmers in our population stated that they had more than one occupation."

Authors' response: We have now analysed all participants who answered having farming as an occupation together for the association of depression and mortality. We have also done subgroup analysis on the farmers coded with a higher SEP occupation. We have added a subsection on this in the "sensitivity analyses" section, and discussed these findings accordingly.

Why the authors don't include a model 3 in Table 3 - including other available data shown in Table 1 (e.g. smoking, alcoholism, education, etc.) - is a complete mystery to me. Inclusion of these variables might help identify potential mechanisms underlying the link between depression and mortality, which would contribute to study novelty.
There is no theoretical or practical justification for only including three control variables (age, sex, self-reported somatic illness). In an expanded model 3 I would also like to see potential interactions between occupation and these other variables currently omitted from table 3.
The authors even mention in the Discussion that "Lifestyle factors such as the low prevalence of possible alcohol abuse and smoking may contribute to the low mortality of farmers." If this can be directly tested with existing data, and if it falls under the general study objective (which I think it does) then why not test it?

Authors' response: We have made new analyses as requested, making changes to Table 1-3 and subsequently throughout the manuscript, as the new analyses changed the basal study population

due to missing values in several single variables. See Figure 1.
We did not find any significant interactions between occupation and the actual variables, and we have commented this in the results chapter.

Figure 1 would be improved if the authors included an additional box after "study population" to include the number of participants who died, the number lost to follow-up (e.g. due to emigration), and the number who were censored.

Authors' response: The figure has been improved as requested by the reviewer.

Several study limitations are not mentioned. Depression symptoms are only assessed for the previous week, which is a very short temporal interval for symptom assessment. For this reason other validated and reliable depression instruments often use much longer temporal intervals (e.g. past month, past six months, past year). This affects categorization of who is vs. isn't depressed, despite the authors' claim in the Strengths and Limitations bullet points that the study offers "a low risk of misclassification".

Authors' response: We have addressed these comments in the text. The reviewer is right that depression symptoms only are assessed for the previous week, which is a very short temporal interval for symptom assessment, and that this might affect categorization of who is vs. is not depressed. However, we do not claim that we examine the connection between depression as a disease and mortality, but between the occurrence of symptoms of depression and mortality. The relatively strong effects of this measure on mortality, however, show that the variable measures very important health aspects.

Another study limitation not currently mentioned is the use of self-reported illness data, rather than objective biomarkers or clinical measures of illness. This limitation could, in principle affect other self-reported measures used here; for example if you self-report a long-standing illness then you might report "nearly always" or "quite often" being worn out from work even if work is not physically demanding. Have the authors tried to verify the accuracy of these and other self-reports using available data? Indeed, the use of self-reported illness has weaknesses over objective, clinical measures of illness.

Authors' response: We have acknowledged this weakness in our manuscript now. We have not tried to verify the mentioned data ourselves, but several methodological studies have been validating many HUNT variables.

Also, the authors state that a 70.2% response rate is "high" (lines 4-5, pg. 19) when, in fact, many social scientists would disagree with this claim. This limitation should be acknowledged too.

Authors' response: We have rephrased this part of the text. Although, a response rate of 70.2% is fairly high compared to similar conducted studies.

In the Discussion the authors state "Although our results cannot be directly generalised to all countries, they are probably valid for countries with a similar socioeconomic environment and agricultural evolution." Please provide examples.

Authors' response: We have rephrased this section and provided a reference to this statement.

Lastly, there are several redundancies throughout the text that should be eliminated. For example, in the Abstract the Objectives are: "To explore all-cause mortality and the association between

symptoms of depression and all-cause mortality in farmers, and to evaluate this in a socioeconomic setting using a prospective cohort design." Delete "and to evaluate this in a socioeconomic setting". Also, in the Strength and Limitations bullet #3 the word depression appears twice consecutively.

Authors' response: The text has been corrected.

Reviewer #2: Cédric Lemogne Authors' response
The lack of analyses adjusted for behavioral risk factors; in my opinion, such analyses could provide meaningful results.

Authors' response: We have provided these analyses, and made changes accordingly throughout the manuscript.

The lack of sensitivity analyses using a continuous measure of depression; in my opinion, dichotomization may result in both Type I and Type II error. Even if clinicians are more comfortable with ill vs healthy categorization, such approach may not adequately capture the association between depressive symptoms and mortality

Authors' response: We have provided these analyses as sensitivity analyses, and the results are available as supplementary material (Supplementary Table 2).

The interpretation of the main results; in my opinion, confidence interval including 1 may signal a high probability to observe these results under the assumption of a null hypothesis

Authors' response: When confidence intervals include 1, one should of course be careful about conclusions. But confidence intervals only show statistical uncertainty and are dependent on the size of the material. As scientists, we should decide whether we believe we observe real effects or not, despite statistical significance. If we see clear trends in the material or observe dose response effects, this often suggests real effects. This is how we have interpreted the estimates.

VERSION 3 - REVIEW

REVIEWER	Jonathan Stieglitz Institute for Advanced Study in Toulouse, France
REVIEW RETURNED	29-Mar-2016

GENERAL COMMENTS	The paper has again improved since the last submission. My prior comments have been incorporated to my satisfaction.