PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Does occupational distress raise the risk of alcohol use, binge-eating, ill health and sleep problems among medical doctors? A UK cross-sectional study
AUTHORS	Medisauskaite, Asta; Kamau, Caroline

VERSION 1 - REVIEW

REVIEWER	Ruth Riley University of Birmingham, UK
REVIEW RETURNED	01-Nov-2018

GENERAL COMMENTS	BMJ Open Occupational distress raises the risk of alcohol use, binge eating, sleep problems and ill health: cross-sectional study of UK medical doctors Reviewer's comments The authors are to be commended on reporting the findings of this very under-researched topic yet one of importance. General Sample size is very small n=417 compared with the total population of doctors (=over 100K). I think you may just need to state this as a limitation but one which does not necessarily detract from the importance of the results. State how you address selection bias in this study. I suggest making a clearer distinction between prescription drug use and illicit drug use throughout; sometimes you refer to drug use when you mean prescription drugs. Just be clear and consistent throughout. The paper requires a good proof-read Outcome measures Can you say, briefly, why GHQ was chosen as a measure of psychiatric morbidity compared to other measures e.g. PHQ, DASS-21 etc Can you also state how accurate self-report measures are for AUDIT C and drug use etc – are these q'aires anonymised or pseudo anonymised in which case, your pps could still be identified which may therefore have implications for fidelity in participants' self-report Introduction You provide evidence of US doctors taking BZDs and opiates but this may slightly skew the picture since there is currently an epidemic of opiate use in the US so maybe not representative of the global population of doctors. Justification for this study can also be warranted by the impact of reduced wellbeing among doctors on patient outcomes i.e. patient safety etc – see recent paper by Panagioti et al (2018) JAMA –

	particularly for doctors with burnout and those experiencing sleeping difficulties Results It would be helpful to include a table of participant demographics/characteristics, including a breakdown of medical speciality, yeas in training, workplace type i.e hospital, GP Discussion It would be useful to know how your findings on sleep problems compare to the general population. How do your findings compare with existing literature such as the National Training Survey findings which provides an accurate snapshot due to the high response rate. What are the implications of these findings for patient safety etc

REVIEWER	Michael Ireland University of Southern Queensland. Australia
REVIEW RETURNED	06-Nov-2018

GENERAL COMMENTS	The rationale for the study could be strengthened in the introduction by citing any of the methodological advancements the current study makes. The introduction could be strengthened by more detail provided on the existing evidence. For example, the opening sentence reads "Doctors experience higher levels of occupational distress than the general population [1]". More could be said about this, such the actual rates of distress observed (more reflections on effect sizes throughout would be useful) and the types of evidence that support this claim (more critical appraisal of the cited studies would be useful). Some justification needs to be made for each instance a validated instrument is broke up or only partly used. For example, none of the published evidence attesting to the reliability or validity of the Brief COPE questionnaire can be applied to the current use of three items. Explanation and justifications are needed as to the creation of binary outcomes. How was this done? What cut-pouts were used? How can the information loss be justified? Given a key aim "This study shows the prevalence of occupational distress and health problems such as ill health symptoms, and health-related problems" The contribution of this type of study will depend on the generalisability of the sample. More attention needs to be given to this in the recruitment and sample section of the methods. How representative is the sample? What guarantees this representativeness? When discussing the logistic regression results in the discussion the effect sizes must be interpreted directly and the interpretations must take these into account. The implications of the data collected is not fully realised so it is hard to be motivated to believe the study really makes substantive contribution to the literature. It is important for the authors to make as clear as possible what the current study demonstrates that is not currently known.

VERSION 1 – AUTHOR RESPONSE

Reviewer: 1 The authors are to be commended on reporting the findings of this very under-researched topic yet one of importance. General Sample size is very small n=417 compared with the total population of doctors (=over 100K). I think you may just need to state this as a limitation but one which does not necessarily detract from the importance of the results.	Thank you very much for your comment. We have acknowledged this in the limitation section: "Future research should explore whether our findings replicate in a larger sample because the current sample size is relatively small, in comparison to the whole population of UK medical doctors." See p.16, line 18
Reviewer: 1 State how you address selection bias in this study.	One way we dealt with selection bias was randomly selecting trusts, inviting medical Royal Colleges and members of the BMA research panel creating an opportunity for doctors from various regions, specialties and at different stages of their career to take part in this study: "To prevent selection bias based on specialty, we invited all Royal Colleges from which 9 agreed to distribute information about the research to their members. To prevent selection bias due to an NHS trust's reputation we randomly selected 25% NHS trusts. From the ones possible to reach, 9 distributed invitations to this study. The third source of data was the BMA's research panel. The majority of UK medical doctors are members of the BMA and any of them can join the BMA's research panel; therefore, this panel represented a broad spectrum of doctors. All medical doctors working in the UK were included in the baseline data analysis." See p.7, line 27 We also added a table with doctors in this study demographic characteristics and compared these characteristics with doctors on the List of Registered Medical Practice: "We compared demographic characteristics of doctors in this study to doctors on the List of Registered Medical Practitioners (LRPM)[39,40] and the comparison showed that the current sample largely mirrors the demographics of UK medical doctors in terms of the proportion of doctors by gender, age, grade and specialty, except that there was

	a higher representation of consultants and public health doctors than the proportions within the LRPM (Table 2)." See p.9, line 33 Table: See p.26
Reviewer: 1 I suggest making a clearer distinction between prescription drug use and illicit drug use throughout; sometimes you refer to drug use when you mean prescription drugs. Just be clear and consistent throughout.	We have clarified the type of drugs we are referring to throughout the piece: "The coping function of alcohol/drug use among doctors, as with the general population, is plausible because alcohol and many drugs have psychoactive properties, e.g. prescription drugs such as benzodiazepine and opiates; illicit drugs such as LSD. Occupational distress is known to predict alcohol misuse in the general population [8] but little is known about whether, for example, doctors with high levels of burnout are at greater risk of using alcohol or drugs (including prescription or legally purchased drugs), and whether other types of occupational distress (e.g. psychiatric morbidity, negative coping strategies) have similar effects." See p.5, line 5 "This study aims to assess whether doctors suffering from occupational distress have an increased risk of (i) using alcohol or drugs (illicit, non-illicit);" See p.6, line 27 "In terms of illicit/non-illicit drug use, 44% of doctors used some type of drugs but almost all were non-illicit drugs: 3% of doctors used prescription opioids, 2% used benzodiazepines, 5% used sleep medication, 5% smoked tobacco, 7% used herbal or homeopathic remedies and 35% used over-the-counter medicines." See p.10, line 16 "44% of doctors use some type of drugs, but mostly over-the-counter medications (35%). Prescription drug use was rare (3% use opioids and 2% use benzodiazepines), suggesting that the proportion of doctors getting drug treatment for anxiety is lower than the proportion of doctors with anxiety (14.7%).[41]" See p.13, line 20

	"The results also show that – for UK doctors – occupational distress as a syndrome has no significant effect on legal or illegal drug use." See p.13, line 27
Reviewer: 1 The paper requires a good proof-read	We have proof-read our manuscript.
Reviewer: 1 Outcome measures Can you say, briefly, why GHQ was chosen as a measure of psychiatric morbidity compared to other measures e.g. PHQ, DASS-21 etc	We have clarified our choice of the GHQ-12: "This measure of psychiatric morbidity was chosen instead of others (e.g. Patient Health Questionnaire) because GHQ-12 was previously extensively used to examine working populations (e.g.[3]), including doctors (e.g.[6])." See p.8, line 27
Reviewer: 1 Can you also state how accurate self-report measures are for AUDIT C and drug use etc – are these q'aires anonymised or pseudo anonymised in which case, your pps could still be identified which may therefore have implications for fidelity in participants' self-report	We have added the following information about self-reporting on sensitive issues in the limitations: "We are also mindful that some participants might not have been comfortable answering some sensitive questions (e.g. about illicit drug or alcohol use) but the risk of response bias was mitigated by allowing doctors to complete confidential self-reported questionnaires. This is a recognised methods of measuring health and health-related behaviours (e.g. alcohol intake [59])." See p.16, line 20
Reviewer: 1 Introduction You provide evidence of US doctors taking BZDs and opiates but this may slightly skew the picture since there is currently an epidemic of opiate use in the US so maybe not representative of the global population of doctors.	We were not able to find another citation on opioids or benzodiazepines use to illustrate the situation in the UK. However, based on our study results we acknowledge that the prevalence of these drugs use is very different among UK doctors. "44% of doctors use some type of drugs, but mostly over-the-counter medications (35%). Prescription drug use was rare (3% use opioids and 2% use benzodiazepines), suggesting that the proportion of doctors getting drug treatment for anxiety is lower than the proportion of doctors with anxiety (14.7%).[41]" See p.13, line 20
Reviewer: 1 Justification for this study can also be warranted by the impact of reduced wellbeing among doctors on patient outcomes i.e. patient safety etc – see recent	Thank you for this comment. We have re-written the introductory paragraph: "Distress suffered by doctors has significant consequences for patient care.

paper by Panagioti et al (2018) JAMA – particularly for doctors with burnout and those experiencing sleeping difficulties	A recent meta-analysis of 47 studies found that burned-out doctors were more likely to provide poor quality care because of reduced professionalism, and they were more likely to be associated with poor patient satisfaction and incidents that jeopardise patient safety.[1] There is a high prevalence of distress among doctors in the United Kingdom (UK), with a systematic review of 30 studies showing that 17% to 52% of doctors have psychiatric morbidity,[2] higher than the prevalence rate of 19% in the general population,[3] and 31% to 54% of doctors have a type of burnout called emotional exhaustion.[2] Little is known, however, about whether occupational distress raises the risk of health problems (e.g. insomnia, binge-drinking) that might compel doctors to be absent from work or take sick leave resulting in under-staffing and a risk to patient safety. ” See p.4, line 3
Reviewer: 1 Results It would be helpful to include a table of participant demographics/characteristics, including a breakdown of medical speciality, yeas in training, workplace type i.e hospital, GP	We have added such table and included information about doctors’ gender, age, year of experience in medicine, grade, workplace, specialty and working hours. See p.26
Reviewer: 1 Discussion It would be useful to know how your findings on sleep problems compare to the general population.	We have added this information: “However, fewer doctors have insomnia (11%) or are alcohol dependent (5%) compared to the general population whereby 37% have insomnia and 13.9-29.1% have alcohol use disorder.[44]” See p.13, line 26
Reviewer: 1 How do your findings compare with existing literature such as the National Training Survey findings which provides an accurate snapshot due to the high response rate.	We have compared our findings with the NTS and the BMA survey results: “The results from this study also revealed a higher prevalence of burnout than the National Trainee Survey (NTS) which reported that 23.8% of medical trainees have burnout [45] but this could be because the current study had a high representation of consultants (49%). The current study supports the recent BMA annual survey reporting that 61% of doctors feel that their stress levels have increased over the last year [46] by showing that a similar proportion of doctors (55.3%) have a type of burnout called emotional exhaustion.”

	See p.13, line 28
Reviewer: 1 What are the implications of these findings for patient safety etc	We have added more information on the implications for patient care, doctors' decision-making and judgment: "Occupational distress among doctors has a detrimental effect on the quality of care and patient safety.[1] This study has revealed that occupational distress also increases the risk of doctors suffering from health problems. The impact of occupational distress on ill health could increase levels of sickness-absence among doctors, thus reducing patient safety because of under-staffing. Likewise, the impact of occupational distress on substance use and sleep problems could mean that distress indirectly impairs doctors' fitness to practice, judgement or decision-making because of being intoxicated, hung-over or having disturbed sleep." See p.16, line 27
Reviewer: 2 The rationale for the study could be strengthened in the introduction by citing any of the methodological advancements the current study makes.	We have added to our introduction that the unique aspect of our study is the analysis of a broad spectrum of health consequences which previously were not explored: "This is the first study to examine such a broad spectrum of health consequences among doctors. The other innovation of the current study is that, whereas many previous studies have sampled US doctors [5] or individual specialties such as oncology [6] and surgery,[7] this study sheds new light on the impact of occupational distress on health problems among doctors in the UK." See p.4, line 19 The other strength of this study is an experimental design. However, we analysed just the baseline data for this paper and, therefore, mention this strength just in the strengths and weaknesses section: "An additional strength of this study is that it is part of a bigger study using "a gold standard" design (a randomised controlled trial) and it was piloted by consulting doctors to ensure that the types of occupational distress and health problems measured were relevant."

	See p.16, line 6
Reviewer: 2 The introduction could be strengthened by more detail provided on the existing evidence. For example, the opening sentence reads “Doctors experience higher levels of occupational distress than the general population [1]”. More could be said about this, such the actual rates of distress observed (more reflections on effect sizes throughout would be useful) and the types of evidence that support this claim (more critical appraisal of the cited studies would be useful).	We have re-written the introductory paragraph adding more information about the types of evidence and the actual rates of distress: “Distress suffered by doctors has significant consequences for patient care. A recent meta-analysis of 47 studies found that burned-out doctors were more likely to provide poor quality care because of reduced professionalism, and they were more likely to be associated with poor patient satisfaction and incidents that jeopardise patient safety.[1] There is a high prevalence of distress among doctors in the United Kingdom (UK), with a systematic review of 30 studies showing that 17% to 52% of doctors have psychiatric morbidity,[2] higher than the prevalence rate of 19% in the general population,[3] and 31% to 54% of doctors have a type of burnout called emotional exhaustion.[2] Little is known, however, about whether occupational distress raises the risk of health problems (e.g. insomnia, binge-drinking) that might compel doctors to be absent from work or take sick leave resulting in under-staffing and a risk to patient safety. ” See p.4, line 3
Reviewer: 2 Some justification needs to be made for each instance a validated instrument is broke up or only partly used. For example, none of the published evidence attesting to the reliability or validity of the Brief COPE questionnaire can be applied to the current use of three items.	Thank you for your comment. We have explained why we chose to break up the questionnaires: “Doctors completed a variety of questionnaires assessing health problems. In some cases, we analysed responses to individual items in addition to the average of whole questionnaires because this offered richer insights into the prevalence of certain types of substance misuse (e.g. binge-drinking), different symptoms of ill health (e.g. backache), different signs of binge-eating (e.g. uncontrollable eating) and sleep disturbances (e.g. trouble falling asleep). This also enabled a better understanding of what aspects of health problems are predicted by occupational distress.“ See p.8, line 2
Reviewer: 2 Explanation and justifications are needed as to the creation of binary outcomes. How was this done?	We have added this explanation to the statistical methods section:

What cut-pouts were used? How can the information loss be justified?	"Outcome variables were re-coded into binary variables in order to distinguish between the presence and absence of a health problem and to identify factors that raise the risk of the symptom being present. The cut-off points where possible were identified from the literature (see citations within Table 1). For sleep problems, we determined the cut-off points based on the guidance for the ISI.[32] For each item doctors reporting moderate or severe sleep problems were noted as having the symptom present, and reporting no/mild sleep problems were noted as having the symptom absent. Doctors' responses about using substances to cope with stress were dichotomised as either not doing so, or using substances a little bit, to a medium extent or a lot. Frequency of drinking was divided into drinking alcohol less than 2-3 times a week and more than 2-3 times a week. Ill health symptoms were recoded such that the presence of each symptom meant experiencing once/twice per week or more often. See Table 1 for more details." See p.9, line 16 Table 1 also reports what cut-off points were used and cites, where possible, the source of such division. See p.22
Reviewer: 2 Given a key aim "This study shows the prevalence of occupational distress and health problems such as ill health symptoms, and health-related problems" The contribution of this type of study will depend on the generalisability of the sample. More attention needs to be given to this in the recruitment and sample section of the methods. How representative is the sample? What guarantees this representativeness?	We have expanded on the recruitment in the participants section: "Doctors took part in this study online (Qualtrics). They learnt about the study through medical Royal Colleges, NHS trusts, and the British Medical Association (BMA). To prevent selection bias based on specialty, we invited all Royal Colleges from which 9 agreed to distribute information about the research to their members. To prevent selection bias due to an NHS trust's reputation we randomly selected 25% NHS trusts. From the ones possible to reach, 9 distributed invitations to this study. The third source of data was the BMA's research panel. The majority of UK medical doctors are members of the BMA and any of them can join the BMA's research panel; therefore, this panel

	represented a broad spectrum of doctors. All medical doctors working in the UK were included in the baseline data analysis." See p.7, line 26 We also added a table with doctors in this study demographic characteristics and compared these characteristics with doctors on the List of Registered Medical Practice: "We compared demographic characteristics of doctors in this study to doctors on the List of Registered Medical Practitioners (LRPM)[39,40] and the comparison showed that the current sample largely mirrors the demographics of UK medical doctors in terms of the proportion of doctors by gender, age, grade and specialty, except that there was a higher representation of consultants and public health doctors than the proportions within the LRPM (Table 2)." See p.9, line 33. Table: See p. 26 We have acknowledged this in the limitation section: "It was not possible to calculate the response rate because it was not clear whether all NHS trusts and medical Royal Colleges who agreed to take part in this study actually distributed the invitation and to how many doctors. Future research should explore whether our findings replicate in a larger sample because the current sample size is relatively small, in comparison to the whole population of UK medical doctors." See p.16, line 16
Reviewer: 2 When discussing the logistic regression results in the discussion the effect sizes must be interpreted directly and the interpretations must take these into account.	We have highlighted variables with large effect sizes: "Doctors who cope with stress by using substances have a higher risk of alcohol dependence, binge-drinking, drinking larger amounts of alcohol, and a higher risk to a great extent of using alcohol more frequently." See p.13, line 13 "This means that the risk of sleep problems or insomnia exists even if

	doctors are suffering from just one of these types of occupational distress which is in particularly noticeable with the increase in psychiatric morbidity." See p.15, line 18 "These findings extend previous studies which investigate how the working conditions experienced by doctors relate to ill health [25] by showing that occupational distress, especially psychiatric morbidity, increases the risk of physical health problems." See p.15, line 30
Reviewer: 2 The implications of the data collected is not fully realised so it is hard to be motivated to believe the study really makes substantive contribution to the literature. It is important for the authors to make as clear as possible what the current study demonstrates that is not currently known.	We have explored the importance of this topic to patient care and as mentioned previously noted that in our introduction. We also added more information on the implications for patient care, doctors' decision-making and judgment: "Occupational distress among doctors has a detrimental effect on the quality of care and patient safety.[1] This study has revealed that occupational distress also increases the risk of doctors suffering from health problems. The impact of occupational distress on ill health could increase levels of sickness-absence among doctors, thus reducing patient safety because of under-staffing. Likewise, the impact of occupational distress on substance use and sleep problems could mean that distress indirectly impairs doctors' fitness to practice, judgement or decision-making because of being intoxicated, hung-over or having disturbed sleep." See p.16, line 27

VERSION 2 – REVIEW

REVIEWER	Ruth Riley University of Birmingham
REVIEW RETURNED	03-Dec-2018

GENERAL COMMENTS	I am satisfied that you have addressed my comments.

REVIEWER	Michael Ireland USQ Australia
REVIEW RETURNED	20-Dec-2018

GENERAL COMMENTS	Thank you for your hard work responding to my queries. Your responses have spurred some additional thoughts. Please see these under the headings “My Response:” below. My Initial Query: The rationale for the study could be strengthened in the introduction by citing any of the methodological advancements the current study makes. Author Response: We have added to our introduction that the unique aspect of our study is the analysis of a broad spectrum of health consequences which previously were not explored: “This is the first study to examine such a broad spectrum of health consequences among doctors. The other innovation of the current study is that, whereas many previous studies have sampled US doctors [5] or individual specialties such as oncology [6] and surgery,[7] this study sheds new light on the impact of occupational distress on health problems among doctors in the UK.” See p.4, line 19 The other strength of this study is an experimental design. However, we analysed just the baseline data for this paper and, therefore, mention this strength just in the strengths and weaknesses section: “An additional strength of this study is that it is part of a bigger study using “a gold standard” design (a randomised controlled trial) and it was piloted by consulting doctors to ensure that the types of occupational distress and health problems measured were relevant.” See p.16, line 6 My Response: It is good to see the introduction strengthened. However, I suggest removing reference to the larger study as a strength of the current study, since it has no bearing on it. That it is part of an experimental design does not in any way strengthen the methodology used for this study. My Initial Query: The introduction could be strengthened by more detail provided on the existing evidence. For example, the opening sentence reads “Doctors experience higher levels of occupational distress than the general population [1]”. More could be said about this, such the actual rates of distress observed (more reflections on effect sizes throughout would be useful) and the types of evidence that support this claim (more critical appraisal of the cited studies would be useful). Author Response: We have re-written the introductory paragraph adding more information about the types of evidence and the actual rates of distress: “Distress suffered by doctors has significant consequences for patient care. A recent meta-analysis of 47 studies found that burned-out doctors were more likely to provide poor quality care because of reduced professionalism, and they were more likely to be associated with poor patient satisfaction and incidents that jeopardise patient safety.[1] There is a high prevalence of distress among doctors in the United Kingdom (UK), with a systematic review of 30 studies showing that 17% to 52% of doctors have psychiatric morbidity,[2] higher than the prevalence rate of 19% in the general population,[3] and 31% to 54% of doctors have a type of burnout called emotional exhaustion.[2] Little is known, however, about whether occupational distress raises the risk of health problems (e.g. insomnia, binge-drinking) that might compel doctors to be absent from work or take sick

	leave resulting in under-staffing and a risk to patient safety. ” See p.4, line 3 My Response: This has been addressed well. My Initial Query Some justification needs to be made for each instance a validated instrument is broke up or only partly used. For example, none of the published evidence attesting to the reliability or validity of the Brief COPE questionnaire can be applied to the current use of three items. Author Response Thank you for your comment. We have explained why we chose to break up the questionnaires: “Doctors completed a variety of questionnaires assessing health problems. In some cases, we analysed responses to individual items in addition to the average of whole questionnaires because this offered richer insights into the prevalence of certain types of substance misuse (e.g. binge-drinking), different symptoms of ill health (e.g. backache), different signs of binge-eating (e.g. uncontrollable eating) and sleep disturbances (e.g. trouble falling asleep). This also enabled a better understanding of what aspects of health problems are predicted by occupational distress.“ See p.8, line 2 My Response: The table is very informative. Perhaps it should be mentioned in the limitations section that by editing previously validated instruments their validity may have been compromised and the extent that they are valid and reliable in their current form is unclear, My Initial Query Explanation and justifications are needed as to the creation of binary outcomes. How was this done? What cut-pouts were used? How can the information loss be justified? Author Response We have added this explanation to the statistical methods section: “Outcome variables were re-coded into binary variables in order to distinguish between the presence and absence of a health problem and to identify factors that raise the risk of the symptom being present. The cut-off points where possible were identified from the literature (see citations within Table 1). For sleep problems, we determined the cut-off points based on the guidance for the ISI.[32] For each item doctors reporting moderate or severe sleep problems were noted as having the symptom present, and reporting no/mild sleep problems were noted as having the symptom absent. Doctors’ responses about using substances to cope with stress were dichotomised as either not doing so, or using substances a little bit, to a medium extent or a lot. Frequency of drinking was divided into drinking alcohol less than 2-3 times a week and more than 2-3 times a week. Ill health symptoms were recoded such that the presence of each symptom meant experiencing once/twice per week or more often. See Table 1 for more details.” See p.9, line 16 Table 1 also reports what cut-off points were used and cites, where possible, the source of such division. See p.22 My Response: Thank you for this information.

	My Initial Query Given a key aim “This study shows the prevalence of occupational distress and health problems such as ill health symptoms, and health-related problems” The contribution of this type of study will depend on the generalisability of the sample. More attention needs to be given to this in the recruitment and sample section of the methods. How representative is the sample? What guarantees this representativeness? Author Response We have expanded on the recruitment in the participants section: “Doctors took part in this study online (Qualtrics). They learnt about the study through medical Royal Colleges, NHS trusts, and the British Medical Association (BMA). To prevent selection bias based on specialty, we invited all Royal Colleges from which 9 agreed to distribute information about the research to their members. To prevent selection bias due to an NHS trust’s reputation we randomly selected 25% NHS trusts. From the ones possible to reach, 9 distributed invitations to this study. The third source of data was the BMA’s research panel. The majority of UK medical doctors are members of the BMA and any of them can join the BMA’s research panel; therefore, this panel represented a broad spectrum of doctors. All medical doctors working in the UK were included in the baseline data analysis.” See p.7, line 26 We also added a table with doctors in this study demographic characteristics and compared these characteristics with doctors on the List of Registered Medical Practice: “We compared demographic characteristics of doctors in this study to doctors on the List of Registered Medical Practitioners (LRPM)[39,40] and the comparison showed that the current sample largely mirrors the demographics of UK medical doctors in terms of the proportion of doctors by gender, age, grade and specialty, except that there was a higher representation of consultants and public health doctors than the proportions within the LRPM (Table 2).” See p.9, line 33. Table: See p. 26 We have acknowledged this in the limitation section: “It was not possible to calculate the response rate because it was not clear whether all NHS trusts and medical Royal Colleges who agreed to take part in this study actually distributed the invitation and to how many doctors. Future research should explore whether our findings replicate in a larger sample because the current sample size is relatively small, in comparison to the whole population of UK medical doctors.” See p.16, line 16 My Response: It is important to keep in mind that among modest samples (those that do not come close to the population size) actual sample size has nothing to do with representativeness. That is, increasing sample size via potentially biased (e.g., convenience) sampling does not improve representativeness so I would remove recommendation for larger samples as a means to address issues with representativeness. My Initial Query When discussing the logistic regression results in the discussion the effect sizes must be interpreted directly and the interpretations must take these into account. Author Response We have highlighted variables with large effect sizes: “Doctors who cope with stress by using substances have a higher risk of alcohol dependence, binge-drinking, drinking larger amounts of alcohol, and a higher risk to a great extent of using alcohol more frequently.” See p.13, line 13 “This means that the risk of sleep

	problems or insomnia exists even if doctors are suffering from just one of these types of occupational distress which is in particularly noticeable with the increase in psychiatric morbidity." See p.15, line 18 "These findings extend previous studies which investigate how the working conditions experienced by doctors relate to ill health [25] by showing that occupational distress, especially psychiatric morbidity, increases the risk of physical health problems." See p.15, line 30 My Response: This is a good start but I meant more inclusion of effect size information in the general statements about the effects. Take the following statement as an example "…also raise the odds of doctors suffering from 30 frequent ill health" it begs the question 'raise it by what degree?' is this a small, medium, or large effect? It is important to keep in mind that statistical significance is not informative as to the importance or strength of an effect and trivial effects that are statistically significant remain meaningless. My Initial Query The implications of the data collected is not fully realised so it is hard to be motivated to believe the study really makes substantive contribution to the literature. It is important for the authors to make as clear as possible what the current study demonstrates that is not currently known. Author Response We have explored the importance of this topic to patient care and as mentioned previously noted that in our introduction. We also added more information on the implications for patient care, doctors' decision-making and judgment: "Occupational distress among doctors has a detrimental effect on the quality of care and patient safety.[1] This study has revealed that occupational distress also increases the risk of doctors suffering from health problems. The impact of occupational distress on ill health could increase levels of sickness-absence among doctors, thus reducing patient safety because of under-staffing. Likewise, the impact of occupational distress on substance use and sleep problems could mean that distress indirectly impairs doctors' fitness to practice, judgement or decision-making because of being intoxicated, hung-over or having disturbed sleep." See p.16, line 27 My Response: Good work.

VERSION 2 – AUTHOR RESPONSE

Comment	Response
Reviewer: 1 I am satisfied that you have addressed my comments.	We thank the reviewer for their contributions.
Reviewer: 2 It is good to see the introduction strengthened. However, I suggest removing reference to the larger study as a strength of the current study, since it has no bearing on it. That it is part of an experimental design does not in any way strengthen the methodology used for this study.	Thank you for your comment. We have removed the information about the larger study. We have deleted this sentence from the strengths and weaknesses section: "An additional strength of this study is that it is part of a bigger study using "a gold standard" design (a randomised controlled trial) and it was piloted by consulting doctors to ensure that the

	types of occupational distress and health problems measured were relevant." Please see page 16 line 9 within the document showing track changes.
Reviewer: 2 The table is very informative. Perhaps it should be mentioned in the limitations section that by editing previously validated instruments their validity may have been compromised and the extent that they are valid and reliable in their current form is unclear,	We have added an acknowledgement about this within the strength and weaknesses section: "We also recognise that using individual items from previously validated questionnaires (rather than all items) might have an impact on the validity and reliability of the measurement methods." Please see page 16, line 30
Reviewer: 2 It is important to keep in mind that among modest samples (those that do not come close to the population size) actual sample size has nothing to do with representativeness. That is, increasing sample size via potentially biased (e.g., convenience) sampling does not improve representativeness so I would remove recommendation for larger samples as a means to address issues with representativeness.	Thank you for your comment. We have deleted our comment about small sample size: "Future research should explore whether our findings replicate in a larger sample because the current sample size is relatively small, in comparison to the whole population of UK medical doctors." Please see p.16, line 24
Reviewer: 2 This is a good start but I meant more inclusion of effect size information in the general statements about the effects. Take the following statement as an example "…also raise the odds of doctors suffering from frequent ill health" it begs the question 'raise it by what degree?' is this a small, medium, or large effect? It is important to keep in mind that statistical significance is not informative as to the importance or strength of an effect and trivial effects that are statistically significant remain meaningless.	We have acknowledged the reviewer's point by clarifying the odds ratios within the discussion (please see page 14 to 16), allowing readers to see the extent of the odds that we are discussing without having to refer back to the tables/results. Odds ratios are considered to be effect sizes but range from 0 to infinity, unlike Cohen's d which ranges from -1 to 1. We agree that it is useful to refer to the size of the effect but we feel that interpreting an odds ratio as small, medium or large will need further research, e.g. a meta-analysis or experiments. Although it is possible to convert odds ratios into Cohen's d using a formula, for the conversion to be accurate we would need to assume that the variables are continuous and similarly distributed in doctors with/without occupational distress (or other predictors). We felt that converting the odds ratios to Cohen's d would not have been meaningful in the case of dichotomous variables (e.g drug use) or in cases where the variables did not have a similar or normal distribution when comparing doctors with/without occupational distress. We have addressed the reviewer's point about effect sizes by adding an acknowledgement about this within the strengths and limitations section: "We need future experimental research to test causation between occupational distress and

	health problems using a longitudinal design, and an evaluation of effect sizes using indicators such as Cohen's d. It was not possible to reliably convert odds ratios into Cohen's d within the current study because of the limitations of interpreting Cohen's d from data with dichotomous outcome variables." Please see page 16 line 13.