

## 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

<b>TITLE (PROVISIONAL)</b>	What's up doc? A national cross-sectional study of psychological wellbeing of hospital doctors in Ireland
<b>AUTHORS</b>	Hayes, Blanaid; Prihodova, Lucia; Walsh, Gillian; Doyle, Frank; Doherty, Sally

### VERSION 1 – REVIEW

<b>REVIEWER</b>	Max Henderson Leeds & York Partnership NHS Foundation Trust Leeds UK
<b>REVIEW RETURNED</b>	27-Jun-2017

<b>GENERAL COMMENTS</b>	<p>well designed and well presented cross-sectional survey. The relatively high response rate is a strength. Doctor's health is an area of increasing importance and this study adds to the growing body of literature showing high levels of psychological distress in the medical profession</p> <p>the abstract is clear and is representative of the paper. i think it was a study of a randomised sample rather than a randomised population</p> <p>The background is helpful and appropriate in length. The relevant literature is reviewed, along with its many limitations.</p> <p>The methodology is appropriate. But i have some questions which i anticipate will be easily addressed. The sample size was based on a pre-test hypothesis of 20% prevalence of what? How was it ascertained that subjects worked "almost exclusively" in hospitals clinics or institutions? Are the authors clear that their methodology would not systematically exclude certain groups? Do all doctors have emails? Are all doctors likely to have up to date postal addresses? Possibly not junior doctors? Would this methodology have captured doctors on long term sick leave? This may be a small point but it would be good to be clear that it had been considered.</p> <p>The tools chosen are appropriated, although i am never clear what measuring 'stress' adds when depression and anxiety are already assessed. How were participants asked about hours worked? I am not sure i could just pick a number....</p>
-------------------------	---

	<p><b>Results</b></p> <p>There was a high response rate. 85% were of Irish nationality. I suspect a similar UK study would find a lower percentage of UK nationals.</p> <p>Given that there were clear grade differences in a number of the study's findings its would be helpful to know the response rate by grade.</p> <p>Univariable analyses are described, showing consistent differences between consultants and training grades. Both groups have problems, but more trainees are distressed or depressed. The authors describe how these findings are adjusted for demographic variables. Given that the consultants worked less hours it would be helpful if the findings of the psychological health questions were also adjusted for hours worked. If the consultants are consistently better off then it is worth trying to find out why. Age is a reasonable question, but so is hours. If neither of these then the answer is likely to be found in more subtle factors related to the psychosocial work environment such as autonomy, role clarity, decision latitude etc.</p> <p><b>Discussion</b></p> <p>There is a fair summary of the paper The strengths and limitations are well set out</p> <p>The Implications section is my major gripe with this paper. The other sections are considered and set squarely within the evidence available. In the implications sections the authors step away from this. They suggest primary preventive strategies, though don't call them this. They then call for the work environment to be improved, although no evidence has been produced as to the role played by the work environment in the findings described, and if so which element. This seems to sit ill with the rest of the paper. In the same paragraph the authors leap quickly without a clear link to talk about treatment. Treatment is of course relevant but who? how? by whom? Then the authors return to primary prevention strategies in medical school regarding self stigma. In the final paragraph the authors rightly contrast the ease with which individual level responses are offered compared to organisational level responses. They state that these individual level approaches ("counselling") can be "very effective" but cite no evidence. I am not sure there is any. Overall I think this section needs cutting back sharply and to be based much more firmly in the evidence both from this study and the wider literature. This would improve the overall quality of the paper</p>
--	--

<b>REVIEWER</b>	Stephen Stansfeld Queen Mary University of London, United Kingdom
<b>REVIEW RETURNED</b>	30-Jun-2017

<b>GENERAL COMMENTS</b>	<p>This is an important study documenting the prevalence of psychological ill-health in Irish hospital doctors by employment grade.</p> <p>Page 2, line 24: 'psychological distress' is this measured by the GHQ? It probably should be mentioned in this sentence.</p> <p>Page 5, line 44: Did you have any specific hypotheses about the differences in prevalence of disorders by employment grade?</p> <p>Page 6, line 6: Do you need to mention that this covered the whole of Eire?</p> <p>Page 6, Data Collection: Some more justification of why you chose all these psychological measures would be helpful.</p>
-------------------------	---

	<p>Page 9, line 30: You adjusted for gender as a confounder but did you look at interactions by gender or even proportions of probable cases by gender across employment grades?</p> <p>Page 10, table 4: The table needs a more informative heading – e.g. ‘prevalence of wellbeing by employment grade’ Also are the prevalence values adjusted for confounding factors or unadjusted?</p> <p>Page 11, Table 5 This table also needs a more informative heading – e.g. ‘Odds ratios and confidence intervals for wellbeing...’ This is not a table of prevalence. It is not clear what the reference category is here – is this training employment grades odds for psychological distress relative to consultants? The odds ratios look enormous – are they correct? Also are these adjusted for confounding factors?</p> <p>Page 12, line 42: ‘largely representative’ is probably fairer. 55% is a good response rate for this type of survey nowadays but it still means 45% did not respond.</p> <p>Page 13, line 12: Survey results prior to the economic collapse are not a fair comparison.</p> <p>Page 13, line 33: This comment on generation Y is verging on victim blaming. It also rather undermines what I believe is a valid argument you put forward that the working conditions and restrictions may be responsible for these high rates of psychiatric disorders. I suggest you leave this comment out on generation Y out.</p> <p>Page 14, line 9: Which ‘population controls’ are being referred to here?</p>
--	---

### VERSION 1 – AUTHOR RESPONSE

Reviewer: 1

Reviewer Name: Max Henderson

Institution and Country: Leeds & York Partnership NHS Foundation Trust, Leeds, UK Please state any competing interests or state ‘None declared’: No competing interests

Comment 1:

well designed and well- presented cross-sectional survey. The relatively high response rate is a strength.

Doctor's health is an area of increasing importance and this study adds to the growing body of literature showing high levels of psychological distress in the medical profession the abstract is clear and is representative of the paper. I think it was a study of a randomised sample rather than a randomised population

Response:

Thank you for your comment on the importance of the study. We agree with your comment on randomisation and have amended the text accordingly.

We changed the text as follows: Design: National cross-sectional study of randomised sample of hospital doctors.

The background is helpful and appropriate in length. The relevant literature is reviewed, along with its many limitations. The methodology is appropriate. But I have some questions which I anticipate will be easily addressed.

Comment 2:

The sample size was based on a pre-test hypothesis of 20% prevalence of what?

Response:

we agree that this requires clarification

We changed the text as follows: The sample size was calculated for a 95% confidence interval, an acceptable margin of error of +/- 5% and an accepted prevalence of psychological distress of 20%.

Comment 3:

How was it ascertained that subjects worked “almost exclusively” in hospitals clinics or institutions?

Response:

all doctors working within the publicly funded health sector in Ireland are employed by a single national healthcare provider - the Health Service Executive. Physicians, surgeons, emergency medicine doctors, anaesthetists, paediatricians, obstetricians and pathologists work exclusively in hospitals. Psychiatrists are attached either to hospitals or institutions (e.g. residential) or attend at clinics in the community, which are separate from primary care centres. Ophthalmic surgeons are attached exclusively to hospitals while ophthalmic physicians work in a variety of locations (personal communication). Some work in hospitals, some as specialists within primary care centres and some in community health clinics alongside services provided by public health nurses. All of these are publicly funded community clinics. The process of delivery of community ophthalmic care is evolving and is the subject of a recently published report.

We changed the text as follows: see 2nd paragraph of section entitled ‘Sample’

In order to be invited to participate in this study, the participants had to work mainly in hospitals, publicly funded clinics or residential institutions.

Comment 4:

Are the authors clear that their methodology would not systematically exclude certain groups?

Response:

We took great care to avoid systematic exclusions apart from those outlined in the ‘limitations’ section. We excluded those working in private practice unless this was combined with working in the public sector as their work environment is very different. It is possible that a very small number of consultants on permanent contracts are un-affiliated to their post-graduate body. We were reliant on the post-graduate training bodies to access our target population, as data protection legislation precluded us from directly accessing their databases. It transpired that initial estimates greatly overstated the numbers as post-graduate training bodies are not necessarily kept up to date with their members’ status (e.g. retirement, sick leave, maternity leave). However, when respondents indicated that they were on leave, or that they had retired, we excluded their responses and removed them from the denominator

We changed the text as follows: See last sentence of ‘Sample’

Those working exclusively in private practice and those who were retired or on sick leave / maternity leave at the time of the survey were excluded. The denominator was adjusted accordingly (see Supplementary File 4).

Comment 5:

Do all doctors have emails? Are all doctors likely to have up to date postal addresses? Possibly not junior doctors?

Response:

in order to overcome this potential source of bias, we used both postal and electronic methods to reach our respondents for both the first and third wave. Where postal questionnaires were ‘returned to sender’, we followed up with the relevant training body, which was generally able to account for this either by providing a more up-to-date address or by indicating that the trainee was ‘out of

programme'. Those who were out of programme or otherwise unaccounted for (e.g. repeated 'return to sender') were removed from the denominator (see Supplementary File 4).

Comment 6:

Would this methodology have captured doctors on long term sick leave? This may be a small point but it would be good to be clear that it had been considered.

Response: we excluded doctors on long term sick leave or maternity leave and those who were on secondment (e.g. abroad). This information was disclosed by respondents at the part of the questionnaire which dealt with work hours over the previous month. In some instances, further elaboration was given in the free-text section of the questionnaire.

We changed the text as follows: See comment R1-4

Comment 7:

The tools chosen are appropriate, although I am never clear what measuring 'stress' adds when depression and anxiety are already assessed.

Response: we chose the DASS partly because it measures the 3 separate entities of depression, anxiety and stress. For further explanation of rationale, see R2-5.

Comment 8:

How were participants asked about hours worked? I am not sure I could just pick a number....

Response:

participants were asked how many hours they had worked in two consecutive working weeks over the previous month

We changed the text as follows: See second paragraph of Data Collection

Participants provided data on demographics (age, sex, nationality, employment stage/ grade), specialty and workload as measured by the question 'how many hours per week did you work over two consecutive working weeks in the past month'.

Results

Comment 9

There was a high response rate. 85% were of Irish nationality. I suspect a similar UK study would find a lower percentage of UK nationals.

Response: According to the Medical Council's Medical Workforce Intelligence Report (new reference 42), in 2014, 68.7% of hospital consultants and 68% of trainees were Irish university graduates. While country of graduation is not the same as nationality, it serves as a proxy in the absence of a more specific comparison. The fact that 85% of respondents were Irish suggests that those of Irish nationality were more likely to respond to the questionnaire.

We changed the text as follows: Respondents held predominantly Irish nationality (85%) and though there was no sex preponderance overall, consultants were predominantly male (61%) and trainees predominantly female (Table 2). According to a workforce intelligence report on the healthcare workforce in 2014 (42), 69% of trainee and consultant doctors were Irish graduates. While nationality is not synonymous with country of graduation, this suggests that respondents were more likely to be Irish.

Response:

We also inserted an amendment as follows to the Limitations section (2nd paragraph):

Another limitation is that the percentage of respondents who were Irish nationals was higher than the number of Irish graduates working in hospitals in a contemporaneous report.

Comment 10:

Given that there were clear grade differences in a number of the study's findings it would be helpful to know the response rate by grade.

Response: We have now included this in the manuscript.

Response:

We changed the text as follows: 1749 doctors participated [overall response rate = 55%, consultants = 60%, trainees = 51%, (range 33-63% between specialties)].

Comment 11:

Univariable analyses are described, showing consistent differences between consultants and training grades. Both groups have problems, but more trainees are distressed or depressed. The authors describe how these findings are adjusted for demographic variables. Given that the consultants worked less hours it would be helpful if the findings of the psychological health questions were also adjusted for hours worked. If the consultants are consistently better off than it is worth trying to find out why. Age is a reasonable question, but so is hours. If neither of these then the answer is likely to be found in more subtle factors related to the psychosocial work environment such as autonomy, role clarity, decision latitude etc.

Response:

Thank you for your comment, indeed – work hours could be considered an indicator of work conditions and contribute to distress. As outlined in Table 3, trainees reported significantly higher mean work hours than consultants. We have repeated the analyses and controlled for mean work hours as suggested. Consequently, Table 4 and the entire result section was updated to reflect the effect of age on the dependent variables.

Furthermore, a few extra sentences were added to the text to reflect the finding that work hours were associated with measures of distress

We changed the text as follows: See penultimate sentence of 1. Statistical analyses, 2. Self-rated health, 3. Subjective wellbeing, 4. Psychological distress, 5, 6, 7. Each element of Mental health 1. General linear models (GLM) were used to analyse the differences between employment groups adjusting for demographic and work variables (age, sex, marital status and mean hours worked (MHW)).

2. In addition, lower MHW was significantly associated with higher SRH ( $B = -0.01$ ,  $p \leq .001$ ).

3. In addition, lower MHW were significantly associated with higher subjective wellbeing ( $B = -.23$ ,  $p \leq .001$ )

4. In addition, higher MHW were significantly associated with higher psychological distress ( $B = .07$ ,  $p \leq .001$ ).

5. In addition, higher MHW were significantly associated with higher levels of depression ( $B = .07$ ,  $p \leq .001$ ).

6. In addition, higher MHW were significantly associated with higher levels of anxiety ( $B = .05$ ,  $p \leq .001$ ).

7. ....though males were less likely to have high scores for stress than females ( $B = -1.07$ ,  $p \leq .05$ ) and higher MHW were significantly associated with higher stress scores ( $B = 1.0$ ,  $p \leq .001$ ).

Comment 12:

Discussion

There is a fair summary of the paper

Response:

The strengths and limitations are well set out

Comment 13:

The Implications section is my major gripe with this paper. The other sections are considered and set squarely within the evidence available. In the implications sections the authors step away from this. They suggest primary preventive strategies, though don't call them this. They then call for the work environment to be improved, although no evidence has been produced as to the role played by the work environment in the findings described, and if so which element. This seems to sit ill with the rest of the paper. In the same paragraph the authors leap quickly without a clear link to talk about treatment. Treatment is of course relevant but who? how? by whom? Then the authors return to primary prevention strategies in medical school regarding self stigma. In the final paragraph the authors rightly contrast the ease with which individual level responses are offered compared to organisational level responses. They state that these individual level approaches ("counselling") can be "very effective" but cite no evidence. I am not sure there is any. Overall I think this section needs cutting back sharply and to be based much more firmly in the evidence both from this study and the wider literature. This would improve the overall quality of the paper

Response:

We acknowledge that this section is a bit muddled and have now made substantial changes to it. The first paragraph calls for raising awareness of mental health issues and available supports across all levels (students, trainees and senior doctors) and incorporates the need for addressing self-stigma which was observed at all levels in this study.

The second paragraph points to a need for clear pathways to care which are evident both to those in need and to those who manage or work with them. The third paragraph comments on specific working conditions (longer hours) which may partly explain the grade differential in measures of wellbeing / distress, cautioning that the implementation of EWTD has had unforeseen consequences for teams. This sets the scene for a call to employers to explore how systemic organisational problems may be addressed and to create organisations conducive to health.

We changed the text as follows:

Implications

This study paves the way for further work to be done in Ireland at the level of both inquiry and intervention.

In the first instance, medical schools, post-graduate training bodies and senior clinicians need to tackle self-stigmatising attitudes to mental ill health which were evident at all grades in this cohort, by embedding in training and professional development information and tools on how to maintain good mental health and on supports available.

For those in difficulty and those who manage them, there is a need for clear pathways and easy access to appropriate support and confidential care, such as own general practitioner, quality occupational health services and support in returning after illness to one's professional role.

Most importantly, the employer needs to prioritise the welfare of its staff, by addressing deep rooted systemic problems contributing to the challenging work environment, such as low staff numbers, long work hours, work organisation and poor people management.(44). As longer working hours were

found to contribute to poor personal wellbeing in this study, and were particularly evident in trainees, we encourage employers to continue working towards achieving compliance with EWTD while also monitoring the unintended consequences such as the break-up of teams and poor quality handover with its implications for patient care.(44)

There is a need for further research to identify strategies to improve physician wellness with particular emphasis on organisational responsibility to create an environment and culture conducive to health, efficiency and meaning in work.(58) An exploration of doctors' own views on pathways to mental health care would help to elucidate what might be favoured by potential users.

Reviewer: 2

Reviewer Name: Stephen Stansfeld

Institution and Country: Queen Mary University of London, United Kingdom Please state any competing interests or state 'None declared': None declared. Except that I was her thesis examiner.

Comment 1:

This is an important study documenting the prevalence of psychological ill-health in Irish hospital doctors by employment grade.

Response: Thank you.

Comment 2:

Page 2, line 24: 'psychological distress' is this measured by the GHQ? It probably should be mentioned in this sentence

Response:

We have now listed the scale names in the abstract. This is further elaboration under 'Data collection' where the instrument is described

Comment 3: Page 5, line 44: Did you have any specific hypotheses about the differences in prevalence of disorders by employment grade?

Response:

There was no specific hypothesis at the outset. However, due to the large variation in psychological distress in doctors in the international literature, we decided to adopt exploratory rather than hypothesis confirming design.

Comment 4: Page 6, line 6: Do you need to mention that this covered the whole of Eire?

Response:

this study was set in the Republic of Ireland, which is now clarified in the first sentence of the design section

We changed the text as follows: Page 5, Line 56

The study was a national cross sectional survey of hospital doctors working in the Republic of Ireland.

Comment 5:

Page 6, Data Collection: Some more justification of why you chose all these psychological measures would be helpful.

Response:

We used the GHQ 12 in order to compare our population with other studies of doctors internationally as this instrument has been widely used. Recognising its limitations however, we were keen to



explore other parameters of mental health and wellbeing. The DASS 21 is a relatively recent instrument and has not been used on doctors previously. However, it was attractive to us because of its facility to measure 3 separate states i.e. core symptoms of depression, anxiety and tension (stress) over previous week. The original confirmatory factor analysis aimed to discriminate between depression and anxiety but the third factor of stress was identified in the process. It is used in both population sampling and clinical research and we felt that it would be useful to pick up on lower levels of stress which would not give rise to the more serious anxiety or depression assessed by specific scales for these entities. The WHO-5 likewise is little used in doctors but we chose it because of its brevity and in order to ensure that we were not exclusively focused on negative states. Finally, the single item measure of self-rated health was chosen simply to enable comparison with national population health surveys.

We changed the text as follows: See 2nd paragraph (3rd sentence) of Data collection  
The General Health Questionnaire (GHQ 12) was included in order to allow for comparison with internationally reported rates of psychological distress. The 21 item Depression, Anxiety, Stress Scale (DASS 21), though not previously used in doctors, was attractive to us because of its facility to measure 3 separate states i.e. core symptoms of depression, anxiety and tension (stress). The WHO-5 likewise is little used in doctors but we chose it because of its brevity and in order to ensure that we were not exclusively focused on negative states. Internal consistency was satisfactory on all scales (Cronbach  $\alpha=0.80-0.93$ ). Two single question items on self-rated health and self-stigma were included and both have previously been used in surveys of population health.

Comment 6:

Page 9, line 30: You adjusted for gender as a confounder but did you look at interactions by gender or even proportions of probable cases by gender across employment grades?

Response:

Thank you for your comment, which prompted us to run additional analyses on this. While there was a degree of variation in probable caseness in the individual wellbeing indicators across grades and sex, these were not consistent and would warrant further elaboration, exceeding the word count of the manuscript.

Comment 7:

Page 10, table 4: The table needs a more informative heading – e.g. ‘prevalence of wellbeing by employment grade’

Response:

we agree

We changed the title of Table 4 as follows: Prevalence and non-adjusted (ANOVA, Chi-Square) and adjusted (GLM) comparisons of wellbeing scales by employment grade.

Comment -8:

Also are the prevalence values adjusted for confounding factors or unadjusted?

Response:

The reported prevalence values are adjusted for age, sex, marital status and mean weekly hours worked.

No change to text

Comment 9:

Page 11, Table 5 This table also needs a more informative heading – e.g. ‘Odds ratios and confidence intervals for wellbeing...’ This is not a table of prevalence.

Response:

We have removed this table as it was superfluous.

Comment -10:

It is not clear what the reference category is here – is this training employment grades odds for psychological distress relative to consultants? The odds ratios look enormous – are they correct? Also are these adjusted for confounding factors?

Response:

We have removed this table as it was superfluous.

Comment 11:

Page 12, line 42: 'largely representative' is probably fairer. 55% is a good response rate for this type of survey nowadays but it still means 45% did not respond.

Response:

we agree

We changed the text as follows: The results can be taken as largely representative as.....

Comment 13:

Page 13, line 12: Survey results prior to the economic collapse are not a fair comparison.

Response:

we agree with this cautionary observation. However, the study we referenced for 2007 has not been repeated.

We changed the text as follows: This survey used the same instrument, albeit that it was undertaken at a time prior to the country's economic collapse in 2008. Subsequent national surveys have elected to use alternative measures which are not directly comparable.

Comment 14:

Page 13, line 33: This comment on generation Y is verging on victim blaming. It also rather undermines what I believe is a valid argument you put forward that the working conditions and restrictions may be responsible for these high rates of psychiatric disorders. I suggest you leave this comment out on generation Y out.

Response:

agreed

We changed the text as follows: We deleted the sentence 'Alternatively, it may reflect the aforementioned generational difference between the 'baby boomers' occupying consultant posts and 'generation y' and 'millennials' who largely occupy the training grades' towards end of discussion as suggested.

Comment 15:

Page 14, line 9: Which 'population controls' are being referred to here?

Response:

this refers to the general population survey of 2017 (ref 29). We agree, the use of the term 'population controls' is inappropriate here.

We changed the text as follows: Considerably more doctors in this study in comparison to the general population (29) perceived stigma in relation to mental health and.....

Table 3. Prevalence and non-adjusted (ANOVA, Chi-Square) and adjusted (GLM) comparisons of mean weekly hours worked and self-stigma by employment grade (as measured by a single item). Consultants HST BST Total ANOVA (F) GLM (B) – BST1 GLM (B) – HST2

	Consultants	HST	BST	Total	ANOVA (F)	GLM (B) – BST1	GLM (B) – HST2
n							
mean %							
(SD) n							
mean %							
(SD) n							
mean %							
(SD) n							
Mean weekly hours worked							
	54.2	15.1	61.1	15.5	59.6	13.0	57.0 15.1 38.4***
Self-stigma							
Strongly disagree	20	2.1	11	2.6	10	2.7	41 2.4
Disagree	134	14.1	51	12.0	43	11.5	228 13.1
Neutral	166	17.5	61	14.4	56	15.0	283 16.2
Agree	438	46.3	177	41.7	173	46.4	788 45.2
Strongly agree	189	19.9	124	29.2	91	24.4	404 23.2
Mean	2.32	1.01	2.17	1.06	2.22	1.03	2.26 1.03 3.68* -.09ns -.04ns
ns = not significant; * = p ≤ .05; ** = p ≤ .01; *** = p ≤ .001							
Fp=ANOVA; 1GLM (B) = BST adjusted for sociodemographic variables & mean hours worked; 2GLM (B) = HST adjusted for sociodemographic variables & mean hours worked; 1 & 2 – Reference category: Consultant							

Table 4: Prevalence and non-adjusted (ANOVA, Chi-Square) and adjusted (GLM) comparisons of wellbeing scales by employment grade (GHQ-121, DASS-212, WHO-53, SRH4). Consultants HST BST Total χ21 ANOVA (F)2 GLM (B) – BST3 GLM (B) – HST4

	Consultants	HST	BST	Total	χ21 ANOVA (F)2	GLM (B) – BST3	GLM (B) – HST4
N/							
mean %/ SD N/							
mean %/ SD N/							
mean %/ SD N/							
mean %/ SD							
Self-rated health (SRH)							
Poor	3	0.3	9	2.1	13	3.5	25 1.4
Fair	102	10.8	53	12.5	59	15.8	214 12.2
Good	302	32	160	37.7	135	36.0	597 34.1
Very good	352	37.3	140	33.0	118	31.5	610 34.9
Excellent	185	19.6	62	14.6	49	13.1	296 16.9
Mean score	3.65	.925	3.46	.959	3.35	1.01	3.54 .960 15.5*** .45*** .25**
Subjective wellbeing (WHO-5)	66.38**						
Likely depression	169	17.8	117	27.6	102	27.2	388 22.2
Low mood	215	22.7	137	32.3	124	33.1	476 27.3

Normal 563 59.5 170 40.1 149 39.7 882 50.5  
 Mean score 53.4 21.3 44.7 19.8 44.4 20.5 49.3 21.2 39.1\*\*\* 8.29\*\*\* 1.92ns  
 Psychological distress (GHQ-12) 47.2\*\*\*  
 No evidence of mental ill health 379 40.8 109 26.1 91 24.7 579 33.8  
 Less than optimal mental health 269 29 149 35.7 122 33.1 540 31.5  
 Probable mental ill health 281 30.2 159 38.1 156 42.3 596 34.8  
 Mean score 2.6 3.3 3.4 3.4 3.7 3.6 3.0 3.4 16.5\*\*\* -2.05\*\* -.96\*  
 Depression (DASS-21) 51.96\*\*\*  
 Normal 761 80.5 288 67.9 254 68.3 1303 74.8  
 Mild 74 7.8 39 9.2 35 9.4 148 8.5  
 Moderate 67 7.1 60 14.2 38 10.2 165 9.5  
 Severe 20 2.1 23 5.4 29 7.8 72 4.1  
 Extremely severe 23 2.4 14 3.3 16 4.3 53 3.0  
 Mean score 5.2 7.1 7.7 8.0 8.1 8.7 6.4 7.8 27.3\*\*\* -2.5\*\* -.45ns  
 Anxiety (DASS-21) 100.41\*\*\*  
 Normal 828 89.0 310 74.5 250 67.0 1388 80.7  
 Mild 30 3.2 24 5.8 30 8.0 84 4.9  
 Moderate 44 4.7 46 11.1 52 13.9 142 8.3  
 Severe 11 1.2 17 4.1 14 3.8 42 2.4  
 Extremely severe 17 1.8 19 4.6 27 7.2 63 3.7  
 Mean score 2.6 4.7 5.0 6.3 6.4 6.9 4.0 5.9 67.2\*\*\* -3.13\*\*\* -1.09\*  
 Stress (DASS-21) 37.31\*\*\*  
 Normal 709 75.9 271 65.8 226 60.8 1206 70.2  
 Mild 76 8.1 53 12.9 55 14.8 184 10.7  
 Moderate 74 7.9 41 10.0 49 13.2 164 9.5  
 Severe 55 5.9 31 7.5 31 8.3 117 6.8  
 Extremely severe 20 2.1 16 3.9 11 3.0 47 2.7  
 Mean score 10.8 8.6 12.8 9.3 13.2 9.2 11.8 8.9 13.6\*\*\* -1.49ns -.41ns  
 1χ<sup>2</sup> = categorical group differences; 2F = ANOVA (continuous variables); 3GLM (B) = BST adjusted for sociodemographic variables & mean hours worked; 4GLM (B) = HST adjusted for sociodemographic variables & mean hours worked; 3 & 4 – Reference category: Consultant  
 ns = not significant; \* = p ≤ .05; \*\* = p ≤ .01; \*\*\* = p ≤ .001;  
 SRH = Self Rated Health (single item); WHO 5 = World Health Organisation Wellbeing scale; GHQ-12 = General Health Questionnaire (12 item); DASS 21 = Depression Anxiety Stress Scale (21 item)

### VERSION 2 – REVIEW

<b>REVIEWER</b>	Max Henderson Leeds & York Partnership NHS Foundation Trust UK
<b>REVIEW RETURNED</b>	29-Aug-2017

<b>GENERAL COMMENTS</b>	Thank you for asking me to review this revision. The authors have addressed all my concerns. I think the paper is substantially improved as a result of the amendments made in response to each reviewer
-------------------------	--

<b>REVIEWER</b>	Stephen Stansfeld Queen Mary University of London United Kingdom
<b>REVIEW RETURNED</b>	14-Aug-2017

**GENERAL COMMENTS**

The authors have satisfactorily addressed my comments.