

## 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>	Changes in job satisfaction among doctors in Norway from 2010 to 2017.A study based on repeated surveys
<b>AUTHORS</b>	Rosta, Judith; Aasland, Olaf; Nylenna, Magne

## VERSION 1 - REVIEW

<b>REVIEWER</b>	Tamara Ritsema George Washington University, Washington DC, USA
<b>REVIEW RETURNED</b>	29-Nov-2018

<b>GENERAL COMMENTS</b>	<p>This is an interesting topic and the approach of serial surveys is novel. Well done. While the article is very well written for a person who is a non-native English speaker, it could use a little bit of polishing from a friendly native English speaker. (Of course, your English is vastly superior to my Norwegian, so all respect to the team!)</p> <p>Page 2, line 9. I don't know what "doctors in different job categories" means. Is this specialty? Length of time in practice? Employer? Explain a bit more.</p> <p>Page 4, line 17. – I would add "on job satisfaction" after "health care reform" in this sentence. You want to link these changes in the environment to your outcome variable.</p> <p>Page 5 and 6 – the Methods section should be a narrative. Right now the dependent and independent variables sections are just lists. Maybe write a narrative and then use a table to display your variables? Also, as a person who has published using the JSS before, I would encourage you to make it more clear in the methods section that you asked the participants to rate each of the 10 items on a 7 point scale and then divided by 10 to get the mean. Previous literature has reported the total score on the JSS. I think it is completely fine to report the mean score, but just be clear that is how you did it for those not familiar with the JSS. Also, just note that the JSS has sometimes been administered using a 4 point Likert scale, in which case the maximum score for the JSS was 40.</p> <p>Page 7 Line 2 – I would report this information about interns with your inclusion and exclusion criteria (under participants)</p>
-------------------------	---

	<p>Page 7, line 10 – Says “units with missing data were excluded”. Does this mean that participants were excluded if they did not reply at each time point? Or does it mean that the individual time point was not included in the analysis if they didn’t respond to that particular survey? Or does it mean that if the participants chose not to answer one item of the JSS that this was excluded from the calculation of the mean? Be specific here.</p> <p>Page 7 , line 25 – I am confused which Table 1 the authors are discussing. There appear to be 2 Table 1s – one with specialty breakdown and one with response rates, etc. Please make sure that the table you want first is the one labeled “Table 1” (I think that is the second table in the PDF).</p> <p>Page 8, line 17; page 9, line 1; page 9, line 12 – I have a personal dislike of reporting “nearly significant” results. I will leave it to the editor of the journal (they may have a policy on this) as to whether to leave it in, but I feel like the assumptions of our statistical tests mean we should report significant or not significant. We are implying we know the direction of a potential association even though the testing did not reveal significance and I think this leads readers to assume unproven associations. The current replication crisis has begun to even undermine our faith in the almighty <math>p &lt; 0.05</math> – implying association in the setting of a higher <math>p</math> value than that is dangerous.</p> <p>Page 12, lines 8-10 – I would just remove. You have done a nice job of international comparisons with other health professionals. This just jumps out at the end as incongruous with the rest of your nice discussion.</p> <p>Where is your limitations section? I don’t see it.</p> <p>Table 1 (the second table in the PDF) – are these percentages representative of all doctors in Norway? For example, there are roughly 3 times as many hospital doctors as GPs in the 2014 administration. Are there 3 x as many hospital doctors as GPs in Norway? Are there 5 x as many GPs as doctors in academia in Norway? Did any particular group have a higher response rate than others?</p>
--	---

<b>REVIEWER</b>	Katja Goetz Institute of Family Medicine
<b>REVIEW RETURNED</b>	04-Dec-2018

<b>GENERAL COMMENTS</b>	Thank you for this manuscript. I have already read to other papers from this series and I miss the added value. I miss a methodological discussion about the use of the job satisfaction scale.
-------------------------	---

<b>REVIEWER</b>	Michael H. Freitag Department of Health Services Research Division of General Practice/Family Medicine University of Oldenburg, Germany
<b>REVIEW RETURNED</b>	29-Jan-2019

<b>GENERAL COMMENTS</b>	<p>Rosta and colleagues report a repeated survey of physician job satisfaction in Norway from 2010-2017. Job satisfaction is an important issue given the shortage of physicians in Europe and its association with patient outcomes.</p> <p>The paper is well written. However, the main point is whether the reported differences in job satisfaction over time are meaningful and real. The first sentence of the discussion section of the main article says "...JSS remained high" whereas in the abstract it says "...decreased significantly".</p> <p>It is intriguing that there was no change among those participants who participated in all four surveys over time. That implies that the (small) changes over time are driven by the differences of sample composition.</p> <p>Interestingly, the proportion of females increased from 37.4% in 2010 to 52.9% in 2016/2017 while the mean age decreased from 50.7 to 47.7 years. There is an international trend that the younger generation is less willing to work long hours and female physicians are more likely to consider time constraints by child care.</p> <p>Therefore, the authors should report their results stratified by age and gender. First, it would be interesting to see age- and gender-specific job satisfaction and second, this would clarify how much of the change over time can be attributed to the increasing proportion of females and younger physicians.</p> <p>The authors could add a sentence on how physician job satisfaction compares to other job categories such as lawyers or teachers and how job satisfaction in other professions changed over the same time period.</p> <p>Minor points: Page 2, line 23: Conclusions (spelling) Page 5, line 2: emphasis (spelling)</p>
-------------------------	---

<b>REVIEWER</b>	A/Prof Danny Hills Federation University Australia Australia
<b>REVIEW RETURNED</b>	31-Jan-2019

<b>GENERAL COMMENTS</b>	<p>Job satisfaction in the medical profession is an important issue for the profession, health services and the community. This manuscript attempts to document trends in medical practitioner job satisfaction in Norway, in light of recent changes in health politics and health care reforms.</p> <p>There are some key issues that need to be addressed in this manuscript before it could be considered ready for publication.</p>
-------------------------	--

	<p>Overall, the manuscript needs to flow better and with appropriate use of past tense throughout.</p> <ol style="list-style-type: none"><li>1. The introduction could flow better. A better approach might be to commence with why the construct of job satisfaction is important and then why it is important in medical practitioners. A more rounded explanation of the construct could follow, including how it has been measured more broadly and in medical practitioners more specifically. What have been the trends and existing evidence on job satisfaction in medical practitioners could follow, including with a little more detailed explanation of the reforms etc., finishing with the aim statement.</li><li>2. The ethics section needs to additionally provide a little more information about the original ethical approval, and how participants provide consent etc when recruited to the panel (perhaps including what is provided under ethics approval in the Declarations section).</li><li>3. The description of the short form of the Job Satisfaction Scale (JSS) used in the study is insufficient. It is unclear that the 10 items chosen actually constitute a shortened version and there is absolutely no validation information on this particular modification of the original scale. S the authors should be well aware, the original scale had two factors - intrinsic and extrinsic satisfaction - described. It is quite possible that this short-form version of the JSS cannot be treated as a single scale and might retain the factor structure of the original JSS. Indeed, the overall mean sum score that has been calculated could be considered invalid. If you have evidence of its validation, then some appropriate reference to the results of validation studies on this particular short-form need to be briefly described and cited. It may also be worthwhile to look more closely at some of the numerous MABEL papers on job satisfaction in medicine (details available on the study website).</li><li>4. More information on the representativeness of the sample (page 8), perhaps a further summary sentence, would be helpful.</li><li>5. The Discussion section is constructed quite atypically. The strengths and limitations would be better placed at the end of this section, prior to a Conclusions section (which is not explicit). The discussion needs to be better integrated, drawing the main outcomes of the current study together with the contextual information of health reforms etc and outcomes described in the literature.</li><li>6. A separate Conclusions section would be useful, which would mainly but not only include the information in the section called 'Policy implications and further research'.</li><li>7. In Table 1, it would be helpful for the reader if the data on job categories included proportions for each year in addition to the frequencies</li><li>8. Some of the references are older than 10 years, some even older than 15 years. Some updating of more contemporary references needs to be considered.</li></ol>
--	---

## VERSION 1 – AUTHOR RESPONSE

Reviewer: 1

Reviewer Name: Tamara Ritsema

Institution and Country: George Washington University, Washington DC, USA

This is an interesting topic and the approach of serial surveys is novel. Well done. While the article is very well written for a person who is a non-native English speaker, it could use a little bit of polishing from a friendly native English speaker. (Of course, your English is vastly superior to my Norwegian, so all respect to the team!)

Response:

The manuscript has now been copy-edited by a person with English as first language.

Reviewer: 1

Page 2, line 9. I don't know what "doctors in different job categories" means. Is this specialty? Length of time in practice? Employer? Explain a bit more.

Response:

Page 2, line 9-10

We now write:

"Doctors working in different job positions (hospital doctors, GPs, private practice specialists, doctors in academia)."

Reviewer: 1

Page 4, line 17. – I would add "on job satisfaction" after "health care reform" in this sentence. You want to link these changes in the environment to your outcome variable.

Response:

Revised.

Page 4, line 15-16:

"The impact of these reforms on doctors' job satisfaction has so far been insufficiently explored."

Reviewer: 1

Page 5 and 6 – the Methods section should be a narrative. Right now the dependent and independent variables sections are just lists. Maybe write a narrative and then use a table to display your variables? Also, as a person who has published using the JSS before, I would encourage you to

make it more clear in the methods section that you asked the participants to rate each of the 10 items on a 7 point scale and then divided by 10 to get the mean. Previous literature has reported the total score on the JSS. I think it is completely fine to report the mean score, but just be clear that is how you did it for those not familiar with the JSS. Also, just note that the JSS has sometimes been administered using a 4 point Likert scale, in which case the maximum score for the JSS was 40.

Response:

We agree. We made the description of the JSS more clear.

Page 8, line 15-17:

"We asked the doctors to score each of the ten items on a seven point Likert scale from 1 (very dissatisfied) to 7 (very satisfied). An unweighted mean sum score was calculated, as well as analyses of single items."

To accommodate the requests of Reviewer 2 and 4, we have expanded the description of the original job satisfaction scale by Warr-Cook-Waal and the ten item job satisfaction scale by Cooper in the "Dependent variable" section (page 7-8). We have also added methodological discussion about the use of the ten item job satisfaction scale both in the "Dependent variable" section and in the "Strengths and limitations" section (page 17, line 14-21), and included reasons for use of this JSS version in the "Introduction" section (page 5, line 18-22).

We would like to keep the list of dependent and independent variables. We think they give a good overview over the JSS items and the different job positions.

Reviewer: 1

Page 7 Line 2 – I would report this information about interns with your inclusion and exclusion criteria (under participants)

Response:

We reported this information under "inclusion and exclusion criteria".

Page 6, line 15-16:

"Inclusion and exclusion criteria

Since interns were only identified in data from 2016-17, this category is excluded in this paper."

Reviewer: 1

Page 7, line 10 – Says "units with missing data were excluded". Does this mean that participants were excluded if they did not reply at each time point? Or does it mean that the individual time point was not included in the analysis if they didn't respond to that particular survey? Or does it mean that if the participants chose not to answer one item of the JSS that this was excluded from the calculation of the mean? Be specific here.

Response:

This is now clarified.

Page 9, line 17-22:

"Three different samples are analysed. The first consists of all respondents at all times, the unbalanced cohort. Here respondents with missing data on gender or age or all JSS items were excluded. The second sample comprises doctors with defined job positions in one of four categories: GPs, specialists in private practice, hospital doctors, and doctors in academia in minimum one survey. The third, longitudinal sample are the doctors who responded at all four points in time. A subsample here are the doctors who did not change job position during the observational period. "

Reviewer: 1

Page 7 , line 25 – I am confused which Table 1 the authors are discussing. There appear to be 2 Table 1s – one with specialty breakdown and one with response rates, etc. Please make sure that the table you want first is the one labeled "Table 1" (I think that is the second table in the PDF).

Response:

Sorry for unclear headings.

The first Table in PDF was a supplement to Figure 1.

To avoid further confusion, we combined our table to figure 1 to the uploaded figure 1.

Reviewer: 1

Page 8, line 17; page 9, line 1; page 9, line 12 – I have a personal dislike of reporting "nearly significant" results. I will leave it to the editor of the journal (they may have a policy on this) as to whether to leave it in, but I feel like the assumptions of our statistical tests mean we should report significant or not significant. We are implying we know the direction of a potential association even though the testing did not reveal significance and I think this leads readers to assume unproven associations. The current replication crisis has begun to even undermine our faith in the almighty  $p < 0.05$  – implying association in the setting of a higher  $p$  value than that is dangerous.

Response:

We agree.

We now write:

Page 11, line 6:

"... and non-significantly different from 2012 ..."

Page 11, line 18:

"... there was a non-significantly change in JSS..."

Page 12, line 3:

"... were found to change non-significantly ..."

Reviewer: 1

Page 12, lines 8-10 – I would just remove. You have done a nice job of international comparisons with other health professionals. This just jumps out at the end as incongruous with the rest of your nice discussion.

Response:

We should like to keep this section, because Reviewer 3 suggests additional information on "how job satisfaction in other professionals changed over the same time period."

Reviewer: 1

Where is your limitations section? I don't see it.

Response:

Our limitation section is on page 17.

Reviewer: 1

Table 1 (the second table in the PDF) – are these percentages representative of all doctors in Norway? For example, there are roughly 3 times as many hospital doctors as GPs in the 2014 administration. Are there 3 x as many hospital doctors as GPs in Norway? Are there 5 x as many GPs as doctors in academia in Norway? Did any particular group have a higher response rate than others?

Response:

Our data are representative of all doctors in Norway, with only negligible deviations.

Yes, there were roughly 3 times as many hospital doctors as GPs in the 2014 administration in Norway: n=4782 GPs and 12 897 hospital doctors (Statistics of members of The Norwegian Medical Association in 2014: <https://legeforeningen.no/Emner/Andre-emner/Legestatistikk/> ).

Yes, there are roughly 3 times as many hospital doctors as GPs in Norway. For example in 2017, 4709 GPs and 13 077 hospital doctors (Statistics of members of The Norwegian Medical Association in 2017. <https://legeforeningen.no/Emner/Andre-emner/Legestatistikk/Yrkesaktive-leger-i-Norge/Stillingsgrupper/yrkesaktive-leger-70-ar-i-norge-fordelt-pa-stilling-2017-12/> )

There are roughly 7 times as many GPs as doctors in academia in Norway (Statistics of members of The Norwegian Medical Association in 2017, <https://legeforeningen.no/Emner/Andre-emner/Legestatistikk/Yrkesaktive-leger-i-Norge/Stillingsgrupper/> ).



Unfortunately, we have no data on response rates by different job positions. This is because there are several changes in job positions between each round. Some respondents retire, withdraw from the panel or change their specialty or positions (e.g. hospital doctor, GP, specialist in private practice, researcher or administrative position).

We included proportions for data on job positions for each year in the Table 1 (page 23) to show that the distribution of different job positions are comparable during the study period. We also included a sentence in "RESULTS/Respondents" section.

Page 10, line 22-23:

"The distribution of doctors in different job positions are comparable over the study period (Table 1)."

We revised the "METHODS/Participants" section to make more clear the sample selection and representativeness of the panel of active Norwegian doctors.

Page 6, line 2-11:

"Design and participants

Since 1994 the Institute for Studies of the Medical Profession (LEFO, [www.legeforsk.org](http://www.legeforsk.org)) has regularly surveyed a representative panel of active Norwegian doctors biannually with postal questionnaires. The original panel was based on an invitation to 2,000 randomly selected active Norwegian doctors in 1993. The 1,272 doctors who agreed to participate were representative of the total doctor work force in terms of age, sex, specialty and place of work. Since then approximately 540 doctors have left the panel due to retirement, death, or voluntary withdrawal. Therefore, the panel was supplemented with approximately 400 young doctors in 2000, 250 young doctors in 2008, 300 in 2012 and 650 doctors in 2016-17, maintaining the representativity (41 42) With this in and out pattern our cohort constitutes what may be called an unbalanced cohort. "

Reviewer: 2

Reviewer Name: Katja Goetz

Institution and Country: Institute of Family Medicine

Reviewer: 2

I have already read to other papers from this series and I miss the added value.

Response:

We have added a separate "Conclusions" section to hopefully make more clear, what our study adds.

Page 17, line 23 – page 18, line 4:

"Conclusion

Job satisfaction for Norwegian doctors remained relatively high, but with a downward trend over the last eight years, where the decrease was statistically significant for GPs and hospital doctors. Private practice specialists were most satisfied, followed by GPs and hospital doctors. The difference between the GPs and the private practice specialists increased over time. While no significant changes were found in the ten JSS items for private practice specialists and doctors in academia, satisfaction with "working hours", "payment", "recognition for good work" and "freedom to choose methods of work" declined significantly both among GPs and hospital doctors. Several health care reforms and regulations over the last decade and changes in the professional culture may explain some of the reduced satisfaction."

Reviewer: 2

I miss a methodological discussion about the use of the job satisfaction scale.

Response:

We have expanded the description of the original job satisfaction scale by Warr-Cook-Waal and the ten item job satisfaction scale by Cooper in the "Dependent variable" section. We have also added methodological discussion about the use of the ten item job satisfaction scale both in the "Dependent variable" section and in the "Strengths and limitations" section, and included reasons for use of this JSS version in the "Introduction" section.

"Introduction" section

Page 5, line 18-22:

" There are several instruments to measure job satisfaction, including single items and multi-item scales (6 36 37). A widely used instrument in health care settings is the ten item version (38) of the Warr-Cook-Wall job satisfaction scale (JSS) (6). assessing both total job satisfaction and satisfaction with different aspects of the job (19 21 26 28-30 38-40) It also allows for good national and international comparisons."

"Dependent variable" section

Page 7, line 1 – page 8, line 3:

" Dependent variable

Total job satisfaction and satisfaction with different aspect of the job were measured with the ten item modified version (38) of the "Job Satisfaction Scale" by Warr, Cook and Wall (6)

The original scale included one item assessing the overall satisfaction and fifteen items assessing two factors related to job satisfaction, the intrinsic factor (seven items on attitudes towards personal

achievement and task success: freedom to choose your own method of working; recognition you get for good work, the amount of responsibility you are given, your opportunity to use your abilities, your chance of promotion, the attention paid to suggestions you make, the amount of variety in your job) and extrinsic factor (eight items on attitudes on working conditions: physical working conditions, your fellow workers, your immediate boss, your rate of pay, industrial relations between management and workers in your firm, the way your firm is managed, your hours of work, your job security). The total job satisfaction was calculated as the sum of all separate items. The scale was tested for validity and reliability in blue-collar male workers employed full-time in a manufacturing industry in the United Kingdom (6)

The ten item form of the original scale was devised by Cooper-Rout-Faragher in 1989 to study job satisfaction, mental health and stress among general practitioners in England. Five items were removed from the original scale that were not relevant for the general practitioners population: "your immediate boss", "industrial relations between management and workers in your firm", "your chance of promotion", "the way your firm is managed" and "your job security". The scale was not tested for validity and reliability by Cooper, Rout and Faragher (38) A validation study of this scale was done in a cohort of Australian clinical medical workforce by Hills, Joyce and Humphries in 2012, where the original seven point Likert scale was reduced to five point Likert scale from 0 (very dissatisfied) to 4 (very satisfied). Factor analytic and internal reliability tests did not support differentiating intrinsic and extrinsic factors. They supported the use of the ten item instrument as a single-factor scale and the use of a composite job satisfaction score (43).

The ten item version (38) of the original Warr-Cook-Wall questionnaire with seven point Likert scale (6) was used extensively in doctors population in Norway and elsewhere (19-22 26 28-30 39 40). We applied this instrument in our study to allow comparisons across countries and over time. The ten item were:"

"Strengths and limitations"

Page 17, line 14-21:

"The ten item Warr-Cook-Wall scale for job satisfaction was specifically designed for GPs in solo practice in the UK (6 38), but it has been used extensively to describe total job satisfaction and satisfaction of different level of work in doctor populations both in Norway and elsewhere (19-22 28-30 40). The validation of this ten item job satisfaction scale in a cohort of Australian medical practitioners provided validity evidence for a single-factor solution and for a use of a composite job satisfaction score. However, it was suggested to include other job-specific items in the scale, especially for doctors having employee status or working in organisational settings (43)."

Reviewer: 3

Reviewer Name: Michael H. Freitag

Institution and Country: Department of Health Services Research, Division of General Practice/Family Medicine, University of Oldenburg, Germany

Please leave your comments for the authors below Rosta and colleagues report a repeated survey of physician job satisfaction in Norway from 2010-2017. Job satisfaction is an important issue given the shortage of physicians in Europe and its association with patient outcomes.

Reviewer: 3

The paper is well written. However, the main point is whether the reported differences in job satisfaction over time are meaningful and real. The first sentence of the discussion section of the main article says "...JSS remained high" whereas in the abstract it says "...decreased significantly".

Response:

The first sentence of the discussion section has now been revised.

Page 13, line 1:

"From 2010 to 2016-17, JSS for all doctors decreased significantly."

Reviewer: 3

It is intriguing that there was no change among those participants who participated in all four surveys over time. That implies that the (small) changes over time are driven by the differences of sample composition.

Response:

In the longitudinal sample (doctors who responded at all four points in time) and in the longitudinal subsample (doctors did not change position at all four points in time), we found some changes over time. We now discuss figure 2 in more detail.

Page 12, line 1-12

"Changes in JSS in the longitudinal sample

All doctors

Using paired sample t-tests, JSS scores were found to change non-significantly from 2010 to 2012 (5.30 vs. 5.34;  $t=1.43$ ;  $p=0.152$ ), to increase significantly from 2012 to 2014 (5.34 vs. 5.41;  $t=2.19$ ;  $p=0.029$ ) and then to decrease significantly from 2014 to 2016-17 (5.41 vs. 5.34;  $t=-2.03$ ;  $p=0.043$ ).

Hospital doctors, GPs and private practice specialists

From 2010 to 2016-17 in the longitudinal subsample, the JSS scores for GPs decreased steadily. A significant increase in JSS was found for specialists in private practice from 2010 to 2012, and for hospital doctors from 2012 to 2014. For all three job positions, there was a non-significant decline in JSS from 2014 to 2016-17. At any point in time, private practice specialists were the most satisfied. GPs were more satisfied than hospital doctors, but the difference between these groups decreased. (Figure 2)."

In the discussion section, we discuss reasons for non-significant changes.

Page 16, line 17-20:

"A possible reason for this stability in JSS may be a combination of the adaption of health care regulations over time and the selection of doctors. The most satisfied doctors are more likely to remain in their current job position."

Reviewer: 3

Interestingly, the proportion of females increased from 37.4% in 2010 to 52.9% in 2016/2017 while the mean age decreased from 50.7 to 47.7 years. There is an international trend that the younger generation is less willing to work long hours and female physicians are more likely to consider time constraints by child care.

Response:

In a previous study with data from 1994 to 2014 we found that the total weekly working hours remained the same for all categories of doctors (doctors in hospital, GPs, specialists working in private practice, community medical officers, doctors in administrative positions), except those working in academia.

(Rosta J, Aasland OG. Legers arbeidstid og tid til pasientarbeid [Doctors' Working Time and Time Spent in Patient Care]. Tidsskrift for Den norske legeforening, 2016;136:16:1355-9.).

In another study based on data from 1994 to 2012 there was no significant difference in weekly working hours between senior doctors and junior doctors (except for 2000). Over the whole period female hospital doctors worked significantly fewer hours than their male colleagues (except for 2004), but this difference decreased with time. The weekly working hours decreased non-significantly from 45 hours in 1994 to 44 hours in 2012 for female doctors, while staying at 46 hours for male doctors.

(Rosta J, Aasland OG. Weekly working hours for Norwegian hospital doctors since 1994 with special attention to postgraduate training, work-home balance and the European Working Time Directive. A longitudinal study. BMJ Open, 2014;4:10 e005704.)

Reviewer: 3

Therefore, the authors should report their results stratified by age and gender. First, it would be interesting to see age- and gender-specific job satisfaction and second, this would clarify how much of the change over time can be attributed to the increasing proportion of females and younger physicians.

Response

In our study, we have analyzed the effect of age and gender on estimated means of JSS. We found that JSS increased with increasing age. There were no gender differences.

Previous studies with data from 1994 to 2008 and the present study with data from 2010 to 2017 show no gender differences but increase with age.

(Aasland OG, Rosta J, Nylenna M. Health care reforms and job satisfaction among doctors in Norway. *Scandinavian Journal of Public Health*, 2010; 38(3):253-8.

Nylenna M, Aasland OG. Job satisfaction among Norwegian doctors. *Tidsskr Nor Laegeforen*. 2010 May 20;130(10):1028-31. doi: 10.4045/tidsskr.09.0955.

Nylenna M, Gulbrandsen P, Førde R, et al. Unhappy doctors? A longitudinal study of life and job satisfaction among Norwegian doctors 1994-2002. *BMC Health Services Research* 2005;Jun 8;5:44.)

According to the Norwegian "Working environment and living conditions survey", JSS among all employees did not differ between gender and it did increase with age.

(Ref.: NOA. Jobbtilfredshet (Job satisfaction) Oslo: NOA (Nasjonal overvåking av arbeidsmiljø). <https://noa.stami.no/arbeidsmiljoindikatorer/psykososialtorganisasjon/jobbtilfredshet1/tilfredshet/>; 2016)

In our analyses, we used General Linear modelling (GLM) controlling for gender and age to estimate the JSS means at four points in time: 2010, 2012, 2014, and 2016-17. Using the same modelling did allow comparison with previous studies on doctors` job satisfaction in Norway, too.

Reviewer: 3

The authors could add a sentence on how physician job satisfaction compares to other job categories such as lawyers or teachers and how job satisfaction in other professions changed over the same time period.

Response:

We added a sentence on how job satisfaction among all employees in Norway changed over the same time period. Unfortunately, no analyses has been performed on changes from 2010 to 2016 by occupational groups (personal communication with "STAMI – The National Institute of Occupational Health" that published data from 2016 on job satisfaction among all employees and by occupational groups in Norway. Dato:12 March 2019).

Page 14, line 23 – page 15, line 2:

"In the Norwegian "Working environment and living conditions survey" from 2010 to 2016, there were no changes in JSS as measured by a five point Likert scale from very dissatisfied to very satisfied. About 90% of the employees reported that they are "quite" or "very satisfied" with their job. In data from 2016, top managers (97%), farmers/fisherman (95%) or physiotherapists (95%) reported a higher level and nurses (89%), policeman (88%) or customer service occupations (82%) a lower level of satisfaction than the doctors (90%). JSS among employees did not differ between gender and it did increase with age (53) "

Reviewer: 3

Minor points:

Page 2, line 23: Conclusions (spelling)

Page 5, line 2: emphasis (spelling)

Response:

Revised.

Reviewer: 4

Reviewer Name: A/Prof Danny Hills

Institution and Country: Federation University Australia, Australia

Please leave your comments for the authors below Job satisfaction in the medical profession is an important issue for the profession, health services and the community. This manuscript attempts to document trends in medical practitioner job satisfaction in Norway, in light of recent changes in health politics and health care reforms.

There are some key issues that need to be addressed in this manuscript before it could be considered ready for publication. Overall, the manuscript needs to flow better and with appropriate use of past tense throughout.

Reviewer: 4

1. The introduction could flow better. A better approach might be to commence with why the construct of job satisfaction is important and then why it is important in medical practitioners. A more rounded explanation of the construct could follow, including how it has been measured more broadly and in medical practitioners more specifically. What have been the trends and existing evidence on job satisfaction in medical practitioners could follow, including with a little more detailed explanation of the reforms etc., finishing with the aim statement.

Response:

We revised the "Introduction" section. This section includes information on:

- importance of job satisfaction among employees,
- importance of job satisfaction in medical practitioners,
- description of the two important health reforms in Norway at the beginning of the 21 century and their impact on doctors' job satisfaction, trends on job satisfaction in medical practitioners,
- description of the three new organizational reforms in Norwegian health care and the lack of their effect on doctors' job satisfaction,

-measurement of job satisfaction more broadly and in medical practitioners, reason for use of the 10 item version of the Warr-Cook-Wall job satisfaction scale,

-aim statement.

Reviewer: 4

2. The ethics section needs to additionally provide a little more information about the original ethical approval, and how participants provide consent etc when recruited to the panel (perhaps including what is provided under ethics approval in the Declarations section).

Response:

We have revised the "Ethics section".

Page 6, line 18-26:

"Ethical approval

According to the Regional Committee for Medical Research Ethics, the study based

on "Norwegian Physician Survey - A bi-annual prospective questionnaire survey to a representative sample of Norwegian physicians" is exempt from review in Norway, cf. §§ 4 of The Act. The project can be implemented without the approval by the Regional Committee for

Medical Research Ethics (IRB 0000 1870). All invitees received a letter with a description of the "Norwegian Physician Survey" aim. It was also explained that the participation is voluntary and the data would be handled confidentially. All participants signed informed written consent before the start of the survey."

Reviewer: 4

3. The description of the short form of the Job Satisfaction Scale (JSS) used in the study is insufficient. It is unclear that the 10 items chosen actually constitute a shortened version and there is absolutely no validation information on this particular modification of the original scale. Since the authors should be well aware, the original scale had two factors - intrinsic and extrinsic satisfaction - described. It is quite possible that this short-form version of the JSS cannot be treated as a single scale and might retain the factor structure of the original JSS. Indeed, the overall mean sum score that has been calculated could be considered invalid. If you have evidence of its validation, then some appropriate reference to the results of validation studies on this particular short-form need to be briefly described and cited. It may also be worthwhile to look more closely at some of the numerous MABEL papers on job satisfaction in medicine (details available on the study website).

Response:

We have expanded the description of the original job satisfaction scale by Warr-Cook-Waal and the ten item job satisfaction scale by Cooper in the "Dependent variable" section. We have also added methodological discussion about the use of the ten item job satisfaction scale both in the "Dependent variable" section and in the "Strengths and limitations" section, and included reasons for use of this JSS version in the "Introduction" section.



"Introduction" section

Page 5, line 18-22:

" There are several instruments to measure job satisfaction, including single items and multi-item scales (6 36 37). A widely used instrument in health care settings is the ten item version<sup>38</sup> of the Warr-Cook-Wall job satisfaction scale (JSS) (6). assessing both total job satisfaction and satisfaction with different aspects of the job (19 21 22 26 28-30 38-41) It also allows for good national and international comparisons."

"Dependent variable" section

Page 7, line 1 – page 8, line 3:

" Dependent variable

Total job satisfaction and satisfaction with different aspect of the job were measured with the ten item modified version (38) of the "Job Satisfaction Scale" by Warr, Cook and Wall (6)

The original scale included one item assessing the overall satisfaction and fifteen items assessing two factors related to job satisfaction, the intrinsic factor (seven items on attitudes towards personal achievement and task success: freedom to choose your own method of working; recognition you get for good work, the amount of responsibility you are given, your opportunity to use your abilities, your chance of promotion, the attention payed to suggestions you make, the amount of variety in your job) and extrinsic factor (eight items on attitudes on working conditions: physical working conditions, your fellow workers, your immediate boss, your rate of pay, industrial relations between management and workers in your firm, the way your firm is managed, your hours of work, your job security). The total job satisfaction was calculated as the sum of all separate items. The scale was tested for validity and reliability in blue-collar male workers employed full-time in a manufactural industry in the United Kingdom (6)

The ten item form of the original scale was devised by Cooper-Rout-Faragher in 1989 to study job satisfaction, mental health and stress among general practitioners in England. Five items were removed from the original scale that were not relevant for the general practitioners population: "your immediate boss", "industrial relations between management and workers in your firm", "your chance of promotion", "the way your firm is managed" and "your job security". The scale was not tested for validity and reliability by Cooper, Rout and Faragher (38) A validation study of this scale was done in a cohort of Australian clinical medical workforce by Hills, Joyce and Humphries in 2012, where the original seven point Likert scale was reduced to five point Likert scale from 0 (very dissatisfied) to 4 (very satisfied). Factor analytic and internal reliability tests did not support differentiating intrinsic and extrinsic factors. They supported the use of the ten item instrument as a single-factor scale and the use of a composite job satisfaction score (43).

The ten item version (38) of the original Warr-Cook-Wall questionnaire with seven point Likert scale (6) was used extensively in doctors population in Norway and elsewhere (19-22 26 28-30 39 40). We applied this instrument in our study to allow comparisons across countries and over time. The ten item were:"

"Strengths and limitations"

Page 17, line 14-21:

"The ten item Warr-Cook-Wall scale for job satisfaction was specifically designed for GPs in solo practice in the UK (6 38), but it has been used extensively to describe total job satisfaction and satisfaction of different level of work in doctor populations both in Norway and elsewhere (19-22 28-30 40). The validation of this ten item job satisfaction scale in a cohort of Australian medical practitioners provided validity evidence for a single-factor solution and for a use of a composite job satisfaction score. However, it was suggested to include other job-specific items in the scale, especially for doctors having employee status or working in organisational settings (43)."

Reviewer: 4

4. More information on the representativeness of the sample (page 8), perhaps a further summary sentence, would be helpful.

Response:

We included proportions for data on job positions for each year in the Table 1 (page 23) to show that the distribution of different job positions are comparable during the study period. We also included a sentence in "RESULTS/Respondents" section.

Page 10, line 22-23:

"The distribution of doctors in different job positions are comparable over the study period (Table 1)."

We revised the "METHODS/Participants" section to make more clear the sample selection and representativeness of the panel of active Norwegian doctors.

Page 6, line 2-11:

"Design and participants

Since 1994 the Institute for Studies of the Medical Profession (LEFO, [www.legeforsk.org](http://www.legeforsk.org)) has regularly surveyed a representative panel of active Norwegian doctors biannually with postal questionnaires. The original panel was based on an invitation to 2,000 randomly selected active Norwegian doctors in 1993. The 1,272 doctors who agreed to participate were representative of the total doctor work force in terms of age, sex, specialty and place of work. Since then approximately 540 doctors have left the panel due to retirement, death, or voluntary withdrawal. Therefore, the panel was supplemented with approximately 400 young doctors in 2000, 250 young doctors in 2008, 300 in 2012 and 650 doctors in 2016-17, maintaining the representativity (41 42) With this in and out pattern our cohort constitutes what may be called an unbalanced cohort. "

Reviewer: 4

5. The Discussion section is constructed quite atypically. The strengths and limitations would be better placed at the end of this section, prior to a Conclusions section (which is not explicit). The discussion needs to be better integrated, drawing the main outcomes of the current study together with the contextual information of health reforms etc and outcomes described in the literature.

Response:

The "Strengths and limitations" section has been placed as suggested.

The "Discussion" section has been revised. We moved the description of the two important health care reforms at the beginning of the 21 century in Norway and their impact on job satisfaction for doctors to the "Introduction" section. The impact of the three new major health care reforms over the last decade on job satisfaction of doctors is now more clear in "Discussion" section.

Reviewer: 4

6. A separate Conclusions section would be useful, which would mainly but not only include the information in the section called 'Policy implications and further research'.

Response:

We added a separate "Conclusion" section.

Page 17, line 23 – page 24, line 4:

"Conclusion

Job satisfaction for Norwegian doctors remained relatively high, but with a downward trend over the last eight years, where the decrease was statistically significant for GPs and hospital doctors. Private practice specialists were most satisfied, followed by GPs and hospital doctors. The difference between the GPs and the private practice specialists increased over time. While no significant changes were found in the ten JSS items for private practice specialists and doctors in academia, satisfaction with "working hours", "payment", "recognition for good work" and "freedom to choose methods of work" declined significantly both among GPs and hospital doctors. Several health care reforms and regulations over the last decade and changes in the professional culture may explain some of the reduced satisfaction."

Reviewer: 4

7. In Table 1, it would be helpful for the reader if the data on job categories included proportions for each year in addition to the frequencies

Response:

We have included proportions for data on job categories for each year.

Reviewer: 4

8. Some of the references are older than 10 years, some even older than 15 years. Some updating of more contemporary references needs to be considered.

Response:

We have removed references on job satisfaction of doctors that are older than 15 years.

Smith R. Why are doctors so unhappy? *BMJ* 2001;322(7294):1073-74.

Zuger A. Dissatisfaction with medical practice. *N Engl J Med* 2004(350):69-75.

We have updated the paper with more contemporary references.

Hills D, Joyce C, Humphreys J. Validation of a Job Satisfaction Scale in the Australian Clinical Medical Workforce. *Evaluation & the Health Professions* 2012;35(1):47-76.

Bakotic D. Relationship between job satisfaction and organisational performance. *Economic Research-Ekonomska Istraživanja* 2016;29(1):118-30.

Jaarsveld Dv, Keyser E. The Moderating Role of Job Satisfaction on Workplace Absenteeism and Substance use amongst the Employees at a Power Utility in Mpumalanga. *Journal of Economics and Behavioral Studies* 2018(2):219-30.

Scanlan JN, Still M. Relationships between burnout, turnover intention, job satisfaction, job demands and job resources for mental health personnel in an Australian mental health service. *BMC Health Services Research* 2019;19(1):62-62.

Coetzee N, Maree DJ, Smit BN. The relationship between chronic fatigue syndrome, burnout, job satisfaction, social support and age among academics at a tertiary institution. *International Journal of Occupational Medicine and Environmental Health* 2019;32(1):75-85.

Casalino L, Crosson F. Physician Satisfaction and Physician Well-Being: Should Anyone Care? *Professions and Professionalism* 2015;5(1):1-12.

Angerer P, Weigl M. Physicians' psychosocial work conditions and quality of care: A literature review. *Professions and Professionalism* 2015;5(1):1-20.

McKinlay J, Marceau L. New wine in an old bottle: does alienation provide an explanation of the origins of physician discontent? *International Journal of Health Services* 2011;41 (2):301-35.

Hora GPR, Ribas Júnior R, de Souza MA. State of the Art of Job Satisfaction Measures: A Systematic Review. *Trends in Psychology* 2018;26:971-86.

Konrad T. Measures, Methods, and Models of Doctor Satisfaction: Future Research Challenges. *Professions and Professionalism* 2015;5(1):1-16.

Goetz K, Campbell SM, Steinhäuser J, et al. Evaluation of job satisfaction of practice staff and general practitioners: an exploratory study. *BMC Family Practice* 2011;12(1):137.

## VERSION 2 – REVIEW

<b>REVIEWER</b>	Tamara Ritsema George Washington University
<b>REVIEW RETURNED</b>	15-Apr-2019

<b>GENERAL COMMENTS</b>	<p>The authors have substantially improved this paper. Good work.</p> <p>I still don't see a very robust limitations section. I would encourage them to state what the limitation is and which way it would influence the results (toward the hypothesis or toward the null)</p>
-------------------------	--

<b>REVIEWER</b>	Michael H. Freitag, Prof. Dr. med. Carl von Ossietzky University Oldenburg, Germany Department of Health Services Research Division of General Practice/Family Medicine
<b>REVIEW RETURNED</b>	14-May-2019

<b>GENERAL COMMENTS</b>	Thanks for the changes made.
-------------------------	------------------------------

<b>REVIEWER</b>	A/Prof Danny Hills Federation University Australia
<b>REVIEW RETURNED</b>	12-Apr-2019

<b>GENERAL COMMENTS</b>	<p>A much improved paper. Important it is published. There are just a couple of things that can be fixed prior to acceptance for publication.</p> <ol style="list-style-type: none"><li>1. On page 6, the following sentence is incomplete - it seems to need an clause at the end to indicate the effect of the white paper. "The Free Choice of Hospital Reform" in 2015 gave users a free choice of hospital,<sup>34</sup> and the white paper on "The Future Primary Care – Proximity and Comprehensiveness" in 2015, which aimed to improve user involvement, availability, prevention, proactivity and collaboration between multidisciplinary teams.</li><li>2. Line 8 on page 8 should read: "... with different aspects of the job ..."</li><li>3. On that same page there is non-English commentary that needs to be removed.</li><li>4. Line 27 on page 8, 'Five' is incorrectly spelled.</li><li>5. From about page 13, the term 'job satisfaction' is replaced by 'JSS', which does not make sense, since it refers to the Job Satisfaction Scale. This needs to be systematically corrected from this point forward.</li></ol>
-------------------------	--

## VERSION 2 – AUTHOR RESPONSE

Reviewer: 1

Reviewer Name: Tamara Ritsema

Institution and Country: George Washington University, Washington DC, USA

Reviewer 1

Good work. I still don't see a very robust limitations section. I would encourage them to state what the limitation is and which way it would influence the results (toward the hypothesis or toward the null).

Response

The most important limitation of the study is a possible skewness in response rate related to level of satisfaction. This is now discussed more thoroughly.

Page 17, line 10 – page 18, line 9

"Strengths and limitations

The main strength of this study is that it allows for generalisation to the whole population of doctors in Norway. Similarities in survey methods and repeated measures over time is another advantage. Furthermore, the response rates were fairly good, ranging from 67% and 75%, which are higher than for other surveys of the medical profession, but do not rule out the possibility of non-response bias.44 It is possible that the doctors with a particularly heavy work burden and therefore a probable lower job satisfaction to a lesser degree than others responded to the questionnaire, leading to an overestimation of satisfaction level. On the other hand, doctors who are dissatisfied with their working conditions might to a larger degree want to express their opinion, which could lead to lower satisfaction scores. A study based on two cross-sectional surveys among English GPs from 2004 and 2005 supported an association between response and satisfaction, respectively less satisfied GPs were more likely to response (65). However, our follow-up of the unbalanced cohort showed changes in the partly overlapping samples of doctors over time, which give us valid data of changes in satisfaction. When it comes to measuring subjective satisfaction, there is in general no alternative to survey individuals in a random sample (66). We do not know whether there is a tendency in our sample towards over- or underestimation of the satisfaction levels with various components of working conditions, or whether there are job-category or medical-discipline-specific differences in the self-reporting. Again, an unbalanced cohort design with follow-up of the partly overlapping sample of doctors, gives us robust data of changes in satisfaction. Another concern is that there is no gold standard of measuring doctors` job satisfaction with a global, check-list based measure (39) The ten item Warr-Cook-Wall scale for job satisfaction was specifically designed for GPs in solo practice in the UK (6, 40), but it has been used extensively to describe total job satisfaction and satisfaction of different level of work in doctor populations both in Norway and elsewhere (19-22, 28-30, 42). The validation of this ten item job satisfaction scale in a cohort of Australian medical practitioners provided validity evidence for a single-factor solution and for a use of a composite job satisfaction score. However, it was suggested to include other job-specific items in the scale, especially for doctors having employee status or working in organisational settings (45). Because job satisfaction varies with personality (67), well-being (7, 8), mental and physical health status (3-5, 40), it is also important to include these co-variates in future analyses."

Reviewer 4

Reviewer Name: A/Prof Danny Hills

Institution and Country: Federation University Australia, Australia

Reviewer 4

1. On page 6, the following sentence is incomplete - it seems to need an clause at the end to indicate the effect of the white paper.

"The Free Choice of Hospital Reform" in 2015 gave users a free choice of

hospital,<sup>34</sup> and the white paper on "The Future Primary Care – Proximity and Comprehensiveness" in 2015, which aimed to improve user involvement, availability, prevention, proactivity and collaboration between multidisciplinary teams.

Response

Revised.

Page 5, line 14-17

"The Free Choice of Hospital Reform" in 2015 gave the users a free choice of hospital (34), and the white paper on "The Future Primary Care – Proximity and Comprehensiveness" in 2015 was implemented to improve user involvement, availability, prevention, proactivity and collaboration between multidisciplinary teams (35). Recent surveys on hospital doctors` and GP` working conditions documented high workload and considerable growth in work demand (36 37) The impact of the latest se reforms on doctors' job satisfaction has so far been insufficiently explored."

Reviewer 4

2. Line 8 on page 8 should read: "... with different aspects of the job ..."

Response

Revised.

Page 7, line 1

"...with different aspects of the job..."

Reviewer 4

3. On that same page there is non-English commentary that needs to be removed.

Response

We did not find non-English commentary on page 8.

Reviewer 4

4. Line 27 on page 8, 'Five' is incorrectly spelled.

Response

We did not find the incorrectly spelled "Five" on page 8, line 27.

Five is correctly spelled on:

Page 7, line 18: "... Five items were..."

Page 7, line 24: "... reduced to five point..."

Page 14, line 24: "...measured by a five point Likert..."

Reviewer 4

5. From about page 13, the term 'job satisfaction' is replaced by 'JSS', which does not make sense, since it refers to the Job Satisfaction Scale. This needs to be systematically corrected from this point forward.

Response

Revised.

Page 12, line 9: "... increase in JSS scores..."

Page 12, line 19: "...and overall job satisfaction..."

Page 13, line 3: "From 2010 to 2016-17, job satisfaction for all doctors decreased significantly."

Page 13, line 18: "In Norway, job satisfaction was stable and high..."

Page 14, line 4: "In our study, job satisfaction decreased..."

Page 16, line 15-16: "... estimated job satisfaction..."

Page 16, line 18: "... stability in job satisfaction may..."

Page 17, line 13-14: "... differences in the estimated means of the job satisfaction scale items between..."

Page 17, line 28: "... found in the ten job satisfaction scale items..."

### VERSION 3 – REVIEW

<b>REVIEWER</b>	Danny Hills Federation University, Australia
<b>REVIEW RETURNED</b>	16-Jul-2019

<b>GENERAL COMMENTS</b>	Much improved version. Comments in mark up are no longer apparent.
-------------------------	--