

## 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>	Disrupted care during the COVID-19 state of emergency and productivity loss attributed to presenteeism in workers: a nationwide cross-sectional study
<b>AUTHORS</b>	Ishimaru, Tomohiro; Tsuno, Kanami; Hori, Ai; Okawara, Makoto; Yasuda, Yoshino; Fujino, Yoshihisa; Tabuchi, Takahiro

### VERSION 1 – REVIEW

<b>REVIEWER</b>	Ruhle, Sascha Heinrich Heine University Düsseldorf, Business and Economics
<b>REVIEW RETURNED</b>	14-Apr-2021

<b>GENERAL COMMENTS</b>	<p>Thank you for the opportunity to review your article “Disrupted care during the COVID-19 state of emergency and subsequent presenteeism in workers: a nationwide cross-sectional study”. Using nationwide collected, cross-sectional data, you analyze the relationships between disrupted care and subsequent presenteeism. While your study is generally well written, I currently find that the weaknesses of the manuscript do not fulfill the high standards of the BMJOpen. Especially, I would recommend grounding your research more strongly in the literature on sickness presenteeism, as currently the conceptual background, as well as the discussion, is underdeveloped. It could be a fruitful avenue to align your study with results on general (subjective) health and its role in the formation of presenteeism (see below). Despite being a cross-sectional study, I think that a valuable contribution could be made by your study.</p> <p>Major</p> <p>1. I found the introduction is well written and interesting (aside from the minor aspect below). However, the way the introduction is written, I was expecting a stronger connection towards previous results in the field of presenteeism (e.g., Miraglia &amp; Johns 2016, Ruhle et al. 2020). More specifically, I was expecting information on how presenteeism could be related to disrupted care. What are your conceptual thoughts on that relationship? Does disruptive care follow one of the dual paths? Via general health? Or is it related more to a process perspective that focuses on the individuals' personal experiences? It could be useful to include recent developments in the field of presenteeism into these thoughts. Overall, I agree that there might be a plausible relationship, yet, please share your ideas with the reader here. Otherwise, concerning the STROBE checklist (which I appreciate), I find the objectives and specific hypotheses too weak to be considered for publication.</p>
-------------------------	---

2. Overall, I would expect you to think about individuals' health as an important predictor of presenteeism more in detail. This is related to your conceptual background, in which health as a driver of sickness presenteeism should be included, and in your studies design. Do you have any information on the individual's (subjective) health? Otherwise, one could critically state that you found that predictors of reduction of health are related to sickness presenteeism, which is already known (Johns, 2010). If you don't have the respective information, please nonetheless consider the role of health and add it as a rather problematic limitation of your study.

3. Method:

I generally enjoyed reading your method section, as it was mostly clear. Yet, some aspects should be additionally considered:

a. First, did you analyze your drop-out? Or what were the invalid responses you used to remove these individuals from the sample? Please make sure to test whether the reduction was random, and not based on some important constructs in your model (e.g., (un-)healthy individuals). If you decide to not include participants in the primary analysis, please make sure that the reasons for the exclusion are made transparent (and justified).

b. Why did you separate labor type into "manual work, desk work, and other."? I am unfamiliar with this separation and not sure what each of these would encompass. Especially as "other" is almost a fourth of your sample, I would like to have more information on this decision. More specifically, there are a lot of different dimensions of labor (e.g., physical, mental, skilled, unskilled, manual, etc.). While I don't expect that you delve deeper into these differences (e.g., Narotzky 2018), I would expect a clearer description of your decision, especially as we know that sickness presenteeism is related to different aspects of labor, such as e.g. the level of education (Gustafsson & Marklund, 2011) or the nature of work (Hansen & Andersen, 2008).

c. How did you handle missing values? Did any occur?

d. Further, did you test the assumptions of the multinomial logistic regression? Did you perform any additional analysis to provide evidence for the robustness of your effects? Overall, I would like to see more statistical work to better understand the robustness of your results. Please draw on adequate statistical literature to justify and clarify your approach (e.g. Stoltzfus 2011).

4. The discussion is rather weak. In my view, this is again related to the missing inclusion of general health as a mechanism that explains how disrupted care is related to sickness presenteeism. Again, stating that "Presenteeism was associated with chronic illnesses even among workers whose care continued." is not surprising, as we know that health is an important aspect for the vulnerability of an individual (Gerich, 2015). In line with this reasoning, I found your explanation of the exacerbation of underlying disease as an argument for the different results from models 1 and 2 not convincing. Further, I stumbled over the mention of "unexpected symptoms" as one identified aspect of your study on page 12. You state that "This study identified that unexpected symptoms or illnesses during the pandemic were associated with presenteeism, especially where care had been

disrupted.” However, unexpected symptoms are part of what you describe as a factor of disrupted care. Again, the conceptual background for this was not clear.

5. Finally, the contribution of the manuscript needs to be clarified. You state that you “showed that disrupted care may influence workers’ subsequent performance”. However, you did not measure productivity/performance as an outcome. Further, that dealing with acute illness will be beneficial for the health of an employee is well known, and that this will reduce the likelihood of sickness presenteeism is also established (e.g., Johns, 2010; Ruhle et al. 2020).

#### Minor

1. Abstract: The description of the “Primary and secondary outcome measures” is unfitting. Please revise and make sure it does not overlap with the Methods. Better, clearly describe the outcomes you were interested in.

2. Related to your conceptual background, you might want to clarify that you are talking about sickness presenteeism. Recent developments suggest that there might be different types of presenteeism.

3. P.7: “The total score, therefore, ranged from 7 to 35 points. Higher scores indicate worse presenteeism.” I would restrain from using “worse” here, as it is higher, but not necessarily worse in all dimensions. Karanika-Murray and Biron (2020) make a compelling case that more sickness presenteeism is not always worse for the individual.

4. There are some smaller errors in language and grammar. While I am not a native speaker, I would suggest rigorous proofread.

#### Reference

(please note that these references are suggestions that might help to deal with the described challenges and should not be interpreted as “please cite”)

Gerich, J. (2015). Sick at work: Methodological problems with research on workplace presenteeism. *Health Services and Outcomes Research Methodology*, 15, 37–53.

Gustafsson, K., & Marklund, S. (2011). Consequences of sickness presence and sickness absence on health and work ability: a Swedish prospective cohort study. *International journal of occupational medicine and environmental health*, 24(2), 153-165.

Hansen, C. D., & Andersen, J. H. (2008). Going ill to work—What personal circumstances, attitudes and work-related factors are associated with sickness presenteeism?. *Social science & medicine*, 67(6), 956-964.

Johns, G. (2010). Presenteeism in the workplace: A review and research agenda. *Journal of Organizational Behavior*, 31, 519–542.

Karanika-Murray, M., & Biron, C. (2020). The health-performance framework of presenteeism: Towards understanding an adaptive behaviour. *Human Relations*, 73(2), 242-261.

Miraglia, M., & Johns, G. (2016). Going to work ill: a meta-analysis of the correlates of presenteeism and a dual-path model. *Journal of Occupational Health Psychology*, 21, 261-283.

	<p>Narotzky, S. (2018). Rethinking the concept of labour. Journal of the Royal Anthropological Institute, 24(S1), 29-43.</p> <p>Ruhle, S. A., Breitsohl, H., Aboagye, E., Baba, V., Biron, C., Correia Leal, C., ... &amp; Yang, T. (2020). "To work, or not to work, that is the question"—Recent trends and avenues for research on presenteeism. European Journal of Work and Organizational Psychology, 29(3), 344-363.</p> <p>Stoltzfus, J. C. (2011). Logistic regression: a brief primer. Academic Emergency Medicine, 18(10), 1099-1104.</p> <p>For full transparency, I cite one (out of nine) studies that I have co-authored (Ruhle et al. 2020). However, I neither state that this paper is relevant beyond the respective information it is cited for nor urge the reader to include it in the manuscript.</p> <p>I hope my comments will help you to improve the manuscript and wish you the best of luck with your project.</p>
--	---

<b>REVIEWER</b>	Vietri, Jeffrey Kantar Health Inc
<b>REVIEW RETURNED</b>	17-Apr-2021

<b>GENERAL COMMENTS</b>	<p>This is an interesting paper that uses a large internet survey sample to explore the association between disrupted care during the Covid-19 emergency in Japan and subsequent presenteeism. In general the paper is clear, but there are a few areas which would benefit from more detail. One is selection of the participants</p> <p>- I was confused by the different levels, and a figure depicting what selection / sampling method was used at each level would be helpful. Also, some explanation of what is meant by 'invalid responses' should be provided, as nearly 10% of the sample was excluded for this reason. Also, there is a mention of the exchange rate between dollars and yen that is unnecessary, as dollars are not reported in the paper.</p> <p>Lastly, the paper would benefit from additional description and justification of the treatment of the variables and different multi variable models. For instance, why adjust for the different types of disruption in model 2 instead of creating an index? Are the results consistent if the measure of presenteeism is used as continuous rather than dichotomized?</p>
-------------------------	--

<b>REVIEWER</b>	Jetha, Arif Institute for Work & Health
<b>REVIEW RETURNED</b>	19-Apr-2021

<b>GENERAL COMMENTS</b>	<p>Dear Authors,</p> <p>Thank you for the opportunity to review the paper entitled: "Disrupted care during the COVID-19 state of emergency and subsequent presenteeism in workers: a nationwide cross-sectional study"</p> <p>Your paper tackled an interesting question regarding the indirect socioeconomic implications of the COVID-19 pandemic on people living with chronic health conditions. Indeed, examining and supporting presenteeism of workers with chronic conditions has important personal, organizational and societal implications. There were some methodological and analytical considerations that require greater elaboration.</p>
-------------------------	--

	<ul style="list-style-type: none"> <li>• The author uses the term “coronavirus disease 2019”. I would ensure throughout the papers that labels “Novel coronavirus”, “COVID-19” and “SARS-COV-2” are used consistently and align with published literature and media.</li> <li>• There is repetition in the abstract between setting, participants and methods subheadings. I suggest being more concise.</li> <li>• Also, I think the strengths of the study could be enhanced to fully capture the innovative aspects of this study.</li> <li>• Perhaps it would be helpful to briefly define chronic disease with regards to this specific study. Are you referring to conditions that are both physical and mental? Would a person who has experienced persistent COVID-19 symptoms (e.g., COVID long-hauler) be captured in your definition? Additional details are required on how people with chronic disease were recruited. Did the authors rely on self-reports?</li> <li>• It would also be worthwhile to briefly present research showing the relationship between clinical care and presenteeism. The authors could also briefly introduce other factors that could play a role in presenteeism (e.g., nature of the health condition, organizational factors, availability of legislative supports).</li> <li>• A strength of your study was the large sample. However, the response rate was 12.5%. Could this be a source of bias and the potential for the healthy worker effect.</li> <li>• Few details were presented on the informed consent process.</li> <li>• One critique is the temporality of the survey and the research questions. The survey was conducted between August and September 2020. However, questions on presenteeism ask about the last 30 days. Meanwhile, questions on healthcare disruption focus on the period of April to May 2020. The difference in the period seems like a critical limitation of the survey questions.</li> <li>• Relatedly, the authors excluded those who were unemployed. Past research shows a relationship between presenteeism and loss of employment. Would there be any way to account for those who might have suffered presenteeism at the beginning of the pandemic then lost their job at the time of the survey?</li> <li>• Greater clarity on the labor type and employment pattern variables are required. These variables were unfamiliar to the reviewer. Perhaps a few more details would be helpful on how and why they were coded in the manner reflected by the authors.</li> <li>• The authors talk about measuring sex, which is a biological variable. How was this assessed? Perhaps the authors actually measured gender.</li> <li>• I am curious as to why the authors chose to rely on p values when they produced confidence intervals.</li> <li>• The authors make reference to high and low presenteeism cutoffs that were established based on past research. Perhaps a few more details could be provided to help the reader of this</li> </ul>
--	---

	<p>specific paper understand how these cutoffs were established and implemented</p> <ul style="list-style-type: none"> <li>• Additional clarity on the difference between Model 1 and Model 2 would be helpful. Based on the description provided by the authors in the text and table, it's not completely clear to the reviewer. Also, were the authors able to establish population weights and integrate the weights in the model?</li> <li>• During the COVID-19 pandemic most workers experienced lost productivity (regardless of having a chronic health condition). Would it be possible to tease apart presenteeism related to a chronic health condition vs. at-work lost productivity related to the stress of the pandemic that may not be attribute to a chronic condition? Similarly could the authors account for chronic conditions that were acquired during the pandemic (e.g., anxiety, depression)</li> <li>• Perhaps some details on the representatives of the sample that was recruited would be helpful for the reader.</li> </ul>
--	---

### VERSION 1 – AUTHOR RESPONSE

Reviewer #1: Dr. Sascha Ruhle, Heinrich Heine University Düsseldorf

We are grateful for your guidance and encouragement for improving our article. Especially because the first author spent his childhood in Düsseldorf for his parents' work, we are pleased to have the guidance of a prominent researcher in this city.

#### Major comments

1) I found the introduction is well written and interesting (aside from the minor aspect below). However, the way the introduction is written, I was expecting a stronger connection towards previous results in the field of presenteeism (e.g., Miraglia & Johns 2016, Ruhle et al. 2020). More specifically, I was expecting information on how presenteeism could be related to disrupted care. What are your conceptual thoughts on that relationship? Does disruptive care follow one of the dual paths? Via general health? Or is it related more to a process perspective that focuses on the individuals' personal experiences? It could be useful to include recent developments in the field of presenteeism into these thoughts. Overall, I agree that there might be a plausible relationship, yet, please share your ideas with the reader here. Otherwise, concerning the STROBE checklist (which I appreciate), I find the objectives and specific hypotheses too weak to be considered for publication.

2) Overall, I would expect you to think about individuals' health as an important predictor of presenteeism more in detail. This is related to your conceptual background, in which health as a driver of sickness presenteeism should be included, and in your studies design. Do you have any information on the individual's (subjective) health? Otherwise, one could critically state that you found that predictors of reduction of health are related to sickness presenteeism, which is already known (Johns, 2010). If you don't have the respective information, please nonetheless consider the role of health and add it as a rather problematic limitation of your study.

Thank you for these valuable comments and for introducing several important references. We have clarified the position of our study in the conceptual model of presenteeism. Specifically, the current study defined presenteeism as reduced performance at work. We then hypothesized that disrupted

care has a negative impact on workers' health status, resulting in reduced performance at work. We decided to rewrite the full manuscript, including the "Introduction" section, and to revise the statistical analyses to better fit the conceptual model of presenteeism in this study. Please see the attached manuscript for these revisions.

3) Methods: First, did you analyze your drop-out? Or what were the invalid responses you used to remove these individuals from the sample? Please make sure to test whether the reduction was random, and not based on some important constructs in your model (e.g., (un-)healthy individuals). If you decide to not include participants in the primary analysis, please make sure that the reasons for the exclusion are made transparent (and justified).

Thank you for this comment. Because this was a cross-sectional study, we did not analyze dropout as would be done in a cohort study. We have added Figure 1—a flow diagram to help the readers understand the reasons for excluding participants. Please see the attached manuscript.

Specifically, we excluded participants who provided incorrect answers to a question intended to identify fraudulent responses ( $n = 1,955$ ), those who selected all medication uses ( $n = 422$ ), and those who selected all underlying diseases ( $n = 141$ ). We considered it clear that these individuals provided unreliable responses.

Additionally, we excluded 10,028 persons who were currently unemployed. Unemployed participants were not asked the questions related to presenteeism because presenteeism can only be measured among people who are working.

4) Methods: Why did you separate labor type into "manual work, desk work, and other."? I am unfamiliar with this separation and not sure what each of these would encompass. Especially as "other" is almost a fourth of your sample, I would like to have more information on this decision. More specifically, there are a lot of different dimensions of labor (e.g., physical, mental, skilled, unskilled, manual, etc.). While I don't expect that you delve deeper into these differences (e.g., Narotzky 2018), I would expect a clearer description of your decision, especially as we know that sickness presenteeism is related to different aspects of labor, such as e.g. the level of education (Gustafsson & Marklund, 2011) or the nature of work (Hansen & Andersen, 2008).

Thank you for this comment. We apologize for this wording being misleading. The question on labor type referred to blue-collar, white-collar, and pink-collar jobs. We asked this question because of differences in the mechanisms of presenteeism by job type. For example, physical health issues are more apparent in blue-collar jobs. We have revised the relevant part of the "Methods" section as follows:

"Job type was categorized as blue-collar, white-collar, or other jobs. Other jobs mainly comprised "pink-collar" jobs such as customer service, retail, and nursing care work.<sup>24</sup>)"

5) Methods: How did you handle missing values? Did any occur?

Thank you for this comment. Because this was an online survey and all questions had to be answered to complete the survey, there were no missing values.

6) Methods: Further, did you test the assumptions of the multinomial logistic regression? Did you perform any additional analysis to provide evidence for the robustness of your effects? Overall, I

would like to see more statistical work to better understand the robustness of your results. Please draw on adequate statistical literature to justify and clarify your approach (e.g. Stoltzfus 2011).

We appreciate the reviewer's advice here and have therefore revised the relevant parts of the "Methods" and "Results" sections as follows:

Methods: "Goodness of fit was assessed using the Hosmer–Lemeshow test."

Results: "The Hosmer–Lemeshow test confirmed the goodness of fit of the adjusted model ( $P > 0.20$ )."

7) The discussion is rather weak. In my view, this is again related to the missing inclusion of general health as a mechanism that explains how disrupted care is related to sickness presenteeism. Again, stating that "Presenteeism was associated with chronic illnesses even among workers whose care continued." is not surprising, as we know that health is an important aspect for the vulnerability of an individual (Gerich, 2015). In line with this reasoning, I found your explanation of the exacerbation of underlying disease as an argument for the different results from models 1 and 2 not convincing. Further, I stumbled over the mention of "unexpected symptoms" as one identified aspect of your study on page 12. You state that "This study identified that unexpected symptoms or illnesses during the pandemic were associated with presenteeism, especially where care had been disrupted." However, unexpected symptoms are part of what you describe as a factor of disrupted care. Again, the conceptual background for this was not clear.

Thank you for this valuable comment. In this study, we hypothesized that disrupted care has a negative impact on workers' health status, resulting in lower productivity. According to this hypothesis, we have changed the statistical analyses and rewritten the "Discussion" section. Although our findings are consistent with Gerich's study, as a novel finding, our study revealed the negative impact of disrupted care on presenteeism for non-COVID-19 patients. Therefore, we revised the relevant part of the "Discussion" section as follows:

"The current study found that experiencing disruptions in routine and non-routine clinical settings and running out of drugs were strongly associated with exacerbation of underlying disease. Furthermore, workers experiencing the exacerbation of underlying disease were much more likely than workers without this experience to subsequently show reduced performance at work (presenteeism). This finding is consistent with a previous study: Gerich showed that presenteeism is strongly influenced by the frequency of health events.<sup>27</sup> Our study provides insight into the possible harmful impact of reduced accessibility of care on presenteeism for non-COVID-19 patients."

We have also added Reference 27: Gerich J. Sick at work: methodological problems with research on workplace presenteeism. *Health Services and Outcomes Research Methodology* 2015;15:37-53.

In addition, we apologize that the phrase "unexpected symptoms" was misleading. In this study, we did not include COVID-19 patients or emergency care patients, so "unexpected symptoms" was intended to imply mild illnesses. We have revised this phrase and added the following information in relevant parts of the "Methods" section:

"In this study, we hypothesized that disrupted care has a negative impact on workers' health status, resulting in increased presenteeism. Here, the scope of care includes all physical and mental illnesses except for COVID-19 infection. We asked about three types of events related to disrupted care during the COVID-19 state of emergency (April and May 2020) and one event related to the negative impact on self-reported health status. The three types of events related to disrupted care were as follows:

1) disruptions in non-routine clinical settings (mostly not medical emergencies but minor illnesses, such as slight fever, wounds, or diarrhea);"



8) Finally, the contribution of the manuscript needs to be clarified. You state that you “showed that disrupted care may influence workers’ subsequent performance”. However, you did not measure productivity/performance as an outcome. Further, that dealing with acute illness will be beneficial for the health of an employee is well known, and that this will reduce the likelihood of sickness presenteeism is also established (e.g., Johns, 2010; Ruhle et al. 2020).

Thank you for this comment. Because the current study defined presenteeism as reduced performance at work, we have carefully rewritten the conclusion paragraph of the “Discussion” section as follows:

“In conclusion, our study showed that workers who experienced disrupted care were much more likely than others to subsequently show increased presenteeism, defined as reduced performance at work. Exacerbation of underlying disease is one possible pathway through which disrupted care could affect presenteeism. Our study provides evidence of the importance of early diagnosis and continuous treatment of non-COVID-19 patients to enable them to remain healthy and continue to work during the pandemic. Increasing accessibility of care for patients, for example by offering telemedicine appointments and drug delivery, could help workers to maintain their performance at work.”

#### Minor comments

9) Abstract: The description of the “Primary and secondary outcome measures” is unfitting. Please revise and make sure it does not overlap with the Methods. Better, clearly describe the outcomes you were interested in.

Thank you for this valuable suggestion. This point is our mistake. We have deleted the subheadings of “Settings,” “Participants,” and “Primary and secondary outcome measures” from the abstract.

10) Related to your conceptual background, you might want to clarify that you are talking about sickness presenteeism. Recent developments suggest that there might be different types of presenteeism.

Thank you for this comment. As you point out, this paper focuses on presenteeism, defined as reduced performance at work. The “Introduction” section has been revised as a whole to make the concept of presenteeism in this study clearer. In addition, we have revised the relevant part of the “Methods” section as follows:

“The WFun, which evaluates “the degree to which the ability to function at work is impaired by health problems,”<sup>19</sup> was originally developed in Japan and has shown good correlation with measures of different types of presenteeism that have been proposed by scholars in recent years.<sup>19-21</sup>”

11) P.7: “The total score, therefore, ranged from 7 to 35 points. Higher scores indicate worse presenteeism.” I would restrain from using “worse” here, as it is higher, but not necessarily worse in all dimensions. Karanika-Murray and Biron (2020) make a compelling case that more sickness presenteeism is not always worse for the individual.

Thank you for pointing this out. We agree with your opinion and have revised the relevant parts of the “Methods” section as follows:

“Total WFun scores therefore range from 7 to 35 points. Higher scores indicate lower performance at work.”

12) There are some smaller errors in language and grammar. While I am not a native speaker, I would suggest rigorous proofread.

Thank you for your comments. Our manuscript was checked by a professional English editing service before submission, and the fact that there were many grammatical errors in the submitted manuscript is a critical problem for us. We therefore sent the reviewer's valuable comment to the English editing service directly and requested that they recheck the entire revised manuscript.

Reviewer #2: Dr. Jeffrey Vietri, Kantar Health Inc

We very much appreciate your comments and suggestions for improving our article. We have made substantial revisions to the manuscript in accordance with your suggestions. Please see our following responses for a detailed description of these changes.

1) I was confused by the different levels, and a figure depicting what selection / sampling method was used at each level would be helpful.

Thank you for this suggestion. We have added a flow diagram depicting the sample selection process as Figure 1. Please see the attached manuscript.

In addition, we have revised the relevant part of the "Methods" section as follows:

"The population comprised panelists aged 15 to 79 years who were registered with an Internet research company (approximately 2.2 million people). Simple random sampling was used to select a survey population of 223,389 people, who were invited via e-mail to complete the survey. Participation was fully voluntary. After transitioning to the linked survey site, participants provided informed consent. We recruited participants in clusters by sex, age (in 5-year age bands), and prefecture (47 regions) to reflect nationally representative statistics.<sup>18</sup> We stopped recruitment when the target numbers of participants for each sex, age, and prefecture category were reached. A total of 28,000 participants responded to the survey (response rate = 12.5%)."

2) Some explanation of what is meant by 'invalid responses' should be provided, as nearly 10% of the sample was excluded for this reason.

Thank you for this valuable comment. We have added information on invalid responses in Figure 1. Please see the attached manuscript.

3) There is a mention of the exchange rate between dollars and yen that is unnecessary, as dollars are not reported in the paper.

Thank you for this suggestion. Upon review, we decided to delete this information from the manuscript.

4) Lastly, the paper would benefit from additional description and justification of the treatment of the variables and different multi variable models. For instance, why adjust for the different types of disruption in model 2 instead of creating an index? Are the results consistent if the measure of presenteeism is used as continuous rather than dichotomized?

Thank you for this valuable comment. In this study, we hypothesized that disrupted care has a negative impact on workers' health status, resulting in lower performance at work. According to this hypothesis, we have changed the statistical analysis. Specifically, we performed two types of analyses to make the concepts tested in this study clearer: First, univariate and multiple logistic regression analyses were used to investigate the associations among variables related to health status, variables related to disrupted care, and WFun score. Second, the same statistical techniques were used to evaluate the association between disrupted care and health status. Please see the attached manuscript.

To address the latter part of your comment, we also conducted a statistical analysis on a continuous scale and obtained robust results. Because this survey focused on people whose work was impaired by reduced productivity, we selected a binary variable. We have now added this information in the "Methods" section as follows:

"WFun score was classified into two groups, in line with a previous study<sup>25</sup>): 7 to 20 points was considered low presenteeism, and 21 to 35 points was considered high presenteeism. A WFun score of 21 or higher requires consideration of accommodations and adjustments in the workplace for workers' illnesses,<sup>25</sup>) and a score of 25 or higher increases the risk of workers taking sick leave.<sup>26</sup>."

Reviewer #3: Dr. Arif Jetha, Institute for Work & Health

We very much appreciate your comments and suggestions for improving our article. We have made substantial revisions to the manuscript in accordance with your suggestions. Please see our responses below for details on the revisions made.

1) The author uses the term "coronavirus disease 2019". I would ensure throughout the papers that labels "Novel coronavirus", "COVID-19" and "SARS-COV-2" are used consistently and align with published literature and media.

Thank you for this comment. We think it is important to note that this virus has many different names. We used the term "coronavirus disease 2019 (COVID-19)" because this is the official name issued by the World Health Organization. The current study did not assess pathology; therefore, we did not use the term for the pathogen causing the disease, SARS-CoV-2.

2) There is repetition in the abstract between setting, participants and methods subheadings. I suggest being more concise.

Thank you for this valuable suggestion. This point is our mistake. We have deleted the subheadings of "Settings," "Participants," and "Primary and secondary outcome measures" from the abstract.

3) Also, I think the strengths of the study could be enhanced to fully capture the innovative aspects of this study.

Thank you for this valuable suggestion. We have now added the following information at in the "Strengths and limitations of this study" section:

"This study provides evidence of the importance of early diagnosis and continuous treatment of non-COVID-19 patients to enable them to remain healthy and continue to work during the pandemic."

4) Perhaps it would be helpful to briefly define chronic disease with regards to this specific study. Are you referring to conditions that are both physical and mental? Would a person who has experienced persistent COVID-19 symptoms (e.g., COVID long-hauler) be captured in your definition? Additional details are required on how people with chronic disease were recruited. Did the authors rely on self-reports?

Thank you for this suggestion. We have revised the relevant part of the “Methods” section as follows: “In this study, we hypothesized that disrupted care has a negative impact on workers’ health status, resulting in increased presenteeism. Here, the scope of care includes all physical and mental illnesses except for COVID-19 infection. We asked about three types of events related to disrupted care during the COVID-19 state of emergency (April and May 2020) and one event related to the negative impact on self-reported health status.”

5) It would also be worthwhile to briefly present research showing the relationship between clinical care and presenteeism. The authors could also briefly introduce other factors that could play a role in presenteeism (e.g., nature of the health condition, organizational factors, availability of legislative supports).

Thank you for this suggestion. We decided to rewrite the full manuscript, including the “Introduction” section, to introduce other factors that could play a role in presenteeism. Please see the attached manuscript.

In addition, we have now added the following information in the relevant part of the “Methods” section: “WFun score was classified into two groups, in line with a previous study<sup>25</sup>): 7 to 20 points was considered low presenteeism, and 21 to 35 points was considered high presenteeism. A WFun score of 21 or higher requires consideration of accommodations and adjustments in the workplace for workers’ illnesses,<sup>25</sup>) and a score of 25 or higher increases the risk of workers taking sick leave.<sup>26</sup>)”

6) A strength of your study was the large sample. However, the response rate was 12.5%. Could this be a source of bias and the potential for the healthy worker effect.

Thank you for this suggestion. We have now added this information in the paragraph discussing study limitations as follows:

“Additionally, the response rate of this study was relatively low (12.5%). The results should also be interpreted carefully because of the healthy worker effect.”

7) Few details were presented on the informed consent process.

Thank you for this suggestion. We have now added this information in the “Methods” section as follows:

“Simple random sampling was used to select a survey population of 223,389 people, who were invited via e-mail to complete the survey. Participation was fully voluntary. After transitioning to the linked survey site, participants provided informed consent.”

8) One critique is the temporality of the survey and the research questions. The survey was conducted between August and September 2020. However, questions on presenteeism ask about the last 30 days. Meanwhile, questions on healthcare disruption focus on the period of April to May 2020. The difference in the period seems like a critical limitation of the survey questions.

Thank you for your suggestion. We have now added this information in the paragraph discussing study limitations as follows:

“Because of the study’s cross-sectional design, causal relations among disrupted care, health status, and presenteeism (defined as reduced performance at work) could not be established. To cope with this limitation, we asked about each issue using different time periods: experience of disrupted care during the COVID-19 state of emergency (April and May 2020) and presenteeism 3 months later (August and September 2020). However, recall bias may have been an issue. Furthermore, workers who experienced disrupted care might have been more likely to remember health problems, which may have caused an overestimation of presenteeism. In addition, cases of resumed care during this 3-month period may have attenuated the effects of disrupted care.”

9) Relatedly, the authors excluded those who were unemployed. Past research shows a relationship between presenteeism and loss of employment. Would there be any way to account for those who might have suffered presenteeism at the beginning of the pandemic then lost their job at the time of the survey?

Thank you for this comment. The current study evaluated presenteeism, defined as reduced work performance. Presenteeism can only be measured in people who are working; thus, we did not analyze unemployed persons in this study.

10) Greater clarity on the labor type and employment pattern variables are required. These variables were unfamiliar to the reviewer. Perhaps a few more details would be helpful on how and why they were coded in the manner reflected by the authors.

Thank you for this comment. We apologize for this wording being misleading. The question on labor type referred to blue-collar, white-collar, and pink-collar jobs. We asked this question because of differences in mechanisms of presenteeism by job type. For example, physical health issues are more apparent in blue-collar jobs. We have revised the relevant part of the “Methods” section as follows: “Employment pattern was categorized as permanent employee, company executive, temporary employee, part-time employee, or self-employed. New employment patterns are emerging as the labor market changes. Therefore, we asked about both classic employee patterns (e.g., permanent employment) and new employment patterns (e.g., temporary employment and self-employment). Job type was categorized as blue-collar, white-collar, or other jobs. Other jobs mainly comprised “pink-collar” jobs such as customer service, retail, and nursing care work.<sup>24</sup>) We also asked about 16 types of illnesses, including hypertension, diabetes, and asthma (listed in full in Table 1).”

11) The authors talk about measuring sex, which is a biological variable. How was this assessed? Perhaps the authors actually measured gender.

Thank you for pointing this out. We have now changed the term “sex” to “gender” throughout the text.

12) I am curious as to why the authors chose to rely on p values when they produced confidence intervals.

Thank you for this suggestion. There are two ways to view a statistical hypothesis test: One is through a P-value (of the test), and the other is through a CI (of a parameter). The P-value and CI are complementary because they attempt to do the same thing (or similar things): The P-value quantifies

how “significant” the association/difference is, whereas the CI quantifies how “precise” the estimation is and identifies the plausible values. A few years ago, some journals instructed authors to eliminate P-values and use only 95% CIs, but journals are now increasingly using both. Reporting the P-values is also potentially useful for future meta-analyses.

13) The authors make reference to high and low presenteeism cutoffs that were established based on past research. Perhaps a few more details could be provided to help the reader of this specific paper understand how these cutoffs were established and implemented.

Thank you for this suggestion. We have now added the following information in the “Methods” section: “WFun score was classified into two groups, in line with a previous study<sup>25</sup>): 7 to 20 points was considered low presenteeism, and 21 to 35 points was considered high presenteeism. A WFun score of 21 or higher requires consideration of accommodations and adjustments in the workplace for workers’ illnesses,<sup>25</sup>) and a score of 25 or higher increases the risk of workers taking sick leave.<sup>26</sup>)”

14) Additional clarity on the difference between Model 1 and Model 2 would be helpful. Based on the description provided by the authors in the text and table, it’s not completely clear to the reviewer. Also, were the authors able to establish population weights and integrate the weights in the model?

Thank you for this valuable comment. In this study, we hypothesized that disrupted care has a negative impact on workers’ health status, resulting in presenteeism, defined as reduced performance at work. According to this hypothesis, we have changed the statistical analyses. Specifically, we performed two types of analyses to make the concepts tested in this study clearer: First, univariate and multiple logistic regression analyses were used to investigate the associations among variables related to health status, variables related to disrupted care, and WFun score. Second, the same statistical techniques were used to evaluate the association between disrupted care and health status. Please see the attached manuscript.

Regarding the latter part of your comment, this may overlap with our response to the editor’s second question, which reads as follows:

“Propensity score methods are normally used for estimating the effect of a binary intervention indicator, accounting for a response variable and background observed covariates. Thus, we did not perform propensity score matching because this study attempted to compare the three groups of “no illness,” “illness without event,” and “illness with event” for each explanatory variable. We did attempt propensity score matching between the “illness without event” and “illness with event” groups, but the areas under the curve were less than 0.8, making our data ineligible for this type of analysis.”

If not, we would appreciate it if you would let us know so that we can provide another response.

15) During the COVID-19 pandemic most workers experienced lost productivity (regardless of having a chronic health condition). Would it be possible to tease apart presenteeism related to a chronic health condition vs. at-work lost productivity related to the stress of the pandemic that may not be attribute to a chronic condition? Similarly could the authors account for chronic conditions that were acquired during the pandemic (e.g., anxiety, depression).

Thank you for this suggestion. We considered it difficult to assess the psychological effects of the COVID-19 pandemic separately from the psychological effects of disrupted care, so we did not adjust for psychological factors in this study. We have now added this information in the paragraph describing the study limitations as follows:

“Another limitation is that we did not specify which underlying disease was associated with disrupted care. We also did not evaluate whether underlying diseases developed before or after the outbreak of the COVID-19 pandemic. We adjusted for potential causes in the statistical analysis, but future studies should consider this issue to clarify the relationship between underlying diseases and disrupted care.”

16) Perhaps some details on the representatives of the sample that was recruited would be helpful for the reader.

Thank you for this suggestion. We have added a flow diagram to help the readers understand the representativeness of the sample (Figure 1). Please see the attached manuscript.

In addition, we have revised part of the “Methods” section as follows:

“The population comprised panelists aged 15 to 79 years who were registered with an Internet research company (approximately 2.2 million people). Simple random sampling was used to select a survey population of 223,389 people, who were invited via e-mail to complete the survey. Participation was fully voluntary. After transitioning to the linked survey site, participants provided informed consent. We recruited participants in clusters by sex, age (in 5-year age bands), and prefecture (47 regions) to reflect nationally representative statistics.<sup>18</sup> We stopped recruitment when the target numbers of participants for each sex, age, and prefecture category were reached. A total of 28,000 participants responded to the survey (response rate = 12.5%).”

#### VERSION 2 – REVIEW

<b>REVIEWER</b>	Ruhle, Sascha Heinrich Heine University Düsseldorf, Business and Economics
<b>REVIEW RETURNED</b>	15-Jul-2021

<b>GENERAL COMMENTS</b>	<p>Thank you for the opportunity to review the revise of your article “Disrupted care during the COVID-19 state of emergency and subsequent presenteeism in workers: a nationwide cross-sectional study”. As this is a revision, I will focus on my previous comments that I feel need more explanation, my evaluation of the changes in the manuscript, as well as the added information.</p> <p>As I will not remark on the well-revised comments, let me start by saying that I liked the thorough approach you have taken towards the revision. A lot of aspects are now much clearer to me and you managed to address several of the issues raised by me and the other reviewer. So consider the comments not addressed as very fittingly revised.</p> <p>Overall, your article still has some minor flaws that need to be addressed. In the following, first I will draw on the unfinished comments from the previous round (R&amp;R Comment #) before addressing the new aspects in order of importance:</p> <p>R&amp;R Comment 3: You are correct in pointing out that dropout is related to longitudinal studies. I actually meant the excluded individuals. Figure 1 now clearly describes the exclusion process. However, I would not use the term “invalid response” in the figure, as you describe in your response that they are unreliable responses. That would be the more fitting term here.</p>
-------------------------	--

	<p>R&amp;R Comment 4: Thank you! Have you more information on that, or is it just a broad umbrella term? Because some of these jobs again have specific circumstances related to presenteeism (e.g., caring for others). I am not familiar with the term (there is no such term in German) which seems to be related to jobs that are primarily done by women (in the past). Just make sure that this term is still adequate and not discriminating.</p> <p>Major</p> <p>1. I would still urge you to clarify your understanding of presenteeism. (This is still related to my previous comments). You start your manuscript with “ Presenteeism, the phenomenon of attending work despite being ill is closely related to work performance.” Here, you clearly and correctly separate presenteeism (attending work while ill) and its consequences (Productivity loss attributable to presenteeism). Therefore, please use the term productivity loss attributable to presenteeism or productivity loss in short, but not presenteeism. Otherwise, you will add to the existing blurriness of the construct. I know that you try to add the information (sometimes in the text, sometimes in the [brackets]). But just switch to productivity loss attributable to presenteeism. Then, you don't have this problem. For me, this is a very major and very important point. As you now have written, you measure productivity loss and not presenteeism!</p> <p>2. I still would like to see a little bit more explanation on disrupted care. For the unfamiliar reader, you don't describe what disrupted care is, or explain the relationships you think exist in detail. So please, provide additional information and compelling arguments here. To some extent, you provide that in the Independent variables section, but I was not able to understand your reasoning behind the relationships (e.g., why should “running out of drugs” impact the productivity loss due to presenteeism?)</p> <p>Again, I hope my comments will help you to improve the manuscript and wish you the best of luck with your project.</p>
--	--

## VERSION 2 – AUTHOR RESPONSE

Reviewer #1: Dr. Sascha Ruhle, Heinrich Heine University Düsseldorf

We very much appreciate your comments and suggestions for improving our article. We have made substantial revisions to the manuscript in accordance with your suggestions. Please see our responses below for details on the revisions made.

1)R&R Comment 3: You are correct in pointing out that dropout is related to longitudinal studies. I actually meant the excluded individuals. Figure 1 now clearly describes the exclusion process. However, I would not use the term “invalid response” in the figure, as you describe in your response that they are unreliable responses. That would be the more fitting term here.



Thank you for this valuable advice. We have revised the term “invalid response” to “unreliable responses” throughout the text and in the figure.

2)R&R Comment 4: Have you more information on that, or is it just a broad umbrella term? Because some of these jobs again have specific circumstances related to presenteeism (e.g., caring for others). I am not familiar with the term (there is no such term in German) which seems to be related to jobs that are primarily done by women (in the past). Just make sure that this term is still adequate and not discriminating.

Thank you for this suggestion. In response to the previous comment, we have revised the category of job type as blue-collar, white-collar, or other jobs. These jobs were adjusted as covariates because it was possible that these different job categories would show different patterns of presenteeism, as you have pointed out.

In line with the reviewer’s suggestion, we again investigated the international and academic use of these terms. We found that these terms are treated as non-discriminatory both internationally and academically, although there are regional differences: for example, “blue collar” has a negative connotation in several countries. However, some blue-collar jobs require a national qualification. For example, in Japan and Germany, skilled blue-collar workers are perceived to have an equal or higher status compared with white-collar workers, and in communist countries, blue-collar jobs are considered to be dignified. For these reasons, we have decided to maintain the current terminology.

3)I would still urge you to clarify your understanding of presenteeism. (This is still related to my previous comments). You start your manuscript with “Presenteeism, the phenomenon of attending work despite being ill is closely related to work performance.” Here, you clearly and correctly separate presenteeism (attending work while ill) and its consequences (Productivity loss attributable to presenteeism). Therefore, please use the term productivity loss attributable to presenteeism or productivity loss in short, but not presenteeism. Otherwise, you will add to the existing blurriness of the construct. I know that you try to add the information (sometimes in the text, sometimes in the [brackets]). But just switch to productivity loss attributable to presenteeism. Then, you don’t have this problem. For me, this is a very major and very important point. As you now have written, you measure productivity loss and not presenteeism!

We appreciate the reviewer’s advice here and have revised the terminology as “productivity loss attributable to presenteeism” or “productivity loss” throughout the text and tables.

4)I still would like to see a little bit more explanation on disrupted care. For the unfamiliar reader, you don’t describe what disrupted care is, or explain the relationships you think exist in detail. So please,

provide additional information and compelling arguments here. To some extent, you provide that in the Independent variables section, but I was not able to understand your reasoning behind the relationships (e.g., why should “running out of drugs” impact the productivity loss due to presenteeism?)

Thank you for this valuable comment. We have revised and added the following information in relevant parts of the Introduction section:

“We hypothesized that disrupted care during the state of emergency had a negative impact on workers’ health status and presenteeism, resulting in productivity loss. For example, workers with back pain are often faced with productivity loss owing to their pain.<sup>18</sup>) However, if they exhaust their supply of painkillers, the pain may worsen and further reduce their productivity. Another example is that if depressed workers have been unable to see a doctor, they may continue to work as their condition worsens owing to a lack of medical advice about sick leave. However, little is known about the relationship between disrupted care and productivity loss attributed to presenteeism during the COVID-19 pandemic. Therefore, we aimed to investigate whether disrupted care during the COVID-19 state of emergency was associated with health status and productivity loss.”

We hope that the manuscript is now suitable for publication in BMJ Open.

Again, thank you for your valuable comments and suggestions for improving our article.