PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	The association between mental health trajectories and somatic symptoms following a second lockdown in Israel: A longitudinal study
AUTHORS	Ben-Ezra, Menachem; Hamama-Raz, Yaira; Leshem, Elazar; Levin, Yafit; Goodwin, Robin

VERSION 1 – REVIEW

REVIEWER	Anne Duffy Queens University, Psychiatry
REVIEW RETURNED	23-Apr-2021

GENERAL COMMENTS	Study sought to describe the trajectory of problems with anxiety disorders from before to following a second lockdown owing to the COVID-19 pandemic. Study sampled 1029 adults 18 years+ in Israel from an internet pool of 100,000 and collected self-report data at T1 prior to the second lockdown and followed up at T2 post-second lockdown (76% of the original sample n=764). The study also explored the relationship between anxiety trajectories (mental health) and somatic symptoms. The impact of COVID-19 on health and mental health of the general and high-risk populations is of great interest currently. That said, this study has several limitations which diminish enthusiasm for publication in the BMJ Open and suggest the paper might be better suited to a journal with more focused readership ie psychology. Specific limitations include: selection bias in terms of recruitment from the internet and no methods to secure a representative sample from the referent general population; limited (non-robust) measurement of the main mental health outcomes and no baseline prior to the pandemic itself. These limitations are substantial and not rectifiable. In addition, the authors have not made a strong case as to the novelty and impact of this study. Other comments for the authors to consider include: Adjustment disorder while common is not strongly linked to well-being or health outcomes. A stronger design would be to focus on screen positive for anxiety disorder using valid measures and stratify by lifetime history of mental disorder. What is the collinearity between somatic symptom ratings and GADS-7 – I would suspect that somatic symptoms are simply a barometer for level of anxiety and unclear as to what the main analysis between anxiety and somatic scores really adds? I also feel the Title, Abstract and Conclusions go beyond what one can conclude from the study given the limitations and that the language/writing could be made more clear and accurate overall.

REVIEWER	Robert Stewart Institute of Psychiatry, Section of Epidemiology (Box 60)
REVIEW RETURNED	28-Apr-2021

GENERAL COMMENTS	This paper reports a 2-wave panel survey of a community cohort in Israel in which the authors investigated changes in anxiety/adjustment disorder before and after a second lockdown, and associations with somatic symptoms at the end of that period. My comments are as follows: 1. I think the title might be amended to provide a more precise definition of the study, although I think this is more a matter for the Editorial Office and journal house style. 2. Given the design of the study, I think that 'somatic toll of a second lockdown' is too strong a term in the conclusions section of the abstract (and elsewhere, including the title). Although the study investigated this particular period of time, there is nothing in the design that permits causality to be assumed in relation to lockdown, as the results might well have been the same from an identical study carried out at a different time period. This is not intended to detract from the quality or utility of the study; it's just that I think the conclusions are over-stated. 3. The third paragraph of the Introduction drifts into considering the methodology of the study about to be described, which feels out of place. 4. The reader really needs more detail on the source population for sampling. Currently we are simply told that there are 100,000 members 'designed to be representative'. I would have expected some description of how the iPanel company identified and approached its members, and the achieved level of national representativeness (I assume there must be some public domain data on this). Otherwise we really don't know what the analysed sample represents (although I accept that some comparisons are made between the analysed sample and national data). 5. In relation to the earlier point about causation, the IADQ instrument for adjustment disorder is described as including a stressors list covering different aspects of life. It would be helpful if the authors could clarify whether that stressors list was adapted in any way to be specific to lockdown-related stressors (as their conclusions seem to indicate that the anxiety exposure is reflecting life under lockdown). Or are these generic stressors (in which case, how is the lockdown-related conclusion justified)? 6. There seems to be no description of other variables in the Methods section. For example, I can't find any description of what 'a little below average' income means and whether this was applied as a definition to stated income level or was how a participant referred to their income level (and, if so, how they would be expected to know the average. 7. Also, what does the Covid risk group variable mean, and how was it calculated? 8. In the Discussion (2nd paragraph) there is a mention of another study having presented pandemic-related changes in anxiety/depression as a single slope. I have to say that this seems a more logical approach than the 4 trajectory categories imposed by this paper (which necessarily conflate substantial changes in symptoms with small changes around the threshold points). It might help, therefore, if the authors provided more clarity on why they feel their approach is advantageous. 9. The authors do acknowledge the single measure of somatic symptoms as a limitation in their Discussion; however, this again

	necessitates much more cautious language about conclusions in other areas of the paper. A 'somatic toll of a second lockdown' is simply not being conclusively demonstrated here - the somatic symptoms might have been present all along, and the association between anxiety and somatic symptoms might have been present in any similar cohort at any time, regardless of the lockdown occurring. Finally, there are relatively few covariates considered and residual confounding is a potential concern.

VERSION 1 – AUTHOR RESPONSE

Reviewer 1

Reviewer: 1: Dr. Anne Duffy, Queens University

Comment 1:" Specific limitations include: selection bias in terms of recruitment from the internet and no methods to secure a representative sample from the referent general population".
Response: Following the Reviewer's comment, we want to clarify that we took measures to obtain a representative sample from the general population. We used quota sampling for this purpose. The sample is representative to the Israeli general population based on the Israeli Bureau of statistics Census for age and sex. The iPanel is a probability based panel (Bodas et al., 2018) adheres to the stringent standards of the world association for market, social, and opinion researchers (ESOMAR).Moreover, selection bias in terms of recruitment from the internet is reduced to a minimum as internet penetration in Israel is 88% and the number of mobile connections in Israel is 116.9% of the population (more than one mobile phone per capita). See https://datareportal.com/reports/digital-2021-israel for more information. This is important as the gap between the general population and internet users in some countries is being closed over time. It is expected that in the next couple of years some countries will reach 100% internet penetration thus making internet panels and surveys as the first choice for population-based studies as the probability of individual to be selected will be the same as traditional sampling methods.
This is important as the gap between the general population and internet users in some countries is being closed over time. It is expected that in the next couple of years some countries will reach 100% internet penetration thus making internet panels and surveys as the first choice for population-based studies as the probability of individual to be selected will be the same as traditional sampling methods.
Comment 2: …"limited (non-robust) measurement of the main mental health outcomes".
Response: Following the Reviewer's comments, the measurements we have used are widely used in Psychiatry and Medicine. These instruments are considered robust indices and outcomes in Psychiatry and Medicine.
GAD-7 is widely used in Psychiatry and Medicine. Please see: Recent research in the BMJ Open (https://bmjopen.bmj.com/content/11/1/e045794), The Lancet (https://www.thelancet.com/journals/lancet/article/PIIS0140-6736(17)32133-5/fulltext), JAMA (https://jamanetwork.com/journals/jama/fullarticle/2771608?resultClick=1) and NEJM (https://www.nejm.org/doi/full/10.1056/NEJMoa1811424).
IADQ is the current standard to measure adjustment disorder based on the newly released ICD-11. It has been used in Psychiatry in the British Journal of Psychiatry (https://www.cambridge.org/core/journals/the-british-journal-of-psychiatry/article/abs/network-structure-of-icd11-adjustment-disorder-a-crosscultural-comparison-of-three-african-countries/676804370018D387F6FBF68549061794#supplementary-materials).
SSS-8 is used in Medicine and published in the JAMA Internal Medicine (https://jamanetwork.com/journals/jamainternalmedicine/article-abstract/1783305) and in the BMJ Open (https://bmjopen.bmj.com/content/9/1/e025200.abstract).

Comment 3: …"and no baseline prior to the pandemic itself".
Response: following the Reviewer's comment, we agree with the reviewer that we do not have a baseline prior to the pandemic itself. However, research published in leading psychiatric and medical journals suffers from the same limitation. We address this limitation in the limitation's sections. Furthermore, we aimed to examine before and after second lockdown, and not to infer causality between the period before-COVID-19 and after it. For this reason we found it important to omit any indication of causality in the manuscript.
Comment 4:…"the authors have not made a strong case as to the novelty and impact of this study".
Response: Following the reviewer's comment, we respectfully disagree. The study has three main strengths: 1st: This is the first study that estimated mental health indices before and after a second lockdown. 2nd: This is one of the first studies to measure trajectories of adjustment disorder based on the newly published ICD-11. 3rd: This is one of the first studies to measure the association between trajectories of mental health and somatic symptoms.
Comment 5: "Other comments for the authors to consider include: Adjustment disorder while common is not strongly linked to well-being or health outcomes".
Response: Adjustment disorder is related to health outcomes. Moreover, two of the stressors mentioned in the IADQ measure are illnesses, experienced by the participant or by significant others. As for references from literature, adjustment disorder was associated with an increased risk of Parkinson Disease (Svensson et al., 2016). Another example is stroke that was also found to be associated with adjustment disorder (Mitchell et al., 2017). Adjustment disorder was also shown to be as a consequence of a Pulmonary Arterial Hypertension diagnosis (Winter et al., 2020). In regard to well-being, adjustment disorder is related to well being (Killikelly et al., 2019; Hamama-Raz et al., 2021). Finally, adjustment disorder as an outcome in clinical study (Eimontas et al., 2017).
Comment 6: "…A stronger design would be to focus on screen positive for anxiety disorder using valid measures and stratify by lifetime history of mental disorder".
Response: Following the reviewer's comment, we wish to emphasize that the suggestion of the reviewer is not in the focus of the current study as many studies have done this before. We wish to clarify that we used screening measures. We have measured the trajectory of anxiety disorder using the GAD-7 which is a valid screening measure, as the reviewer suggested. For example: BMJ open (https://bmjopen.bmj.com/content/11/4/e049653) and (https://bmjopen.bmj.com/content/11/5/e0453250).
Comment 7: "…What is the collinearity between somatic symptom ratings and GADS-7 – I would suspect that somatic symptoms are simply a barometer for level of anxiety and unclear as to what the main analysis between anxiety and somatic scores really adds?".

Response: The measures used in the current study are distinct for their high construct validity. There is no direct resemblance in symptoms used in the different measures of anxiety, adjustment disorder and somatization. To demonstrate the level of overlap between the constructs, we took two approached, one that relies on the continuous scores that represent severity of symptoms and the second that relies on dichotomous scores that represent probable diagnoses.

As for the first, we present the effect sizes (r^2 = explained variance) of the simple bivariate correlations between the constructs. According to Cohen (1988, 1992), the effect size is low if the value of r varies around 0.1, medium if r varies around 0.3, and large if r varies more than 0.5.

Correlations showed significant correlations between Anxiety at T1 and Somatic symptoms at T2 ($r = .464$ $p < .001$, which is 21.5% variance explained), and between adjustment disorder symptoms at T1 and Somatic symptoms at T2 ($r = .456$ $p < .001$ which is 20.8% variance explained). These are equivalent to low-medium effect sizes. The cross-sectional correlations at T2 between anxiety ($r = .620$ $p < .001$) and adjustment disorder ($r = .620$ $p < .001$) symptoms on the one hand, and somatic symptoms on the other hand were significant. Importantly, these are equivalent to 38.4% variance explained which also are equivalent to medium effect sizes.

From a medical point of view, we employed a Chi square test to examine the convergence or the level of agreement between diagnoses of probable somatic symptoms at T2 and anxiety at both T1-T2, and between diagnoses of probable somatic symptoms at T2 and adjustment disorder at both T1-T2. Out of the participants diagnosed with somatic symptoms at T2, 47.9% were diagnosed with adjustment disorder at T1 (χ^2 (1) = 80.24 $p < .001$), and 49.3% were diagnosed with adjustment disorder at T2 (χ^2 (1) = 97.65 $p < .001$). Out of the participants diagnosed with somatic symptoms at T2, 36.8% were diagnosed with probable anxiety at T1 (χ^2 (1) = 68.24 $p < .001$), and 38.9% were diagnosed with probable anxiety at T2 (χ^2 (1) = 101.91 $p < .001$).

In sum, there is no evidence of multicollinearity.

Comment 8: "…I also feel the Title, Abstract and Conclusions go beyond what one can conclude from the study given the limitations and that the language/writing could be made clearer and more accurate overall".
Response: Following the reviewer's comment, we have toned down our claims and proof-read the paper for clarity and accuracy.

Reviewer 2

Reviewer: 2: Dr. Robert Stewart, Institute of Psychiatry
Comment 1: "1. I think the title might be amended to provide a more precise definition of the study, although I think this is more a matter for the Editorial Office and journal house style."
Response: Following the reviewer's comment, we have revised the title. Please see revised title of the manuscript.
Comment 2: "2. Given the design of the study, I think that 'somatic toll of a second lockdown' is too strong a term in the conclusions section of the abstract (and elsewhere, including the title). Although the study investigated this particular period of time, there is nothing in the design that permits causality to be assumed in relation to lockdown, as the results might well have been the same from an identical study carried out at a different time period. This is not intended to detract from the quality or utility of the study; it's just that I think the conclusions are over-stated".
Response: following the reviewer's comment, we amended the manuscript accordingly toning down and removed any indication to causality. Please see revised manuscript.

Comment 3: 3. The third paragraph of the Introduction drifts into considering the methodology of the study about to be described, which feels out of place."
Response: following the reviewer's comment, we have now moved this to the Methods section.
Comment 4: "4. The reader really needs more detail on the source population for sampling. Currently we are simply told that there are 100,000 members 'designed to be representative'. I would have expected some description of how the iPanel company identified and approached its members, and the achieved level of national representativeness (I assume there must be some public domain data on this). Otherwise, we really don't know what the analysed sample represents (although I accept that some comparisons are made between the analysed sample and national data). "
Response: Following the reviewer's comments and those of reviewer 1, we have overhauled the sampling section in order to present clearly the Panel and it's representativeness for the Israeli 1 population. In addition, see our comment to reviewer 1. In addition, please see our table in the manuscript comparing our sample to the Israeli Bureau of statistics on age and sex.
Comment 5: 5. In relation to the earlier point about causation, the IADQ instrument for adjustment disorder is described as including a stressors list covering different aspects of life. It would be helpful if the authors could clarify whether that stressors list was adapted in any way to be specific to lockdown-related stressors (as their conclusions seem to indicate that the anxiety exposure is

reflecting life under lockdown). Or are these generic stressors (in which case, how is the lockdown-related conclusion justified)? “.
Response: The stressors list was adapted to be specific to COVID-19, and the time frame was designated at the lockdown as the general list covers some of the lockdown factors. We wanted to stick to the original list of stressors as the IADQ is used relatively frequently although related to the new ICD-11 new conceptualization. The use of stressors related to COVID-19 as a mean to estimate the lockdown stressors is used in previous research in the UK (Chandola et al., 2020). The mental health impact of COVID-19 and lockdown-related stressors among adults in the UK. Psychological medicine, 1-10.), The authors conducted their research in UK and estimated the effects of lockdown. They conceptualized stressor variables in terms of factors that are important for mental health that may have changed during the COVID-19 pandemic. These include COVID-19-specific stressors, and more indirect stressors arising from the UK lockdown conditions. COVID-19-specific stressors included reports of symptoms of COVID-19 (respondents were asked if they had ‘experienced symptoms that could be caused by COVID-19’) and reported testing for COVID-19 (no tests, tested negative/inconclusive/waiting for results and positive tests). Additional stressors included health treatment-related, family roles-related, economic, financial and psychological stressors. Eventually, the stressors list was similar to the IADQ stressors list.

Comment 6: “6. There seems to be no description of other variables in the Methods section. For example, I can't find any description of what 'a little below average' income means and whether this was applied as a definition to stated income level or was how a participant referred to their income level (and, if so, how they would be expected to know the average.”
Response: Following the reviewers' comment, we have added a description of demographics in the methods section.

Comment 7: “7. Also, what does the Covid risk group variable mean, and how was it calculated?”.
Response: Following the reviewer's comment, the COVID risk group was a dichotomized variable based on the CDC & World Health Organization definitions of risk groups for COVID-19. Participants who present one of the medical conditions that qualifies them to be a part of the COVID-19 risk groups were defined as "yes" all the rest were defined as "no". Please see similar studies used the categorization. This was used before (See Ben-Ezra et al., 2020).
Comment: “8. In the Discussion (2nd paragraph) there is a mention of another study having presented pandemic-related changes in anxiety/depression as a single slope. I have to say that this seems a more logical approach than the 4 trajectory categories imposed by this paper (which necessarily conflate substantial changes in symptoms with small changes around the threshold points). It might help, therefore, if the authors provided more clarity on why they feel their approach is advantageous.”
Response: The person-oriented approach seeks to match theories and methods that portray development as a holistic and individualized process. This approach has gained popularity as model-based varieties of person-oriented methods have emerged. It allows for heterogeneity in the growth trajectories so that a sample is not forced to be portrayed in a single slope. Since theory (as presented in the manuscript, and for example Bonanno, 2004) supports the four trajectories, we decided to rely on this perspective. It is widely accepted that long-term reactions to stress are highly heterogeneous labile and demonstrate a highly complex and fluctuating course over the life span (e.g., Bonanno et al., 2012). It will be misleading to present the population in a single slope. Our analysis underscored the complex and non-homogenous reactions to lockdowns. It allows the variation that is needed to capture the full picture in the general population.
Comment 9: “9. The authors do acknowledge the single measure of somatic symptoms as a limitation in their Discussion; however, this again necessitates much more cautious language about conclusions in other areas of the paper. A 'somatic toll of a second lockdown' is simply not being conclusively demonstrated here - the somatic symptoms might have been present all along, and the association between anxiety and somatic symptoms might have been present in any similar cohort at any time,

regardless of the lockdown occurring. Finally, there are relatively few covariates considered and residual confounding is a potential concern.”.

Response: Following the reviewer's comments, we have toned down our claims and mention only the associations found in the longitudinal study. Furthermore, the risk of confounding was reduced as each instrument has core symptoms that does not overlap with other instruments (symptoms). The measures used in the current study are distinct with high construct validity, that not once were showed to be correlated and yet distinguished. There is no direct resemblance in symptoms used in the different measures of anxiety, adjustment disorder and somatization. To demonstrate the level of overlap between the constructs, we took two approaches, one that relies on the continuous scores that represent severity of symptoms and the second that relies on dichotomous scores that represent probable diagnoses. In addition, see our response to comment 7 made by reviewer 1.

VERSION 2 – REVIEW

REVIEWER	Robert Stewart Institute of Psychiatry, Section of Epidemiology (Box 60)
REVIEW RETURNED	22-Jul-2021

GENERAL COMMENTS	I appreciate the efforts taken by the authors to improve this paper and feel that my previous comments have been adequately addressed.