Dimensions and modulators of behavioural and mental-health changes during the Covid-19 pandemic: an N=343,017 study

Adam Hampshire 1*, Peter Hellyer 1 2, Eyal Soreq 1, Mitul A. Mehta 2, Konstantinos Ioannidis 3 4, William Trender 1, Jon E. Grant 5, Samuel R. Chamberlain 4

1. Imperial College London 2. King's College London 3. Cambridgeshire and Peterborough NHS Foundation Trust. 4. Department of Psychiatry, University of Cambridge, UK. 5. Department of Psychiatry, University of Chicago, USA.

**Corresponding author a.hampshire@imperial.ac.uk*

Acknowledgements:

Dr Hampshire is supported by the UK Dementia Research Institute and Biomedical Research Centre at Imperial College London with technology development supported by EU-CIG EC Marie-Curie CIG and NIHR grant II-LB-0715-20006. Dr Soreq's role was supported by MRC grant MR/R005370/1. William Trender is supported by the EPSRC Center for Doctoral Training in Neurotechnology. Dr Chamberlain's role in this study was funded by a Wellcome Trust Clinical Fellowship (Reference 110049/Z/15/Z). Mitul Mehta is in part supported by the National Institute for Health Research (NIHR) Biomedical Research Centre at South London and Maudsley NHS Foundation Trust and King's College London. The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health and Social Care. We would like to acknowledge COST Action CA16207 'European Network for Problematic Usage of the Internet', supported by COST (European Cooperation in Science and Technology); and the support of the National UK Research Network for Behavioural Addictions (NUK-BA).

Disclosures: Dr. Chamberlain consults for Promentis, and receives stipends for journal editorial work from Elsevier. Dr. Grant has received research grants from the TLC Foundation for Body-Focused Repetitive Behaviors, Biohaven, Promentis, and Avanir Pharmaceuticals. Prof. Mehta has received grant income from Takeda Pharmaceuticals, Johnson & Johnson and Lundbeck.

We thank Zoe Heron, Catherine Abbott and everyone involved at BBC2 Horizon, as well as Gaby Hornsby, who together made a demographically balanced study at this high population scale possible in such a short timeframe.

Abstract

BACKGROUND

The effects of the Covid-19 pandemic on mental health remain unclear. To mitigate the risks and capitalise on opportunities for positive change, we must understand how the impact has been mediated by sociodemographics, mental disorders, personality traits, life circumstances and the coping measures people choose to take.

METHODS

Data were collected from 376,987 members of the general public, predominantly in the UK, between late December 2019 and May 2020. Interaction models examined differences in sociodemographic distributions of mood and anxiety for 233,268 people in January vs. 109,749 in May. Factor analysis of a comprehensive instrument determined the dimensionality of self-perceived pandemic-driven change in wellbeing, outlook and behaviour for 74,830 participants in May. Linear modelling identified demographic, contextual, clinical, and trait predictors of pandemic impact. Topic modelling distilled prevalent advice from free-text responses.

RESULTS

Anxiety, depression and insomnia changed markedly in demographically-mediated ways. Untoward changes were larger for older adults. Benefits were greater for younger adults. Social connectedness was negatively affected across most mental and neurological conditions. There were disorder-specific changes in other domains, e.g., heightened conflict at home for attention-deficit hyperactivity disorder and heightened anxiety for obsessive-compulsive disorder. Psychiatric symptoms, personality traits, occupational variables and living conditions were amongst the strongest predictors of pandemic impact. Frontline health workers, carers of vulnerable older adults, and disabled or sheltered adults were disproportionately affected. Fifty advice topics were identified from free-text, the prevalence of which covaried with subpopulation, context and traits.

CONCLUSIONS

The general public report positive and negative consequences of the pandemic. Particular subsets of people have heightened risk of untoward effects whereas other groups appear resilient. To be valid and effective, studies seeking to quantify, predict or mitigate the impact of pandemics on mental health should apply holistic approaches, combining multiple psycho-socio-economic factors.

Key words: pandemic, coronavirus, Covid-19, mental health, wellbeing, mood, impulsive, compulsive, predictors, population.

Introduction

The coronavirus disease 2019 (Covid-19) pandemic has brought about unprecedented change in peoples' lives due to direct and indirect consequences of the illness, physical distancing and socio-economic restructuring. These changes will have affected mood, anxiety, and other aspects of mental health in widespread, profound but idiosyncratic ways that depend on individual circumstances (Pfefferbaum & North, 2020). Expert groups have highlighted that the impact on mental health and well-being are likely to relate to (i) aspects of demographics such as age or ethnicity, (ii) social networks, (iii) financial/occupational circumstances (iv), being shielded or having carer responsibilities, (v) preexisting mental health symptoms, (vi) maladaptive online technology use, (vii) personality traits and (viii) tendency towards compulsive behaviors (Holmes et al., 2020; Kiraly et al., 2020; Pareek et al., 2020).

Urgent calls have been made to study these relationships because they are critical to inform policy and healthcare decisions, and to guide researchers and clinicians; however, to date, there is little published information. Indeed, knowledge about the impact of pandemics on mental health is limited, with many studies focusing on small rarefied samples, not examining temporal dynamics of change pre- to post-lockdown (e.g. see Brooks et al., 2020, for review), nor integrating the diverse psycho-socio-economic variables relevant to mental health (Ioannidis et al. 2020).

To address these issues, we applied a combination of multivariate-analysis and machine learning methods to a unique large-scale dataset, comprising questionnaires from 376,987 people, collected between December 2019 and May 2020, in collaboration with BBC2 Horizon, predominantly within the UK. We determined: (i) differences in the population

distributions of depression, anxiety, and sleep problems between pre-Covid and peak UK lockdown timepoints; (ii) the dimensionality and statistical predictors of self-perceived pandemic impact, with a particular focus on sociodemographic characteristics, presence of mental/neurologic disorders, personality traits, and technology use; and (iii) individually tailorable advice topics learnt from free text describing peoples' lived experiences.

Methods

RECRUITMENT

Starting from December 26th 2019, participants were recruited to the study website, where they completed cognitive tests and a detailed questionnaire. Articles describing the study were placed on the BBC2 Horizon, BBC Home page, BBC News Home page and circulated on mobile news meta-apps from January 1st. To maximise representativeness of the sample there were no inclusion/exclusion criteria. Analyses here exclude data from participants under 16 years old, as they completed a briefer questionnaire, and those who responded to the questionnaire unfeasibly fast (<4 minutes). Cognitive test data will be reported separately. The study was approved by the Imperial College Research Ethics Committee (17IC4009).

DATA COLLECTION

Data were collected via our custom server system, which produces study-specific websites (https://gbws.cognitron.co.uk) on the Amazon EC2. Questionnaires and tests were programmed in Javascript and HTML5. They were deliverable via personal computers, tablets and smartphones. The questionnaire included scales quantifying sociodemographic, lifestyle, online technology use, personality, and mental health (supplementary **Appendix 1**). Participants could enroll for longitudinal follow up, currently scheduled for 3, 6 and 12 months. People returning to the site outside of these timepoints were navigated to a different URL. On May 2nd 2020, the questionnaire was augmented - in light of the Covid-19 pandemic - with an extended mood scale, and an instrument comprising 47 items quantifying self-perceived effects on mood, behaviour and outlook (Pandemic General Impact Scale PD-GIS-11). Questions regarding pre-existing psychiatric and neurological conditions, lockdown context, having the virus, and free text fields were added. This coincided with further promotion via BBC2 Horizon and BBC Homepage.

DATA PROCESSING & STATISTICAL ANALYSIS

Analyses were conducted in MATLAB R2020a. Participants with missing data were retained as some questions were contingent on others; therefore, observations per analysis vary with data availability. The questionnaire was organised into the following scales: demographics/lifestyle, online technology use, mood, personality, compulsive traits, and pandemic impact. Where appropriate, scales were summarised in the following steps. Agree-disagree and frequency items were filtered for missing data casewise within scale, converted to numeric, rank transformed to normality, and subscale scores estimated using principal component analysis. Components with eigenvalues>1 were varimax rotated and component scores estimated by regression. PCA models are in **Appendix 2**.

Cohort demographics were segmented into Pre-Pandemic, Early-Pandemic, and Mid-Pandemic epochs **(Figure 1a)** . General linear modelling tested how sociodemographic variables statistically predicted differences in mood, anxiety, and sleep, between the densely sampled and demographically matched Pre-Pandemic and Mid-Pandemic epochs.

Subscales of the pandemic impact instrument (PD-GIS-11 **Appendix 3**) quantified self-perceived impact across 11 psycho-socio-economic dimensions. They were cross-validated against the mental health self-assessment using CCA and Pearson's correlation. General linear modelling (GLM) determined the relationship of sociodemographic variables, home context, cohabitees and work arrangements to the PD-GIS-11. Further GLMs examined carers and people reporting psychiatric and neurologic diagnoses (for groups with N>90 members) with the sociodemographics factored out. Due to the expected shared variance between online technology use, personality traits and compulsivity, their multivariate relationships with the PD-GIS-11 were quantified via canonical correlation analysis (CCA).

Latent Dirichlet Allocation (LDA) (Blei, et al., 2003) extracted common topics from free text responses to the questions "What has been most POSITIVE about the lockdown?", "What has been most NEGATIVE about the lockdown?" and "What have you done that you would recommend to others because it has helped you during the lockdown?" Optimal numbers of topics were estimated for each question as follows. LDA models of different complexity were trained on 2/3rds of the participants' data. The trained models were applied to the remaining data. The lowest resultant perplexity value, quantifying fit of theoretical and observed topic

word distributions, was identified and the LDA model retrained on all data at the corresponding level of complexity. Inter-individual differences in topic mixture estimates were estimated and analysed with reference to mood, PD-GIS-11 and at-risk sub-populations using CCA and t-tests.

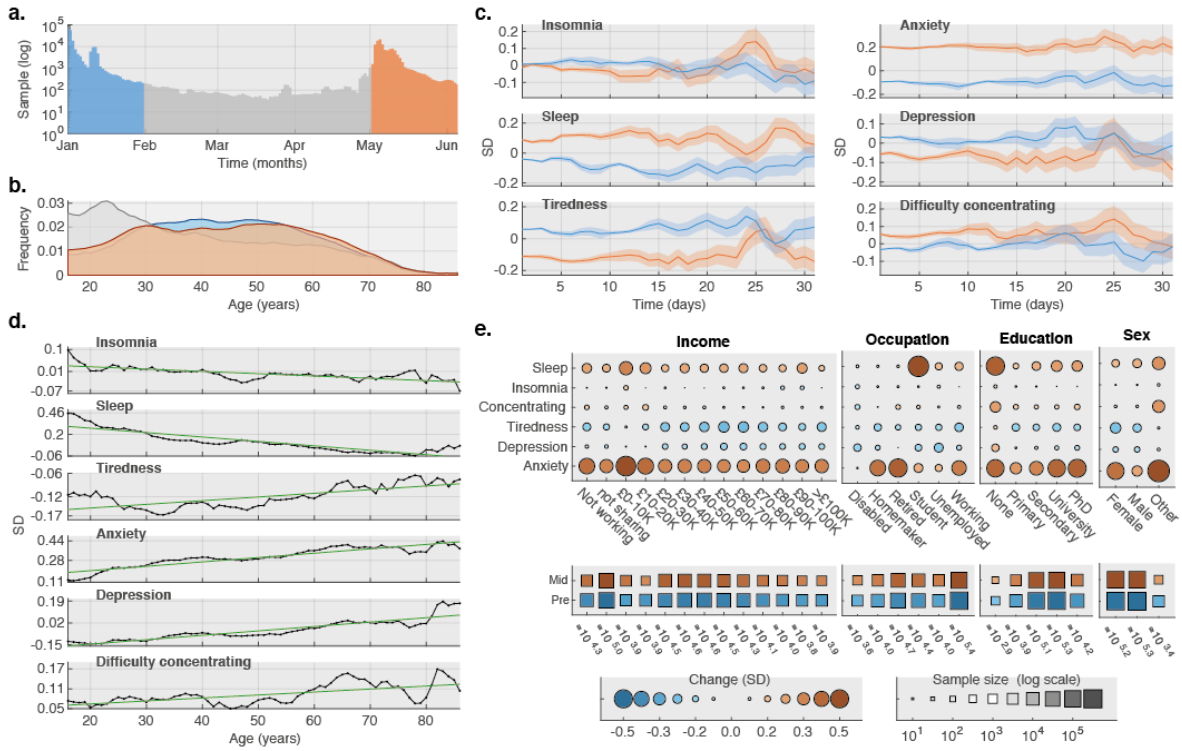
Results

Differences in the population distributions of depression, anxiety, and sleep problems Pre-vs Mid-Pandemic

Mid-Pandemic and Pre-Pandemic epochs were closely sociodemographically matched, whereas the sparsely sampled intervening months were skewed towards young adults (**Figure 1b**). Therefore, differences in mood assessment scores were calculated for the Mid minus Pre-Pandemic Epochs, with demographics factored out (**Figure 1c)**. During this time, the incidence of anxiety increased ~0.3 standard deviation units (SDs) whereas depression decreased ~0.08 SDs. Tiredness decreased ~0.16SDs. Problems concentrating increased marginally ~0.07SDs. Problems with sleep stayed approximately the same. Mean reported hours slept per night increased ~0.19SDs.

GLMs with interaction terms showed these small but significant overall differences were amplified in select sociodemographic sub-populations (**Figure 1d-e**). Older adults showed the greatest increase in anxiety incidence, (0.4-0.5SD for 60-80 year-olds). Depression was lower for younger adults and higher for older adults (-0.2SDs to 0.1SDs). Females (0.32SDs) had a greater increase in anxiety than Males (0.21SDs) but less than those identifying as Other (0.38SDs). Increased anxiety incidence was greater for retired people (0.38SDs), followed by homemakers (0.32SDs) and workers (0.29SDs). People who identified gender as Other reported the greatest increase in problems concentrating (0.22SDs). These differences across sociodemographic axes indicated profound and disproportionate pandemic effects on mental health for some people, supporting the next step of fractionating the impact of the pandemic across domains important in day-to-day life.

Figure 1. Sociodemographic mediated differences in national mood scores during the Pre- and Mid-Pandemic epochs

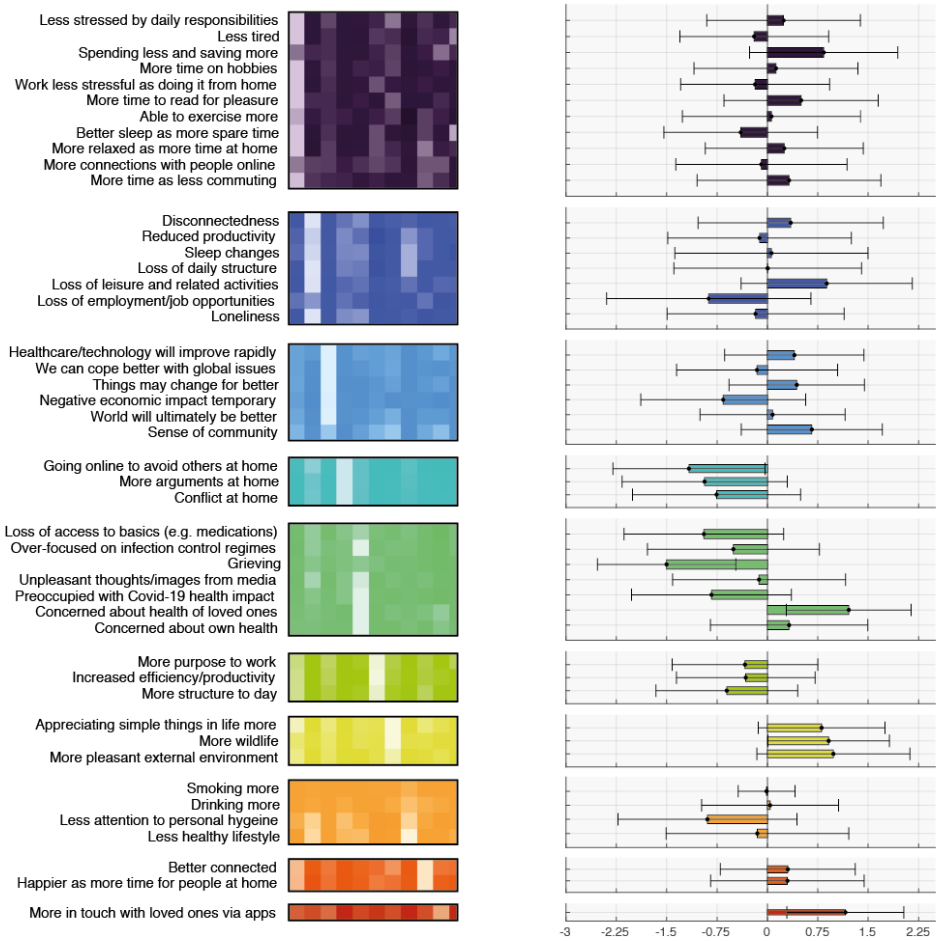


1a. Data were collected throughout the first 5 months of 2020. Sampling was concentrated in early January and early May, when BBC2 Horizon, News and Homepage promoted the study. 1b. Sociodemographic distributions, including age, were closely matched Pre-Pandemic (January) and Mid-Pandemic (May). Early-Pandemic (February-April) was more sparsely sampled and captured people from a different age distribution. 1c. Daily mean with SEM mood scores of the population, calculated separately for each of the 31 days post promotion launches Mid-Pandemic and Pre-Pandemic. Age, gender, handedness, first language, country of residence occupational status, and earnings are factored out. Significantly scaled differences are reproducibly evident across the days, most notably increased prevalence of mood-anxiety, but also improved sleep and tiredness scores. 1d. Differences in mean mood scores Mid-Pandemic minus Pre-Pandemic related to age. Older adults had a greater increase in anxiety. Younger adults increased sleep. Younger adults were less depressed whereas older adults were more depressed. 1e. Sub-population counts (middle squares - size represents log N per sub population) and the corresponding scale/valence of mood-score change (upper circles - size represents SD units and colour direction of change). Substantial differences were evident as a function of sociodemographic sub-groups, with heightened anxiety particularly in retired people, workers, homemakers, low income earners, and for gender Other vs Female vs Male.

Dimensionality & statistical predictors of self-perceived pandemic impact

Analysis of responses to individual items of the Pandemic General Impact Scale (PD-GIS-11) for 74,830 people in May 2020 showed high population variability with strong overall agree/disagree biases for a subset of positive and negative statements (**Figure 2a)**. Principal Component Analysis identified 11 latent variables underlying differences in peoples' responses (**Figure 2b**).

Figure 2. Individual item analysis and principal component analysis for the PD-GIS-11



2a. Some of the strongest agreement was with positive statements, e.g., improved natural environment, enjoying the simpler things in life, spending less, saving more and greater sense of community. Amongst the strongest agreement with negative statements was concern for health of loved ones, which was higher than concern about one's own health, and loss of leisure/health activities. The strongest disagreement was with statements regarding loss of employment, increased conflict at home, preoccupation with infection and loss of access to basics. There was strong agreement that technology/science/healthcare will advance more rapidly and things will change but not necessarily for the worse, and strong disagreement that economic impact will be temporary. People agreed that communication apps helped stay in touch with loved ones. 2b. PCA identified 11 components underlying PD-GIS-11 responses. These correlated with the mood items (**Appendix 3**) and were labelled 1: positive perception of more free time, less stress and reduced tiredness. 2: loneliness/ disruption. 3: positive outlook. 4: increased conflict at home. 5: increased health obsession. 6: increased work engagement/efficiency. 7: improved environment. 8: concern about unhealthier lifestyle. 9: time for people. 10: connectedness. 11 Better sleep.

Relationships of sociodemographic variables, home context, cohabitees and work arrangements, with the PD-GIS-11 are in **Figure 3a & Appendix 3**. Work arrangements were amongst the most substantial predictors. Healthworkers, particularly those on the frontline with Covid-19 patients had less time and were less relaxed relative to those who had been furloughed (1.5SDs), but also reported sleeping better (0.34SDs), having greater engagement in work (0.67SDs), and were the most likely to agree that the environment had changed for the better (0.25SDs).

The strongest contextual predictors of increased conflict in the home were cohabiting with preschool or school children (0.32SDs & 0.3SDs) or living with parents (0.38SDs)/in-laws (0.16SDs). People who had no, or unpleasant, outside space were less likely to report reduced tiredness/stress than those with relaxing outside space (0.31SDs), were more worried about negative health implications of the lockdown (0.43SDs), and less likely to report improved natural environment (0.43SDs). People from minority ethnic backgrounds tended to report higher health concerns (0.15-0.3SDs), but also were more likely to have a positive outlook (0.05-0.27SDs). The 8,347 participants who indicated they were looking after vulnerable older adults had higher scores for health obsession (0.17SDs) than the cohort average. People identifying as disabled or shielded reported substantially higher health obsession (0.5SDs), were less likely to report better connectedness (-0.57SDs) and were amongst the least likely to report feeling less stressed and having more time (-0.46SDs).

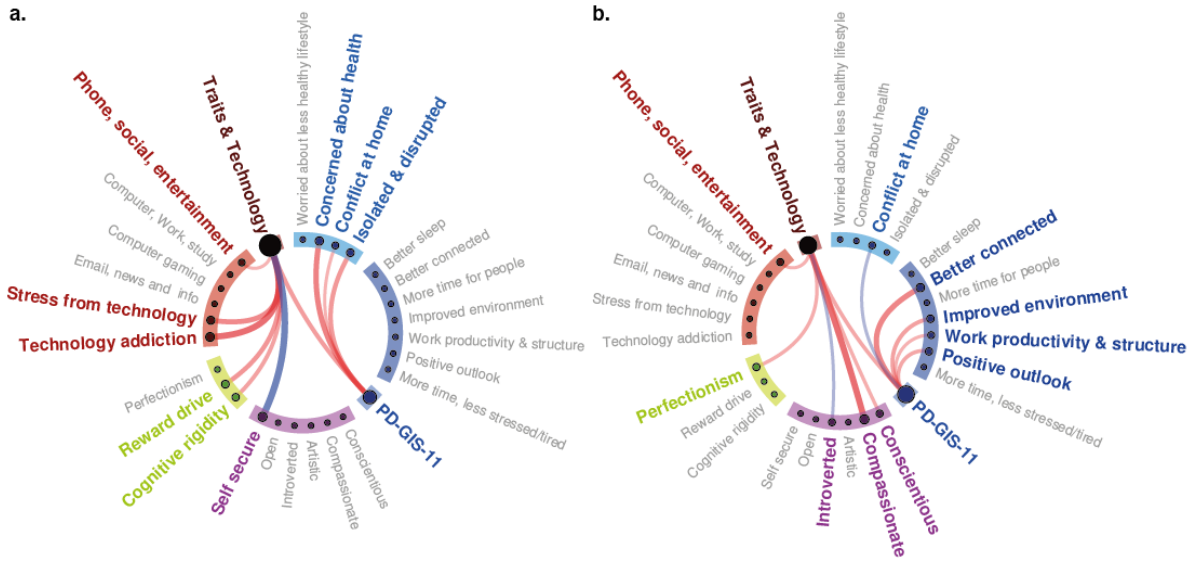
PD-GIS-11 scores and incidences for people reporting diagnoses of different mental and neurological conditions are in (**Appendix 3**). As expected, depression (10,526) and anxiety (10,525) overlapped in terms of comorbidity (6,305). After factoring out demographic and contextual variables, and relative to the population mean, individuals with anxiety disorder and/or obsessive compulsive disorder reported greater increases in health obsessions relative to those reporting no conditions (0.25SDs). Those with attention-deficit hyperactivity disorder (ADHD) reported greater increase in conflict at home (0.17SDs), individuals with depression were more worried about having unhealthier lifestyles (0.18SDs). All clinical groups, apart from People with Parkinson's, scored lower in terms of increased connectedness (0.05-0.21SDs).

Figure 3. Impact of the Covid-19 pandemic as a function of key demographic, social, and mental health variables

a. There were large and disproportionate impacts of the pandemic on the wellbeing of healthcare workers and disabled/shielded people. b. Conflict at home was predicted by type of cohabitees. c. Ethnicity predicted increased health concerns. d. People with anxiety and obsessive compulsive disorder (OCD) were disproportionately concerned about health. e. People with attentional deficit hyperactivity disorder were more likely to report increased conflict at home. All effects reported in standard deviation units. Full GLM results are in Appendix 3.

CCA quantifying associations between online technology use, personality traits, compulsivity, and Covid-19 impact (**Figure 4**) showed 11 statistically significant correlation modes. The first two had substantial canonical r values (r=0.41 & r=0.31). Cross-validation analysis confirmed the model was not overfitted (**Appendix 3**). Back projecting the first CCA mode onto the PD-GIS-11 showed strong associations with adverse impact, including isolation/disruption (r=0.44), increased health obsession (r=0.58) and increased conflict at home (r=0.27). Back projection on the other side of the first mode showed strong associations with technology addiction (r=0.67), technology stress (r=0.52), compulsivity subscale reward drive (r=0.45) and cognitive rigidity (r=0.38), and negative association with the personality trait 'self security' (r=-0.75). The PD-GIS-11 side of the second mode summarised positive pandemic impact, including better connectedness (r=0.55), improved environment (r=0.39), positive outlook (r=0.35), and greater work engagement (r=0.27), with a negative relationship to conflict at home (r=-0.26). The other side of the second mode associated positively with the traits compassion (r=0.69), conscientiousness (r=-0.31) and perfectionism (r=0.32), and negatively with introversion (r=-0.28).

Figure 4. Canonical Correlation of Trait and Technology vs PD-GIS-11



Correlation weights of the trait & technology vs PD-GIS-11 canonical variates with each other and their back projections, thresholded at | r | >0.25 (full results in Appendix 3). **a.** Mode 1 associated negative aspects of pandemic impact with technology stress and addiction, compulsivity traits and insecurity. **b.** Mode 2 associated positive aspects of pandemic impact with positive personality traits and perfectionism, a compulsivity trait. Red lines = positive correlations. Blue lines = negative correlations. Line weight = correlation r values.

Data-driven analysis of common impact and advice topics from free text

Exemplars of common topics from the free text are in **Appendix 4a**. Descriptions of the main negatives from 49,482 participants (**Figure 5a**) were best summarised by just five topics. In prevalence order, these were being unable to see relatives (especially older relatives), impracticalities of working or schooling from home, disruption of social and recreational activities, physical and mental health worries, and frustration with the media or government.

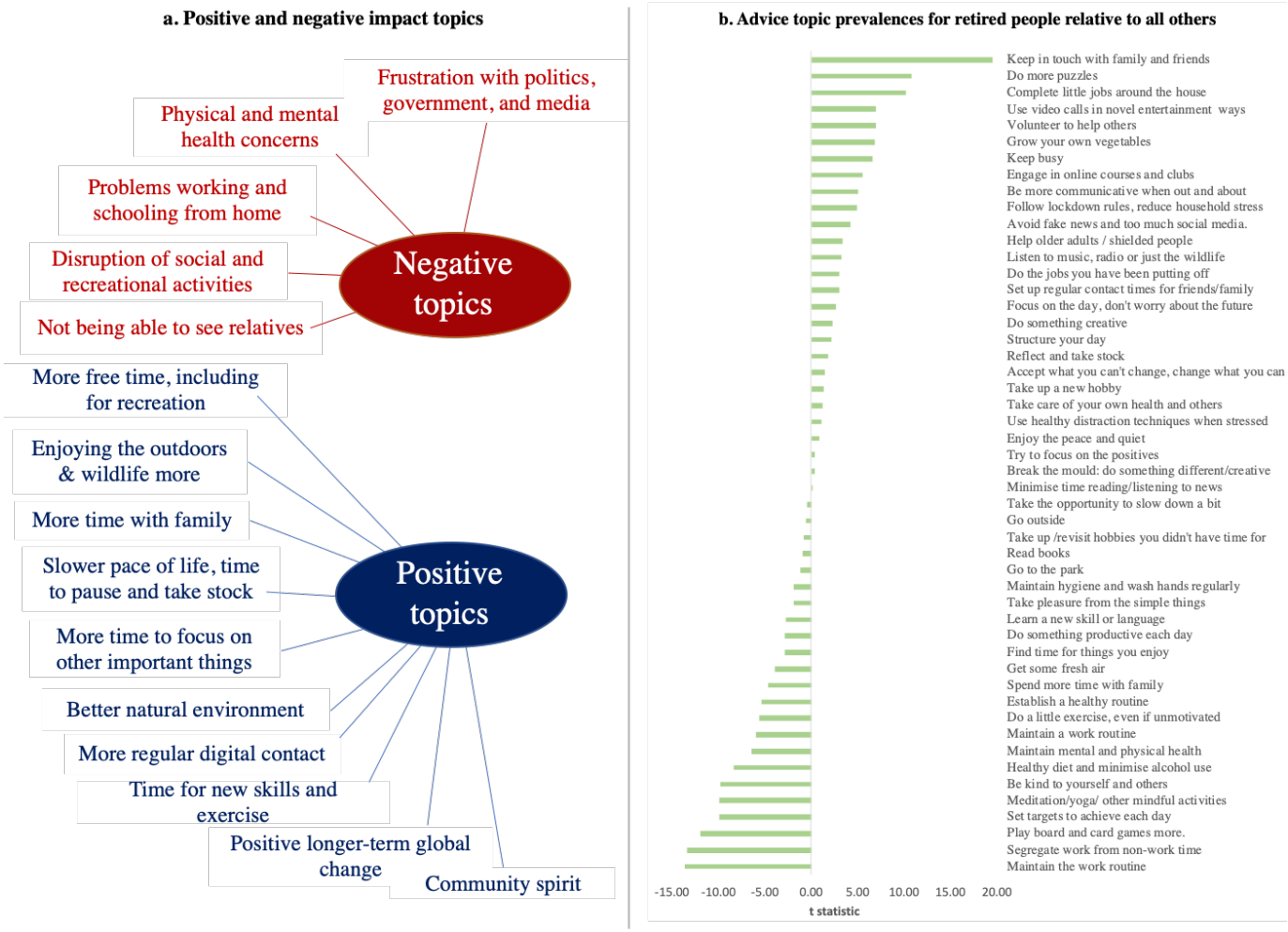
Descriptions of the main positives from 48,315 people (**Figure 5a**) were summarised by 10 topics, these being more free/recreational time, improved environment, time for important things, slower pace of life, positive long-term change, more time with family, learning new skills/expertise, more regular digital contact with friends/family, and enjoying the outdoors.

Advice from 44,376 people (**Figure 5b**) was much more variable, being optimally summarised by ~50 topics, the most prevalent being establishing a healthy routine, video conferencing to stay in touch, meditation/yoga, regular exercise, time outside, keeping busy, going for regular walks and planning a healthy diet.

CCA showed substantial shared variance between peoples' topic mixtures (**Figure 5b**) and both the mood self assessment (Appendix 4b) and the PD-GIS-11 (Appendix 4c). The prevalence of topics varied significantly amongst sub-populations most affected by the pandemic relative to the broader cohort (Appendix 4d reports in full). For example, 'be more communicative with people' ($t=4.703$, $p<0.001$) and 'take the time to reflect' ($t=4.253$ $p<0.001$) were more prevalent amongst disabled/sheltered people. 'Maintain personal hygiene/handwashing' ($t=3.199$ $p=0.001$), and 'segregate work and home time' ($t=3.102$ $p=0.002$) were more prevalent amongst frontline health workers. Retired people were more likely to recommend 'keeping in touch with friends and family' ($t=19.5056$ $p<0.001$) and 'doing more puzzles' (10.729 $p<0.001$). People who had been diagnosed with anxiety or depression were more likely to recommend 'take the time to be kind to yourself' (anx $t=4.996$ $p<0.001$; dep $t=7.863$ $p<0.001$) and to 'listen to music or appreciate wildlife' (anx $t=3.742$ $p<0.001$; dep $t=4.3207$ $p<0.001$). Those with anxiety also were more likely to recommend meditation or yoga (anx $t=4.552$ $p<0.001$; dep $t=-0.278$ $p=0.787$), whereas those with depression recommended 'taking pleasure in simple things' (anx $t=1.687$ $p=0.0917$; dep

t=3.868 p<0.001). Thus, pragmatic measures people chose to apply differed significantly for sub-populations whose mental health was most at risk.

Figure 5. Topic modelling of free text responses



a. Latent Dirichlet Allocation (LDA) identified 5 negative topics and 10 positive topics from the free text responses describing the main negative and positive impacts of the Covid-19 pandemic. **b.** LDA identified ~50 topics from free text describing what people would recommend to others based on their lived experiences of the pandemic. The prevalence of the topics varied significantly across sub populations. Above, we present differences in the advice topic prevalence for retired older adults relative to the rest of the cohort.

DISCUSSION

Our results provide converging evidence at large population scale that by May 2020 (peak UK lockdown), the Covid-19 pandemic had profound but idiosyncratic effects on mental health and wellbeing in the UK. The effects were complex, comprising multiple dimensions of change, some being positive. The substantial variability in these dimensions of impact was associated with a combination of clinical/biological and psycho-socio-economic variables. Furthermore, there were statistically robust relationships between these variables and what people reported, in their own words, had helped them during the pandemic.

Differences in the population distributions of depression, anxiety, and sleep were observed from pre-pandemic to the time of maximal UK lockdown, being indicative of more pronounced untoward effects in particular groups, such as in older adults. However, such macro-data overlooks considerable nuanced variations in the nature and extent of pandemic impact across individuals. Demographic characteristics, work, environment and social circumstances had robust associations, varying in scale from small to large, with the nature and extent of impact. In some ways this was anticipated based on extant literature (Kirby, 2020; Public Health England, 2020; Chowkwanyun & Reed, 2020; Lai et al., 2020; Rossi et al., 2020), but in others it was unexpected. Most notably, health workers showed very large differences to the broader population, being less relaxed and having less free time, but also reporting better sleep and greater work engagement. People from minority ethnic backgrounds were more worried about the impact of the pandemic on health, but also had more positive outlooks.

It was predicted that the impact of the Covid-19 pandemic would be substantially influenced by mental health and neurologic disorders (Pfefferbaum & North, 2020; Holmes et al., 2020) and dimensional traits (Chamberlain et al., 2018). Notably, although overall, people with psychiatric and neurologic conditions were less likely to report increased connectedness, these generalised associations were small scaled. More selective associations were observed for particular disorders: e.g. elevated health worries in anxiety disorders / obsessive-compulsive disorder, and increased conflict at home in adults with attention-deficit hyperactivity disorder (Cortese & Coghill, 2018; Moukaddam & Shah, 2020; Posner et al., 2020). However, trait and personality scales collectively were more predictive: negative pandemic impact was associated with compulsivity and insecurity, and positive impact with

compassion, conscientiousness, perfectionism, and extrovertedness. This accords with the notion that certain personality traits are prominent in shaping resilience, whereas others engender vulnerability (Barton & Rogerson, 2018; Grant et al., 2020).

Online technology can provide a powerful mediator of positive and negative impact during the pandemic. Prior studies proposed that a subset of people develop problematic usage of online technology (Ioannidis et al., 2016; Tiego et al., 2019; Kiraly et al., 2020). Accordingly, negative Covid-19 impact was strongly linked not with time using online technology *per se*, but with maladaptive online behaviours. Conversely, the benefits of using technology to stay connected were prominent in the questionnaire and free text analyses. One might conclude that limiting screen time could be counterproductive; more nuanced approaches to develop healthy online behaviours are warranted.

The profound impact of the pandemic was best captured in the data-driven analysis of free text describing lived-experience. Crucially, ~50 topics optimally summarised the practical measures people said helped them during the pandemic. The fact that the prevalences of these topics differed as a function of participant characteristics makes intuitive sense; some measures are irrelevant or impractical, whereas others are more germane, depending on one's circumstances. We believe there is potential to identify pragmatic individually-tailored advice based on this novel intersection of sociodemographic, contextual and personality variables with the corpus of peoples' self expressed expert experiences.

Taken together, these results demonstrate the importance of measuring multiple dimensions when quantifying pandemic impact on mental health, and the necessity of incorporating the broad psycho-socio-economic context when seeking to understand, predict or mitigate such impact. Indeed, the largest effect sizes explaining pandemic impact related to occupation and home context, followed by medium associations with personality traits, compulsivity and maladaptive technology use, and smaller but still notable effects for other disorders and demographic characteristics. This complexity in turn necessitates careful study design to account for confounding factors.

We will recontact this cohort at 3, 6 and 12 months to plot change in the idiosyncratic impact of the pandemic and its aftermath in vivo (Wang et al., 2020; Fancourt et al., 2020), including more detailed assessment of identified vulnerable sub-groups. Finally, the current paper

focused on data from people aged 16+. It is vital that research also examines younger people using this and related methodologies adapted for that purpose (Holmes et al., 2020).

REFERENCES

Blei DM, Ng AY, Jordan MI, Lafferty J. Latent Dirichlet allocation. Journal of Machine Learning Research. 2003 Jan; 3: 993–1022.

Chamberlain SR, Stochl J, Redden SA, Grant JE. Latent traits of impulsivity and compulsivity: toward dimensional psychiatry. Psychol Med. 2018 Apr;48(5):810-821.

Chowkwanyun M, Reed AL Jr. Racial Health Disparities and Covid-19 - Caution and Context. N Engl J Med. 2020 May 6. doi: 10.1056/NEJMp2012910. Epub ahead of print. PMID: 32374952.

Cortese S, Coghill D. Twenty years of research on attention- deficit/hyperactivity disorder (ADHD): looking back, looking forward. Evid Based Ment Health. 2018 Nov;21(4):173-176.

Fancourt D, Steptoe A, Bu F. Trajectories of depression and anxiety during enforced isolation due to COVID-19; longitudinal analyses of 59,318 adults in the UK with and without diagnosed mental illness. 2020, Jul. Online pre-print, prior to peer review, from medRxiv.

Grant JE, Pinto A, Chamberlain SR. Obsessive-Compulsive Personality Disorder. American Psychiatric Association Publishing. Oct 2019.

Holmes EA, O'Connor RC, Perry VH, Tracey I, Wessely S, Arseneault L, Ballard C, Christensen H, Cohen Silver R, Everall I, Ford T, John A, Kabir T, King K, Madan I, Michie S, Przybylski AK, Shafran R, Sweeney A, Worthman CM, Yardley L, Cowan K, Cope C, Hotopf M, Bullmore E. Multidisciplinary research priorities for the COVID-19 pandemic: a call for action for mental health science. Lancet Psychiatry. 2020 Jun;7(6):547-560.

Ioannidis K, Chamberlain SR, Treder MS, Kiraly F, Leppink EW, Redden SA, Stein DJ, Lochner C, Grant JE. Problematic internet use (PIU): Associations with the impulsive-compulsive spectrum. An application of machine learning in psychiatry. J Psychiatr Res. 2016 Dec;83:94-102.

Király O, Potenza MN, Stein DJ, King DL, Hodgins DC, Saunders JB, Griffiths MD, Gjoneska B, Billieux J, Brand M, Abbott MW, Chamberlain SR, Corazza O, Burkauskas J, Sales CMD, Montag C, Lochner C, Grünblatt E, Wegmann E, Martinotti G, Lee HK, Rumpf HJ, Castro-Calvo J, Rahimi-Movaghar A, Higuchi S, Menchon JM, Zohar J, Pellegrini L, Walitza S, Fineberg NA, Demetrovics Z. Preventing problematic internet use during the COVID-19 pandemic: Consensus guidance. Compr Psychiatry. 2020 May 12;100:152180.

Kirby T. Evidence mounts on the disproportionate effect of COVID-19 on ethnic minorities. Lancet Respir Med. 2020 May 8:S2213-2600(20)30228-9. Doi: 10.1016/S2213-2600(20)30228-9. Epub ahead of print.

Ioannidis K, Askelund AD, Kievit RA, van Harmelen AL. The complex neurobiology of resilient functioning after childhood maltreatment. BMC Med. 2020 Feb 13;18(1):32.

Lai J, Ma S, Wang Y, Cai Z, Hu J, Wei N, Wu J, Du H, Chen T, Li R, Tan H, Kang L, Yao L, Huang M, Wang H, Wang G, Liu Z, Hu S. Factors Associated With Mental Health Outcomes Among Health Care Workers Exposed to Coronavirus Disease 2019. JAMA Netw Open. 2020 Mar 2;3(3):e203976.

Lim GY, Tam WW, Lu Y, Ho CS, Zhang MW, Ho RC. Prevalence of Depression in the Community from 30 Countries between 1994 and 2014. Sci Rep. 2018 Feb 12;8(1):2861.

Moukaddam N, Shah A. Psychiatrists Beware! The Impact of COVID-19 and Pandemics on Mental Health. Psychiatric Times. 2020, 37(3).

Oshio A, Taku K, Hirano M, Saeed G. Resilience and Big Five personality traits: A meta-analysis. Personality and Individual Differences. June 2018. 127, 54-60.

Posner J, Polanczyk GV, Sonuga-Barke E. Attention-deficit hyperactivity disorder. Lancet. 2020 Feb 8;395(10222):450-462.

Pfefferbaum B, North CS. Mental Health and the Covid-19 Pandemic. N Engl J Med. 2020 Apr 13. doi: 10.1056/NEJMp2008017. Epub ahead of print.

Pareek, M, Bangash MN, Pareek, N, Pan, D, Sze, S, Minhas, JS, Hanif, W, Khunti, K. Ethnicity and COVID-19: an urgent public health research priority Lancet. 2020. April. https://doi.org/10.1016/S0140-6736(20)30922-3

Public Health England. Disparities in the risk and outcomes of COVID-19. 2020 Jun. www.gov.uk/phe

Remes O, Brayne C, van der Linde R, Lafortune L. A systematic review of reviews on the prevalence of anxiety disorders in adult populations. Brain Behav. 2016 Jun 5;6(7):e00497.

Rossi R, Socci V, Pacitti F, Di Lorenzo G, Di Marco A, Siracusano A, Rossi A. Mental Health Outcomes Among Frontline and Second-Line Health Care Workers During the Coronavirus Disease 2019 (COVID-19) Pandemic in Italy. JAMA Netw Open. 2020 May 1;3(5):e2010185.

Tiego J, Lochner C, Ioannidis K, Brand M, Stein DJ, Yücel M, Grant JE, Chamberlain SR. Problematic use of the Internet is a unidimensional quasi-trait with impulsive and compulsive subtypes. BMC Psychiatry. 2019 Nov 8;19(1):348.

Wang C, Pan R, Wan X, Tan Y, Xu L, McIntyre RS, Choo FN, Tran B, Ho R, Sharma VK, Ho C. A longitudinal study on the mental health of general population during the COVID-19 epidemic in China. Brain Behav Immun. 2020 Apr 13:S0889-1591(20)30511-0. doi: 10.1016/j.bbi.2020.04.028. Epub ahead of print. PMID: 32298802; PMCID: PMC7153528.

APPENDICES

Appendix 1 - Study Questionnaires

(a) Demographic and other contextual information

The following background information was collected from participants: age, gender, ethnicity, country of residence, level of education, occupational status, and income.

(b) Mood, anxiety, and sleep

Mood and anxiety symptoms were recorded using the extensively validated Patient Health Questionnaire 2 (PHQ-2) and GAD-7 respectively (Kroenke et al., 2003; Spitzer et al., 2006). The PHQ-2 and GAD-7 ask about symptoms over the preceding two weeks, and each question is answered on a 4-point scale, from 0 (not at all) to 3 (nearly every day). Additionally, we asked how many hours on average participants slept per night.

(c) Personality traits, and compulsivity

Personality traits were quantified using the extensively validated Big-5 Inventory, which comprises 44 questions (John et al., 2008). Each question is a short phrase and is answered on a 5-point rating scale from 1 (strongly disagree) to 5 (strongly agree). Aspects of personality classically reflect extraversion, agreeableness, conscientiousness, neuroticism and openness to experience (John et al., 2008). Based on prior factor analysis of data from 60,000 participants, we used an abbreviated version, comprising 18 questions with a data-driven structure of 6 components. These are reported in the factor analysis in **Appendix 2**.

Compulsivity is a trans-diagnostic concept representing the tendency towards repetitive habits, and was measured using the Cambridge-Chicago Compulsivity Trait Scale (CHI-T) (Chamberlain & Grant, 2018). This is a 15-item questionnaire that is answered on a 4-point rating scale ranging from 1 (strongly disagree) to 4 (strongly agree). The CHI-T is sensitive to compulsivity across a range of disorders (Chamberlain & Grant, 2018; Albertella et al., 2019).

(d) Impact of the pandemic

The Pandemic General Impact Scale (PD-GIS-11) was developed for the current study to quantify the self-reported negative and positive impacts of the Covid-19 pandemic, on multiple levels of psycho-socio-economic investigation. The instrument comprised 47 questions, relating to potential negative and positive aspects of the situation, and longer term outlook. Each item is answered on a 5-point scale ranging from 1 (strongly disagree) to 4 (strongly agree). Negative impact questions covered areas of concern for health (own health and that of others), being concerned with the consequences of contracting Covid-19, loneliness, conflict at home, negative emotions from reading/listening to news, grieving, loss of employment/job/income, loss of leisure and well being activities, loss of daily structure, disruption of sleep patterns, less healthy lifestyles, less focus on personal hygiene, loss of productivity, social disconnection, life being dominated by infection control routines, loss of important goods/medication/services, more arguments in the household, and going on the internet to avoid people at home. Positive impact questions covered less commuting time, more structure to the day, joy at being able to spent more time with people at home, more connections with people online, sense of shared community, more efficient/productive work, being more relaxed due to more time at home, better sleep due to spare time, greater sense of purpose in work, greater opportunity to exercise, improved natural environment, time to read for pleasure, work less stressful due to doing it from home, spending more time on hobbies, spending less and saving money, more social contact outside of the home, feeling less tired, feeling better connected with people at home, more wildlife, taking greater appreciation for the simple things in life, and being less stressed by daily responsibilities.

(e) Online Technology use

Technology use was quantified by asking about frequency of use of the following, over the previous 4 week period: Smart Phone, Computer (Desktop/Laptop), Tablet Device, Gaming Console, Email, Social Media, reading the news, playing computer games, online gambling, working, learning/studying, shopping, streaming films or music, and searching for information online. Each question was responded to on a 7-point scale, from 0 (never) to 7 (more often than hourly every day).

(f) Stress from online technology

Stress from online technology was measured by asking the participants the following questions, regarding the past 4 weeks: When you checked Email, did it tend to make you feel stressed/unhappy or relieved/happy? When you used social media, did it tend to make you feel stressed/unhappy or relieved/happy? When you read the news, did it tend to make you feel stressed/unhappy or relieved/happy? When you played computer games, did it tend to make you feel stressed/unhappy or relieved/happy? The response options for each question were: "Mostly stressed/unhappy","Mostly relieved/happy","Both", or "Neither".

(g) Maladaptive ('Addictive') use of online technology

Maladaptive use of online technology was quantified using the following questions, which were based on expert consensus amongst the study team in the field of Problematic Usage of the Internet: How often did you check email or social media accounts after you went to bed? How often did you use internet related activities to block out disturbing thoughts or soothe yourself? How often did you choose to spend time on internet related activities to battle loneliness or boredom? How often did you suffer from negative financial consequences because of an online activity? How often did you check your email or social media account or equivalent before something else that you needed to do? How often did you try to stop an excessive online activity but feel a compulsion to continue? How often did you try to cut down the amount of time you spend on-line and fail? The questions asked about these areas over the preceding 4 week period. For the first question (using technology before bed), response options were 1 (never) to 5 (daily). For the other questions, response options were: 1 (never) to 7 (more than hourly every day).

References

Albertella L, Chamberlain SR, Le Pelley ME, Greenwood LM, Lee RSC, Den Ouden L, Segrave RA, Grant JE, Yücel M. Compulsivity is measurable across distinct psychiatric symptom domains and is associated with familial risk and reward- related attentional capture. CNS Spectr. 2019 Oct 24:1-8.

Brooks SK, Webster RK, Smith LE, Woodland L, Wessely S, Greenberg N, Rubin GJ. The psychological impact of quarantine and how to reduce it: rapid review of the evidence. Lancet. 2020 Mar 14;395(10227):912-920.

Chamberlain SR, Grant JE. Initial validation of a transdiagnostic compulsivity questionnaire: the Cambridge-Chicago Compulsivity Trait Scale. CNS Spectr. 2018 Oct;23(5):340-346.

John OP, Naumann LP, Soto CJ. Paradigm shift to the integrative Big-Five trait taxonomy: History, measurement, and conceptual issues. In O. P. John, R. W. Robins, & L. A. Pervin (Eds.), Handbook of personality: Theory and research (3rd ed., pp. 114–158). New York: Guilford Press. 2008.

Spitzer RL, Kroenke K, Williams JB, Löwe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. Arch Intern Med. 2006 May 22;166(10):1092-7.

Kroenke K, Spitzer RL, Williams JB. The Patient Health Questionnaire-2: validity of a two-item depression screener. Med Care. 2003 Nov; 41(11):1284-92.

Appendix 2 - Principal Component Analyses of Questionnaire Instrument scales are here:
https://www.dropbox.com/s/5ugvo7skf45gevy/Appendix%202%20-%20FactorModels.pdf?dl=0

Appendix 3a - PD-GIS individual item and PCA analyses are here:
https://www.dropbox.com/s/e6rqsazb3z90pth/Appendix%203%20-%20PD_GIS.pdf?dl=0

Appendix 3b - Canonical correlation analysis between Mood State and PD-GIS-11 is here:
https://www.dropbox.com/s/e6rqsazb3z90pth/Appendix%203%20-%20PD_GIS.pdf?dl=0

Those who reported greater improvement in sleep during the pandemic had a reduced global incidence of problems getting or staying asleep. Those indicating heightened conflict at home had higher irritability scores. Increased loneliness and depression, increased conflict at home, better sleep, more structured work and increased obsession with health had the most shared variance with items of the wellbeing scale. Therefore, the PD-GIS-11 scale provided finer grained information on how the pandemic affected the mood of the nation.

Appendix 3c - Demographic predictors of the PD-GIS-11 is here:
https://www.dropbox.com/s/e6rqsazb3z90pth/Appendix%203%20-%20PD_GIS.pdf?dl=0

Appendix 3d - Rates of previously diagnosed clinical conditions reported in the sample is here:
https://www.dropbox.com/s/e6rqsazb3z90pth/Appendix%203%20-%20PD_GIS.pdf?dl=0

Appendix 3e - GLM for clinical conditions vs PD-GIS-11 subscales is here:
https://www.dropbox.com/s/e6rqsazb3z90pth/Appendix%203%20-%20PD_GIS.pdf?dl=0

Appendix 3f - sub-sampling and permutation analyses to characterise overfit and reproducibility of the canonical correlation analysis of PD-GIS-11 items vs traits and technology use are available here:
https://www.dropbox.com/s/e6rqsazb3z90pth/Appendix%203%20-%20PD_GIS.pdf?dl=0

Results show test-train modelling, whereby the CCA model is trained on data from one set of data, and then applied to predict PD-GIS-11 scores in the remaining subset from the corresponding trait and technology use/misuse measures. Curves are means for 100 repeats for each subsample sze conducted ranging up in N=1000 steps from 100. Top, left is the true data canonical R values for the trained data. Top right is the same sub-sampled data but with permuted labels on the X side of the CCA. Bottom left is the true canonical R values for the predicted/ held out data. Top right is for the same held out data but with the labels permuted. As can be seen from the curves, at this population scale, there is practically no overfit, with the trained and test canonical R values being near identical. Furthermore, permuted canonical R values are near 0 even for the data the model is trained on.

Appendix 4a - topics and their top ten exemplars are available here:
https://www.dropbox.com/s/jgis9cyngjzzsa1/Appendix%204%20-%20NLP.pdf?dl=0

A mixture of ~ 5 topics were identified as being optimal for summarising the free text regarding negatives of the pandemic, having the minimal perplexity index, when applied to held out data. ~10 topics were optimal for positives, and ~50 topics for advice.

Appendix 4b - CCA of free text topic estimates vs mood self assessment item scores
https://www.dropbox.com/s/jgis9cyngjzzsa1/Appendix%204%20-%20NLP.pdf?dl=0

Appendix 4c - CCA of free text topic estimates vs items of the PD-GIS-11

https://www.dropbox.com/s/jgis9cyngjzzsa1/Appendix%204%20-%20NLP.pdf?dl=0

Appendix 4d - Topics, their relative prevalence in select sub-populations and top ten exemplars selected as having the highest topic mixtures are available here.
https://www.dropbox.com/s/jgis9cyngjzzsa1/Appendix%204%20-%20NLP.pdf?dl=0

Despite the intrinsically noisy nature of topic modelling, string associations were evident, whereby certain at-risk sub-populations were significantly more likely to report some topics helped.