Learning the Mental Health Impact of COVID-19 in the United States with Explainable Artificial Intelligence: Observational study

Indra Prakash Jha, Raghav Awasthi, Ajit Kumar, Vibhor Kumar, Tavpritesh Sethi

Submitted to: JMIR Mental Health
on: October 17, 2020

Table of Contents

Learning the Mental Health Impact of COVID-19 in the United States with Explainable Artificial Intelligence: Observational study

Indra Prakash Jha[1*] MCA; Raghav Awasthi[1*] MSc; Ajit Kumar[2] MBA; Vibhor Kumar[1] PhD; Tavpritesh Sethi[1] PhD

[1]Indraprastha Institute of Information technology NEW DELHI IN
[2]Adobe , Noida, Inida Noida IR
[*]these authors contributed equally

Corresponding Author:
Tavpritesh Sethi PhD
Indraprastha Institute of Information technology
Room-309. R &D building
IIIT Campus , Okhla Phase 3,
NEW DELHI
IN

Abstract

Background: : COVID-19 pandemic has deeply affected the health, economic, and social fabric of nations. Identification of individual-level susceptibility factors may help people in identifying and managing their emotional, psychological, and social well-being.

Objective: This work is focused on learning a ranked list of factors that could indicate a predisposition to a mental disorder during the COVID pandemic.

Methods: In this study, We have used a survey of 17764 adults in the USA at different age groups, genders, and socioeconomic statuses. Through initial statistical analysis followed by Bayesian Network inference, we have identified key factors affecting Mental health during the COVID pandemic. Integrating Bayesian networks with classical machine learning approaches lead to effective modeling of the level of mental health.

Results: Overall, females are more stressed than males, and people of age-group 18-29 are more vulnerable to anxiety than other age groups. Using the Bayesian Network Model, we found that people with chronic medical condition of mental illness are more prone to mental disorders during the COVID age. The new realities of working from home, home-schooling, and lack of communication with family/friends/neighbors induces mental pressure. Financial assistance from social security helps in reducing mental stress during COVID generated economic crises. Finally, using supervised ML models, we predicted the most mentally vulnerable people with ~80% accuracy.

Conclusions: Multiple factors such as Social isolation, digital communication, working, and schooling from home, were identified as crucial factors of mental illness during Covid-19. Regular non-virtual communication with friends and family, healthy social life and social security are key factors and especially taking care of people with mental disease history appear to be even more important. Clinical Trial: ...

(JMIR Preprints 17/10/2020:25097)
DOI: https://doi.org/10.2196/preprints.25097

Preprint Settings

1) Would you like to publish your submitted manuscript as preprint?
Please make my preprint PDF available to anyone at any time (recommended).
Please make my preprint PDF available only to logged-in users; I understand that my title and abstract will remain visible to all users.
Only make the preprint title and abstract visible.
✓ **No, I do not wish to publish my submitted manuscript as a preprint.**
2) If accepted for publication in a JMIR journal, would you like the PDF to be visible to the public?
✓ **Yes, please make my accepted manuscript PDF available to anyone at any time (Recommended).**

Yes, but please make my accepted manuscript PDF available only to logged-in users; I understand that the title and abstract will remain v
Yes, but only make the title and abstract visible (see Important note, above). I understand that if I later pay to participate in <a href="http

Original Manuscript

Learning the Mental Health Impact of COVID-19 in the United States with Explainable Artificial Intelligence: Observational study

Indra Prakash Jha MCA [1,*], Raghav Awasthi MSc [1,*], Ajit Kumar MBA [2], Vibhor Kumar PhD [1,#], Tavpritesh Sethi PhD[1,#]

1 Indraprastha Institute of Information Technology, Delhi (IIIT-D)
2 Adobe
co-corresponding author
* authors contributed equally to this work

Corresponding author Address:

Tavpritesh Sethi -
Phone: 011 26907533
Email: tavpriteshsethi@iiitd.ac.in
Office: A-309, R&D Block, IIIT Delhi, Delhi India, 110020

Abstract:

Background:

COVID-19 pandemic has deeply affected the health, economic, and social fabric of nations. Identification of individual-level susceptibility factors may help people in identifying and managing their emotional, psychological, and social well-being.

Objective:

This work is focused on learning a ranked list of factors that could indicate a predisposition to a mental disorder during the COVID pandemic.

Methods:

In this study, We have used a survey of 17764 adults in the USA at different age groups, genders, and socioeconomic statuses. Through initial statistical analysis followed by Bayesian Network inference, we have identified key factors affecting Mental health during the COVID pandemic. Integrating Bayesian networks with classical machine learning approaches lead to effective modeling of the level of mental health.

Results:

Overall, females are more stressed than males, and people of age-group 18-29 are more vulnerable to anxiety than other age groups. Using the Bayesian Network Model, we found that people with the chronic medical condition of mental illness are more prone to mental disorders during the COVID age. The new realities of working from home, home-schooling, and lack of communication with family/friends/neighbors induces mental pressure. Financial assistance from social security helps in reducing mental stress during COVID-generated economic crises. Finally, using supervised ML models, we predicted the most mentally vulnerable people with ~80% accuracy.

Conclusions:

Multiple factors such as Social isolation, digital communication, working, and schooling from home, were identified as crucial factors of mental illness during Covid-19. Regular in-person communication with friends and family, healthy social life, and social security are key factors, and especially taking care of people with mental disease history appears to be even more important.

Keyword:

COVID-19; Mental Health; Machine Learning; Bayesian Network; Explainable Artificial Intelligence

Introduction:

After 7 months of initial reporting, the coronavirus pandemic continues to rage across the world. The mental health consequences of the COVID-19 pandemic are profound. More than half a million lives and more than 400 million jobs have been lost [1] causing a considerable degree of fear, worry, and concern. These effects are seen in the population at large and may be pronounced among certain groups in particular, such as youth, frontline workers [2], caregivers, and people with chronic medical conditions. The new world order has introduced unprecedented interventions of country-

wide lockdowns that are necessary to control the spread but have led to increased social isolation. Loneliness, depression, harmful alcohol, drug use, and self-harm or suicidal behavior are also expected to rise.
The Lancet Psychiatry[3] recently highlighted the needs of vulnerable groups during this time, including those with severe mental illness, learning difficulties, and neurodevelopmental disorders, as well as socially excluded groups such as prisoners, the homeless, and refugees. Calls to action highlighting the need for engaging more early-career psychiatrists [4,5], technology such as Telepsychiatry, and highlighting the high susceptibility of frontline medical workers themselves [6] highlights the magnitude of the problem. Further, interventions are expected to have a gender-specific impact with women more likely to be exposed to additional stressors related to informal care, already existing economic disparity, and school closures. Similarly, age and comorbidity status may have a direct impact on susceptibility to mental health challenges due to their relationship with COVID-19 morbidity and mortality. Indeed, it has been established that emotional distress is ubiquitous in affected populations — a finding certain to be echoed in populations affected by the Covid-19 pandemic [7]. Finally, the role of social media [8, 9] is complex with some research indicating the association of social media exposure with a higher prevalence of mental health problems [10].
However, most of these effects have been studied in isolation with a lack of modeling the collective impact of such factors. This study addresses this gap through the use of Bayesian networks, an explainable artificial intelligence approach that captures the joint multivariate distribution underlying a large survey data collected across the United States. We also address the gap of prediction of vulnerability to mental health events such as anxiety attacks using supervised machine learning models.

Methods:

Datasets

We extracted data of 17764 adults [11] from two weekly surveys (20 April 2020 - 26 April 2020 and 4 May 2020 - 10 May 2020) of the U.S. adult household population nationwide for 18 regional areas including 10 states (CA, CO, FL, LA, MN, MO, MT1, NY, OR, TX) and 8 Metropolitan Statistical Areas (Atlanta, Baltimore, Birmingham, Chicago, Cleveland, Columbus, Phoenix, Pittsburgh). Two rounds of data collection were available at the time of this analysis and both rounds of data until May 25, 2020, were included in this analysis. The details of the original data are available elsewhere [12]. To summarize, the dataset comprised variables on physical health, mental health, Insurance related Policy, economic security, and social dynamics. Figure-1 shows the socio-demographic characteristics of respondents participating in the Survey.

Analysis

Figure 2(a) shows the flow diagram for the analyses conducted. The survey questions were classified into several types of indicators such as mental-health, work-from-home, communication, covid-symptoms, chronic-medical-conditions, behavioral-aspects, insurance-assistance, and many others *(see* . Supplementary Table-1).

Item Reliability Analysis

We constructed a model for mental-health indicators, attribute *soc5a* (Felt nervous, anxious, or on edge), attribute *soc5b* (Felt depressed), attribute *soc5c* (Felt lonely), attribute *soc5d* (Felt hopeless about the future), and *soc5e*(sweating, trouble breathing, pounding heart, etc, in the last seven days) as outcome variables. Hence we first evaluated the consistency in answers to mental-health questions using Item Reliability Analysis**.** A scale for reliability measure of internal consistency, **Cronbach's alpha**, was calculated using the *Psych* package in R [13].

Test of independence among mental health indicator and other indicators

Thereafter, a pairwise chi-square test of independence was performed to examine associations of *mental health indicators* and other variables and a ***P*** of less than 0.05 was taken as the cut-off for significance.

Data-Driven Bayesian Network Analysis

Since mental health variables may have complex dependencies with potential confounding factors, mediation, and inter-causal dependency, we extended our association analysis with data-driven Bayesian network (BN) structure learning. The structure of the learned Bayesian network was made robust through bootstrapping and ensemble averaging of edge directions. Hill climbing optimizer [14] with the Akaike Information Criterion(AIC) [15] based score was used to select the best PGM (Probabilistic Graphical Model) that explained the data. Bootstrapped learning and majority voting over 101 BNs were done. Exact Inference using the belief propagation algorithm [16] was learned in order to quantify the strength of learned associations. The Analysis was performed in the R using the package *wiseR* [17].

Mental health prediction using supervised Machine Learning

Next, the *Markov-blanket* [18] of *mental-health-indicators* was extracted to select features that may predict responses to the *mental-health-indicators*. Data were partitioned into training (80%) and testing (20%) sets and the class imbalance was corrected using the Synthetic Minority Oversampling Technique(SMOTE) [19]. Different supervised machine learning models - Random-Forest(RF), Support vector machine (SVM), logistic, naive-Bayes, were learned for predicting the response to mental health indicators using the Scikit-learn library [20] in Python.

Results:

Item Reliability Analysis

Attribute *soc5a* (Felt nervous, anxious, or on edge), attribute *soc5b* (Felt depressed), attribute *soc5c* (Felt lonely), attribute *soc5d* (Felt hopeless about the future) achieved a Cronbach's alpha approximating 0.8 [Figure 2(b)], thus confirming their internal consistency and suitability for modeling.

Gender and age-related variation in mental-health indicators

Gender and age-specific difference was observed in attribute *soc5a*, with females having a higher incidence than males (two proportion z-test, ***P*** <0.001) (Figure 3a), and young adults in the 18-29 age group (***P*** <0.001) (Figure 3b). Age group 18 - 29 in both genders was most vulnerable to mental stress > 5 days in a week, thus indicating that COVID-19 may have disproportionately affected the mental health of youth due to a variety of factors.

Associations of anxiety in the United States

A Chi-square test revealed many significant associations of the mental health variables (Supplementary Figure 1). However, this analysis does not account for potential confounding or "explaining away" effects.

Data-Driven Bayesian Network Analysis

Hence a data-driven Bayesian network structure-learning exercise was carried out and revealed interesting findings. From the learned structure, attribute *soc5a* (Felt nervous, anxious, or on edge in the last seven days) was found to be the parent variable for other mental health indicators in almost 100% of the bootstrapped networks, represented as the strength of the edges [Figure 4(b)]. Being a driver variable in the structure, attribute *soc5a* was taken as the primary dependent variable for downstream modeling analysis.

Impact of social life and work-related stressors.

Our analysis using network inference via the Exact Inference algorithm showed a clear impact of in-person social communication on the reduction of anxiety levels. A strong (>5% with CI ~1% at both sides) and (>6.5% with CI ~1.5% at both sides) monotonic increase between control of anxiety and frequency of speaking with neighbors *(*attribute *soc2a,* attribute *soc2b)* were observed. This effect was weaker (~1.5% with a wide confidence interval) with digital communication with friends and family conducted over phone, text, email, or other internet media (attribute *soc3a*, attribute *soc3b*). This finding underscores the importance of social communication while maintaining the appropriate measures such as masks and social distancing in order to maintain mental health during such isolating times. We also observed that the presence of kids in the house reduces the probability of depression by >11% with CI ~2% on both sides. Further, Exact Inference upon the network revealed an increase in the conditional probability of anxiety (attribute *soc5a*) arising from canceled or postponed work (>4% with CI ~1.4%), canceled, or postponed school (7% with CI ~1.5%), worked from home (>5% with CI ~1.3) and studied from home (>7% with CI ~1.8). Interestingly, although 83% of all volunteers chose to wear the mask, 77% avoided restaurants, 83% avoided public and crowded places these measures were not found to be associated with a significant change in anxiety levels as inferred from our model. These inferences are summarized in figure-5.

Impact of symptoms and comorbidities

We also investigated the relationship between mental stress and COVID symptoms indicators. WHO recommends contacting health service providers if any COVID symptoms (attributes *phys1a* to *phys1q*) are experienced within the last seven days. Our network did not indicate any significant impact of these responses on mental health (attribute *soc5a*), the conditional probability of which remained unchanged (62.2%) across the responses. Surprisingly, although medical conditions (attributes *phys3a* to *phys3m*) are known to increase the risk of serious illness from COVID-19, our model showed that suffering from cancer (attribute *phys3k*) and hypertension (attribute *phys3b*) had a reverse impact on anxiety levels. Those suffering from cancer had approximately 8.3% (with ~2% CI) higher conditional probability of having less than one anxiety-ridden day in a week (>7% effect for hypertension with CI ~1.5%). Additionally, Cystic-fibrosis (attribute *phys3i*) and Liver-disease (attribute *phys3j*) have wide confidence intervals with non-significant differences in mean values. (Figure-5).

Impact of economic factors

Receiving income assistance through Social Security improved the conditional probability of less than one day of anxiety in a week by 10.4% (with CI ~1.5%) as compared with the segment of people who did not apply or receive it. Just applying for income assistance led to a 4% improvement

(Figure-5). Similarly, Supplemental Social Security (~5.5 with CI ~4%) and Health insurance (~5% with CI ~2%) also led to similar results.
In addition to this, Old age people (age more than 60) found health insurance more relaxing than young age people. COVID has also severely affected the financial condition of individuals. That may also lead to mental stress.

Predictive modeling for susceptibility to anxiety attacks

Our supervised modeling approach used the Markov blanket of the attribute *soc5a* variable, i.e. age (attribute *age4)*, physical symptoms in the last seven days (attribute *phys7_4*), stay at home (attribute *phys2_18*), and prior clinical diagnosis of any mental health condition (attribute *phys3h*) as predictors.

The following three prediction scenarios were considered-

1. Mental issues *less than one day* in a week (class 1) vs. Mental issues *more than **one** day* in a week (class 0)
2. Mental issues *less than one day* in a week (class 1) vs. Mental issues *more than **three** days* in a week (class 0)
3. Mental issues *less than one day* in a week (class 1) vs. Mental issues *more than **five** days* in a week (class 0)

Random-Forest(RF) models achieved the best performance in comparison with SVM, logistic regression and naive-Bayes models on the basis of standard model performance indicators (accuracy, sensitivity, specificity, AUROC), summarized in table-1. We observed a decay (Accuracy 0.80 to 0.64, also confidence intervals mentioned in the table-1) in model predictability as we moved from high risk of depression (case-3) to low risk of depression (case-1) [table-1]. Such a trend was visible with all 4 machine learning techniques we used.

Table 1: Model Performance Indicators of the Supervised Model for Prediction of Stress.

Scenarios	RF	SVM	Naive Bayes	Logistic
Mental issues less than one day in a week (class 1) Vs Mental issues more than five days in a week (class 0)				
Accuracy (+- CI)	.80(+-.016)	.80(+-.016)	.77(+-.017)	.77(+- .017)
Sensitivity (+- CI)	.59(+-.063)	.56(+-.063)	.59(+-.063)	.59(+-.063)
Specificity (+- CI)	.82(+-.016)	.82(+-.016)	.79(+-.017)	.78(+-.017)
AU Roc (+- CI)	.71(+-.026)	.69(+-.026)	.69(+-.025)	.68(+-.025)

Mental issues less than one day in a week (class 1) Vs Mental issues more than three days in a week (class 0)				
Accuracy (+- CI)	.72(+- .018)	.72(+- .018)	.74(+-.017)	.73(+-.018)
Sensitivity (+- CI)	.6(+-.041)	.6(+-.041)	.56(+-.041)	.57(+-.041)
Specificity (+- CI)	.75(+-.018)	.75(+-.018)	.78(+-.017)	.76(+-.018)
AU Roc (+- CI)	.68(+-.022)	.67(+-.022)	.67(+-.022)	.67(+-.022)
Mental issues less than one day in a week (class 1) Vs Mental issues more than one day in a week (class 0)				
Accuracy (+- CI)	.66(+- .019)	.66(+- .019)	.65 (+-.019)	.62 (+- .019)
Sensitivity (+- CI)	.48(+-.027)	.49(+-.027)	.45(+-.026)	.61(+-.026)
Specificity (+- CI)	.77(+-.018)	.76(+-.018)	.77(+-.018)	.64(+-.020)
AU Roc (+- CI)	.62(+-.019)	.62(+-.019)	.61(+-.020)	.62(+-.018)

Discussion:

Mental Health is a serious public health concern. Mood disorders and suicide-related outcomes have increased significantly over the last decade among all age groups and genders [21 22]. The rapid spread of coronavirus infection forced governments across the world to close public gathering places, schools, colleges, restaurants, and industries. Social isolation, digital communication, working, and schooling from home have become the new normal and many jobs have been lost. Collectively, this has triggered a high level of anxiety, stress, and depression, globally. We did not find studies that have used models to not just predict but also to explain the subtle effects of life-situations on mental health. An explainable probabilistic graphical modeling approach with bootstraps and exact inference allowed us to capture many of these effects in a robust manner. Our study revealed that individuals having a prior diagnosis of any mental illness are the most vulnerable for mental illness during the COVID-19 phase, which recommends building national-level policies to regularly track their mental status and treat them accordingly. Most importantly, our results re-iterate the economic underpinnings of collective mental health response. Income assistance via Social Security or Supplemental Social Security had a demonstrable effect on the alleviation of anxiety as inferred from our model, which provides the first scientific evidence, to the best of our knowledge, proving the utility of such efforts. The effect of the extent of such measures may be captured in such modeling studies conducted in various parts of the globe, with widely varying assistance structures during this time.

Our findings from the United States can also stimulate further cultural and social research in other geographies with similar or different social structures. For example, the effects of in-person communication, as opposed to digital connectedness, may be different in countries where community living and joint families are still commonplace, e.g. India. Digital connectedness was not as effective as talking to a neighbor, at least in the United States highlighting that these are fundamentally different influences on mental health and need to be further explored in systematic studies. We conjecture that such differences may arise from the evolutionary mechanisms that have shaped human societies to live and share in close physical connectedness. Such an effect has been previously shown in primates kept in isolation who display depressive symptoms [23 24]. Similarly, parenting and its association with neuropeptide hormones may partially explain [25] our results that the presence of kids reduces anxiety levels. Interestingly, the COVID-19 pandemic has created a unique natural experiment on the collective mental health response of individuals to a health emergency.

The life-cycle of such a response may need to be further studied as the world goes through various phases of the pandemic to its resolution. However, our study indicates that the mental health impact is observable within a span of a few months, especially on young individuals. Further research will be needed, ideally in a longitudinal setting, where the same individuals can be surveyed again to understand the dynamics of the collective mental health response.

Our results also highlight that modern technological development in virtual communication is not able to replace natural socializing. Hence it becomes imperative to design better and empathetic technological tools that may shape a society that prevents isolation and alienation even while maintaining physical distancing and preventive measures for limiting spread. Personalization and contextualization of such measures will also be important as our results indicate that persons with previous mental health conditions may be disproportionately affected.

Finally, our results indicate that it may be possible to identify people at the highest risk of developing mental health disturbances. Our model achieved its best performance for those who were most vulnerable (having mental stress more than five days in a week) vs. least vulnerable people (having no stress or less than one-day stress in a week). This can help in the segmentation of vulnerable populations such as front-line healthcare workers and who are facing disproportionately higher levels of stress during this time.

A key factor in clinical and public health models is transparency and explainability in the face of complex interactions. Mental health variables are expected to have complex dependencies with potential confounding factors, mediation, and inter-causal dependency, therefore we extended our association analysis with data-driven structure learning of Bayesian network. We preferred this approach over black-box machine learning and standard statistical modeling because of several reasons. Structure learning allows us to discover and model confounding factors transparently whereas black-box machine learning models such as random forests and gradient boosted machines are not best suited for transparent reasoning. Standard statistical approaches make it humanly impossible to model interactions among hundreds of variables. Structure learning allows discovery and dissection of interactions into mediation, confounding, and inter-causal effects. The challenge of incorrect learning is addressed by ensembling many Bayesian networks (101 in our case) and choosing the ensemble voted structure. Our AI approach has earlier been validated for public health problems [26 27] and the current study demonstrates the under-explored potential of such an approach in complex mental health scenarios.

Our study has a few limitations. Establishing causal inference in cross-sectional data is nearly impossible and we acknowledge the possibility of confounding. However, this was precisely the reason we chose the structure learning approach as some of the confounding influences can be transparently discovered and explained. The ensemble voted structure over the sufficiently large number of bootstrapped structures is expected to be robust as a set of 101 BNs was found to be sufficiently large for the current study, to address the challenge of incorrect learning. Our approach is best suited as a probabilistic reasoning model explaining mental health determinants, and to make predictions, a useful outcome in the COVID-19 induced mental health morbidity. We could not explain why anxiety levels may be lower in persons with pre-existing cancer or hypertension. This may be a result of reduced work-environment related stress or more contact with family members at home. However, the current dataset is not suited to address this at a finer level of explainability. Also, we could not comment upon the temporality and persistence of these effects. Our results were currently limited to only one geography, i.e. the United States. However, the relatively large sample size and multi-ethnic involvement in the survey makes the model representative for most of the ethnicities and influences across the United States, hence likely to hold true in the United States. Finally, we believe that our study is a step in the use of explainable AI to predict mental health at a population level using survey data, hence making it broadly applicable. Survey datasets are notoriously noisy, and our approach achieved a balance between knowledge discovery and predictive accuracy of 80%, thus establishing a baseline under a novel scenario. Our algorithms can be used as a screening method for identifying individuals who need help and further studies with additional measurements and features may increase the accuracy of predictions. Therefore, predictive models for screening and assessing the mental health impact of COVID-19 is a crucial step towards proactive management and prevention of psychiatric comorbidities as populations continue to fight the pandemic.

Author Contribution:

Study Design : VK,TP,IJ , Dataset: IJ,AK , Data Analysis: IJ,RA , Paper writing: IJ,RA,TP Paper Review: VK,TP

Conflict of Interest:

None

Funding Statement:

The authors did not have specific funding for this work.

References:

1. COVID-19: Stimulating the economy and employment: As jobs crisis deepens, ILO warns of uncertain and incomplete labour market recovery. Accessed August 6, 2020. https://www.ilo.org/global/about-the-ilo/newsroom/news/WCMS_749398/lang--en/index.htm

2. Mrklas K, Shalaby R, Hrabok M, et al. Prevalence of Perceived Stress, Anxiety, Depression, and Obsessive-Compulsive Symptoms in Health Care Workers and Other Workers in Alberta During the COVID-19 Pandemic: Cross-Sectional Survey. *JMIR Ment Health*. 2020;7(9). doi:10.2196/22408

3. Psychiatry TL. Mental health and COVID-19: change the conversation. *Lancet Psychiatry*. 2020;7(6):463. doi:10.1016/S2215-0366(20)30194-2

4. Pereira-Sanchez V, Adiukwu F, Hayek SE, et al. COVID-19 effect on mental health: patients and workforce. *Lancet Psychiatry*. 2020;7(6):e29-e30. doi:10.1016/S2215-0366(20)30153-X

5. Chen Q, Liang M, Li Y, et al. Mental health care for medical staff in China during the COVID-19 outbreak. *Lancet Psychiatry*. 2020;7(4):e15-e16. doi:10.1016/S2215-0366(20)30078-X

6. Mahase E. Covid-19: Mental health consequences of pandemic need urgent research, paper advises. *BMJ*. 2020;369:m1515. doi:10.1136/bmj.m1515

7. Pfefferbaum B, North CS. Mental Health and the Covid-19 Pandemic. *N Engl J Med*. 2020;0(0):null. doi:10.1056/NEJMp2008017

8. Seabrook EM, Kern ML, Rickard NS. Social Networking Sites, Depression, and Anxiety: A Systematic Review. *JMIR Ment Health*. 2016;3(4). doi:10.2196/mental.5842

9. Bruen AJ, Wall A, Haines-Delmont A, Perkins E. Exploring Suicidal Ideation Using an

Innovative Mobile App-Strength Within Me: The Usability and Acceptability of Setting up a Trial Involving Mobile Technology and Mental Health Service Users. *JMIR Ment Health*. 2020;7(9). doi:10.2196/18407

10. Gao J, Zheng P, Jia Y, et al. Mental health problems and social media exposure during COVID-19 outbreak. *PLOS ONE*. 2020;15(4):e0231924. doi:10.1371/journal.pone.0231924

11. Homepage | COVID Impact Survey. Untitled. Accessed May 25, 2020. https://www.covid-impact.org

12. COVID Impact Survey. Untitled. Accessed May 25, 2020. https://www.covid-impact.org/about-the-survey-questionnaire

13. Revelle W. *Psych: Procedures for Psychological, Psychometric, and Personality Research*.; 2020. Accessed June 26, 2020. https://CRAN.R-project.org/package=psych

14. Gámez JA, Mateo JL, Puerta JM. Learning Bayesian networks by hill climbing: efficient methods based on progressive restriction of the neighborhood. *Data Min Knowl Discov*. 2011;22(1):106-148. doi:10.1007/s10618-010-0178-6

15. Bozdogan H. Model selection and Akaike's Information Criterion (AIC): The general theory and its analytical extensions. *Psychometrika*. 1987;52(3):345-370. doi:10.1007/BF02294361

16. Yedidia JS, Freeman WT, Weiss Y. Generalized Belief Propagation. In: Leen TK, Dietterich TG, Tresp V, eds. *Advances in Neural Information Processing Systems 13*. MIT Press; 2001:689-695. Accessed June 28, 2020. http://papers.nips.cc/paper/1832-generalized-belief-propagation.pdf

17. Sethi T, Maheshwari S. *WiseR: A Shiny Application for End-to-End Bayesian Decision Network Analysis and Web-Deployment*.; 2018. Accessed May 8, 2020. https://CRAN.R-project.org/package=wiseR

18. Probabilistic Graphical Models: Principles and Techniques - Daphne Koller, Nir Friedman - Google Books. Accessed August 9, 2020.

19. Chawla NV, Bowyer KW, Hall LO, Kegelmeyer WP. SMOTE: Synthetic Minority Over-sampling Technique. *J Artif Intell Res*. 2002;16:321-357. doi:10.1613/jair.953

20. scikit-learn: machine learning in Python — scikit-learn 0.23.1 documentation. Accessed June 26, 2020. https://scikit-learn.org/stable/

21. Bilsen J. Suicide and Youth: Risk Factors. *Front Psychiatry*. 2018;9. doi:10.3389/fpsyt.2018.00540

22. Brådvik L. Suicide Risk and Mental Disorders. *Int J Environ Res Public Health*. 2018;15(9). doi:10.3390/ijerph15092028

23. Hennessy MB, McCowan B, Jiang J, Capitanio JP. Depressive-like behavioral response of adult male rhesus monkeys during routine animal husbandry procedure. *Front Behav Neurosci*. 2014;8. doi:10.3389/fnbeh.2014.00309

24. McKinney WT. Primate Social Isolation: Psychiatric Implications. *Arch Gen Psychiatry*. 1974;31(3):422-426. doi:10.1001/archpsyc.1974.01760150122018

25. Kohl J, Autry AE, Dulac C. The neurobiology of parenting: A neural circuit perspective. *BioEssays News Rev Mol Cell Dev Biol*. 2017;39(1):1-11. doi:10.1002/bies.201600159

26. Sethi T, Mittal A, Maheshwari S, Chugh S. Learning to Address Health Inequality in the United States with a Bayesian Decision Network. *Proc AAAI Conf Artif Intell*. 2019;33(01):710-717. doi:10.1609/aaai.v33i01.3301710

27. Awasthi R, Patel P, Joshi V, Karkal S, Sethi T. Learning Explainable Interventions to Mitigate HIV Transmission in Sex Workers Across Five States in India. *ArXiv201201930 Cs*. Published online November 30, 2020. Accessed February 1, 2021. accepted at NeurIPS 2020 Workshop on Machine Learning for the Developing World. http://arxiv.org/abs/2012.01930

Figure 1: Socio-demographics of respondents who participated in the survey. It can be seen that there was almost a similar representation from both genders. Age groups 25-75 were predominantly captured in the survey. Most of the respondents had received education BA or above and were nearly equally distributed across geographies within the US**.**

Figure 2: a) Outline of the analytical pipeline. b) Item Reliability Analysis of Mental Health Indicators revealed a high

degree of internal consistency (Cronbach's alpha value >0.70) for most of the psychological variables, thus indicating suitability for the modeling exercise.

Figure 3: (a) Gender-wise and (b) age-wise distribution of Mental Issues (attribute soc5a) variable. Significance was tested using two proportion z-test and chi-square test respectively, showing a higher prevalence of mental issues among youth and in women.

Figure 4. (a) Consensus structure learned through 101 bootstrapped samples. Hill-climbing search along with Bayesian Information Criterion was used to learn the structures and connections having edge strength and direction strength more than 90% are shown. The color of the edges represents the proportion of networks in which that edge was present in the 101 bootstrapped samples, an indicator of confidence; **(b) attribute soc5a was found to be the parent Node of all other Mental Health Variables**, therefore, leading to our choice of this variable as the primary dependent variable.

Figure 5: Inferences from the Bayesian network: The difference in inferred probability was calculated after conditioning on the independent variables. A positive association implies a mental-stress-inducer whereas a negative association implies a mental-stress-reducer. The red circle shows the mean value with green and blue showing confidence intervals.

Supplementary Table 1: variable groups as Indicators

Code	Attributes Description
mental_health_indicators	
(questions related to mental stress in last seven days)	
soc5a	**Felt nervous, anxious, or on edge**
soc5b	**Felt depressed**
soc5c	**Felt lonely**
soc5d	**Felt hopeless about the future**
soc5e	**sweating, trouble breathing, pounding**

	heart, etc
work_from_home_indicators	
(questions related to work and school from home)	
phys2_9	work from home
phys2_10	study from home
phys2_4	postponed work activities
phys2_5	postponed school activities
communication_indicators	
(questions related to communication with friends/neighbor/family)	
soc2a	in-person communication during COVID
soc2b	in-person communication before COVID
soc3a	digital communication during COVID
soc3b	digital communication before COVID
covid_symptoms_indicators	
(questions on covid symptoms Prescribed by WHO)	
phys1a	Fever
phys1b	Chills
phys1c	Runny or stuffy nose
phys1d	Chest congestion
phys1e	Skin rash
phys1f	Cough
phys1g	Sore throat
phys1h	Sneezing
phys1i	Muscle or body aches
phys1j	Headaches
phys1k	Fatigue or tiredness
phys1l	Shortness of breath
phys1m	Abdominal discomfort
phys1n	Nausea or vomiting
phys1o	Diarrhea
phys1p	Changed or lost sense of taste or smell
phys1q	Loss of appetite
clinical_history_indicators	

(questions on prior clinical history of diseases)	
phys3a	**Diabetes**
phys3b	**High blood pressure or hypertension**
phys3c	**Heart disease, attack, stroke**
phys3d	**Asthma**
phys3e	**Chronic lung disease or COPD**
phys3f	**Bronchitis or emphysema**
phys3g	**Allergies**
phys3h	**A mental health condition**
phys3i	**Cystic fibrosis**
phys3j	**Liver disease**
phys3k	**Cancer**
phys3l	**compromised immune system**
phys3m	**obesity**
behavioural_indicators	
(questions on behavioral changes proposed/enforced by government)	
phys2_1	**Canceled a doctor appointment**
phys2_2	**Worn a face mask**
phys2_3	**Visited a doctor or hospital**
phys2_7	**Canceled outside housekeepers**
phys2_8	**Avoided restaurants**
phys2_11	**postponed pleasure, social activities**
phys2_12	**Stockpiled food or water**
phys2_13	**Avoided public or crowded places**
phys2_16	**Washed or sanitized hands**
insurance_assistance_indicators	
(questions on insurance and different assistance)	
econ6a	**Unemployment insurance**
econ6b	**SNAP**
econ6c	**TANF**
econ6d	**Social Security**
econ6e	**Supplemental Social Security**
econ6f	**Health insurance**

econ6g	aid from the government
Other_useful_Indicators	
phys7_4	did not feel hot, chilly, sweating
phys2_18	Stayed home because I felt unwell

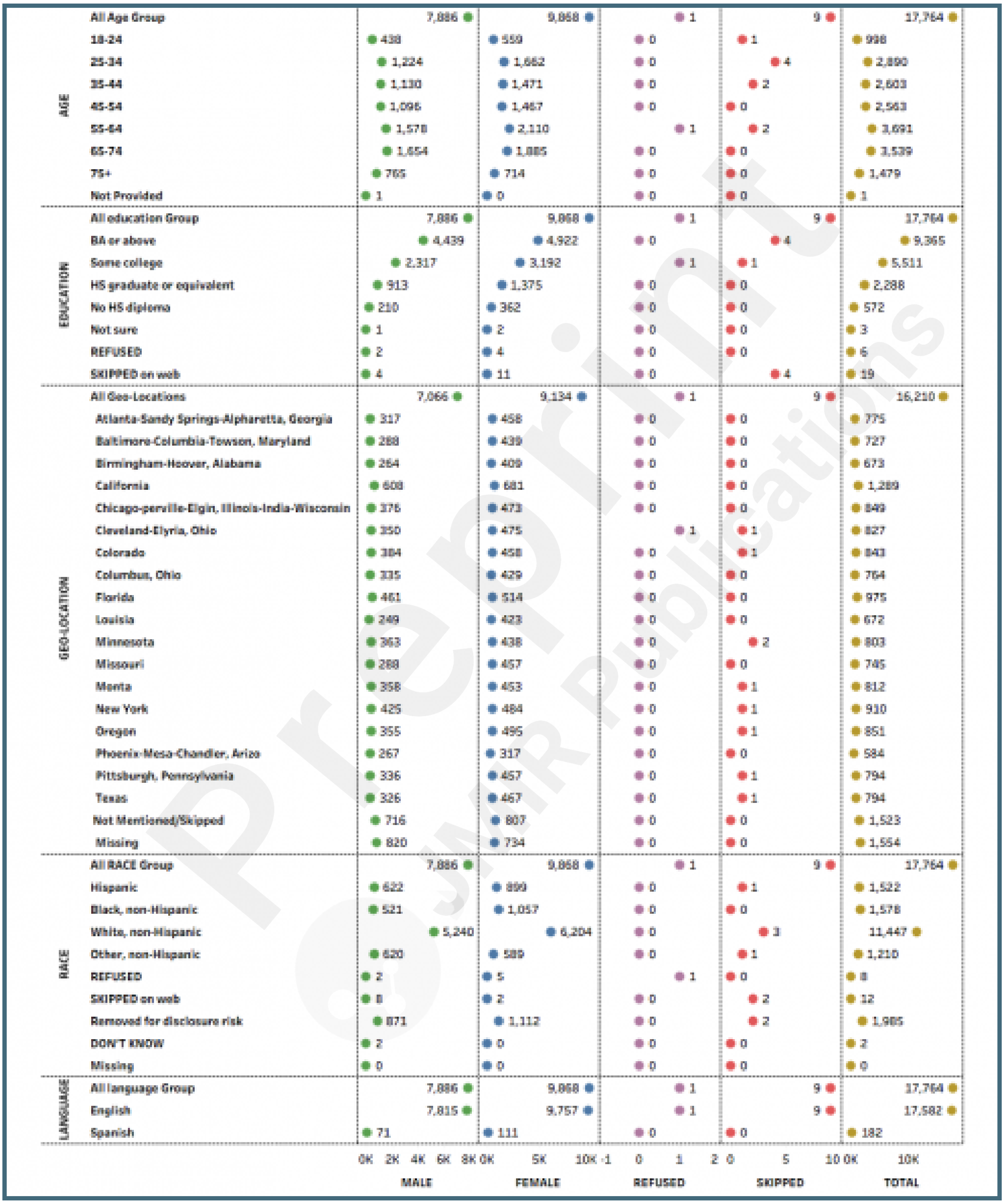
Supplementary Figure 1: Heatmap for negative log p values for the significant association.

Supplementary Figure 2: Bayesian Network topology

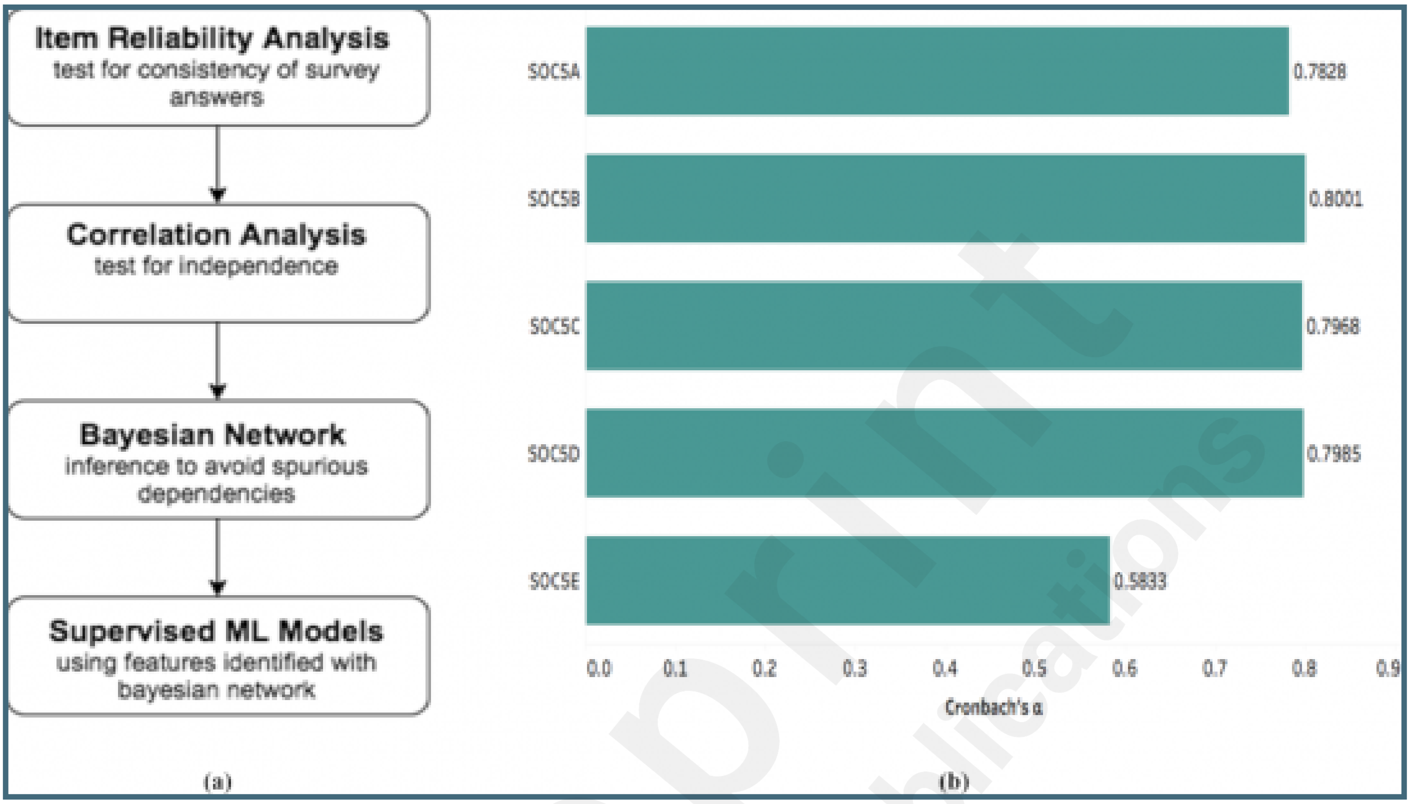
Supplementary Files

Figures

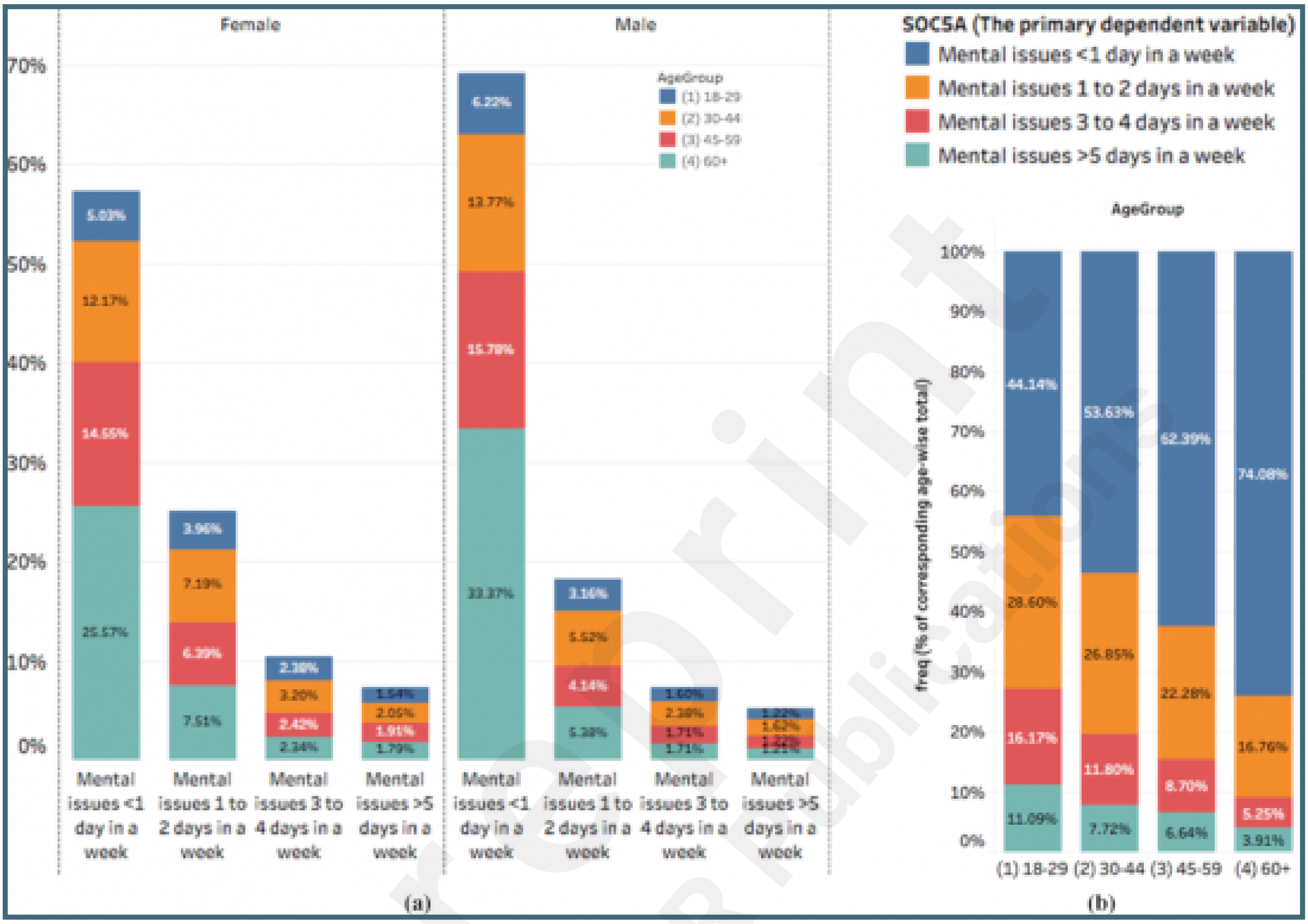
Socio-demographics of respondents who participated in the survey. It can be seen that there was almost similar representation from both genders. Age groups 25-75 were predominantly captured in the survey. Most of the respondents had received education BA or above and were nearly equally distributed across geographies within the US.



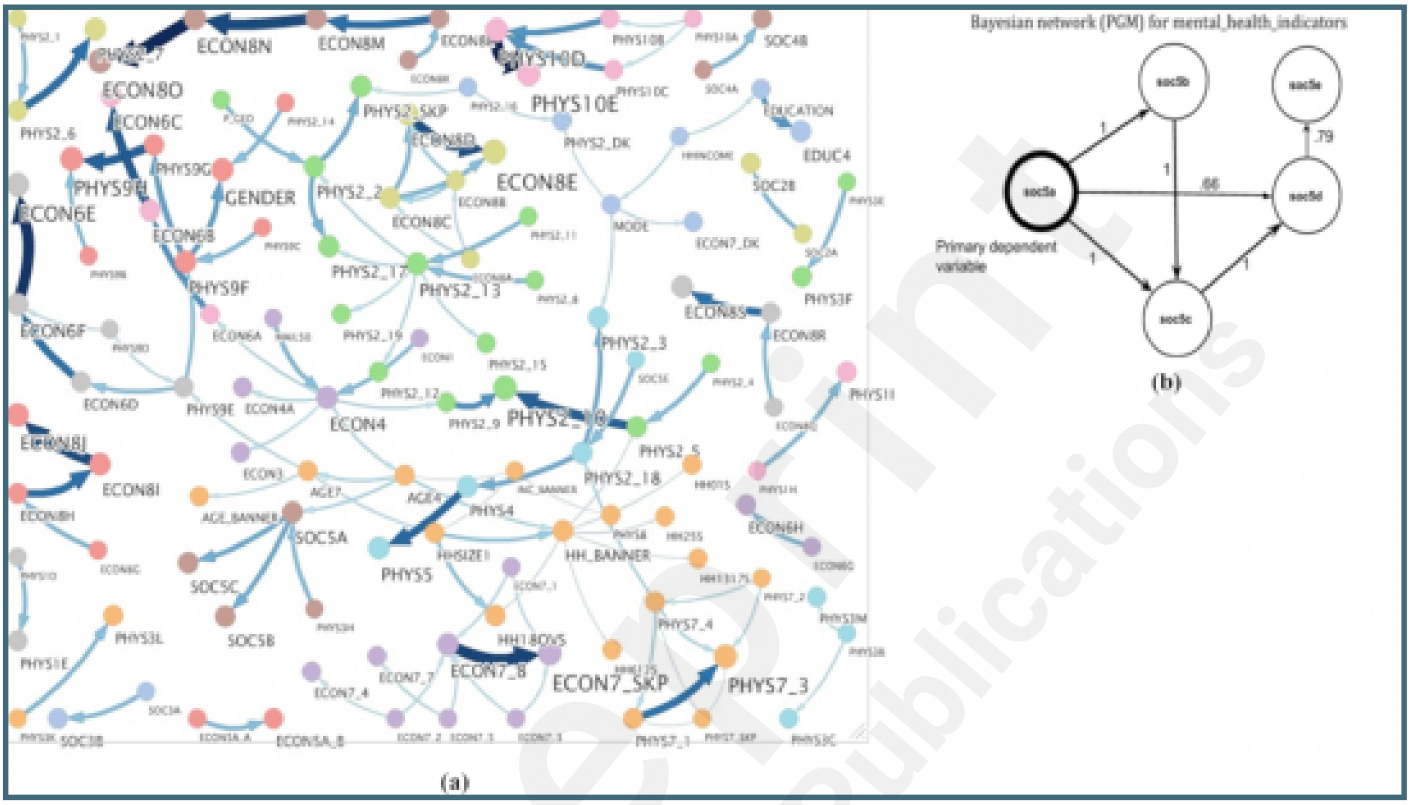
a) Outline of the analytical pipeline. b) Item Reliability Analysis of Mental Health Indicators revealed a high degree of internal consistency (Cronbach's alpha value >0.70) for most of the psychological variables, thus indicating suitability for the modeling exercise.



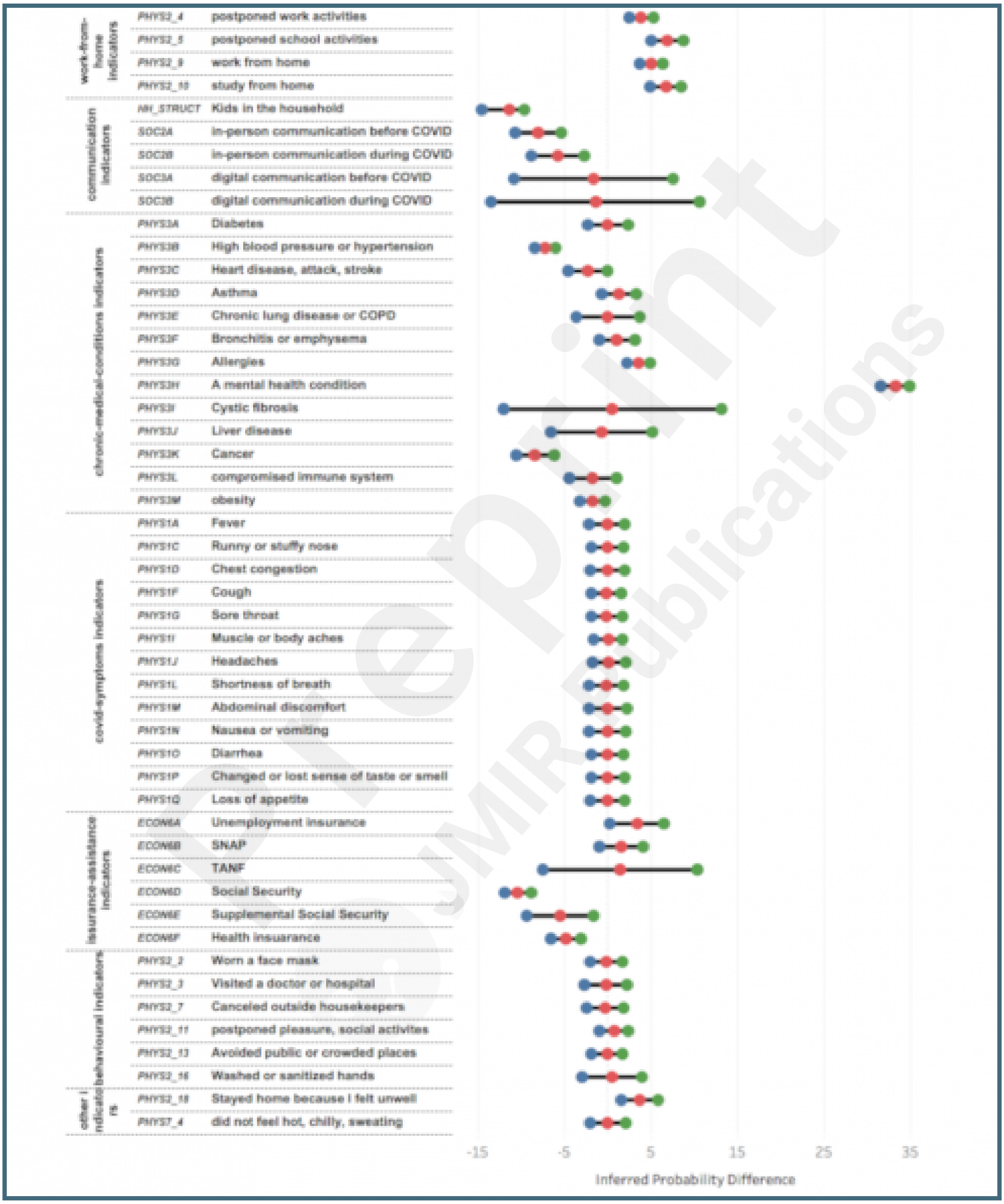
(a) Gender-wise and (b) age-wise distribution of Mental Issues (attribute soc5a) variable. Significance was tested using two proportion z-test and chi-square test respectively, showing a higher prevalence of mental issues among youth and in women.



. (a) Consensus structure learned through 101 bootstrapped samples. Hill-climbing search along with Bayesian Information Criterion was used to learn the structures and connections having edge strength and direction strength more than 90% are shown. The color of the edges represents the proportion of networks in which that edge was present in the 101 bootstrapped samples, an indicator of confidence; (b) attribute soc5a was found to be the parent Node of all other Mental Health Variables, therefore, leading to our choice of this variable as the primary dependent variable.



Inferences from the Bayesian network. The difference in inferred probability was calculated after conditioning on the independent variables. A positive association implies a mental-stress inducing factor whereas a negative association implies mental-stress reduction factor. The red circle shows the mean value with green and blue showing confidence intervals.



Multimedia Appendixes

Table 1: Model Performance Indicators of the Supervised Model for Prediction of Stress.
URL: https://asset.jmir.pub/assets/26001a02a0f631c9e55a0033419da8f2.xlsx

Supplementary Table 1: variable groups as Indicators.
URL: https://asset.jmir.pub/assets/9b64b6f0799b18c03436e872bb7dd4b3.xlsx