Understanding the Pandemic's Effects on Workforce Dynamics

Group Leader: Ada Terzioğlu

Ümmügülsüm Çiftçi | Fatma Deniz Bektaş | Batuhan Ayranlı

Elif Tuğba Kanar | Dide Çalışkan | Şeyha Baran Erdoğan

***Abstract*—** SARS-CoV-2, or coronavirus is a disease that caused a global panic and resulted in an almost two-year quarantine period. The expression COVID-19 represents "coronavirus disease 2019." (University of Virginia Health, n.d.). It is undeniable that the COVID-19 pandemic has both mental and physical effects on humans. Post COVID-19 period resulted with anxiety, depression, reduced physical capacity, pain and more health problems (Shanbehzadeh et al., 2021). It can be anticipated that these problems would cause complications in people's personal and professional lives. This paper searches for professional life complications with exploratory data analysis, using seven different research questions. At the end of the research, it becomes clear that COVID-19 affected humans' mental health in such a way that their performance in their professional lives' changes.

***Keywords*—**COVID-19, stress level, company support level, health impact, remote work, productivity change

1)INTRODUCTION

This report examines COVID-19 dataset covering various aspects of business life statistics in different departments, age groups, gender, work hours etc. The focus is to analyze and understand the patterns and correlations between different economic, socioeconomic, psychological factors. Through this analysis, the report aims to identify significant trends and provide insights that can inform policymaking and societal comprehension of COVID-19 dynamics.

1.1)Data Description

There are a 1000 observations and 15 columns. The variables used in these observations are represented in the table below.

Variable	Description
Sector	The sector of employment (Healthcare, Education, Finance, Technology, Retail, Manufacturing, Other)
Job Role	The individual's job role (Manager, Staff, Freelancer, Intern, Executive)
Employment Type	Employment status (Full-Time, Part-Time, Self-Employed, Unemployed)
Gender	Gender of the individual (Male, Female, Other)
Age Group	Age group of the individual (18-24, 25-34, 35-44, 45-54, 55+)
Remote Work Frequency	Frequency of remote work during the pandemic (Never, Rarely, Occasionally, Often, Always)
Work Hours Increase (%)	Change in work hours compared to pre-pandemic levels
Productivity Change (%)	Change in productivity compared to pre-pandemic levels
Stress Level (1-10)	Self-reported stress level on a scale of 1 to 10
Job Satisfaction Change	Change in job satisfaction on a scale of -5 to 5
Income Change (%)	Change in income compared to pre-pandemic levels

Tech Tools Use Increase (%)	Increase in the use of technological tools compared to pre-pandemic levels
Health Impact (1-10)	Self-reported health impact on a scale of 1 to 10
Family Work Balance Impact (1-10)	Impact on family-work balance on a scale of -5 to 5
Company Support Level (1-10)	Perceived company support level on a scale of 1 to 10

Also, the descriptive statistics to have a clear understanding of the data set were obtained through the phyton codes in data cleaning steps in part 2. The table showing these values is provided in below, Figure1.

	Work hours increase (%)	Productivity change (%)	Stress level (1-10)	Job satisfaction (-5,5)	Income change (%)	Tech tools use increase (%)	Health impact (1-10)	Family work balance impact (-5,5)	Company support level (1-10)
count	1000.000000	1000.000000	1000.000000	1000.000000	1000.00000	1000.000000	1000.000000	1000.000000	1000.000000
mean	9.408310	-2.995480	5.486500	0.308900	-4.98091	30.128310	5.399200	-0.147400	5.476500
std	15.198916	19.699374	3.007842	3.023507	10.28614	12.364212	2.862816	3.149481	2.825541
min	-35.110000	-67.030000	1.000000	-5.000000	-38.48000	-11.170000	1.000000	-5.000000	1.000000
25%	-0.807500	-16.262500	3.000000	-2.000000	-11.96000	22.170000	3.000000	-3.000000	3.000000
50%	9.410000	-2.835000	5.250000	0.000000	-5.01500	30.085000	5.000000	0.000000	5.000000
75%	18.642500	10.240000	8.000000	2.500000	1.87000	38.687500	8.000000	3.000000	8.000000
max	58.650000	59.260000	10.000000	7.000000	27.76000	73.980000	10.000000	5.000000	10.000000

Figure1: *Descriptive statistics of the continuous variables*

Looking at Figure1, it can be seen that the variables with intervals (stress level, job satisfaction, health impact and company support level) have means which are very near to their intervals' middle values. On the other hand, the variables with percentages seem to have a mean smaller than zero, despite the variables Work Hours Increase and Tech Tools Use Increase.

A lot of interpretations can be made looking at this table, which gives a clear understanding of the data set. Overall, COVID-19 seems to have negative effects in general (decrease in productivity, decrease in income, low value of job satisfaction…).

2) DATA TIDYING AND CLEANING

First of all, for easier adjustments the name of the data set was changed to "COVID-19". Anaconda/Jupiter Notebook and phyton codes (NumPy and pandas libraries), was used to clean the data set. The file was uploaded to Anaconda/Jupiter Notebook, read as df. The notation (;) was used as an indicator to separate values while reading.

With head and tail functions the uploaded data was checked and compared to the original data set to control whether we uploaded it correctly.

With info() function, the existence of NA values, headers was checked and we had an overall look of the dataset. The column names were checked and converted into one single format (all lower-case letters).

Using duplicated() function the duplicated values were found and eliminated. (The index values 501 and 521 were dropped.) Later, this elimination was controlled with the slicing method to make sure we did not lose any values.

After that, lstrip() and rstrip() functions were used to get rid of unnecessary white spaces.

The uniqueness of the string values were checked and we came up with problematic values such as ytern'-intern, free-freelancer, o-other… After correct adjustments, all values were unique.

With the type function, data types were checked. To convert the percentage columns into float type "%" was eliminated.

The descriptive statistics were analyzed with the describe() function and for job satisfaction, we encountered with outliers which were converted to the mean value.

The percentage of null values around 2%. Since the percentages were not so high, the null values of the categorical variables were changed into mode and float variables were changed into their mean. For this null() function was used but this function only perceive NA as null and does not see NaN as null values. Therefore, NaN values were converted to mean/mode with the replace() function.

3) EXPLORATY DATA ANAYLSIS

3.1)How has the COVID-19 pandemic differently impacted employee health and job

satisfaction across sectors? (Ümmügülsüm Çiftçi, 2666501)

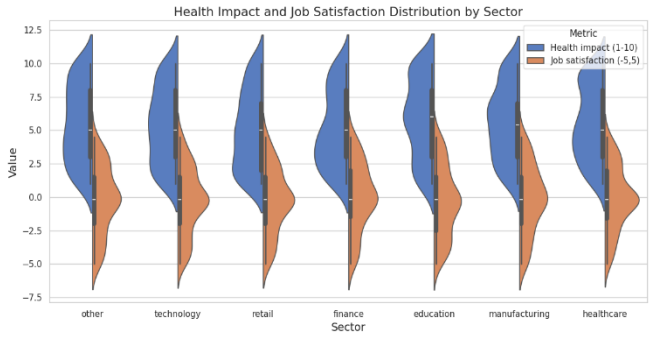

Figure2: *Health Impact and Job Satisfaction Distribution by Sector*

The COVID-19 pandemic has significantly impacted the health and job satisfaction of workers across various sectors. In Figure2, we can see that education sector was the most negatively affected in terms of health, it has the least job satisfaction. Educators faced a lot of issues due to the transition to remote education, the lack of sufficient infrastructure and work in this regard, the decreasing success of students and the increasing workload in this field. Furthermore, they have been exposed to a lot of technological tools. Therefore, they had a lot of physical and psychological ailments, and they were stressing out more. Because of this, their job satisfaction is less than others. On the other hand, retail sector has the least health impact. Normally, it is expected that retail sector has more health impact since this jobs require more physical contact, and people who work in this fields can be affected easily in terms of health. However, contrary to expectations, this sector was not affected that much in terms of health. This may have been due to health precautions in the industry and regulations for employees to do their jobs. However, these results may not completely eliminate the long-term health risks faced by workers in this sector.

In Figure.1, we can see that finance sector has highest level of job satisfaction. This may have been due to the fact that the sector has adapted well to technology and provided a stable working environment during the crisis. Despite sectoral differences, job satisfaction has declined across all sectors during the pandemic. The uncertainties created by the pandemic, the increasing workload and the deterioration in the work-life balance can be negatively affected the overall satisfaction levels of the employees. Accordingly, we can say that the pandemic has created difficulties in the entire business world.

3.2) How do stress levels vary in different sectors during the pandemic? (Ada Terzioğlu, 2666774)

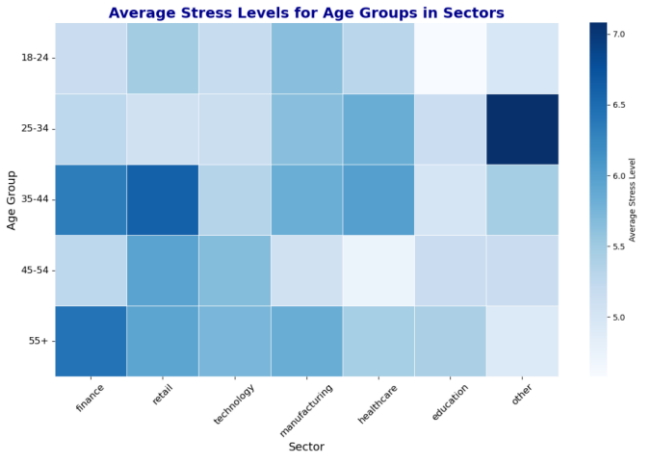

Figure3: *Heatmap showing average stress levels for age groups in sectors*

The Figure3 is a heat map displaying age groups on the y-axis and different sectors on the x-axis. The colors are scaled based on average stress levels, with dark blue indicating higher levels and lighter shades representing lower levels.

As can be observed on the heat map, people in the 35-44 age group rank their stress levels the highest, followed by people over 55.

According to a survey made by the American Psychological Association (APA), the reasons behind these high-stress levels for people between the ages of 35 and 44 are the fear of economic instability and the responsibility for providing for their families. Whereas for people over 55, stress is based on health issues. Being at high risk for a virus during a global lockdown and having other health-related issues due to physical inactivity made them more stressed. (PAR, Inc.,2024)

When looked at these two age groups it can be observed that the finance and retail sectors have the highest average stress level compared to the others. In the retail sector, Shanafelt interpreted that most of the stress was based on being in close contact with the public, putting the workers at high health risks since there was no option of working remotely (2020).

For finance it was the opposite, according to Shanafelt (2020), this blurred the lines between home and work life, resulted in more stress for the workers at that drastically market volatility time.
The expectation was that health workers under intensive working conditions during the COVID period were the most stressful sector group and that young adults were more affected by the global pandemic than other age groups. However, the average stress levels of health workers are mostly aligned with those in other sectors, if not lower. The most stressed group in the dataset is adults aged 25-34 who work in an undefined sector.
The least stressful workers are young adults 18-24 who work in education. Education workers were also the least stressful employees nearly in every age group.

3.3) Does stress level have any effect on health? (Şeyha Baran Erdoğan, 2558435)

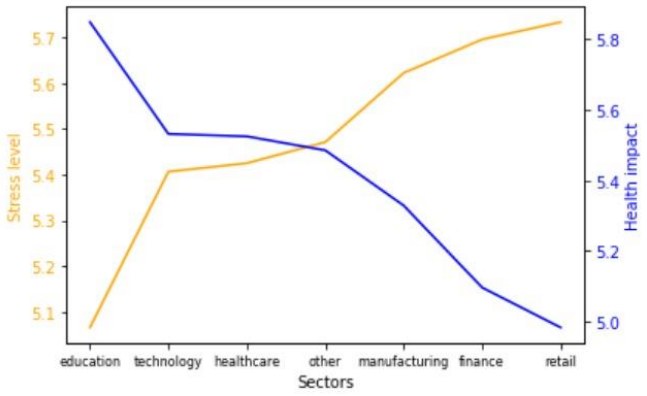

Figure4: *Correlation between stress level and health by sectors*

As it is known, the effect of stress in our lives increases day by day. Moreover, the impact of this has increased even more during the COVID process. During this time, general health parameters also changed noticeably. According to the research made by Schneiderman, stress has a negative impact on health as he said "…if stressors are too strong and too persistent in individuals who are biologically vulnerable because of age, genetic, or constitutional factors, stressors may lead to disease…" (Schneiderman et al., 2004).

To demonstrate this, there are two graphs showing the status of stress levels and health indicators by sector. A negative correlation was expected as seen in these 2 stem-shaped graphs. It would be more reasonable to show this in a single chart to control. Exactly as expected, there is a negative correlation between health and stress.

The image shows the stress levels of employees according to sectors (on a scale of 1-10):

The Highest Stress Level: The financial and retail (avg 5.7).

The Lowest Stress Level: Education sector (avg 5.1).

Technology, healthcare, manufacturing and other sectors have moderate stress levels ranging from 5.3 to 5.6. Dec.

The health data obtained from the second visual reveals the average health scores of the sectors (high scores indicate a more positive health status):

Highest Health Value: The education sector is the most positive sector in terms of health with a value close to 6.

Health Values at an Intermediate Level:
Manufacturing sector: 5.6
Health sector: 5.5
Other sectors: 5.4
The Lowest Health Values:
Technology: 5.2
Finance: 5.1
Retail: 5

Noteworthy points when comparing the stress levels and health values of the sectors:

Education Sector:This sector, where the stress level is the lowest, also has the highest health score. This relationship may reflect the positive effect of low stress levels on health.

Finance and Retail: These sectors with the highest stress levels also have the lowest health values. This suggests that high stress leads to a negative impact on health.

Other Sectors: Moderate levels of stress and health values have been observed in the manufacturing, health and technology sectors. Especially the manufacturing sector attracts attention with its positive health value despite the relatively high stress level.

Conclusion: As a result, research already indicates that there is a clear negative correlation between stress levels and health. Graphics did not surprise It has been observed that health indicators are lower in sectors where stress is high.

3.4) How does a company's support towards their employees affect the productivity changes after COVID-19 in a specific sector? (Dide Çalışkan, 2712933)

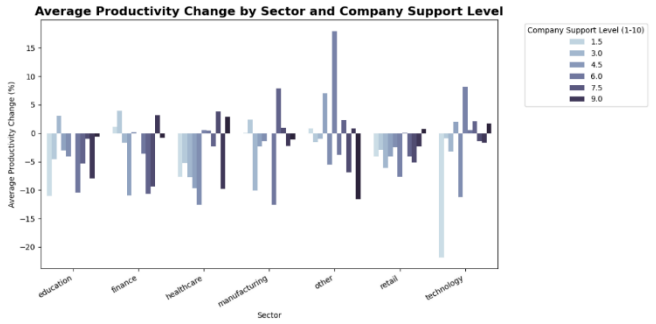

Figure5: *Average productivity change by sector and company support level*

It is expected that as the company support level (1-10) increases productivity will be affected positively since the employees will be in a warm and supportive environment. So, a linear relation between two variables is anticipated. However, after taking COVID-19 pandemic's negative effects on human mentality into consideration, the relation between the variables may not be as linear as expected. To observe this, the bar plot provided in Figure5 was created.

The following graph is a bar chard regarding company support levels and the change in productivity. The x-axis represents different sectors. The y-axis provides the average productivity change as percentages. As the shade of blue gets darker in the bars, company support levels increase.

In general, it is seen that low company support levels result in decreasing productivity changes. However, there are outliers which are finance, manufacturing and others. It was anticipated that as the support level increases productivity would increase as well but this is not easily observed. For example, in the other sector even a support level of 9 could not prevent a sharp decrease in productivity. There is not a linear relationship between support level and productivity change. However, as expected, the sharpest decrease in productivity also has a low support level which is around 1.5. The highest increase in productivity has an average support level, which is around 5.

With all these findings, it can be said that even though company support influenced the productivity of its workers, this was a minor effect. This is probably due to the fact that the COVID-19 pandemic resulted in many mental challenges which affected people's professional life in a negative way.

3.5) How did the frequency of remote work during COVID-19 affect workers' productivity changes? (Fatma Deniz Bektaş, 2666477)

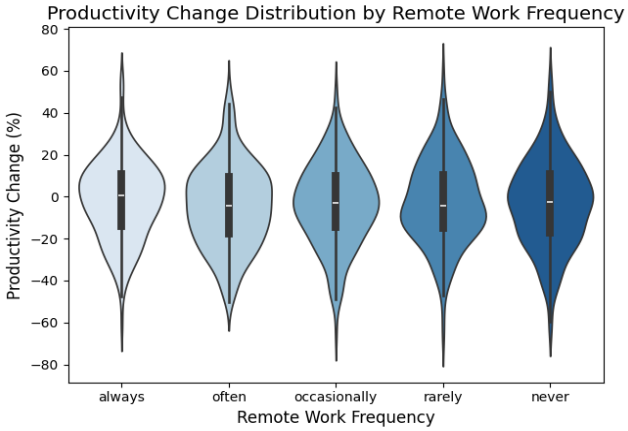

Figure6*: Productivity change distribution by remote work frequency*

Before comparing the variables, productivity change and remote work frequency, it is expected to see more positive changes in employee productivity as the frequency of teleworking increases. As remote working increases, employees may be more productivite due to flexible working hours and reduced distractions. On the other hand, employees falling under "rarely" and "never" categories may experience a decline in productivity as they are not used to working remotely.

The violin plot in Figure6 shows the distribution and density of the data. Wide regions represent more dense data and narrow regions represent sparse data. The line in the middle represents the median and the extremes represent the minimum and maximum values. The symmetry of the figure shows if it is a balanced distribution. Also a lighter shade indicates the frequecny category

“always” and as the shade gets darker frequency gets closer to category “ never”.

This graph presents the change in workers' productivity in relation to COVID-19 according to the frequency of teleworking: always, often, occasionally, rarely, never. These violins provide information on the intensity and distribution of the percentages of change for these productivity categories. The variation in changes for the “always” and “often” telecommuting groups lies mainly between -20% and +20%. In fact, positive productivity changes are more frequent in these groups, implying that working at home can be more productive for these groups.

On the other hand, in the “rarely” and “never” telecommuting groups, productivity change percentages are spread over a wider range (approximately -60% to +20%) and negative changes are more dominant. Most of the time, productivity declines when workers work from home less frequently. This graph reveals the fact that productivity is more stable and positive on days when people work more from home. These findings conclude that increasing frequency of remote work results in an escalation of employees’ productivity, as anticipated.

3.6) How does productivity change percentages vary from age groups in different sectors? (Batuhan Ayranlı, 2649622)

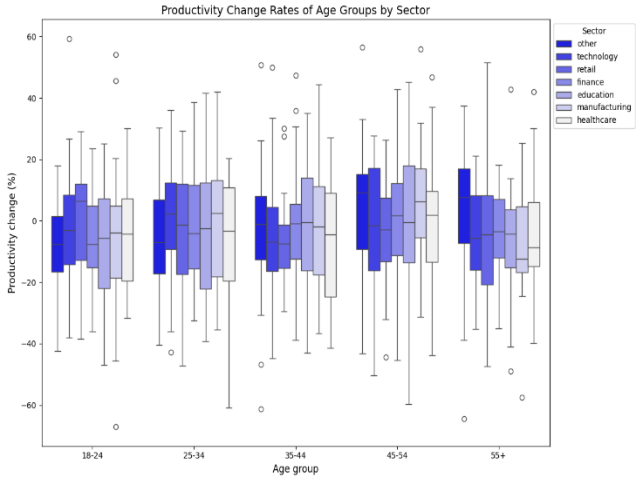

Figure7*: Productivity change in age groups and sectors*

This boxplot shows the distribution of productivity change rates by age groups in different sectors. In the chart, an attempt was made to compare the change rates of the sectors in each age group. There are categorized age groups as 55+, 45-54, 35-44, 25-34 and 18-24, respectively. Efficiency change rates are given as percentages between -60% and +60%. If we examine the graph, the highest productivity increase is observed in the 18-24 age group. In this age group, the maximum values in the box plots are generally higher than in other age groups. In other words, the productivity change rate in younger age groups (18-24, 25-34) has a wider range and more variation. While the overall productivity change in middle age groups (35-44, 45-54) is more balanced and stable compared to young people, productivity remains stable in the older group (55+), although there is a very slight positive change.

Sectorally, the upper limits and outliers of the Technology sector in the box plots are at higher levels compared to other sectors. Likewise, the Retail sector also has a wider distribution. These may indicate that rates of productivity change are more unpredictable or dynamic than in other sectors. There are no excessive positive or negative changes observed between age groups in these sectors, so productivity changes are more balanced and predictable.

3.7) How has COVID-19 pandemic affected income levels across sectors by gender? (Elif Tuğba Kanar, 2666600)

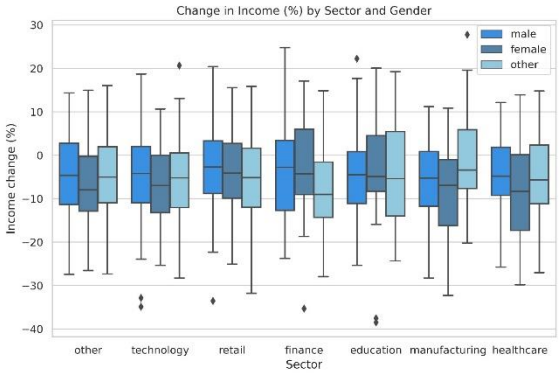

Figure8: *Change in income and sector by gender*

The main goal to create this plot was to observe the percentage of income change after COVID19 in

Technology, Retail, Finance, Education, Manufacturing, Healthcare and Other sectors while distinguishing between genders (male, female and other). A box plot is used due to its ability to show great understanding for 2 categorical (Sector and Gender) and a numerical (Income change (%) variable. The first thing which can be noticed in this figure is that all the median values are within the range of 0 and –10. That is, generally the income levels are affected negatively by COVID19 for all the genders and sectors. This was really shocking to see since many could have guessed a more stable income change in the technology sector due to its significant flexibility or for the healthcare workers to earn more because of their huge efforts. Nevertheless, it is also important to consider whiskers and skewness also affect whether the majority of people in a specific sector and gender deals with positive or negative higher rates of income change. For instance, if the plot is positively skewed like the female workers in education, this means incomes of a few people got affected negatively.

At first glance, in technology, other and retail sectors, their ranges overlap for each gender, whereas for the rest, there are differences. The wider the range, the more variation and the more variation, the less stable income a job offers in a sector. For example, female healthcare workers and others working in education sector may face higher income changes suggesting a gender difference in income levels. Most of the outliers are negative outliers that falls in the range of –30% and –40%, however; there is also some positive outliers too. These outliers may be due to job losses or advancements, but that needs further investigation to totally understand the reasons.

4) CONCLUSION

The aim of this research was to explore the effects of the global pandemic on the professional and personal lives of employees from different sectors and age groups.

Unusual working conditions affected the job satisfaction levels drastically. Especially in the education sector, excessive usage of technology and different teaching techniques caused dissatisfaction among employees.

People in the 35-44 age group who work in finance and retail ranked their stress levels the highest, because of the unstable economy and high health risks in their sector.

The correlation between stress levels and health impact is strongly negative. The data collected from different sectors confirms this relationship.

Company support contributed to the productivity of their employees. However, the correlation between the support levels and productivity change was weak so, it can't be said that it was the only reason behind the productivity changes.

There is a positive correlation between the freshness of remote work and the productivity of the employees. People who do remote work are reported to be more productive than the ones who don't.

The productivity change for different age groups was mostly negative for young adults and more balanced for older age groups.

In every sector, the income of the employees decreased during the pandemic. Women employees were the most affected gender group by this economic crisis. The research highlights that the COVID-19 pandemic has disrupted professional and personal life, with significant variation in the effects on stress, health, job satisfaction, productivity, and income across sectors, age groups, and genders. The key findings underline the critical role of sector-specific challenges, working from home, and organizational support in shaping these outcomes.

5) REFERENCES

PAR, Inc. (2024, July 8). Understanding stress in America. PAR, Inc. https://www.parinc.com/learning-center/par-blog/detail/blog/2024/05/23/understanding-stress-in-america

Schneiderman, N., Ironson, G., & Siegel, S. D. (2004). Stress and health: psychological, behavioral, and biological determinants. Annual Review of Clinical Psychology, 1(1), 607–628. https://doi.org/10.1146/annurev.clinpsy.1.102803.144141

Shanafelt, T., et al. (2020). The Impact of COVID-19 on Employee Stress Levels in the

Financial and Retail Sectors. Journal of Occupational Health Psychology, 25(4), 358-370.

Shanbehzadeh, S., Tavahomi, M., Zanjari, N., Ebrahimi-Takamjani, I., & Amiri-Arimi, S. (2021). Physical and mental health complications post-COVID-19: Scoping review. Journal of Psychosomatic Research, 147, 110525. https://doi.org/10.1016/j.jpsychores.2021.110525

University of Virginia Health. (n.d.). COVID-19 glossary. Retrieved January 16, 2025, from https://uvahealth.com/services/covid19-glossary