Examining START Involvement and Patient Mental Health Condition Trends During the COVID-19 Pandemic

Andrew Cho, Justin Sapun, and Anish Sikhinam

Dartmouth College

This manuscript was compiled on November 23, 2022

How has the COVID-19 pandemic affected mental health patients, and how did the National Center for START Services respond to it? We examine the secondary effects of COVID-19 upon its arrival to the United States by looking at how isolation, boredom, panic, etc. can affect the mental health conditions of START participants. In addition, we analyze how effective the START initiative was during the pandemic by scrutinizing patients' behavior assessment, and the amount of patient involvement time within START. We use the data from the Systemic, Therapeutic, Assessment, Resources, and Treatment (START) initiative to evaluate such questions. From January 2019 to December 2020, we discover that patient scores decreased by 10%, while the average time START staff spends with patients also decreased by 38.46%. When analyzing patients with diagnoses that did not follow this general trend, we find that there were no conclusively significant correlations between the time START has spent with patients and their progressive behavior assessment scores.

Mental Health | National Center for START Services | COVID-19 | IDD

1. Introduction

When COVID-19 first arrived, no one could have expected the impact it would have on the world. The pandemic affected people from around the globe in an unprecedented way, and it had far and wide repercussions on all parts of people's everyday lives. In the United States alone, there have been more than 98 million cases, with 1 million deaths as of writing. However, the effects of the pandemic were felt by everyone, not only those who were just physically infected. Coupled with the rise in popularity of teleworking, remote school, and online shopping, the prevalence of physical, face-to-face interactions during the months of the pandemic had decreased substantially.

This increase in social isolation can, and has had a significant psychological impact on mental health conditions, and there has been a significant amount of literature which examines the COVID-19 pandemic's effects on these issues in great detail. Czeisler et al. 2020 conducted representative panel surveys among adults on mental health conditions across the US during June 24-30, 2020 – they found that 40.9% of respondents had reported at least one adverse mental health or behavioral health condition (1). Brooks et al. 2020 analyzes many different studies focusing on the effects of quarantine on the prevalence of a number of different problems. Of these, they find that low mood and irritability studies stand out and have high prevalence (2). These impacts of COVID-19 are disproportionately amplified in those with pre-existing mental illness, including those who are already suffering from loneliness and isolation prior to the pandemic (3). The World Health Organization (WHO) finds that, in the first year of the COVID-19 pandemic, global prevalence of anxiety and depression increased by 25%, and that young people are disproportionately at risk of self-harming behaviors (4).

Additionally, because of the COVID shutdown, access to mental health resources and services has become greatly limited and inaccessible. The WHO notes that the increase in mental health problems coincides with severe disruptions to mental health services, and that services for mental, neurological, and substance use conditions were the most disrupted over the pandemic (4). Mostly every mental health service has been greatly adapted in order to control infection, and these adaptations for infection-control reasons could have been detrimental to mental health patients whose treatments have been reduced, and to those who have been confined in hospitals with reduced therapeutic services, according to Moreno et al. 2020 (5).

As the pandemic had greatly affected mental health conditions, as well as availability of mental health services, we wanted to research further into the topic of mental health patients, and their recorded patient conditions over time, particularly during the COVID pandemic. Our main research has been conducted in two main parts – a preliminary study and exploration of general trends, as well as a deeper dive and regression analysis within certain trends we found that did not follow the general trends discovered within the preliminary exploration. We analyzed data on mental health patients from the National Center for START Services clinic, a mental health clinic which focuses on providing resources to those with intellectual and developmental disabilities and mental health needs.

First, we looked at START time involvement of patients in general over the COVID-19 pandemic. As the pandemic decreased the availability of much needed resources, we wanted to find out if this had any impact on how much time was spent by patients in the START program throughout the pandemic. Additionally, we asked: how did mental health conditions change for START patients over time during the COVID-19 pandemic? We looked at these trends for all START patients in general, and we also looked at specific differences within the trends by the patient's mental health diagnosis.

It was in this exploration that we found START patients with certain types of diagnoses who had significantly different mental health condition trends over the pandemic. These diagnoses are considered as outliers because one or more of their mental health conditions had increased, or the average START involvement time increased, which was different from the general trends decreasing over the time period of the pandemic. We assumed that the patients with these diagnoses had been highly affected during COVID-19, as their mental health conditions had increased, or their START involvement time had increased. This led us to analyze how effective the START program was helping patients with these diagnoses specifically. To be precise, we wanted to answer the question: Is there a relationship between the amount of time the patients with these outlier diagnoses spent in the START program, and their mental health condition trends throughout the pandemic?

Overall, we hope to supplement existing analysis and literature about the COVID-19 pandemic, and its effects on mental health conditions of patients with pre-existing diagnoses. Additionally, we hope to add to the discussion of mental health resources such as START and their effectiveness, especially during the pandemic, in treating these patients and their mental health conditions. Finally, we hope to discover and shed light upon potential specific mental health diagnoses, which may require more attention and support from mental health clinics than what is currently being provided.

2. Data

We retrieve data from the National Center for START Services. The National Center for START Services aims to improve the lives of persons with IDD and mental health needs through the START (Systemic, Therapeutic, Assessment, Resources, Treatment) model, with services and support emphasizing local, person-centered, positive, multidisciplinary, cost-effective, and informed practices (6).

We receive three raw data files, two of which are limited to the group of START participants from January 1, 2019–December 31, 2020. It's important to note that the third file contains 12 sheets of data, on demographic information, Aberrant Behavior Checklist (ABC), Recent Stressors Questionnaire (RSQ), diagnosis and medication data, emergency service utilization, and START tool completion dates for all START participants, ranging from September 2009 to December 2020.

From the START data set, we pulled four sheets to analyze: Time, Demographics, Diagnostic, and ABC Scores (Table 1). We pulled these specific sheets to look at specific behavioral and diagnosis trends in START participants over the past decade and to see if COVID19 had any effect on these preliminary reports. The Time sheet gives the opportunity to look at the 12 types of services they offer and more importantly, their frequency. Some examples include: Clinical Case Consultation, Crisis Follow-up, Cross-Systems Crisis Prevention and Intervention Planning, Specialized Training, and Outreach. The Demographics sheet provides information about the race, gender, presenting problems at enrollment, and general patient information. The Diagnostic sheet displays information about the patient's present illness and their diagnosis by START staff. The ABC Scores sheet provides quantified intake and re-evaluation ABC scores of patient condition, allowing for a time analysis of how patients are evolving throughout their time with START. This ABC assessment tests: Lethargy/Social Withdrawal (scored from 0-48), Irritability/Agitation (scored from 0-45), Stereotypic Behavior (scored from 0-21), Hyperactivity/Noncompliance(scored from 0-48), and Inappropriate Speech (scored from 0-12). All scoring metrics have been normalized to range from 0 to 100 for our analyses. To clarify, lower scores indicate better conditions. All cases of merging for this research was performed using exact

Table 1. Diagnosis That Didn't Follow the General Trend

Sheet Name	Number of Observations	Date Range	Unit of Analysis
time_cleaned	304489	01/01/2019 - 12/31/2020	Types of Services
demographics_cleaned	4986	09/01/2009 - 12/1/2020	Patient Information History
diagnostic_df	30983	NA	Patient Illness / Diagnosis
abc_df	26056	03/02/2006 - 12/20/2021	Intake
Re-evaluation of patient condition			

merging on patient IDs, which was easily accessible. When merging the various sheets, we only kept patients who enrolled between January 1st, 2019 and January 20th, 2020, and continued their enrollment after COVID19 was reported in the United States. This way we could only analyze patients who experienced change between pre-COVID and during COVID. Additionally, we also filtered down to psychiatric patients because it contained the most significant data, as compared to medical or other patients. The result of combining these data sets are data sets that have a tremendous increase in row count because some patients are reported to have multiple diagnoses in the Diagnostics sheet, creating extra non-duplicate entries following the merge. We performed some additional exact merges to set up data frames for further analysis.

One limitation of our data is that we were only able to effectively analyze psychiatric patients because there was not enough data to make a separate analysis of medical and other patients. Another limitation is that there were a lot of single case diagnoses that we had to drop to be able to get adequate results for the overall trend. As such, some patients were not included in the regression trend plot.

3. Methods

A. Data Cleaning. The data cleaning and most of the merging completed in our research was completed in the 00_data_cleaning.ipynb Jupyter notebook. This allowed new cleaned data frames which were used in the analysis to be created and organized separately from the raw files, decreasing the overall processing time. The data cleaning process began with importing the three data files obtained from the START initiative, and separating the full SIRS data set into 12 separate dataframes. From there, we created copies of three dataframes: time_cleaned, demographics_cleaned, and feis_cleaned. Although feis_cleaned was not used in our final research focus, we used it to obtain preliminary findings that you can see in extra_findings. From there, we renamed the Patient ID column in the 12 data frames obtained from the full SIRS data set, to have the same column name: 'Local ID'. Then, we noticed incorrect date entries in the Time dataframe, so we used regular expressions to find them, replaced them with the appropriate date, and finally converted the remaining entries to datetime format. Subsequently, we once again used regular expression to convert the 'Presenting problems at time of enrollment' values in the demographic data set to a list of elements rather than one long string to allow for better interpretation of accurate value counts. In addition, we also combined all instances of 'Other: . . . ' to a combined shorthand of 'Other'. This reduced the total unique count from 1189 to 113 elements. Afterwards, we noticed the FEIS dataframe columns had strange symbols and random occurrences of new lines, so to be able to understand the inquiries, we had to replace these instances with the help of the Functools module. Soon after, we noticed some diagnoses reported were repeated or had similar wording for the same patient so we grouped into one main variable. The number of diagnoses dropped from 1875 to 1666.

B. Data Merging & Filtering.

General Merging & Filtering. Next, once we believed we cleaned the relevant columns of each dataframe, we were able to use exact merging, as the data was consistent with the use of Patient ID indicators. Our first merge began with combining diagnostic and ABC scores dataframes using an inner merge. The number of observations increased by threefold to 94,589 because some patients are reported to have multiple diagnoses, leaving us with a long formatted new_df dataframe. Following the merge, we dropped duplicates and NaNs leaving us with 92,830 rows. Then, we only kept patients who enrolled between January 1st, 2019 and January 20th, 2020, and continued their enrollment after COVID19 was reported in the United States. This left us with the clean_df dataframe with normalized patient ABC scores. The last stage of our data cleaning sequence was to exact merge the Time and diagnostic data set to look at average minutes spent for cohorts of START participants. This second merge increased the number of observations by a similar threefold, revealing 1,190,121 rows for the same reason mentioned above. After that, we filtered the data to only include psychiatric patients who enrolled between January 1st, 2019 and January 20th, 2020, and continued their enrollment after COVID19 was reported in the United States. We also filtered out diagnosis data with a patient count of 1 to limit specific cases. This reduced the rows to 414,785 rows, with a total diagnosis count of 60. This leaves us with a time_diag dataframe which contains a specific occurrence of time being spent at START, the quarter when this time was spent, the type of diagnosis, and the number of minutes spent in START. We concluded by saving the new dataframes to csv files in a subdirectory called 'cleaned'. The time_diag dataframe was saved in the sub directory as "all_time_diag_clean".

Task Specific Merging & Filtering. After general merging, we further merged and filtered for each specific task. In all our analyses, only diagnoses of at least 10 ABC score observations from January 1st, 2019 to January 1st, 2021 were used. We merged our patient condition data set with the time data set and calculated the average per patient START involvement time grouped by diagnosis and quarter. Moreover, we created another dataframe with the same data sets but calculated the average per patient START involvement time by the type of START involvement. For both dataframes, we created an adjusted quarter variation. If adjusted, the current patient condition scores will correspond with START's involvement time from the previous available quarter. This was done to clearly see if there was a causal relationship between START's involvement and patient conditions by aligning current START's involvement with future patient condition scores.

C. Functions. All of our user defined functions were stored in a separate python script utils.py. This allowed for more space and focus centered on the analysis in the Jupyter notebooks. In the utils.py script, we created five user defined functions to perform various tasks throughout our research methodology. We created plot_score_trends() to take a dataframe of a specific group of START patients with mean ABC scores by quarter during COVID, and visually plot their scores over time. We created extract_possible_diagnosis() to detect diagnoses whose patient condition worsened during COVID-19. We created extract_diagnosis_time() to detect diagnoses whose START involvement time increased during COVID-19. For both extract_possible_diagnosis() and extract_diagnosis_time(), the diagnoses were determined if the values from 2020 were higher than the values from 2019 or if there was a sharp increase in value from 2019 Q3 to 2020 Q2. We created addMainIssue() to group each diagnosis with their main category of patient condition. Here is the procedure of the function: for each diagnosis, if condition A worsened from 2019 to 2020, group the diagnosis to condition A. If there is another type(s) of condition that satisfies the previous filter, group the diagnosis to that/those condition type(s) as well. If no condition type satisfies, group the diagnosis to the highest scoring type of condition. The last function, regressionAnalysis() performs OLS least-squares regression.

D. Research Design. As mentioned in the introduction, our research splits into two main components: what were the general trends of patient conditions and START's involvement time during COVID-19? and how effective was START in treating patients who faced possible hardships during COVID-19? To answer the first component, we visually and quantitatively looked at changes between pre-COVID and COVID values for patient conditions and START's involvement time in 02_general_trends.ipynb. To answer the second question, we used two functions (extract_possible_diagnosis() and extract_diagnosis_time()) to detect the diagnoses that didn't follow the general trends in 03_detecting_off_trend.ipynb. Then, we grouped each diagnosis by its primary type of patient condition and performed OLS regressions in 04_regression.ipynb. We performed regressions both on unadjusted and adjusted quarters. We also performed regressions using specific types of START's Involvement as well.

To explore these questions, we needed to have two proper dataframes which meet the ideal conditions. One dataframe should contain average minutes spent at START by diagnosis and quarter. The other dataframe should contain mean ABC scores by diagnosis and quarter. With these dataframes, we will be able to plot and visualize the average amount of time that the patients spent in START quarter by quarter, as well as visualize the patient's mental health conditions, as recorded by ABC scores quarter by quarter over the pandemic.

4. Results

Result - Part 1. As stated in the introduction, our main goals for the first part of our research was to explore general trends during the COVID-19 pandemic time period. This We aimed to explore the question: how did the average time patients spend with START change over the course of the pandemic? How did START patients' mental health conditions, as measured by their ABC scores, change over the course of the pandemic? Are there any general trends to note here?

Average Minutes Spent with START by Quarter by Diagnosis. In order to create the dataframe to analyze average minutes spent with START, we used the all_time_diag_clean dataframe, and grouped it by diagnosis and quarter. Then we ran nunique() on Local ID, to get the patient count per quarter per diagnosis, and we saved this in id_df_all. Afterwards, we used the all_time_diag_clean dataframe and grouped by diagnosis and quarter again, but this time we summed up the minutes per quarter per diagnosis. We saved this in avg_time_df. Next, we merged avg_time_df and id_df_all by diagnosis and quarter to get a dataframe which contained diagnoses, quarters, minutes spent, and patient counts. We added a new column of average minutes spent per quarter per diagnosis by dividing the minutes spent column by the patient count column. This dataframe was the one we needed – it contained average minutes spent at START by diagnosis and by quarter. We saved this dataframe in time_diag_avg_min. To plot time_diag_avg_min, we pivoted this table with the index being diagnosis, columns being quarters, and values being average minutes spent at START. The result of the Average Time Psychiatric Patients Spend with START Staff by Diagnosis is shown in Figure 1.

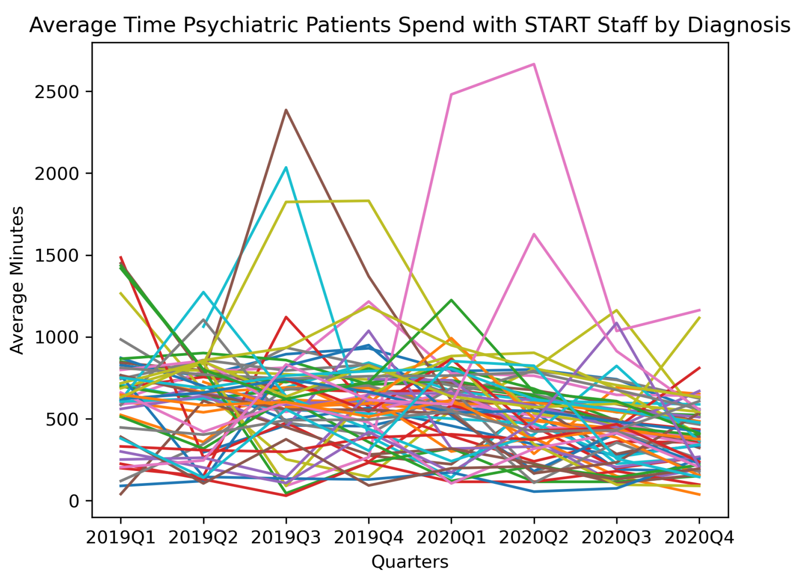

Fig. 1. Average Time Psychiatric Patient Spend with START Staff by Diagnosis

As shown in Figure 1, with patients of most diagnoses, the lines appear to be mostly stable, and there is a clear, although slight, decrease. Towards the end of the period, in 2020 Q4, we can see that the overall trend diverges towards an average of 450 minutes spent. Additionally, there are some clear outliers which can be seen here, where patients with certain diagnoses had a significant increase (upwards to around 500%) in the average amount of time spent at START during the pandemic.

Average Minutes Spent with START by Quarter Across All Diagnoses. Next, we took the pivoted table and calculated the mean time spent across psychiatric patients of all diagnoses. The result of Average Time Psychiatric Patients Spend with START Staff is shown in Figure 2.

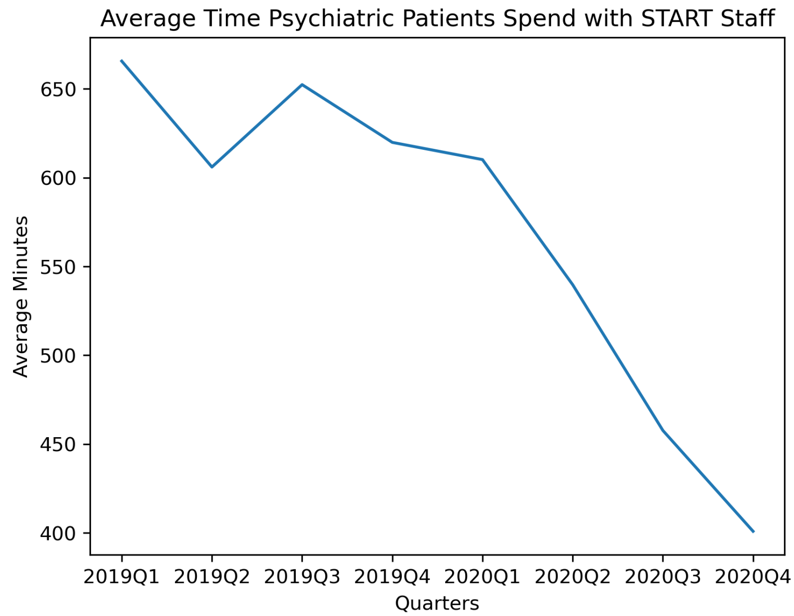

Fig. 2. Average Time Psychiatric Patients Spend with START Staff

Figure 2 is just like the Figure 1, except since it averages the amount of time spent at START across all diagnoses, we can see that it results in just one line. This gives us an approximate idea of what the overall trend is. We can see that over time, there has been a significant decrease in the average amount of time spent at START during the pandemic, dropping from above 650 minutes to approximately 400 minutes over two years – a 38% decrease. This may be due to the reduced availability of mental health services, as well as due to quarantine forcing mental health patients to stay at home, or make less frequent visits.

Average Mental Health Conditions as Measured by ABC Scores by Quarter. In order to create the dataframe to analyze average ABC scores, we used the clean_df dataframe, and filtered out all of the cases between 2019 Q1 and 2020 Q4, as well as all the cases whose diagnoses are more unique, meaning that there are less than 10 cases within the whole data set with that same diagnosis. Then we grouped by diagnosis and quarter, and calculated the mean ABC scores. We saved this dataframe in clean_df_ideal. Because there are five scores to track for each diagnosis, we decided to only plot the average score trends across all diagnoses, rather than the average score trend for each specific diagnosis. Thus, we grouped only by quarter and calculated the mean of each of the 5 scores per quarter. The results of the Average ABC scores for Psychiatric Patients are shown in Figure 3.

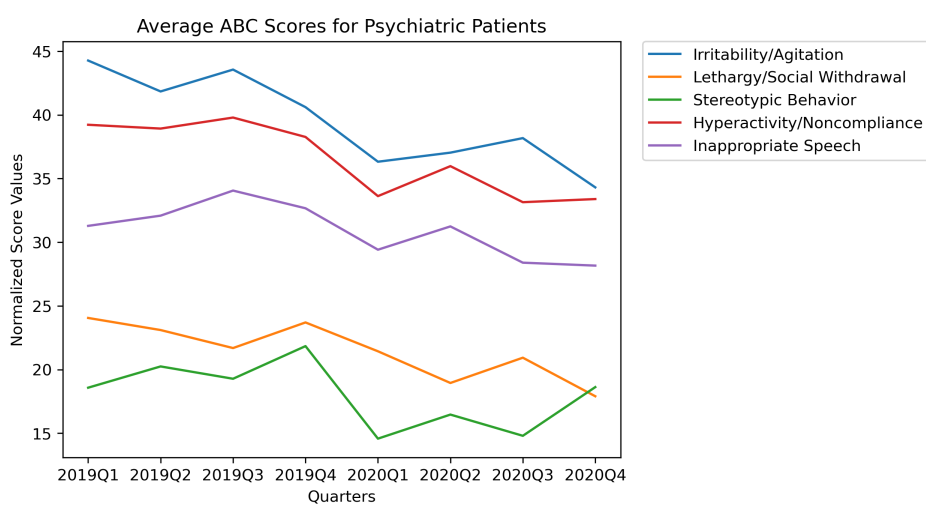

Fig. 3. Average ABC Scores for Psychiatric Patients

In Figure 3, each metric has been scaled differently, and so we have normalized all the data beforehand. Patients who displayed more of each mental health behavior received a higher score. As shown in the graph, almost every score decreased on average throughout the pandemic, with the *Stereotypic Behavior* score greatly decreasing from 2019 to 2020, and a slight uptick towards the end from 2020 Q3 to 2020 Q4. This shows that overall, despite the challenges that the pandemic posed to psychiatric patients within START, their mental health conditions have seen a slight improvement, as the scores have all decreased.

We continued to analyze these ABC scores to see if there were any interesting or relevant trends by looking at different variables. For example, combining the clean_df with demographics_cleaned, we were able to look at how ABC score trends differed by variables such as the patient's race, their primary caregiver, employment status, and by specific diagnoses. All of these findings can be found in the extra_findings subdirectory of the repository. However, there were no significant changes in trends compared to the general trends, except for certain diagnoses. Some patients with certain diagnoses had much different score trends from the general trend over the pandemic, and for this reason, we decided to explore further into score trends by diagnosis in the next section.

Result - Part 2. Trying to achieve a deeper understanding of patient conditions during COVID-19, we conducted random case-studies of diagnoses where we tracked the patient condition of a singular diagnosis over time. Here is the data table of one case-study (Note: for censoring purposes, all values were rounded by 5. Moreover, by column, every value was added a random number between -10 and 10. The values were averaged over each year):

Table 2. Diagnosis That Didn't Follow the General Trend

Date Reviewed	2019	2020
Irritability/Agitation	6.5	14.0
Lethargy/Social Withdrawal	-2.0	8.0
Hyperactivity/Noncompliance	1.0	6.0

As noted in the Table 2, we can see that there was a clear worsening in the patient condition of *Irritability/Agitation*, *Lethargy/Social Withdrawal*, and *Hyperactivity/Noncompliance*. Thus, we realized that not all patients followed the general trend.

Grouping each patient by their diagnosis, we then decided to detect which diagnoses did not follow the general pattern in regards to both patient condition and START's involvement. Using extract_possible_diagnosis() and extract_diagnosis_time(), we extracted the diagnoses that had at least one pattern that didn't follow the general trend. For example, even if 4 out of the 5 categories of patient condition trends followed the general pattern for a given diagnosis, the diagnosis will still be recorded because 1 category did not follow the general trend. From 60 eligible diagnoses after filtering, 22 diagnoses had at least one category of patient condition that didn't follow the trend and 22 diagnoses faced an increase in START involvement time due during COVID-19. Because there were duplicate diagnoses when combining both results, the combined number of diagnoses that didn't follow the general trend on both the metrics was 36. Thus, 36 out of 60 (60%) diagnoses had at least one pattern that didn't follow the general trend based on 6 conditions (5 patient categories and START's involvement time). Assuming patients with these diagnoses faced possible hardships during COVID-19, we aimed to see how effective START was in helping improve their conditions. Understand the importance of each diagnosis pertains to different types of patient condition categories, we grouped each diagnosis with their primary type of patient condition. This step was completed by using our function, addMainIssue(). To look at how effective START's involvement was on the diagnoses that didn't follow the general trend, we performed regression analyses between each type of patient condition and START involvement time. We performed these regressions in both unadjusted and adjusted quarters and in both nonspecific and specific types of START involvement.

OLS Regression of Nonspecific START Involvement. After performing regression analysis on each type of patient condition based on unadjusted quarters, we found zero meaningful relationships. However, when performing regression analysis on adjusted quarters, we found a slight relationship between START Involvement Time and patient condition on two types of condition: *Lethargy/Social Withdrawal* and *Stereotypic Behavior*.

Figure 4 and Table 3 are the regression results between START Involvement Time and Condition Scores of *Lethargy/Social Withdrawal*. A total of 57 observations were made and had a R-squared of 0.032. Moreover, the regression resulted in a correlation of -0.0067 between Lethargy/Social Withdrawal condition scores and quarter-adjusted per Patient START Involvement Time. Although the correlation is not the strongest, it may be possible that START's involvement improved conditions of diagnosis with *Lethargy/Social Withdrawal* during COVID-19.

Figure 5 and Table 4 are the regression results between START Involvement Time and Condition Scores of *Stereotypic Behavior*. A total of 47 observations were made and had a R-squared of 0.039. Moreover, the regression resulted in a correlation of -0.0089 between *Stereotypic Behavior* condition scores and quarter-adjusted per Patient START Involvement Time. Although the correlation is not the strongest, it may be possible that START's involvement improved conditions of diagnosis with *Stereotypic Behavior* during COVID-19.

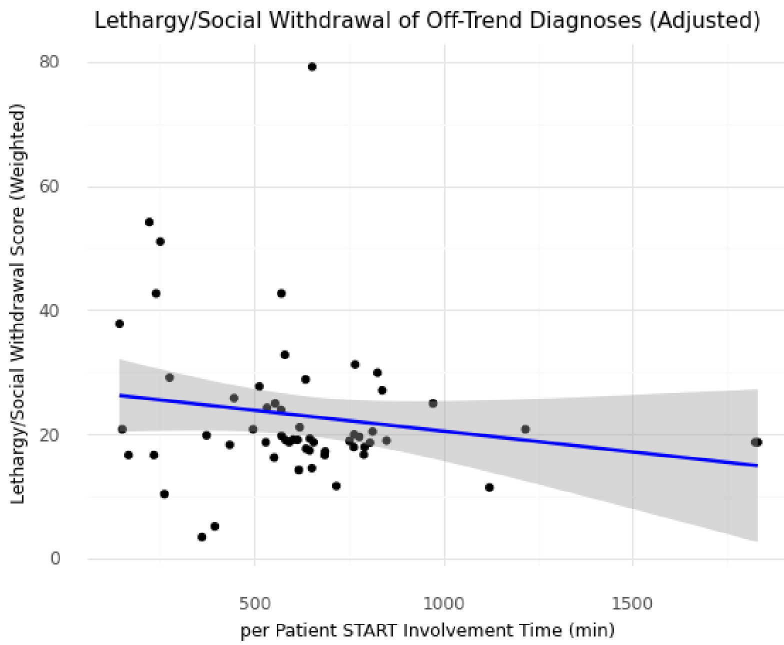

Fig. 4. per Patient START Involvement Time (Adjusted) vs *Lethargy/Social Withdrawal* Patient Condition (Weighted)

Table 3. OLS Regression Summary for *Lethargy/Social Withdrawal* (Adjusted)

Dep. Variable:	Lethargy/Social Withdrawal	**R-squared:**	0.032
Model:	OLS	**Adj. R-squared:**	0.014
Method:	Least Squares	**F-statistic:**	1.824
Date:	Tue, 22 Nov 2022	**Prob (F-statistic):**	0.182
Time:	11:35:48	**Log-Likelihood:**	-221.38
No. Observations:	57	**AIC:**	446.8
Df Residuals:	55	**BIC:**	450.8
Df Model:	1		
Covariance Type:	nonrobust		

	coef	**std err**	**t**	**P$>|t|$**	**[0.025**	**0.975]**
const	27.2306	3.524	7.728	0.000	20.169	34.292
avg minutes	-0.0067	0.005	-1.350	0.182	-0.017	0.003

Omnibus:	46.387	**Durbin-Watson:**	2.221
Prob(Omnibus):	0.000	**Jarque-Bera (JB):**	194.039
Skew:	2.232	**Prob(JB):**	7.33e-43
Kurtosis:	10.859	**Cond. No.**	1.58e+03

Notes:

[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.

[2] The condition number is large, 1.58e+03. This might indicate that there are strong multicollinearity or other numerical problems.

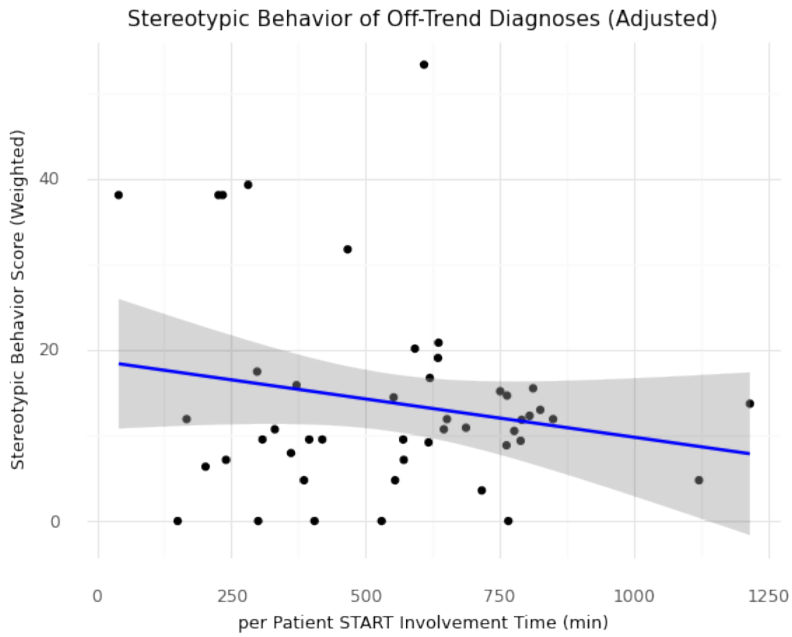

Fig. 5. per Patient START Involvement Time (Adjusted) vs *Stereotypic Behavior* Patient Condition (Weighted)

Table 4. OLS Regression Summary for *Stereotypic Behavior* (Adjusted)

Dep. Variable:	Stereotypic Behavior	**R-squared:**	0.039
Model:	OLS	**Adj. R-squared:**	0.017
Method:	Least Squares	**F-statistic:**	1.813
Date:	Tue, 22 Nov 2022	**Prob (F-statistic):**	0.185
Time:	12:10:05	**Log-Likelihood:**	-180.22
No. Observations:	47	**AIC:**	364.4
Df Residuals:	45	**BIC:**	368.1
Df Model:	1		
Covariance Type:	nonrobust		

	coef	**std err**	**t**	**P> |t|**	**[0.025**	**0.975]**
const	18.7344	4.010	4.672	0.000	10.659	26.810
avg minutes	-0.0089	0.007	-1.346	0.185	-0.022	0.004

Omnibus:	17.868	**Durbin-Watson:**	1.266
Prob(Omnibus):	0.000	**Jarque-Bera (JB):**	23.022
Skew:	1.311	**Prob(JB):**	1.00e-05
Kurtosis:	5.208	**Cond. No.**	1.45e+03

Notes:

[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.

[2] The condition number is large, 1.45e+03. This might indicate that there are strong multicollinearity or other numerical problems.

OLS Regression of Specific START Involvement. Understanding the possible relationships between START Involvement and *Lethargy/Social Withdrawal* & *Stereotypic Behavior* conditions during COVID-19, we looked deeper into the relationships by performing an OLS regression analysis based on the type of START involvement. Of the 12 types of START involvement, we grouped 5 with the lowest observations into one "other" section for each condition metric. Thus, for each of *Lethargy/Social Withdrawal* and *Stereotypic Behavior*, we performed a least-squares OLS regression with 8 quarter-adjusted types of START Involvement and its corresponding patient condition scores.

- For *Lethargy/Social Withdrawal*, we grouped CET/CTE; Clinical Education Team; Comprehensive Service Evaluations; General Administrative; and Individual Specific Travel into "other." Thus, we worked with 8 independent variables (Clinical Case Consultation; Crisis Follow-up; Cross-Systems Crisis Prevention and Intervention Planning; Intake/Assessment Activities; Medical/Psychiatric Consultation; Outreach, Specialized Training and System Linkages; Therapeutic Supports and Services; other) for *Lethargy/Social Withdrawal*.

- For *Stereotypic Behavior*, we grouped CET/CTE; Clinical Education Team; Comprehensive Service Evaluations; Individual Specific Travel; and Medical/Psychiatric Consultation into "other." Thus, we worked with 8 independent variables (Clinical Case Consultation; Crisis Follow-up; Cross-Systems Crisis Prevention and Intervention Planning; Intake/Assessment Activities; General Administrative; Outreach, Specialized Training and System Linkages; Therapeutic Supports and Services; other) for *Stereotypic Behavior*.

The OLS regression result between Specific START Involvement and *Lethargy/Social Withdrawal* are recorded in Table 5. Here are the description of the results:

- ***Medical/Psychiatric Consultation***: Surprisingly, we found a coefficient of 0.0552 for our *Medical/Psychiatric Consultation* dummy variable. Considering the p-value is 0.007, indicating that it is statistically significant, our results indicate that *Medical/Psychiatric Consultation* had negative effects on the *Lethargy/Social Withdrawal* condition.

- ***Outreach, Specialized Training and System Linkages***: We found a coefficient of -0.0395 for our *Outreach, Specialized Training and System Linkage* dummy variable. Although it is not statistically significant as the p-value is 0.121, we might be able to reasonably assume that this variable was the driver that improved the *Lethargy/Social Withdrawal* condition.

- ***All Other Variables***: All other dummy variables seemed to not show any statistically significant correlation.

Table 5. OLS Regression Summary b/w Specific START Involvement & *Lethargy/Social Withdrawal* (Adjusted)

Dep. Variable:	Lethargy/Social Withdrawal	**R-squared:**	0.390
Model:	OLS	**Adj. R-squared:**	0.195
Method:	Least Squares	**F-statistic:**	2.002
Date:	Tue, 22 Nov 2022	**Prob (F-statistic):**	0.0883
Time:	12:30:46	**Log-Likelihood:**	-100.93
No. Observations:	34	**AIC:**	219.9
Df Residuals:	25	**BIC:**	233.6
Df Model:	8		
Covariance Type:	nonrobust		

	coef	**std err**	**t**	**P> |t|**	**[0.025**	**0.975]**
const	26.3146	6.676	3.942	0.001	12.565	40.064
Clinical Case Consultation	0.0184	0.025	0.748	0.461	-0.032	0.069
Crisis Follow-up	-0.0110	0.013	-0.837	0.411	-0.038	0.016
Cross-Systems Crisis Prevention and Intervention Planning	-0.0111	0.014	-0.805	0.429	-0.039	0.017
Intake/Assessment Activities	-0.0068	0.035	-0.196	0.846	-0.078	0.065
Medical/Psychiatric Consultation	0.0552	0.019	2.940	0.007	0.017	0.094
Outreach, Specialized Training and System Linkages	-0.0395	0.025	-1.605	0.121	-0.090	0.011
Therapeutic Supports and Services	-0.0026	0.007	-0.400	0.693	-0.016	0.011
other	0.0026	0.004	0.644	0.525	-0.006	0.011

Omnibus:	2.209	**Durbin-Watson:**	2.292
Prob(Omnibus):	0.331	**Jarque-Bera (JB):**	1.092
Skew:	0.214	**Prob(JB):**	0.579
Kurtosis:	3.767	**Cond. No.**	4.69e+03

Notes:

[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.

[2] The condition number is large, 4.69e+03. This might indicate that there are strong multicollinearity or other numerical problems.

Table 6. OLS Regression Summary b/w Specific START Involvement & *Stereotypic Behavior* (Adjusted)

Dep. Variable:	Stereotypic Behavior	**R-squared:**	0.583
Model:	OLS	**Adj. R-squared:**	0.345
Method:	Least Squares	**F-statistic:**	2.450
Date:	Tue, 22 Nov 2022	**Prob (F-statistic):**	0.0682
Time:	12:54:35	**Log-Likelihood:**	-54.764
No. Observations:	23	**AIC:**	127.5
Df Residuals:	14	**BIC:**	137.7
Df Model:	8		
Covariance Type:	nonrobust		

	coef	**std err**	**t**	**P> |t|**	**[0.025**	**0.975]**
const	0.6811	5.902	0.115	0.910	-11.978	13.340
Clinical Case Consultation	0.0190	0.022	0.872	0.398	-0.028	0.066
Crisis Follow-up	0.0120	0.013	0.897	0.385	-0.017	0.041
Cross-Systems Crisis Prevention and Intervention Planning	0.0033	0.021	0.160	0.875	-0.041	0.048
General Administrative	-0.0369	0.019	-1.984	0.067	-0.077	0.003
Intake/Assessment Activities	0.0568	0.031	1.820	0.090	-0.010	0.124
Outreach, Specialized Training and System Linkages	0.0043	0.017	0.250	0.806	-0.033	0.041
Therapeutic Supports and Services	-0.0023	0.005	-0.462	0.651	-0.013	0.008
other	0.0082	0.004	2.151	0.049	2.28e-05	0.016

Omnibus:	1.635	**Durbin-Watson:**	1.755
Prob(Omnibus):	0.442	**Jarque-Bera (JB):**	1.152
Skew:	0.286	**Prob(JB):**	0.562
Kurtosis:	2.065	**Cond. No.**	5.78e+03

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
[2] The condition number is large, 5.78e+03. This might indicate that there are strong multicollinearity or other numerical problems.

The OLS regression result between Specific START Involvement and *Stereotypic Behavior* are recorded in Table 6. Here are the description of the results:

- ***General Administrative***: We found a coefficient of -0.0369 for our *General Administrative* dummy variable. Its p-value is 0.067, which is not significantly significant, but it has the lowest p-value out of all the other dummy variables (except other). Thus, we might be able to reasonably assume that this variable was the driver that improved the *Stereotypic Behavior* condition.

- ***Intake/Assessment Activities***: Surprisingly, we found a coefficient of 0.0568 for our *Intake/Assessment Activities* dummy variable. Its p-value is 0.090, which is not significantly significant, but it has the second lowest p-value out of all the other dummy variables (except *other*). Thus, we might be able to reasonably assume that Intake/Assessment Activities worsened the *Stereotypic Behavior* condition.

- ***Other***: We found a coefficient of 0.0082 for our *other* dummy variable. Its p-value is 0.049, which is the only statistically significant value in this regression. This variable comprises START Involvement sections of CET/CTE, Clinical Education Team, Comprehensive Service Evaluations, Individual Specific Travel, and Medical/Psychiatric Consultation. Considering that its coefficient is nearly 0, we can reasonably conclude that it had no impact on the *Stereotypic Behavior* condition.

- ***All Other Variables*** All other dummy variables seemed to not show any statistically significant correlation.

5. Discussion and Limitations

Discussion. In Results 1, we found that the overall trend was that average time spent within the START program, as well as mental health behaviors as measured by ABC scores tended to decrease quarter by quarter over the course of the pandemic.

In Results 2, when analyzing diagnoses that did not follow this overall trend, we found that with unadjusted quarters there was no meaningful relationship between average, nonspecific START involvement time and ABC scores. With adjusted quarters, there was a small negative correlation between START involvement times and *Lethargy/Social Withdrawal* Scores, as well as *Stereotypic Behavior* scores. However, the r-squared values are relatively low at 0.032 and 0.039.

When looking at specific START involvement types, the only statistically significant results ($p\text{-value} < 0.05$) we've found are a small positive correlation between *Medical/Physical Consultation* and *Lethargy/Social Withdrawal* Scores. To our contrary belief, our results indicate that *Medical/Physical Consultation* may have worsened *Lethargy/Social Withdrawal conditions*.

While we found an extremely slight positive correlation between *other* and *Stereotypic Behavior* which was statistically significant, the correlation between the variables is so small that it most likely doesn't exist. We also found other correlations, although they were statistically insignificant (p-value > 0.05). For instance, there was a negative correlation between *Outreach, Specialized Training and System Linkages* and *Lethargy/Social Withdrawal* Scores, but its p-value was 0.121. Moreover, there was a negative correlation between *General Administrative* and *Stereotypic Behavior*, but its p-value was 0.067. Although both relationships were statistically insignificant, it may be plausible that *Outreach, Specialized Training and System Linkages* improved *Lethargy/Social Withdrawal* conditions and *General Administrative* improved *Stereotypic Behavior* conditions.

There are many possible explanations for our findings. First, as all of the data that we analyzed looked at psychiatric patients who enrolled before COVID and continued their enrollment through COVID, we can attribute the average decrease in mental health condition scores to the general improvement of these patients, as a result of the START program. Additionally, we can attribute the decrease in average time spent in START to decreases in availability of mental health resources, as well as increased social isolation over the pandemic. As both mental health patients as well as non-essential healthcare workers had to quarantine since the shutdown, opportunities for START to provide care were limited, and most likely had to shift to a remote format. Another reason could be that as mental health conditions improved, there was less of a need for mental health patients to visit the START clinic as often, hence less average time spent.

However, when analyzing patients with diagnoses that didn't improve over COVID, the potential reasons for the mostly inconclusive regression results aforementioned may be due to START not playing a significant role in affecting the behaviors of patients with these diagnoses. It could be possible that START is not as effective at treating patients with these diagnoses. However, there may be other unknown factors which caused these patients to remain stagnant or increase in scores or average time spent at START, such as the quality of care that these patients received before and during COVID, as well as other sources of care these patients received.

Limitations. Our analysis was heavily limited by the amount of data that we had available to use. Because our data analysis focuses heavily on subsetting by diagnosis of the patient, there were some diagnosis categories where there was score data only for a few patients. While we mitigated this issue by only using diagnoses with at least 10 patients, some diagnoses had just few enough data where it came close to this limit, but were still included in the analysis. Moreover, while we have time spent in START data from 2019 onwards, there is nothing provided prior to this date. This means that we only have 2019 time data to analyze, and this may not be the most accurate representation of how much average time is spent at START pre-pandemic.

As New Hampshire has a smaller, more racially homogeneous population, some next steps to take would be to analyze data from other START clinics in different states, as well as do a nationwide analysis. With this, there would be much more cases per diagnosis, as well as longer time data, which could possibly contribute to more conclusive results. We could also look at differences between the pandemic's impact on mental health conditions by race, as there would be more racially diverse cases to analyze.

6. Conclusion

In conclusion, our study finds that during the COVID-19 pandemic, average time spent in the START program decreased. Additionally, we find that the mean mental health conditions of patients (as measured by ABC scores) also decreased during the pandemic. When analyzing the specific diagnoses that did not follow this general trend with adjusted quarters, we find that there are small negative correlations between the nonspecific START involvement time and Lethargy/Social Withdrawal Scores as well as Stereotypic Behavior Scores. When observing specific START involvement times with these two mental health conditions, we did find one statistically significant result, which shows a small positive correlation between Medical/Physical Consultation involvement time and Lethargy/Social Withdrawal Scores. However, this relationship is small, and we find that there are no other statistically significant relationships present.

7. Acknowledgements

We thank professor Jaren Haber for his support and guidance in making our research possible.

8. Works Cited

1. Mark É. Czeisler, et al., "Mental Health, Substance Use, and Suicidal Ideation During the COVID-19 Pandemic".Centers for Disease Control and Prevention (2020).
2. Samantha K. Brooks, et al., "The psychological impact of quarantine and how to reduce it: rapid review of the evidence".The Lancet (2020).
3. Tzung-Jeng Hwang, et al., "Loneliness and social isolation during the COVID-19 pandemic".National Library of Medicine (2020)
4. Alison Brunier, Carla Drysdale. "COVID-19 pandemic triggers 25% increase in prevalence of anxiety and depression worldwide".World Health Organization (2022).
5. Carmen Moreno, at al., "How mental health care should change as a consequence of the COVID-19 pandemic".National Library of Medicine (2020).
6. Center for START Services. "Center for START Services." Accessed November 22, 2022. https://centerforstartservices.org/.