An evaluation of COVID-19 virtual treatment delivery for eating disorders: Assessing eating disorder symptoms among patients across levels of care

by

Cara Anne Lyons

B.A. Religious Studies, University of Pittsburgh, 2018

Submitted to the Graduate Faculty of the

Graduate School of Public Health in partial fulfillment

of the requirements for the degree of

Master of Public Health

University of Pittsburgh

2021

UNIVERSITY OF PITTSBURGH

GRADUATE SCHOOL OF PUBLIC HEALTH

This essay is submitted

by

Cara Anne Lyons

It was submitted on

April 28, 2021

and approved by

Essay Advisor: Nancy W. Glynn, PhD, Associate Professor, Department of Epidemiology, Graduate School of Public Health, University of Pittsburgh

Natacha M. DeGenna, PhD, Assistant Professor, Department of Psychiatry, School of Medicine, and Department of Epidemiology, Graduate School of Public Health, University of Pittsburgh

Rachel P. Kolko Conlon, PhD, Research Instructor, Department of Psychiatry, School of Medicine, University of Pittsburgh

An evaluation of COVID-19 virtual treatment delivery for eating disorders: Assessing eating disorder symptoms among patients across levels of care

Cara Anne Lyons, MPH

University of Pittsburgh, 2021

Abstract

In response to the call-to-action we are facing as an eating disorder (ED) field to provide data on program outcomes and to comply with regulatory standards regarding incorporation of patient-reported data into clinical care processes, the University of Pittsburgh Medical Center's Center for Eating Disorders has administered questionnaires to patients enrolled in treatment to monitor treatment progress. Given the negative effects of the COVID-19 pandemic on patients with EDs and the shift to virtual treatment delivery, data are needed on ED symptoms and changes during treatment among patients receiving ED services during the COVID-19 pandemic and virtual delivery of care. Thus, this study aimed to evaluate changes in symptoms among patients with EDs enrolled in a partial hospitalization program (PHP), intensive outpatient program (IOP), and weekly outpatient group. Following the transition from in-person to virtual delivery due to the COVID-19 pandemic, we administered a web-based version of the Eating Disorders Examination Questionnaire (EDE-Q) to all enrolled patients to track ED symptomology, behaviors, and attitudes over the course of 28 days. We hypothesized that higher baseline EDE-Q scores would be associated with enrollment in a higher level of care, longer length of stay, and greater EDE-Q score reductions during virtual treatment delivery. Patients across levels of care who completed the EDE-Q at least twice were included in analyses (N=46, n=17 in PHP; n=22 in IOP; n=7 in weekly outpatient group). Baseline EDE-Q scores were not related to initial level of care, length of stay, or discharge disposition (82.6% of patients stepped

down to outpatient care; *p*s>.05). Notably, patients reported a significant improvement in global EDE-Q scores ($t(45)=1.98$, $p=0.05$, $d=0.59$) and shape concern subscale scores ($t(45)=2.01$, $p=0.05$, $d=0.60$). Results indicate that patients experience significant improvements in ED symptoms during virtual treatment delivery. Additional work is needed to further understand the impact of ED treatment delivery on patients' functioning and long-term symptom trajectories, including evaluation of treatment response in the context of virtual treatment delivery and the COVID-19 pandemic.

Table of Contents

List of Tables

List of Figures

1.0 Introduction

Eating disorders are a psychiatric disorder associated with an extremely high mortality rate, second only to opioid use disorder, however remain largely under researched. A recent report by the Harvard School of Public Health and the Academy for Eating Disorders found that 10,200 deaths per year are a direct result of an eating disorder (Deloitte Access Economics, 2020). Approximately 30 million Americans will be diagnosed with an eating disorder during their lifetime, though this number is largely thought to be an underestimate of the true prevalence due to the stigmatization of eating disorders in the general population. Prevalence and incidence rates vary based on diagnosis, though updates to the Diagnostic and Statistical Manual of Mental Disorders (DSM) have expanded the amount of individuals that meet criteria for a diagnosis, thus increasing the amount of patients receiving treatment for their illness.

1.1 Epidemiology

1.1.1 Anorexia Nervosa

According to the DSM-5, anorexia nervosa (AN) is diagnosed in individuals that meet the following criteria:

1. Restriction of energy intake relative to requirements, leading to a significantly low body weight in the context of age, sex, developmental trajectory, and physical health.

2. Intense fear of gaining weight or of becoming fat, or persistent behavior that interferes with weight gain, even though at a significantly low weight.

3. Disturbance in the way in which one's body weight or shape is experienced, undue influence of body weight or shape on self-evaluation, or persistent lack of recognition of the seriousness of the current low body weight (American Psychiatric Association, 2013).

AN is further categorized as restricting or binge-eating/purging type. In the former subtype, individuals do not engage in binge eating or purging and exclusively restrict their intake or engage in excessive exercise. In the latter, individuals engage in binge eating and/or purging, such as self-induced vomiting or the misuse of laxatives, diuretics, or enemas (American Psychiatric Association, 2013). AN is associated with the highest mortality rate of eating disorders, with a standardized mortality ratio of 5.86 (Deloitte Access Economics, 2020).

There are few community studies that evaluate the epidemiology of AN in part because of the rarity of the illness. A study conducted by Keski-Rahkonen et. al assessed 2,881 women from birth cohorts of 1975-1979 of Finnish twins. Women who had recovered from AN were compared to their unaffected twins and controls. Lifetime prevalence of AN diagnosed with the DSM-IV was 2.2%. In women ages 15 to 19, the incidence of AN was 270 per 100,000 person-years (Keski-Rahkonen et. al, 2007).

After the diagnostic revision of AN for the DSM-5 in 2013, a study by Mustelin et al. suggested that the prevalence and incidence of this illness was higher than what had been previously reported. With the new diagnostic criteria, this study calculated the lifetime prevalence of AN to be 3.6% and the 15-year incidence rate to be 230 per 100,000 person years. In this study particularly, cases defined according to the DSM-5 had a higher minimum BMI

than those defined by the DSM-IV. Furthermore, the 5-year recovery rate was higher in those cases associated with the DSM-5 at 81% compared to the cases in the DSM-IV cohort, which was 67% (Mustelin et al., 2016).

Prevalence and incidence rates vary among males and non-Whites. There have been few studies that address the epidemiology of AN in men, however, a 1995 study by Hoek et al. estimated this incidence to be approximately 0.5 per 100,000 person years. More recently in 2009, a Finnish study used a twin cohort to further estimate incidence and prevalence of AN in males. Though the sample size was small, the incidence rates were much higher than estimated in the previously mentioned study. Lifetime prevalence was 0.24% and the incidence rate was 15.7 per 100,000 person years (Raevouri et al., 2009).

1.1.2 Bulimia Nervosa

In 2013, the diagnostic criteria for bulimia nervosa (BN) was updated to reflect changes in frequency of behaviors, as well as subtypes of BN. Currently, BN is diagnosed if a patient experiences:

1. Recurrent episodes of binge eating
2. Recurrent and inappropriate compensatory behaviors to prevent weight gain
3. The episodes of binge eating and compensatory actions occur at least once a week for three months
4. Self-image is largely influenced by body shape and weight
5. This disturbance does not occur during episodes of AN

Per the DSM-5, an episode of binge eating is defined as “(1) eating, in a discrete period of time, an amount of food that is definitely larger than what most individuals would eat in a

similar period of time under similar circumstances, and (2) a sense of lack of control over eating during the episode" (American Psychiatric Association, 2013). BN can be categorized as being in partial remission or full remission. Further, severity can range from mild to extreme based on the frequency of behaviors.

Previous estimates reported the lifetime prevalence of BN to be between 0.9% and 3%. Prevalence rates increased after the most recent version of the DSM was published and the requirements for frequency of binge eating episodes and compensatory behaviors were decreased. Since this change, lifetime prevalence estimates increased to a range of 4% to 6.7% (Wade, 2019).

There have been few studies regarding incidence rates of BN, however recent research shows a peak incidence rate of 300 per 100,000 person-years in those 16-20 years of age in females according to DSM-IV criteria (Keski-Rahkohen, et al., 2009). There has been little updated research regarding the true community incidence rates of BN following the updates to the DSM, however small cohort studies suggest an increase in BN diagnosis (Hay, 2020).

1.1.3 Binge Eating Disorder

Diagnostic criteria for binge eating disorder (BED) are similar to that of BN, however they do not include compensatory behaviors. Episodes of binge eating are accompanied by three or more of the following:

1. Eating more rapidly than normal
2. Eating until feeling uncomfortably full
3. Eating large amounts of food when not feeling physically hungry
4. Eating alone because of feeling embarrassed by how much one is eating

5. Feeling disgusted with oneself, depressed, or very guilty

Those with BED experience distress from these episodes and these episodes occur at least once weekly for 3 or more months (American Psychiatric Association, 2013).

Recent estimates reveal the lifetime prevalence of BED to be higher than AN or BN with a lifetime prevalence of 4.5% across men and women, according to DSM-5 criteria (Kornstein, 2017). These prevalence rates are higher than what was estimated under the DSM-4, which were 1.2% and 1.5% for 12-month and lifetime prevalence, respectively. As with AN and BN, BED is more common in females than in men (Hudson, et al., 2007). Of note, though obesity is frequently associated with BED, more than half of patients (57.6%) with BED are clinically obese with a BMI >30 (Montano, Rasgon, and Herman, 2016).

Current research suggests that BED is heritable, similar to AN and BN, though few genetic studies of BED exist as it has only recently been listed as a DSM diagnosis. Heritability of BED ranges between 0.39 and 0.45 and those with BED continue to be recruited for Genome-Wide Association Studies (GWAS) to further understand the genetic factors that influence this disorder (Bulik, Blake, and Austin, 2019).

1.1.4 Avoidant/Restrictive Food Intake Disorder (ARFID)

ARFID commonly develops in infancy or early childhood and can persist into adulthood. ARFID is diagnosed with the following criteria:

1. An eating or feeding disturbance as manifested by persistent failure to meet appropriate nutritional and/or energy needs associated with significant weight loss, significant nutritional deficiency, dependency of enteral feeding/oral nutritional supplements, or marked interference with psychosocial functioning

2. The disturbance is not better explained by lack of available food or by an associated culturally sanctioned practice

3. The eating disturbance does not occur exclusively during the course of AN or BN and there is no evidence of a disturbance in the way in which one's body shape or weight is experienced

4. The eating disturbance is not attributable to a concurrent medical condition or not better explained by another mental disorder

Diagnostic markers of ARFID include malnutrition, low weight, growth delay, and the need for artificial nutrition. Further, those diagnosed with autism spectrum disorder and anxiety disorders may be at an increased risk of ARFID (American Psychiatric Association, 2013).

There are few studies estimating the epidemiology of ARFID, as this diagnosis was newly added in the publication of the DSM-5. One retrospective study reviewed patients attending treatment for eating disorders during 2012 who were diagnosed according to the DSM-IV. Of these patients, 22.5% met criteria for ARFID (Nicely, et al., 2014). A second study found that 13.8% of their patients met criteria for ARFID and were notably younger than their patients with AN or BN (Fisher, et al., 2014).

Currently, there is no empirically supported treatment for ARFID, though treatment generally focuses more on behavioral and nutritional approaches than is used in the treatment of other eating disorders (Zimmeran and Fisher, 2017).

1.1.5 Other Specified Feeding or Eating Disorder (OSFED)

OSFED is diagnosed in those who do not meet full criteria for one of the eating disorders described above, or whose symptoms cause clinically significant distress areas of functioning. Examples of OSFED include:

1. Atypical AN – all criteria of AN are met, except that despite significant weight loss, the individual's weight is within or above normal range

2. BN of low frequency and/or limited duration – all criteria for BN are met, except that the binge eating and compensatory behavior occur, on average, less than once weekly and/or for less than three months

3. BED of low frequency and/or limited duration – all criteria for BED are met, except that the binge eating occurs, on average, less than once a week and/or for less than three months

4. Purging disorder – recurrent purging behavior to influence weight or shape in the absence of binge eating

5. Night eating syndrome – Recurrent episodes of night eating, as manifested by eating after awakening from sleep or by excessive food consumption after the evening meal. There is awareness and recall of the eating. The night eating is not better explained by external influences such as changes in the individual's sleep-wake cycle or by local social norms. The night eating causes significant distress and/or impairment in functioning. The disordered pattern of eating is not better explained by binge-eating disorder or another mental disorder, including substance use, and is not attributable to another medical disorder or to an effect of medication

Lifetime prevalence rates of OSFED are higher than other classified eating disorders, with rates ranging from 2.73% and 3.96%. Comparatively, the proportion of those diagnosed with OSFED is higher than other eating disorders. According to Galmiche et al., 44.2% of women with eating disorders were diagnosed with OSFED and 39.5% of men with eating disorders were diagnosed with OSFED (Galmiche et al., 2019). OSFED also has the highest one-year prevalence rate, at 0.73%.

1.2 Eating Disorders and the COVID-19 Pandemic

The COVID-19 pandemic has presented unprecedented challenges to both those suffering from eating disorders, as well as those treating eating disorders. Preliminary research shows that those with eating disorders, whether in current treatment or not, experienced increased anxiety over both their ED symptoms and their access to food. In a study conducted by Termorschuizen, et al., over 1,000 persons from the US and the Netherlands who self-identified as having an eating disorder were assessed via an online survey asking quantitative questions regarding the impact of COVID-19 and their eating disorder symptoms and treatment. Participants were recruited through various social media outlets. A total of 53% of US participants reported that they are “very concerned” about worsening of their ED due to lack of structure. Regarding their behaviors, 29% of US participants reported that they frequently restricted their intake due to COVID-19 related factors and 29% reported increased anxiety about an inability to exercise due to COVID-19. Further, 40% of participants reported that the quality of their eating disorder treatment has been “somewhat worse than usual” following the transition to telehealth (Termorshuizen, et al., 2020).

1.3 Measures of Eating Disorder Symptomology

Several measures have been adapted to measure eating disorder symptomology, however the Eating Disorder Examination (EDE) has been widely used in clinical practice and research studies. The EDE is administered in an interview format by a trained clinician. The Eating Disorder Examination-Questionnaire (EDE-Q) was adapted from the EDE to be administered as a self-report manner. The EDE and EDE-Q were created by Fairburn & Cooper in 1993 and 1994, respectively, and are widely used to measure ED pathology. Both formats assess restraint, eating concern, weight concern, and shape concern, as well as produce a global score in the form of a mean of the four subscales. The reliability and validity of this measure have been documented in many studies throughout the years (Beumont, Kopec-Schrader, Talbot, & Touryz, 1993; Luce & Crowther, 1999; Mond, Hay, Rodgers, Owen, & Beumont, 2004; Berg, Peterson, Frazier, & Crow, 2012). The EDE-Q is not used exclusively to diagnose an eating disorder, and rather is used in conjunction with guidelines from the DSM and clinical presentation. The EDE-Q, though, can provide insight to the clinical and social impairment of one's disorder on their life. Generally, a global score of 4 or higher has been deemed indicative of a clinically diagnosable eating disorder (Allen, et al., 2011).

The Eating Attitudes Test (EAT) was first developed in 1979 to increase the early identification of AN. It was later adapted to a 26-item test (EAT-26) and is a self-report survey that assess three subscales: dieting, BN and food preoccupation, and oral control. Further, the EAT-26 suffers from the same limitations as other self-report scales, in that the way a patient is asked to complete the survey can impact their scoring – i.e., whether they are administered the survey in a group setting or individually. Further, the EAT-26 showed stronger validity in recognizing symptoms of AN than it did with BN or BED. A comparative study of the EDE-Q,

the EAT-26, and the EDI noted the risk for type-2 error. The EAT-26 did well with correctly rating healthy individuals as low risk, though those with symptoms of AN were not likely to receive as high of a score as would be expected (Birthe & Laberg, 2000).

The Eating Disorder Inventory (EDI) is a third self-report measure commonly used to evaluate eating disorder symptomology consisting of 12 primary scales: drive for thinness, BN, body dissatisfaction, low self-esteem, personal alienation, interpersonal insecurity, interpersonal alienation, interoceptive deficits, emotional dysregulation, perfectionism, asceticism, and maturity fears. The EDI was most recently revised in 2004 (EDI-3) and is comparative to the EDI and the EDE-2. The comparative study mentioned previously found that the EDE-Q was more adept at recognizing symptoms of AN than the EDI. Overall, this study showed the EDE-Q more accurately identifies eating disorder symptoms than the EDI or the EAT (Engelsen & Laberg, 2001).

1.4 Gaps in Knowledge

The COVID-19 pandemic resulted in a universal forced adaptation of eating disorder treatment delivery. To date, there is no one specific treatment modality that is superior in treating AN. Current treatment for AN is most successful when combining cognitive behavioral therapy (CBT) and interpersonal psychotherapy (IPT), however relapse rates among all patients with AN remain high between 35% and 41% (Kass, Kolko, & Wilfley, 2013; Berends, et al., 2016). CBT has been well established as a treatment for both BN and BED, however limited data exist on patients' response to virtual ED treatment delivery during the COVID-19 pandemic. Prior to the pandemic, the intensive treatments such as partial hospitalization and intensive outpatient were

largely in-person. The change in ED treatment delivery and engagement during virtual delivery includes a lack of in-person connection with a treatment team and other patients, challenges with viewing foods prepared and consumed during supervised meals, and needing to send weights from home scales. The impact of these changes has not been evaluated in regards to patients' ED treatment response and symptom improvement during virtual delivery.

1.5 Public Health Significance

Eating disorders have proven to be a significant public health crisis, both economically and medically. Recent reports from the Harvard School of Public Health revealed that eating disorders pose a $64.7 billion cost to the US economy encompassing productivity losses, informal care, efficiency losses, and the health system (Deloitte Access Economics, 2020). Further, despite the lethality of EDs, funding for ED research in the United States is minimal. In 2015, the federal spending support for eating disorder research was equal to $0.73 per affected person. Conversely, research for autism spectrum disorder was supported at rate of $58.65 per affected person and research for schizophrenia was supported at a rate of $86.97 per affected individual (Murray, et al., 2017). With one person dying every 52 minutes as a direct result of an eating disorder, further research into treatment and prevention is critical (Deloitte Access Economics, 2020).

2.0 Objectives

The objectives of this study were to (1) discuss differences in eating disorder symptom assessment and tracking during in-person treatment versus virtual care delivery; (2) identify levels of care for eating disorder treatment and key clinical variables, such as length of stay; and (3) assess improvements in eating disorder symptoms reported by patients enrolled in virtual eating disorder treatment. We hypothesized that patients receiving virtual eating disorder treatment with higher initial global EDE-Q scores will have a greater length in stay. We further hypothesized that patients will exhibit statistically significant improvements in self-reported symptoms during virtual treatment delivery. While the EDE-Q has been administered to assess symptom changes, further data are needed to evaluate those undergoing virtual treatment delivery as a result of the COVID-19 pandemic.

3.0 Methods

3.1 UPMC Center for Eating Disorders Study Population

This study population is derived from the University of Pittsburgh Medical Center's Center for Eating Disorders (UPMC CED) patient population base. Patients at UPMC CED were admitted either to the Partial Hospitalization (PHP), Intensive Outpatient Program (IOP), or Binge Eating Disorder outpatient group, as determined by the clinician completing the intake assessment based on behavior, BMI, and medical acuity. All patients met criteria for an eating disorder and were diagnosed with anorexia nervosa (AN), bulimia nervosa (BN), other specified feeding and eating disorders (OSFED), binge eating disorder (BED) or avoidant restrictive food intake disorder (ARFID) according to the *Diagnostic and Statistical Manual of Mental Disorders 5th edition*. This study population is comprised of 61 patients who sought virtual ED treatment at UPMC CED from April 2, 2020 through October 2, 2020. Of these 61 patients, 6 patients were male and 55 were female. Age and raced/ethnicity information was not readily available for this study.

Of these patients, 24 were initially recommended for treatment at PHP, 25 at IOP, and 12 at the BED group. There were 7 patients who presented to UPMC CED for an initial evaluation, however did not continue with treatment, and therefore did not have an initial diagnosis. Of those diagnosed, 13 patients were treated for AN, 5 patients for BN, 2 patients for BED, 32 patients for OSFED, and 2 patients for ARFID, for a total of 54 study participants (Table 1).

3.2 Eating Disorder Symptomology

Patients were asked to complete the EDE-Q on a weekly basis, beginning at their admission to their assigned program through an online form. The EDE-Q is scored using a 7-point scale, with scores of 4 or higher meeting clinical significance (Citation). Subscales are calculated as a mean of the total items related to the subscale. The global sum score is the average of the 4 subscales. As of July 30, 2020, the EDE-Q was altered within the clinic to reflect a 7-day measure to give an indication of weekly symptom change rather than the traditional 28-day measure time frame and was adjusted accordingly to normative scores to time represented in each scoring bin (Table 2). All subscales and other methods remained the same for computing the subscales and global scores. During the period of assessment for the present study (from April 2, 2020 through October 2, 2020), 36 patients completed their initial EDE-Q using the 28-day format and 25 patients completed their initial EDE-Q with the 7-day format.

3.3 Other Variables of Clinical Significance

3.3.1 CIA and PHQ-9

Due to the biopsychosocial nature of eating disorders, patients at UPMC CED were also presented with the Clinical Impairment Assessment (CIA) and the Patient Health Questionnaire (PHQ-9). The CIA is a 16-item self-report measure that assesses the severity of psychosocial impairment as a result of an eating disorder, such as mood, cognitive function, and work performance. Each item of the CIA is rated with a score of 0, 1, 2, or 3, with options of "not at

all", "a little", "quite a bit", and "a lot", respectively. Global CIA scores are calculated as a sum total of individual ratings with an impairment score of 16 being a cut off for clinically significant impairment (Bohn & Fairburn, 2008).

The PHQ-9 is a self-report diagnostic measure that assesses depressive symptoms in which scores of the 9 DSM-IV criteria for depression are rated on a scale of 0 ("not at all") to 3 ("nearly every day"). Scores for the PHQ-9 range from 0-27.

3.3.2 Length of Stay and Discharge Disposition

Individual length of stay was calculated by the number of days a patient was enrolled in programming at UPMC CED. Some patients remained under the care of UPMC CED while they transitioned from a PHP level to an IOP level. Thus, their total length of stay was the sum of their time in both programs. Discharge disposition was identified on a scale of 0-2 with 0 being a discharge to outpatient services, 1 being a discharge to either a separate IOP or residential treatment center (RTC), and 2 being a discharge against medical advice (AMA). A discharge to outpatient services was indicative of successful treatment completion. A discharge to an IOP or RTC indicated a need for further mental health treatment. An AMA discharge indicated non-compliance with treatment and medical recommendations.

3.4 Statistical Analyses

Data were analyzed utilizing StataSE 16 (StataCorp, 2016). Initial and final EDE-Q sub scores and global scores were compared via paired t-tests. ANOVA analyses and a Bonferroni

correction were used to compare initial global EDE-Q scores to both discharge disposition and initial diagnosis. Pearson correlation values were calculated to assess if there was a relationship between behaviors documented in the EDE-Q and level of care and initial diagnosis. A p-value of <0.05 was considered statistically significant.

4.0 Results

Completed baseline EDE-Q surveys were available for 61 patients, however only 75.4% (n=46) of these patients had a subsequent final EDE-Q survey that could be analyzed to provide metrics of change over time. Of the patients who entered into PHP ($n=24$), IOP ($n=25$), or BED ($n=12$), their mean initial global EDE-Q scores were 3.5, 3.4, and 3.2, respectively. Initial mean EDE-Q sub scores were calculated and are detailed in Table 3. The total average initial score for restraint was lowest of the scores with a mean of 2.08. The score of shape concern was highest with a mean of 4.46. Data were compared for those patients who had completed both an EDE-Q on admission and prior to discharge (n=46). There was a statistically significant improvement in global EDE-Q scores by a mean of 0.39, (p=0.05), as well as shape concern sub scores by a mean of 0.44, (p=0.05) (Figure 1) (Table 5). Improvement in restraint sub scores (p=0.10), eating concern (p=0.19), and weight concern (p=0.10) were not statistically significant (Table 4).

In comparing EDE-Q scores to a patient's diagnosis, 13 patients were diagnosed with AN, 5 patients with BN, 2 patients with BED, and 34 patients with OFSED/ARFID (Table 1). There were 7 patients who did not have an initial diagnosis as they did not continue with treatment after initial evaluation. A one-way ANOVA was conducted to determine if initial global EDE-Q score varied based on diagnosis. Initial analysis revealed that at least one diagnosis group had a statistically different initial EDE-Q score than others (p=0.03), however after applying the Bonferroni correction, there was no significant difference in initial scores across diagnoses (Table 6a). Due to the minimal amount of patients diagnosed with ARFID (n=2), those with ARFID and OSFED were combined into a new category. After further analysis by ANOVA, there was no particular diagnoses with a statistically significant initial global EDE-

Q score (p=0.23) (Table 6b). A one-way ANOVA analysis revealed no significant difference in initial global EDE-Q scores by initial levels of care (p=0.85) (Table 5).

Regarding discharge disposition, of those who submitted a final EDE-Q, patients discharged out to an outpatient level of care ($n=38$), to an IOP or residential program ($n=4$), or were discharged against medical advice (AMA) ($n=9$). Discharge disposition data were unavailable for 3 patients. A one-way ANOVA analysis revealed no significant difference in initial global EDE-Q scores with regard to discharge location (Table 8). Further, there was no significant correlation between initial EDE-Q score and length of stay (Table 7).

Of those who were initially provided a 28-day version of the EDE-Q (n=36), 17 patients had subsequent final EDE-Q scores measured with the same timeline. There were 25 patients who were initially assessed with a 7-day version of the EDE-Q and 29 patients whose final EDE-Q was the 7-day version. There were 7 patients who were initially assessed with a 28-day version of the survey, but completed their final EDE-Q with a 7-day format. When comparing improvement in behaviors according to survey format, these 7 patients were excluded from the analysis. Scoring of the 7-day format was comparative to the 28-day version according to baseline norms (Table 2).

Patients who were administered the 28-day EDE-Q reported a statistically significant improvement in binge eating episodes (p=0.05), however those assessed with a 7-day format did not (p=0.23) (Table 10). Conversely, those assessed with a 7-day measure reported a statistically significant improvement in frequency of episodes of loss of control (p=0.05), whereas the improvement in those completing the 28-day format was not significant (p=0.09). Patients in both groups reported a significant improvement with regard to days with a binge eating episode (p=0.03, p=0.05) (Tables 9 and 10). Improvement in frequency of vomiting episodes was not

statistically significant for the 28-day group or the 7-day group were significant (p=0.07 and p=0.33). Similarly, improvement in frequency of laxative use was not significant in either the 28-day group or the 7-day group, p=0.50 and p=0.33, respectively. Those in the 28-day group did report a significant improvement in frequency of over-exercise episodes (p=0.006), whereas the improvement among the 7-day group was not significant (p=0.458) (Tables 9 and 10). Of note, the effect sizes for all behaviors except laxative use were relatively large in those who completed the 28 day EDE-Q as their initial and final measure (Table 10)

Initial rates of behavior frequencies did not differ based on diagnosis in those who completed an initial EDE-Q in either format (Tables 11 and 12). Among those who completed an initial EDE-Q of the 7-day format, there was a statistically significant difference in frequency of episodes of loss of control between those in the BED and PHP groups (p=0.03) (Table 14). Further, in this group, there was a statistically significant difference in episodes of binge eating between BED and PHP (p=0.0005), as well as BED and IOP (p=0.0005). Among those who completed a 28-day format of the EDE-Q, there was a significant difference between the PHP and BED groups (p=0.003), as well as the IOP and BED groups (p=0.003) with regards to binge eating episodes (Table 13).

Regardless of EDE-Q format, there was no significant correlation between behavior frequency and total length of stay (Table 15). Those who completed an initial EDE-Q of the 28-day format experienced low levels of correlations with behaviors. Frequency of laxative use was the highest correlated behavior to length of stay, though still remained low. (r=0.46, p=0.06). There were even lower correlations of behavior with length of stay among those who were administered a 7-day format of the initial EDE-Q, with the highest correlative behavior being days with episodes of binge eating (r=-0.18, p=0.30).

5.0 Discussion

As hypothesized, there was a clinically significant improvement in self-report scores upon completion of virtual treatment. A previous study by Jennings and Phillips found clinical normative EDE-Q scores to be 3.31, 3.10, 2.95, 3.74, and 3.52 for global score, restraint, eating concern, shape concern, and weight concern in males (Jennings and Phillips, 2017). Among females with clinically diagnosed eating disorders, EDE-Q normative means are 4.0, 3.7, 3.5, 4.7, and 4.1 for global score, restraint, eating concern, shape concern, and weight concern (Dahlgren, Lindvall, and Rø, 2017). Both global EDE-Q score and shape concern subscale scores improved significantly. There was notable improvement in the other subscales, though these scores did not reach statistical significance. This is likely due to the small sample size of patients. The significant improvement in shape concern was unexpected, as traditionally, body image concerns are a continuing symptom throughout recovery and are often the last to subside (Eshkevari, et al., 2014). Further research may aim to investigate the relationship between body image/dysmorphia and continued exposure to a telehealth setting – patients in a virtual setting were required to focus on their appearance through a camera for a longer duration than they would be exposed to their body image in a traditional, in-person treatment setting.

Our secondary hypothesis regarding initial EDE-Q scores and length of stay, however, was rejected. There was little to no correlation between initial scores and length of stay in days. It is possible that other factors, such as rate of treatment progress or change in EDE-Q scores and reported behavior would be more predictive of length of stay in ED programming. Moreover, it is unclear if these results are directly related to the virtual form of treatment delivery, as baseline

data from in-person treatment was not assessed in this study. Therefore, it cannot be concluded that the lack of correlation was a result of telehealth delivery.

Of note, in assessing if there is an association between initial EDE-Q score and diagnosis, initial results were contradictory. When there were four diagnosis groups, initial one-way ANOVA revealed that one diagnosis group was statistically different from another. However, once a Bonferroni correction was applied, there was no significant difference between groups. This was likely due to the small sample size and the few patients who were diagnosed with ARFID. Those with ARFID and OSFED were combined into a similar diagnosis group as neither diagnosis fall under similar categorization as AN, BN, or BED. Once the sample was adjusted accordingly, a one-way ANOVA did not reveal any significant differences with regards to initial EDE-Q scores and diagnoses. This was in contrast to our initial hypothesis that those with BED and BN would have lower initial EDE-Q scores, as previous studies have shown that the EDE-Q has a higher reliability with screening for symptoms of AN than it does BN or BED (Reas, Grilo, and Masheb, 2004) (Binford, Grange, and Jellar, 2005).

There were several limitations present in this study. Most notably is the limited sample size. The clinic is limited in the number of patients that can participate in PHP, IOP, or BED outpatient group at one time, therefore given the short duration of time in which data were collected, there were fewer data points than would have been ideal. The limited sample size also limited the available power to evaluate race, ethnicity, and sex within this study, which can be a direction for future research. Despite this, though, our data still revealed not only clinically significant improvement in symptoms, but statistically significant improvement in symptoms. This study was further limited by the self-report nature of categorical information and timing of the baseline survey. Ideally, each patient's initial EDE-Q would be recorded during their clinical

interview, prior to beginning program. However, there were many patients who first completed the EDE-Q at the completion of their first week of treatment, which would not be a true measure of symptoms of behaviors prior to treatment. Further, not all patients completed their final EDE-Q at the end of their final week of treatment. There were patients who did not complete a final EDE-Q, or whose last completed EDE-Q was prior to the completion of treatment. Thus, these results may not be a true representation of treatment efficacy. Despite these limitations, the study has numerous strengths and the data still revealed both clinically and statistically significant data across the sample and levels of care. The relatively large effect sizes, particularly in behavior improvement among the 28 day EDE-Q cohort indicates that the data not reaching statistical significance was impacted by the small sample size. The data collected represented a clinical sample across diagnoses and levels of care. Further, this study highlighted the benefit of repeated data measurement among PHP, IOP, and BED levels of care, which allowed for evaluation of symptom changes over time in the context of virtual treatment delivery.

Previous studies have used weight gain as a measure of patient improvement. Weight data were not analyzed in this study as not all patients in the clinic required weight restoration. Further, there was concern for the reliability of patient weights, as these were being self-reported by patients due to the virtual nature of delivery. This may be considered as a measurement in future studies.

Overall findings of this study suggest that virtual treatment for eating disorders is effective in reducing ED symptoms and yielding referrals to lower levels of ED care, though future studies should directly compare outcomes of in-person treatment to virtual treatment. The COVID-19 pandemic has introduced a multitude of challenges to delivering healthcare. Eating disorder treatment was uniquely interrupted, and providers quickly adapted treatment delivery to

provide evidence-informed ED care to their patients. Given that EDs are associated with severe complications and high mortality rates, particular attention needs to be focused to the impact of the pandemic on treatment outcomes. The pandemic and its restrictions continue to pose new limitations to this vulnerable population, and research is warranted on the optimization of virtual delivery of care to support patients' reductions of ED symptoms and improved quality of life to stem this major public health concern.

Appendix A Tables

Table 1 Diagnoses and Level of Treatment in the Study Population

	Frequency (%)
AN Diagnosis	24.07
BN Diagnosis	9.26
BED Diagnosis	3.70
ARFID/OSFED Diagnosis	62.96
PHP	39.34
IOP	40.98
BED	19.67

Table 2 Comparing bins from the 28-day EDE-Q to the 7-day EDE-Q

28-day EDE-Q				7-day EDE-Q			
Lower bin (number of days)	**Upper bin (number of days)**	**Percent of timeframe lower bin**	**Percent of timeframe upper bin**	**Lower bin (number of days)**	**Upper bin (number of days)**	**Percent of timeframe lower bin**	**Percent of timeframe upper bin**
0	0	0	0	0	0	0	0
1	5	0.04	0.18	1	1	0.14	0.14
6	12	0.21	0.43	2	3	0.29	0.43
13	15	0.46	0.54	4	4	0.57	0.57
16	22	0.57	0.76	5	5	0.71	0.71
23	27	0.82	0.96	6	6	0.86	0.86
28	28	1	1	7	7	1	1

Table 3 Initial EDE-Q scores per level of care, N=61

	Mean Score ± SD		
	Partial Hospitalization Program	**Intensive Outpatient Program**	**Binge Eating Disorder Outpatient**
Global	3.45 ± 1.24	3.38 ± 1.32	3.20 ± 1.19
Restraint	2.09 ± 1.61	2.22 ± 1.54	1.82 ± 1.46
Eating Concern	3.05 ± 1.32	2.96 ± 1.44	2.65 ± 1.18
Shape Concern	4.58 ± 1.28	4.41 ± 1.21	4.31 ± 1.23
Weight Concern	4.09 ± 1.52	3.92 ± 1.46	4.02 ± 1.49

Table 4 Comparison of initial and final EDE-Q scores global and by sub scores, N=46

	Descriptive Statistics					Paired T-test Statistics					
	Mean	StDev	Lower	Upper	Std Error	Mean Diff	Lower	Upper	t	p-value	Cohen's d
Global											
Initial	3.34	1.30	2.96	3.73	0.19	0.39	-0.01	0.79	1.98	0.05*	0.29
Final	2.95	1.39	2.54	3.37	0.21						
Restraint											
Initial	2.09	1.75	1.57	2.61	0.26	0.42	-0.09	0.92	1.67	0.10	0.24
Final	1.67	1.75	1.15	2.19	0.26						
Eating Concern											
Initial	2.90	1.45	2.46	3.33	0.21	0.34	-0.18	0.86	1.33	0.19	0.23
Final	2.55	1.61	2.07	3.03	0.24						
Shape Concern											
Initial	4.40	1.35	4.00	4.80	0.20	0.44	-0.001	0.87	2.01	0.05*	0.30
Final	3.97	1.50	3.52	4.41	0.22						
Weight Concern											
Initial	3.99	1.54	3.54	4.45	0.23	0.37	-0.08	0.82	1.65	0.11	0.24
Final	3.62	1.50	3.18	4.10	0.22						

* Statistically significant with $p \leq 0.05$

Table 5 One-way ANOVA analysis of variance of initial global EDE-Q scores by initial level of care

Source	df	SS	MS	F	p
Between groups	2	0.53	0.26	0.17	0.85
Within groups	58	92.46	1.59		
Total	60	92.99			

Table 6a One-way ANOVA analysis of variance of initial global EDE-Q scores by diagnosis

Source	df	SS	MS	F	p
Between groups	4	15.51	3.88	2.77	0.04*
Within groups	49	68.57	1.40		
Total	53	84.08			

* Statistically significant with $p \leq 0.05$

Table 6b One-way ANOVA analysis of variance of initial global EDE-Q scores by diagnosis combining ARFID and OSFED

Source	df	SS	MS	F	p
Between groups	3	6.94	2.31	1.50	0.22
Within groups	50	77.14	1.54		
Total	53	84.08			

Table 7 Correlation values of initial EDE-Q scores to length of stay (days), N=52

Score	***r***	*p*
Global	-0.07	0.62
Restraint	-0.19	0.18
Weight Concern	-0.09	0.52
Shape Concern	0.07	0.62
Eating Concern	0.03	0.83

Table 8 One-way ANOVA analysis of variance of initial global EDE-Q scores by discharge disposition

Source	**df**	**SS**	**MS**	**F**	**p**
Between groups	2	1.62	0.81	0.55	0.58
Within groups	48	71.07	1.48		
Total	50	72.69			

Table 9 Comparison of initial and final behavior frequencies – 28 days EDE-Q, N=17

	Descriptive Statistics					Paired T-test Statistics					
	Mean	StDev	Lower	Upper	Std Error	Mean Diff	Lower	Upper	t	p-value	Cohen's d
Frequency of Binge Eating (Episodes)											
Initial	4.12	5.57	1.26	6.98	1.35	2.71	-0.03	5.44	2.10	0.05*	0.68
Final	1.41	0.73	-0.13	2.95	0.73						
Frequency of Loss of Control (Episodes)											
Initial	4.94	6.94	1.37	8.51	1.68	3.29	-0.65	7.23	1.77	0.09	0.61
Final	1.65	3.12	0.04	3.25	0.76						
Frequency of Binge Eating (Days)											
Initial	5.35	7.03	1.74	8.97	1.70	3.88	0.40	7.37	2.36	0.03*	0.72
Final	1.47	3.02	-0.08	3.02	0.73						
Frequency of Vomiting (Episodes)											
Initial	3.88	7.83	-0.15	7.91	1.90	3.47	-0.33	7.27	1.94	0.07	0.61
Final	0.41	1.70	-0.46	1.28	0.41						
Frequency of Laxative Use (Episodes)											
Initial	0.35	0.79	-0.05	0.76	0.19	-0.59	-2.40	1.22	-0.69	0.5	0.23
Final	0.94	3.40	-0.81	2.69	0.82						
Frequency of Over Exercise (Episodes)											
Initial	7.0	8.58	2.59	11.41	2.08	5.82	1.88	9.77	3.13	0.01*	0.93
Final	1.18	2.13	0.08	2.27	0.52						

* Statistically significant with $p \leq 0.05$

Table 10 Comparison of initial and final behavior frequencies – 7 day EDE-Q, N=15

	Descriptive Statistics					Paired T-test Statistics					
	Mean	StDev	Lower	Upper	Std Error	Mean Diff	Lower	Upper	t	p-value	Cohen's d
Frequency of Binge Eating (Episodes)											
Initial	1.13	1.68	0.20	2.07	0.43	0.4	-0.29	1.09	1.25	0.23	0.23
Final	0.73	1.83	-0.28	1.75	0.47						
Frequency of Loss of Control (Episodes)											
Initial	0.73	1.03	0.16	1.31	0.27	0.53	-0.02	1.08	2.09	0.05*	0.64
Final	0.2	0.56	-0.11	0.51	0.14						
Frequency of Binge Eating (Days)											
Initial	0.67	0.82	0.21	1.12	0.21	0.33	-0.01	0.68	2.09	0.05*	0.47
Final	0.33	0.62	-0.01	0.68	0.16						
Frequency of Vomiting (Episodes)											
Initial	0.33	0.90	-0.16	0.83	0.23	0.13	-0.15	0.42	1.00	0.33	0.17
Final	0.20	0.56	-0.11	0.51	0.14						
Frequency of Laxative Use (Episodes)											
Initial	0.07	0.26	-0.08	0.21	0.07	0.07	-0.08	0.21	1.00	0.33	0.38
Final	0	0	0	0	0						
Frequency of Over Exercise (Episodes)											
Initial	0.67	1.23	-0.02	1.35	0.32	-0.20	-0.76	0.36	-0.76	0.46	0.17
Final	0.87	1.06	0.28	1.45	0.27						

* Statistically significant with $p \leq 0.05$

Table 11 One-way ANOVA analysis of variance of initial behavior frequencies by diagnosis – 28 day EDE-Q, N=35

Frequency of Binge Eating (Episodes)					
Source	**df**	**SS**	**MS**	**F**	**p**
Between groups	3	262.47	87.49	1.09	0.37
Within groups	32	2577.83	80.56		
Total	35	2840.31			
Frequency of Loss of Control (Episodes)					
Source	**df**	**SS**	**MS**	**F**	**p**
Between groups	3	47.15	15.72	0.20	0.90
Within groups	32	2571.83	80.37		
Total	35	2618.97			
Frequency of Binge Eating (Days)					
Source	**df**	**SS**	**MS**	**F**	**p**
Between groups	3	14.07	4.69	0.05	0.98
Within groups	32	2904.91	90.78		
Total	35	2918.97			
Frequency of Vomiting (Episodes)					
Source	**df**	**SS**	**MS**	**F**	**p**
Between groups	3	76.52	25.51	0.41	0.75
Within groups	32	1984.48	62.01		
Total	35	2061.00			
Frequency of Laxative Use (Episodes)					
Source	**df**	**SS**	**MS**	**F**	**p**
Between groups	3	18.84	6.28	0.28	0.84
Within groups	32	711.47	22.23		
Total	35	730.31			
Frequency of Over Exercise (Episodes)					
Source	**df**	**SS**	**MS**	**F**	**p**
Between groups	3	153.73	51.24	0.98	0.41
Within groups	32	1670.82	52.13		
Total	35	1824.56			

Table 12 One-way ANOVA analysis of variance of initial behavior frequencies by diagnosis – 7 day EDE-Q, N=17

Frequency of Binge Eating (Episodes)					
Source	**df**	**SS**	**MS**	**F**	**p**
Between groups	2	1.90	0.95	0.27	0.77
Within groups	15	52.55	3.50		
Total	17	54.44			
Frequency of Loss of Control (Episodes)					
Source	**df**	**SS**	**MS**	**F**	**p**
Between groups	2	7.60	3.80	3.13	0.07
Within groups	15	18.18	1.21		
Total	17	25.78			
Frequency of Binge Eating (Days)					
Source	**df**	**SS**	**MS**	**F**	**p**
Between groups	2	8.40	4.20	3.07	0.07
Within groups	15	20.55	1.37		
Total	17	28.94			
Frequency of Vomiting (Episodes)					
Source	**df**	**SS**	**MS**	**F**	**p**
Between groups	2	0.88	0.44	0.62	0.55
Within groups	15	10.73	0.72		
Total	17	11.61			
Frequency of Laxative Use (Episodes)					
Source	**df**	**SS**	**MS**	**F**	**p**
Between groups	2	0.11	0.06	1.0	0.39
Within groups	15	0.83	0.06		
Total	17	0.94			
Frequency of Over Exercise (Episodes)					
Source	**df**	**SS**	**MS**	**F**	**p**
Between groups	2	0.76	0.38	0.27	0.77
Within groups	15	21.52	1.43		
Total	17	22.28			

Table 13 One-way ANOVA analysis of variance of initial behavior frequencies by initial level of care – 28 day EDE-Q, N=35

Frequency of Binge Eating (Episodes)					
Source	**df**	**SS**	**MS**	**F**	**p**
Between groups	2	846.68	423.34	7.01	0.003*
Within groups	33	1993.63	60.41		
Total	35	2840.31			
Frequency of Loss of Control (Episodes)					
Source	**df**	**SS**	**MS**	**F**	**p**
Between groups	2	111.56	55.78	0.73	0.49
Within groups	33	2507.42	75.98		
Total	35	2618.97			
Frequency of Binge Eating (Days)					
Source	**df**	**SS**	**MS**	**F**	**p**
Between groups	2	352.22	176.11	2.26	0.12
Within groups	33	2566.75	77.78		
Total	35	2918.97			
Frequency of Vomiting (Episodes)					
Source	**df**	**SS**	**MS**	**F**	**p**
Between groups	2	174.65	87.32	1.53	0.23
Within groups	33	1886.35	57.16		
Total	35	2061.00			
Frequency of Laxative Use (Episodes)					
Source	**df**	**SS**	**MS**	**F**	**p**
Between groups	2	26.20	13.10	0.61	0.55
Within groups	33	704.10	21.34		
Total	35	730.31			
Frequency of Over Exercise (Episodes)					
Source	**df**	**SS**	**MS**	**F**	**p**
Between groups	2	243.89	121.94	2.55	0.09
Within groups	33	1580.67	47.90		
Total	35	1824.56			

* Statistically significant with $p \leq 0.05$

Table 14 One-way ANOVA analysis of variance of initial behavior frequencies by initial level of care – 7 day EDEQ, N=24

Frequency of Binge Eating (Episodes)					
Source	**df**	**SS**	**MS**	**F**	**p**
Between groups	2	9.20	4.60	1.73	0.20
Within groups	22	58.56	2.66		
Total	24	67.76			
Frequency of Loss of Control (Episodes)					
Source	**df**	**SS**	**MS**	**F**	**p**
Between groups	2	14.08	7.04	4.24	0.03*
Within groups	22	36.56	1.66		
Total	24	50.64			
Frequency of Binge Eating (Days)					
Source	**df**	**SS**	**MS**	**F**	**p**
Between groups	2	18.84	9.42	10.91	0.0005*
Within groups	22	19.00	0.86		
Total	24	37.84			
Frequency of Vomiting (Episodes)					
Source	**df**	**SS**	**MS**	**F**	**p**
Between groups	2	3.51	1.76	2.71	0.09
Within groups	22	14.25	0.65		
Total	24	17.76			
Frequency of Laxative Use (Episodes)					
Source	**df**	**SS**	**MS**	**F**	**p**
Between groups	2	3.08	1.54	2.63	0.09
Within groups	22	12.92	0.59		
Total	24	16.00			
Frequency of Over Exercise (Episodes)					
Source	**df**	**SS**	**MS**	**F**	**p**
Between groups	2	2.78	1.39	0.31	0.74
Within groups	22	99.22	4.51		
Total	24	102.00			

* Statistically significant with $p \leq 0.05$

Table 15 Correlation values of initial behavior frequencies to length of stay (days)

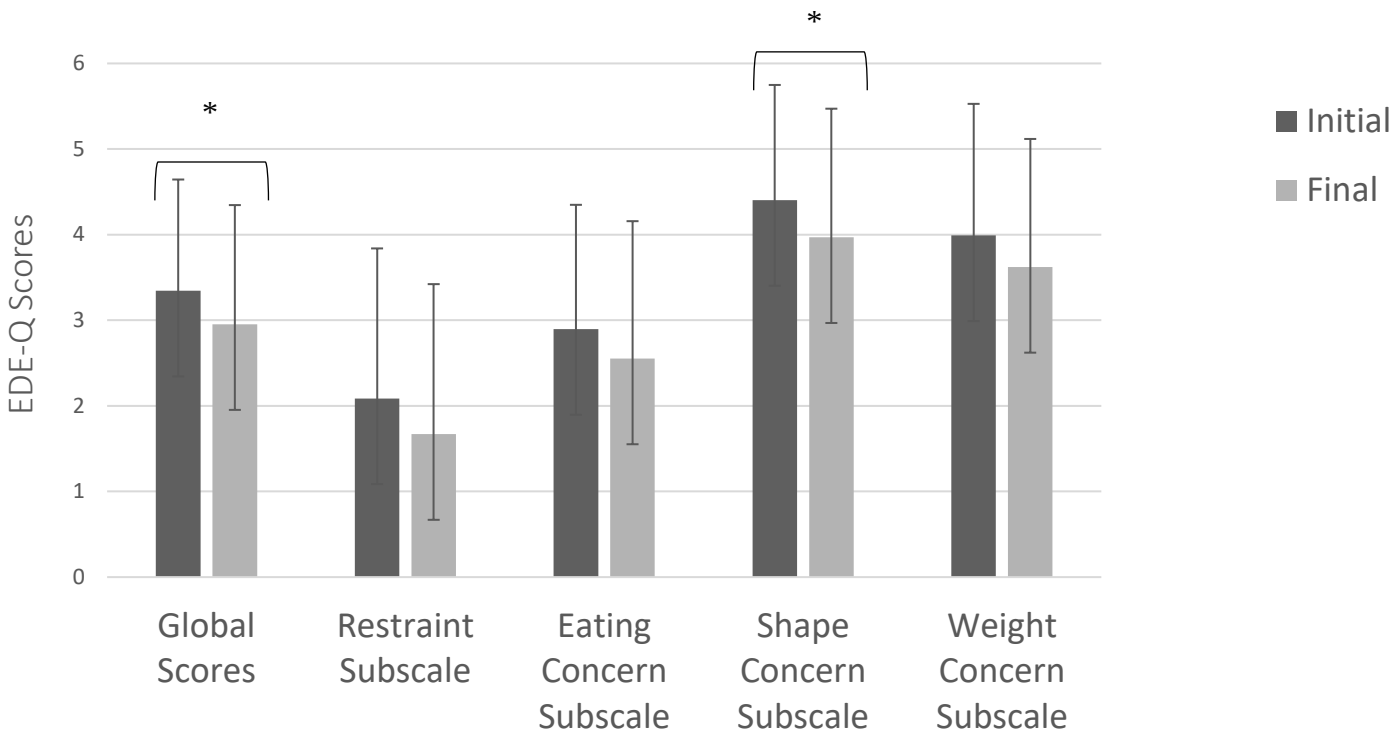
Behavior	**r *, *p*, N=35**	**r **, *p*, N=17**
Frequency of Binge Eating (Episodes)	-0.12, 0.49	-0.34, 0.18
Frequency of Loss of Control (Episodes)	-0.17, 0.33	-0.37, 0.14
Frequency of Binge Eating (Days)	-0.18, 0.30	-0.38, 0.13
Frequency of Vomiting (Episodes)	-0.08, 0.65	-0.33, 0.20
Frequency of Laxative Use (Episodes)	-0.11, 0.53	0.46, 0.06
Frequency of Over Exercise (Episodes)	0.05, 0.78	-0.42, 0.09

* values from the 28 day EDE-Q

** values from the 7 day EDE-Q

Appendix B Figures

Figure 1 Bar chart comparing initial EDE-Q scores to final EDE-Q scores



Bibliography

Allen, K. L., Byrne, S. M., Lampard, A., Watson, H., & Fursland, A. (2011). Confirmatory factor analysis of the eating disorder examination-questionnaire (EDE-Q). *Eating Behaviors, 12*(2), 143-151.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). https://doi.org/10.1176/appi.books.9780890425596.

Berends, T., van Meijel, B., Nugteren, W., Deen, M., Danner, U. N., Hoek, H. W., & van Elburg, A. A. (2016). Rate, timing and predictors of relapse in patients with anorexia nervosa following a relapse prevention program: a cohort study. *BMC psychiatry, 16*(1), 1-7.

Berg, K. C., Peterson, C. B., Frazier, P., & Crow, S. J. (2012). Psychometric evaluation of the eating disorder examination and eating disorder examination-questionnaire: A systematic review of the literature. *International Journal of Eating Disorders, 45*(3), 428-438.

Beumont, P. J. V., Kopec-Schrader, E. M., Talbot, P., & Touyz, S. W. (1993). Measuring the specific psychopathology of eating disorder patients. *Australian and New Zealand Journal of Psychiatry, 27*(3), 506-511.

Binford, R. B., Le Grange, D., & Jellar, C. C. (2005). Eating Disorders Examination versus Eating Disorders Examination-Questionnaire in adolescents with full and partial-syndrome bulimia nervosa and anorexia nervosa. *International Journal of Eating Disorders, 37*(1), 44-49.

Bohn, K., & Fairburn, C. G. (2008). The clinical impairment assessment questionnaire (CIA). Cognitive behavioral therapy for eating disorders, 315-317.

Bulik, C. M., Blake, L., & Austin, J. (2019). Genetics of eating disorders: What the clinician needs to know. *Psychiatric Clinics, 42*(1), 59-73.

Dahlgren, C. L., Stedal, K., & Rø, Ø. (2017). Eating Disorder Examination Questionnaire (EDE-Q) and Clinical Impairment Assessment (CIA): clinical norms and functional impairment in male and female adults with eating disorders. *Nordic journal of psychiatry, 71*(4), 256-261.

Deloitte Access Economics. *The Social and Economic Cost of Eating Disorders in the United States of America: A Report for the Strategic Training Initiative for the Prevention of Eating Disorders and the Academy for Eating Disorders*. June 2020. Available at: https://www.hsph.harvard.edu/striped/report-economic-costs-of-eating-disorders/.

Engelsen, B. K., & Laberg, J. C. (2001). A comparison of three questionnaires (EAT-12, EDI, and EDE-Q) for assessment of eating problems in healthy female adolescents. *Nordic Journal of Psychiatry, 55*(2), 129-135.

Eshkevari, E., Rieger, E., Longo, M. R., Haggard, P., & Treasure, J. (2014). Persistent body image disturbance following recovery from eating disorders. *International Journal of Eating Disorders, 47(*4), 400-409.

Fisher, M. M., Rosen, D. S., Ornstein, R. M., Mammel, K. A., Katzman, D. K., Rome, E. S., ... & Walsh, B. T. (2014). Characteristics of avoidant/restrictive food intake disorder in children and adolescents: a "new disorder" in DSM-5. *Journal of Adolescent Health, 55*(1), 49-52.

Galmiche, M., Déchelotte, P., Lambert, G., & Tavolacci, M. P. (2019). Prevalence of eating disorders over the 2000–2018 period: a systematic literature review. *The American journal of clinical nutrition, 109*(5), 1402-1413.

Hay, P. (2020). Current approach to eating disorders: a clinical update. *Internal medicine journal, 50*(1), 24-29.

Hudson, J. I., Hiripi, E., Pope Jr, H. G., & Kessler, R. C. (2007). The prevalence and correlates of eating disorders in the National Comorbidity Survey Replication. *Biological psychiatry, 61(*3), 348-358.

Jennings, K. M., & Phillips, K. E. (2017). Eating Disorder Examination–Questionnaire (EDE–Q): norms for clinical sample of female adolescents with anorexia nervosa. *Archives of psychiatric nursing, 31(*6), 578.

Kass, A. E., Kolko, R. P., & Wilfley, D. E. (2013). Psychological treatments for eating disorders. *Current opinion in psychiatry, 26*(6), 549.

Keski-Rahkonen, A., Hoek, H. W., Susser, E. S., Linna, M. S., Sihvola, E., Raevuori, A., ... & Rissanen, A. (2007). Epidemiology and course of anorexia nervosa in the community. *American Journal of Psychiatry, 164*(8), 1259-1265.

Keski-Rahkonen, A., Hoek, H. W., Linna, M. S., Raevuori, A., Sihvola, E., Bulik, C. M., ... & Kaprio, J. (2009). Incidence and outcomes of bulimia nervosa: a nationwide population-based study. *Psychological medicine, 39*(5), 823.

Kornstein, S. G. (2017). Epidemiology and recognition of binge-eating disorder in psychiatry and primary care. *The Journal of clinical psychiatry, 78*(suppl 1), 3-8.

Luce, K. H., & Crowther, J. H. (1999). The reliability of the eating disorder examination—Self-report questionnaire version (EDE-Q). *International Journal of Eating Disorders, 25*(3), 349-351.

Mond, J. M., Hay, P. J., Rodgers, B., Owen, C., & Beumont, P. J. (2004). Validity of the Eating Disorder Examination Questionnaire (EDE-Q) in screening for eating disorders in community samples. *Behaviour research and therapy, 42*(5), 551–567.

Montano, C. B., Rasgon, N. L., & Herman, B. K. (2016). Diagnosing binge eating disorder in a primary care setting. *Postgraduate medicine, 128*(1), 115-123.

Murray, S. B., Pila, E., Griffiths, S., & Le Grange, D. (2017). When illness severity and research dollars do not align: are we overlooking eating disorders?. *World Psychiatry, 16*(3), 321.

Mustelin, L., Silén, Y., Raevuori, A., Hoek, H. W., Kaprio, J., & Keski-Rahkonen, A. (2016). The DSM-5 diagnostic criteria for anorexia nervosa may change its population prevalence and prognostic value. *Journal of psychiatric research, 77*, 85-91.

Nicely, T. A., Lane-Loney, S., Masciulli, E., Hollenbeak, C. S., & Ornstein, R. M. (2014). Prevalence and characteristics of avoidant/restrictive food intake disorder in a cohort of young patients in day treatment for eating disorders. *Journal of eating disorders, 2*(1), 1-8.

Raevuori, A., Hoek, H. W., Susser, E., Kaprio, J., Rissanen, A., & Keski-Rahkonen, A. (2009). Epidemiology of anorexia nervosa in men: a nationwide study of Finnish twins. *PloS one, 4*(2), e4402.

Reas, D. L., Grilo, C. M., & Masheb, R. M. (2006). Reliability of the Eating Disorder Examination-Questionnaire in patients with binge eating disorder. *Behaviour research and therapy, 44*(1), 43-51.

StataCorp. 2019. *Stata Statistical Software: Release 16*. College Station, TX: StataCorp LLC.

Termorshuizen, J. D., Watson, H. J., Thornton, L. M., Borg, S., Flatt, R. E., MacDermod, C. M., ... & Bulik, C. M. (2020). Early impact of COVID-19 on individuals with self-reported eating disorders: A survey of~ 1,000 individuals in the United States and the Netherlands. *International Journal of Eating Disorders, 53*(11), 1780-1790.

Wade, T. D. (2019). Recent research on bulimia nervosa. *Psychiatric Clinics, 42*(1), 21-32.

Zimmerman, J., & Fisher, M. (2017). Avoidant/restrictive food intake disorder (ARFID). *Current problems in pediatric and adolescent health care, 47*(4), 95-103.