



## OPEN ACCESS

## EDITED BY

Francesca Pacitti,  
University of L'Aquila, Italy

## REVIEWED BY

Tommaso Barlattani,  
University of L'Aquila, Italy  
Roberto Corrêa Leite,  
DBT Campinas, Brazil

## \*CORRESPONDENCE

Elizabeth Kroll  
✉ Elizabeth.Kroll@charliehealth.com

RECEIVED 30 July 2025

REVISED 29 October 2025

ACCEPTED 05 November 2025

PUBLISHED 10 December 2025

## CITATION

Kroll E, Hessdorf S, Brandis A, Kanopka K,  
Kohlmeier J, Heard H, Zant I,  
Stolzenbach S and Fenkel C (2025) Examining  
caregiver-adolescent symptom agreement  
and its influence on treatment outcomes: a  
quantitative analysis.  
*Front. Psychol.* 16:1676629.  
doi: 10.3389/fpsyg.2025.1676629

## COPYRIGHT

© 2025 Kroll, Hessdorf, Brandis, Kanopka,  
Kohlmeier, Heard, Zant, Stolzenbach and  
Fenkel. This is an open-access article  
distributed under the terms of the [Creative  
Commons Attribution License \(CC BY\)](#). The  
use, distribution or reproduction in other  
forums is permitted, provided the original  
author(s) and the copyright owner(s) are  
credited and that the original publication in  
this journal is cited, in accordance with  
accepted academic practice. No use,  
distribution or reproduction is permitted  
which does not comply with these terms.

# Examining caregiver-adolescent symptom agreement and its influence on treatment outcomes: a quantitative analysis

Elizabeth Kroll<sup>1\*</sup>, Sara Hessdorf<sup>2</sup>, Alyson Brandis<sup>2</sup>,  
Klint Kanopka<sup>3</sup>, Jonathan Kohlmeier<sup>1</sup>, Hannah Heard<sup>1</sup>,  
Izabella Zant<sup>1</sup>, Stephanie Stolzenbach<sup>1</sup> and Caroline Fenkel<sup>1</sup>

<sup>1</sup>Charlie Health, Inc, Bozeman, MT, United States, <sup>2</sup>Department of Practicum Education and Community Partnerships, New York University, New York, NY, United States, <sup>3</sup>Department of Applied Statistics, Social Science, and Humanities, Steinhardt School of Culture, Education, and Human Development, New York University, New York, NY, United States

**Introduction:** Caregivers play a critical role in identifying and interpreting adolescent mental health symptoms, yet their perceptions often differ from those of adolescents themselves. Such discrepancies can influence treatment engagement and outcomes, making it essential to understand the factors that affect caregiver-adolescent agreement in symptom reporting.

**Methods:** This study analyzed baseline depression and anxiety symptom reports from 763 caregiver-adolescent dyads enrolled in a virtual intensive outpatient program. Agreement between adolescent self-reports and caregiver proxy reports was assessed using concordance analyses. Potential moderators included adolescent age, gender identity, race, neurodivergence, caregiver well-being, and family functioning. Additional analyses examined whether agreement predicted treatment outcomes, including symptom improvement and treatment completion.

**Results:** Overall, caregiver-adolescent agreement was low to moderate (CCC = 0.215). Adolescents identifying as non-binary showed the largest discrepancies, typically rating their symptoms as more severe than their caregivers. Higher family functioning was associated with reduced directional disagreement for depression, though this pattern did not hold for anxiety. Dyads in which caregivers reported higher symptom severity than adolescents demonstrated greater symptom improvement by discharge, while agreement levels did not significantly predict treatment completion.

**Discussion:** Informant discrepancies were common and clinically meaningful, particularly among gender-diverse adolescents and families reporting lower functioning. Although overall agreement did not predict treatment completion, caregiver perceptions of greater symptom severity were linked to greater reductions in adolescent depression and anxiety over treatment. These findings suggest that incorporating caregiver perspectives into assessment and treatment planning may enhance outcomes, especially in high-acuity virtual care settings. Future research should investigate mechanisms through which caregiver engagement influences adolescent symptom change over time.

## KEYWORDS

IOP, mental health, depression, anxiety, caregivers, adolescents

# 1 Introduction

## 1.1 Adolescent mental health

The mental health crisis among adolescents and young adults is a significant public health concern, with anxiety and depression being among the most prevalent disorders affecting this demographic (Patel et al., 2007; McGorry and Mei, 2018). The World Health Organization (2024) reports that 4.4% of 10–14-year-olds and 5.5% of 15–19-year-olds experience an anxiety disorder, and an estimated 1.4% of adolescents aged 10–14 years, and 3.5% of 15–19-year-olds experience depression. Within the United States, the Centers for Disease Control and Prevention (2025) state that approximately 13% of adolescents experienced a diagnosed mental or behavioral health condition between 2018 and 2019. For young adults, this figure rises to nearly 25%, reflecting the ongoing challenges they face as they navigate the transition to adulthood.

These conditions may disrupt daily functioning, education, family formation, and damage social relationships (Gustavson et al., 2018). They also can result in life-threatening situations such as suicidal ideation (Windarwati et al., 2022). A study by Hughes et al. (2017) underscores the long-term consequences of untreated mental health disorders, linking early exposure to anxiety and depression with increased risks of substance use, chronic health problems, and poor relationship outcomes in adulthood. They found that those who experienced mental health challenges in adolescence were more likely to experience difficulties in maintaining stable employment and relationships as young adults (Hughes et al., 2017). Children with untreated anxiety can also experience impaired coping such as substance use or suicidal ideation (Bosworth, 2019). Recognizing these multifaceted influences highlights the importance of early identification and intervention, as addressing these factors can significantly improve long-term mental and behavioral health outcomes for adolescents and young adults.

## 1.2 The role of caregivers in adolescent mental health treatment

Caregivers play a crucial role in the mental health of adolescents, as children rely on their caregivers for basic needs, emotional support, and socialization (Center on the Developing Child, 2016). Bowen Family Systems theory frames the family as an interconnected system in which distress or dysfunction in one member can reverberate across the entire unit, sometimes across generations (Erdem and Safi, 2018). Research further supports this system's view: Panter-Brick et al. (2014) found that caregiver mental health scores were predictive of the scores of their adolescents on comparable assessments, suggesting a direct link between parent and child well-being. Furthermore, greater family functioning was correlated with increased parent satisfaction and self-efficacy which is indicative of greater social support from the caregivers (Angle et al., 2015). Together these findings underscore the importance of considering the caregiver context in both assessment and treatment planning for adolescents.

Beyond their role in shaping mental health outcomes, caregivers also serve as gatekeepers of treatment. Caregivers with self-reported mental health issues tend to assess greater need for their adolescents to receive mental health treatment, and are therefore more likely to

encourage treatment. In turn, these adolescents experience more successful treatment, including fewer dropouts and improved treatment outcomes (De Los Reyes et al., 2021). By contrast, caregivers who assess adolescents as having fewer mental health issues are less likely to encourage their child to receive mental health treatment, which may be harmful for those who assess themselves differently than their caregivers (Bajeux et al., 2018).

Caregiver misperception is not uncommon: in a study of young children ages 2–6 who met criteria for a psychiatric diagnosis, only 38.8% were perceived by their parents as needing help (McGinnis et al., 2022). While this sample skews younger, it highlights a persistent challenge across developmental stages, caregiver awareness is essential to recognizing and addressing adolescent needs. As a result mental health care can be delayed which has long been shown to have negative effects on individuals with studies showing increased risk of screening for depression and anxiety (Noel et al., 2025), and increased risk of treatment drop out (Costantini, 2025). For neurodivergent individuals specifically, a delay in diagnosis is associated with poorer mental health in general (Huynh et al., 2024; Holden and Kobayashi-Wood, 2025), and increased medical utilization (Vu et al., 2023).

When caregiver and adolescent views align, treatment is more likely to be timely and effective. However, agreement on symptom severity is often inconsistent. A study examining the level of agreement between parents and adolescents and the need for mental healthcare showed that almost half of either group identified a need for mental health care, but only one third had their needs met (Schnyder et al., 2020). Agreement on whether care was needed was moderate with higher disagreement being linked to the parents' lack of knowledge of the adolescent's emotional state. This emphasizes the importance of strong communication and connection between caregivers and adolescents, as well as the effect these relationships have on mental health and mutual understanding.

## 1.3 Self-reported vs. caregiver-reported

Comparing mental health assessments from adolescent clients and their caregivers offers critical insight into the high rate of discrepancies in adolescent mental health reporting (Bajeux et al., 2018). Informant discrepancies not only offer insight into a caregiver-child relationship, but can also impact the course of an adolescent's mental health treatment. Around 76% of caregiver-child dynamics disagree on a target problem to address during treatment (De Los Reyes and Kazdin, 2005). These disagreements may lead to misaligned treatment goals, diminished therapeutic engagement, or conflicting expectations between the therapist, adolescent, and caregiver.

Informant discrepancies show up for a variety of reasons. Adolescent clients may demonstrate limited objectivity when reporting on themselves, their mental health symptoms, and behaviors, in part due to ongoing identity development and evolving self-awareness (Dang et al., 2020). Additionally, adolescent clients may be more likely to under-report symptoms out of shame or a distrust of authority and how these reports will be used (Tarren-Sweeney, 2019). On the other side, caregivers present their own set of biases when it comes to reporting on adolescent mental health. Caregivers tend to over-report external or behavioral problems, and under-report internal problems for adolescents (Bajeux et al., 2018). As a result, adolescents struggling with

internalizing symptoms may not receive appropriate interventions, while externalizing behaviors may be more likely to drive treatment decisions.

Neurodivergent conditions such as ASD and ADHD may influence the level of agreement between caregiver- and adolescent-report, in part because differential symptom observability (e.g., externalizing vs. internalizing behaviors), reduced adolescent self-insight, and altered social-cognitive processes (e.g., perspective taking, awareness of change) can create systematic informant divergences (Lopez-Espejo et al., 2023; McDonald et al., 2016; Ismail et al., 2017).

Gender can also influence these patterns of perception. Moens et al. (2018) found that caregivers are less likely to report behavioral problems in female adolescents compared to males, potentially reflecting gendered expectations about emotional expression and social behavior. These biases are often influenced by the caregivers' own mental health, parenting styles, and attitudes. Specifically, caregivers with their own mental health symptoms tend to overestimate adolescents' behavioral problems (Ehrlich et al., 2011). Parenting style also plays a role: punitive or authoritarian parents tend to over-pathologize adolescent behavior, whereas more permissive caregivers may underestimate concerns (Bajeux et al., 2018). In short, discrepancies between adolescent and caregiver mental health reports are not random; they are formed by social and psychological factors that can significantly impact assessment accuracy, treatment planning, and the therapeutic alliance.

## 1.4 Study goal

This study examines the relationship between self-reported and caregiver-reported symptom severity for both anxiety and depression, with the goal of identifying patterns in discrepancies between adolescent and caregiver ratings. Additionally, it explores how these differences may influence treatment outcomes. Based on prior literature, we expected to find that high agreement will significantly predict better outcomes for both anxiety and depression symptoms. Additionally, we expected these findings to be moderated by gender. When interpreting results, readers should consider the unique context of a virtual IOP. Our goal was not to evaluate differences between virtual and in-person settings, but rather to examine caregiver-adolescent agreement in the context of virtual care, where the dynamics of reporting and involvement may differ.

## 2 Methods

### 2.1 Setting

Data was collected from adolescent clients, ages 11–17, admitted to Charlie Health and their caregivers. Charlie Health is a virtual intensive behavioral health outpatient treatment program that treats individuals between the ages of 11 to 50 who present with high acuity mental health conditions. Treatment lasts approximately 9–12 weeks. Each week, programming at Charlie Health consists of 9 h of group sessions, 1 h of individual therapy, and an optional hour of family therapy depending on the client's needs. While not required, family participation is encouraged.

### 2.2 Ethics statement

This study was reviewed and approved by NorthStar Institutional Review Board (Protocol #NB500297) as secondary research under the Common Rule 45 CFR 46.104(d)(4)(ii).

### 2.3 Data collection

Client self-reported data were collected between October 2023 and January 2025 at intake and discharge from Charlie Health's treatment program. For intake data, clients were sent an email 3 h prior to their initial admissions assessment, which asked them to fill out the intake survey. For those who did not submit prior to their intake assessment, they were given an intake survey in the first hour of their orientation to group sessions. A Charlie Health staff member joined the group and distributed personalized links to each client. For the small subset of clients who did not complete an intake survey by the end of orientation, emails follow ups were sent with the link for up to 1 week post first appointment. At all points of contact, clients were instructed that the survey was optional and would not impact their admission status.

Discharge surveys were distributed on the clients' last day in group sessions. Clients were informed that the survey was optional and would not affect their discharge status. If a client missed their final group session, they were emailed and texted a personalized link to the discharge survey and prompted to fill it out with a \$5 gift card incentive.

Caregiver data was collected between October 2023 and January 2025 during their client's first full week of admission. The survey was sent via a routine week 1 email to all primary caregivers listed on clients' account. Surveys remained open for 1 week. Caregivers filled this survey out on their own time.

Both Client and Caregiver surveys were filled out on a personal computer and were asked to be taken in a confidential space.

In order to be included in the study both an intake survey and caregiver survey had to be submitted. Additionally, to be included in discharge analyses, a discharge survey had to be submitted by the client. There were no additional exclusion criteria.

### 2.4 Measures

#### 2.4.1 Client measures

##### 2.4.1.1 Depression

Client self-reported depression severity was measured with The Patient Health Questionnaire (PHQ-9) (Johnson et al., 2002). This nine-item measure uses a 0–3 scale (not at all, several days, more than half the days, nearly every day) for all questions and produces a final sum score between 0–27 with scores of 5, 10, 15, and 20 representing mild, moderate, moderately severe, and severe depression, respectively. Respondents are asked to recall how often symptoms have impacted them in the 2 weeks prior to the assessment.

##### 2.4.1.2 Anxiety

Client self-reported anxiety severity was measured with The Generalized Anxiety Disorder-7 (GAD-7) (Spitzer et al., 2006). This

seven-item measure used a 0–3 (not at all, several days, more than half the days, nearly every day) scale for all questions and produced a sum score between 0 and 21 with score cut offs of 5, 10, and 15 representing mild, moderate, and severe anxiety, respectively. Respondents are asked to recall how often symptoms have impacted them in the 2 weeks prior to the assessment.

#### 2.4.1.3 Discharge reason

Reason for discharge was gathered from administrative data and condensed into routine discharge and non-routine discharge. Routine discharge was decided by a client's primary therapist and defined as the patient completing treatment plan goals, or the patient stabilizing to such an extent that completion of the treatment process at a less restrictive level-of-care was therapeutically appropriate. All other discharge reasons were collapsed into non-routine indicating that they were discharged prior to completing treatment.

#### 2.4.1.4 Demographics

Client demographics data was collected at intake. Clients were asked to disclose their age, gender, race, and neurodivergent identity all identity questions were multi-select to allow clients to select all answers that matched their preferred identities. Additionally, all questions were optional and clients could opt to not disclose any demographic information. Demographic categories are reported in Table 1.

### 2.4.2 Caregiver measures

#### 2.4.2.1 Caregiver well-being

Caregiver well-being was collected using the WHO-5 well-being Index. This five-item measure is scored on a 0–5 scale and final scores are calculated by summing all answers and multiplying by four to obtain a score on a scale from 0 to 100 with higher scores indicating

higher well-being. A score below 50 is often considered at-risk for depression (Topp et al., 2015).

#### 2.4.2.2 Child depression

Caregivers were asked to assess the severity of their loved one's depression using the PROMIS Parent Proxy Depressive Symptoms short form (Irwin et al., 2012; Broderick et al., 2013). This six-item measure used a 1–5 scale (never, almost never, sometimes, often, almost always) and sum scores were calculated to obtain a final score between 6 and 30. Caregivers were asked to recall how symptoms impacted their loved one in the 7 days prior.

#### 2.4.2.3 Child anxiety

Caregivers were asked to assess the severity of their loved one's anxiety symptoms using the PROMIS Parent Proxy Anxiety short form (Irwin et al., 2012; Broderick et al., 2013). This eight-item measure used a 1–5 (never, almost never, sometimes, often, almost always) scale and sum scores were calculated to obtain a final score between 8 and 40. Caregivers were asked to recall how symptoms impacted their loved one in the 7 days prior.

#### 2.4.2.4 Family functioning

Family functioning was assessed using the Family Satisfaction Subscale (FSS) of the Family Adaptability and Cohesion Evaluation Scales (FACES), which evaluates overall satisfaction with various aspects of family dynamics. This eight-item measure asks respondents to rate their satisfaction with family-related factors such as emotional closeness, communication quality, time spent together, conflict resolution, and mutual support on a scale of 1 to 5 (never or almost never, once in a while, sometimes, frequently, almost always or always) with a recall time of 1 month prior to assessment.

Items include prompts such as, “The degree of closeness between family members” and, “Your family's ability to resolve conflicts.” Responses were rated on a likert scale, with higher scores on a scale of 8 to 40 indicating greater satisfaction with family functioning. The FSS has been used in both clinical and research contexts to capture individuals' subjective evaluations of family support and cohesion (Olson et al., 2013).

#### 2.4.2.5 Demographics

Caregivers were asked to identify their relationship to the client (Table 2).

## 2.5 Data prep

All data were pulled from their respective locations and merged based on unique client IDs. Client data was deleted from the data set if there was no corresponding caregiver survey. Both client self-reported data and parent proxy scales for depression and anxiety were then standardized using z-score standardization. This process transforms raw scores into standardized values based on the mean and standard deviation of the sample, allowing for direct comparison between measures that originally use different scoring systems.

Level of agreement on depression and anxiety symptoms was calculated by subtracting the client self-reported scaled score from the caregiver proxy scaled score. Negative scores indicated that clients rated their symptoms as more severe than their caregivers; a score of

TABLE 1 Client demographics.

Demographic category	Percentage	<i>n</i>
<b>Race</b>		
Black or African American	7.86%	60
White	70.77%	540
Other	9.70%	74
Missing	11.66%	89
<b>Gender</b>		
Female	52.42%	400
Male	30.41%	232
Non-binary/other	10.09%	77
Missing	7.08%	54
<b>Neurodivergent identity</b>		
Autism	2.62%	20
ADHD	22.54%	172
AuDHD	3.15%	24
Other	6.42%	49
None	40.37%	308
Missing	24.90%	190



TABLE 2 Caregiver and client relationships.

Caregiver relationship	Percentage	<i>n</i>
Parent	93.71%	715
Grandparent	3.54%	27
Other direct relation	1.18%	9
Legal guardian–non family	0.92%	7
Missing	0.66%	5

0 indicated perfect agreement, and positive scores indicated that caregivers scored client symptoms as more severe than the client themselves.

To enable stable analyses, response categories were collapsed into the largest conceptually coherent groupings. This approach reduced small cell sizes while preserving the most meaningful distinctions in the data. Additionally, to enable analyses that required mutually exclusive categories, multi-select responses were recoded into an “other” category across all demographics.

All analyses were conducted in R (Version 4. X. X; [R Core Team, 2013](#)) using the following packages: *dplyr*, *ggplot2*, *ggrepel*, *segmented*, and *rms* ([Wickham et al., 2023](#); [Wickham, 2023](#); [Slowikowski, 2023](#); [Muggeo, 2022](#); [Harrell, 2023](#)).

## 2.6 Data analysis strategy

To assess general trends in symptom agreement, concordance correlation coefficients ([Lin, 1989](#)) were estimated between client-reported and caregiver-reported depression (PHQ-9 vs. PROMIS Parent Proxy Depression) and anxiety (GAD-7 vs. PROMIS Parent Proxy Anxiety) scores. Bland–Altman analyses were performed to assess variation in agreement across the range of observed symptom severities ([Bland and Altman, 1986](#)).

To determine whether client demographics were associated with symptom agreement, a linear regression was conducted to test the relationship between client age and agreement scores for depression and anxiety. Additionally, one-way ANOVAs were performed to assess whether client gender, race, and neurodivergent identity impacted agreement scores, with Bonferroni corrections applied to adjust for multiple comparisons.

Next, the relationship between family functioning and symptom agreement was explored through a series of linear and logistic regressions. Simple linear regressions examined whether family functioning predicted overall symptom agreement for depression and anxiety. Additional models assessed the relationship between family functioning and magnitude of agreement (absolute difference regardless of direction) and family functioning and disagreement direction (whether caregivers rated symptoms higher than clients). Parallel analyses were conducted to investigate whether caregiver well-being was similarly associated with symptom agreement patterns.

To evaluate the impact of symptom agreement on treatment outcomes, linear regressions were conducted to determine whether agreement scores at intake predicted depression and anxiety severity at discharge. Additionally, logistic regressions assessed whether agreement level (overall score, absolute discrepancy, or direction of disagreement) influenced treatment completion (routine vs. non-routine discharge). Finally, based on prior analyses identifying

age, gender, and family functioning as significant predictors of agreement, multiple linear regression models were used to determine whether these variables moderated the relationship between agreement scores and post-treatment symptom severity.

## 3 Results

### 3.1 Descriptives

#### 3.1.1 Demographics

Data were collected from 763 client and caregiver duos between October 2023 and January 2025, which represents approximately 5% of the eligible families within this time period ( $n = 16,797$ ). Within this sample, the majority of clients identified as white (70.77%), female (52.42%), and neurotypical (40.37%). Within the caregiver sample, the majority of respondents identified themselves as the client’s parent (93.71%). Additional demographic information can be found in [Tables 1, 2](#).

At intake, clients presented with moderate depression (PHQ-9  $M = 13.10$ , 95% CI [12.60, 13.57]) and anxiety (GAD-7  $M = 10.73$ , 95% CI [10.37, 11.20]). Internal consistency was excellent for both the PHQ-9 ( $\alpha = 0.88$ ) and GAD-7 ( $\alpha = 0.90$ ) at intake, consistent with the literature. Clients who reported self-harm behavior (38.76%) reported an average of 6.42 days of self-harm in the 30 days leading up to admission. Comparatively, caregiver proxy scores were moderately high for depression ( $M = 19.46$ , 95% CI [19.10, 19.82]) and moderate for anxiety ( $M = 22.27$ , 95% CI [21.84, 22.71]). By discharge, clients reported mild depression ( $M = 6.44$ , 95% CI [5.76, 7.11]) and mild anxiety ( $M = 5.10$ , 95% CI [4.53, 5.67]), and internal consistency remained high (PHQ-9  $\alpha = 0.89$ ; GAD-7  $\alpha = 0.91$ ).

On the WHO-5 Well-Being Index, caregivers self-reported an average score of 50.07 (95% CI [48.51, 51.30]), placing them just above the clinical threshold commonly used to flag risk for depression. This score reflects the caregiver’s own well-being, not their perception of the client’s mental health. Approximately 44% ( $n = 398$ ) of caregivers fall below that 50 point threshold indicating poor well-being in almost half of the sample. Internal consistency remained high for the items in the WHO-5 ( $\alpha = 0.88$ ), consistent with the literature on the WHO-5. Finally, caregivers reported moderately high family functioning ( $M = 29.03$ , 95% CI [28.62, 29.43]) and the FACES measures showed high internal consistency within this sample ( $\alpha = 0.87$ ).

### 3.2 Caregiver-child agreement

#### 3.2.1 Overall agreement rates

To determine general trends in symptom severity agreement between caregivers and their loved one at intake, Lin’s concordance correlation coefficient was calculated between client self-reported depression symptoms and caregiver-reported depression symptoms. A statistically significant, moderate positive correlation was found between the two measures ( $r(770) = 0.215$ ,  $p < 0.05$ , 95% CI [0.1714, 0.2585]). Similar, though slightly lower, results were found between client self-reported anxiety symptoms and caregiver-reported anxiety symptoms ( $r(761) = 0.124$ ,  $p < 0.05$ , 95% CI [0.1714, 0.2585]). Note

that Lin's concordance correlation coefficient contains a penalty that scales with the difference in mean responses. As these ratings are on different scales, this provides an extremely conservative lower bound on agreement. These results indicate that while clients and caregivers have a moderate agreement level regarding symptom severity for both anxiety and depression, they are not a one-to-one match and some discrepancies are present. Figures 1, 2 present Bland–Altman plots for depression and anxiety, respectively, to probe these discrepancies for heterogeneity across symptom severity. In both cases, we observe more variation in agreement near mean levels of symptom severity, and more agreement at more extreme (high or low) levels of symptom severity.

Presenting agreement in a slightly different way, we first find  $z$ -scores of both self-reported and caregiver-reported depression and anxiety scores and compute the difference. Figures 3, 4 show a histogram of these differences for depression and anxiety, respectively. At intake, the average difference in scaled depression scores was  $-0.02$  indicating almost perfect agreement between clients and caregivers on average with a slight tip towards clients rating depression symptoms as more severe. Similarly, the average difference in scaled anxiety symptoms at intake was  $-0.001$ ,

indicating clients and caregivers were using corresponding locations on their respective scales.

### 3.2.2 Agreement by demographic variables

Due to the number of follow-up regressions and subgroup analyses conducted, these analyses should be considered exploratory and interpreted with caution rather than as confirmatory tests.

A linear regression was run to determine the relationship between level of symptom agreement and client age. Age was a significant predictor of depression symptom severity agreement  $b = 0.039$ ,  $t(770) = 2.79$ ,  $p = 0.005$ , indicating that every one-year increase in age is associated with an increase in depression agreement scores of 0.039 points. In other words, the older a client was at the time of data collection, the more the score shifted to caregivers scoring the client's depression symptoms as worse. While the overall model was significant, it only predicted 1% of the variance in agreement (multiple  $R^2 = 0.010$ ; adjusted  $R^2 = 0.009$ ). Similarly, age was a significant predictor of anxiety symptom severity agreement  $b = 0.036$ ,  $t(761) = 2.603$ ,  $p = 0.009$ , indicating that for every 1-year increase in age, anxiety agreement scores increased by 0.036 points. Again, this model only explained approximately 1% of the variance in agreement

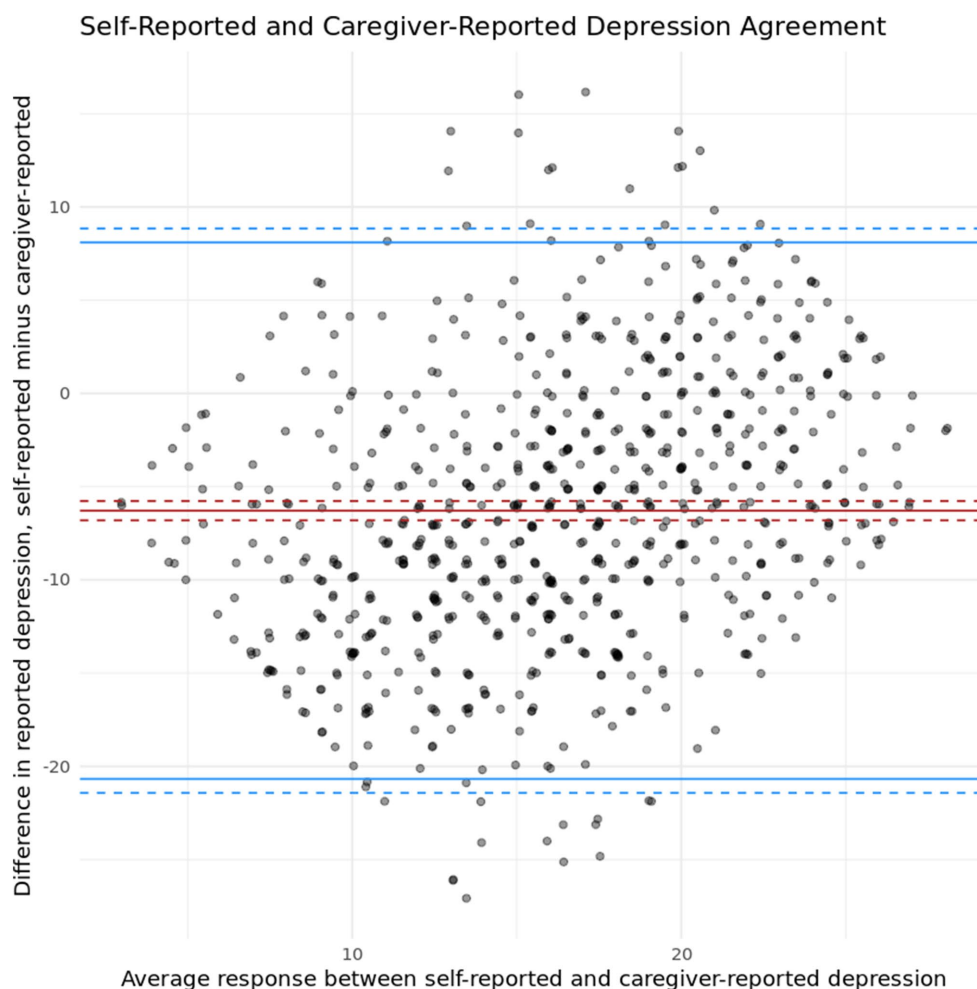


FIGURE 1  
Bland–Altman plot showing agreement between self-reported and caregiver-reported depression symptoms.

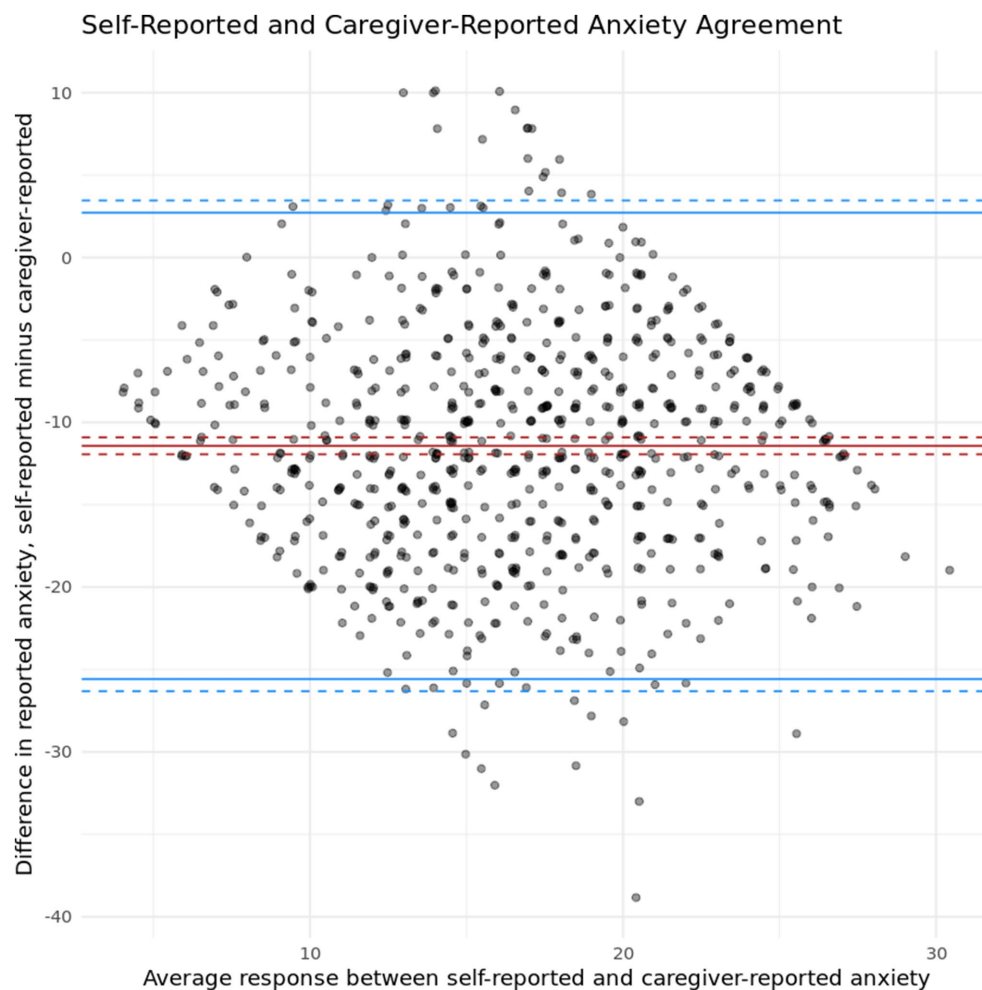


FIGURE 2  
Bland–Altman plot showing agreement between self-reported and caregiver-reported anxiety symptoms.

scores. While both cases were significant, effect sizes indicate that the significance may only be a statistical artifact of sample size, and not a clinically significant finding.

Three ANOVAs were run to assess the patterns of symptom agreement based on client race, gender, and neurodivergent identity. Due to the repeated testing, a Bonferroni correction was applied to each ANOVA test, with a new significance level of  $p = 0.0167$ .

The first ANOVA looked at patterns of depression agreement based on a client's reported gender. A one-way ANOVA indicated a significant main effect of gender on depression symptom agreement levels  $F(2, 613) = 13.45$ ,  $p < 0.001$  and this finding remained significant after the Bonferroni adjustment. *Post-hoc* comparisons using a Tukey HSD test revealed significant differences between all three gender categories with the largest difference in mean agreement scores found between Male clients and Non-Binary/Other clients ( $M = 0.229$  vs.  $M = -0.502$ ). Similarly, a one-way ANOVA indicated a significant main effect of gender on anxiety symptom agreement levels  $F(2, 606) = 4.54$ ,  $p = 0.011$ . This finding also remained significant after the Bonferroni correction. A *post-hoc* Tukey HSD test revealed that significant differences in anxiety agreement levels were found between male and non-binary clients, and female and non-binary clients, but not between male and female clients.

Next, we investigated the relationship between a client's neurodivergent identity and levels of symptom agreement. Based on one-way ANOVAs, depression symptom agreement scores were not associated with a client's neurodivergent identity ( $F(4, 580) = 0.67$ ,  $p = 0.611$ ). Anxiety symptom agreement was significantly associated with a client's neurodivergent identity  $F(4, 573) = 2.92$ ,  $p = 0.021$ , however this model only predicted approximately 2% of the variance, and therefore while statistically detectable, was not clinically meaningful.

Finally, we assessed the impact of client race on symptom agreement levels. Based on one-way ANOVAs, neither depression nor anxiety symptom agreement scores were associated with client racial categories ( $F(2, 587) = 0.35$ ,  $p = 0.71$ ;  $F(2, 582) = 0.58$ ,  $p = 0.56$ ).

### 3.2.3 Family functioning and symptom agreement

To determine if the level of family functioning impacts symptom agreement, regressions were performed looking at overall symptom agreement for both anxiety and depression, as well as magnitude of disagreement (regardless of direction), and sign of disagreement (regardless of magnitude). A simple linear regression on family functioning and depression agreement indicated a statistically significant relationship between the two variables ( $\beta = -0.036$ ,

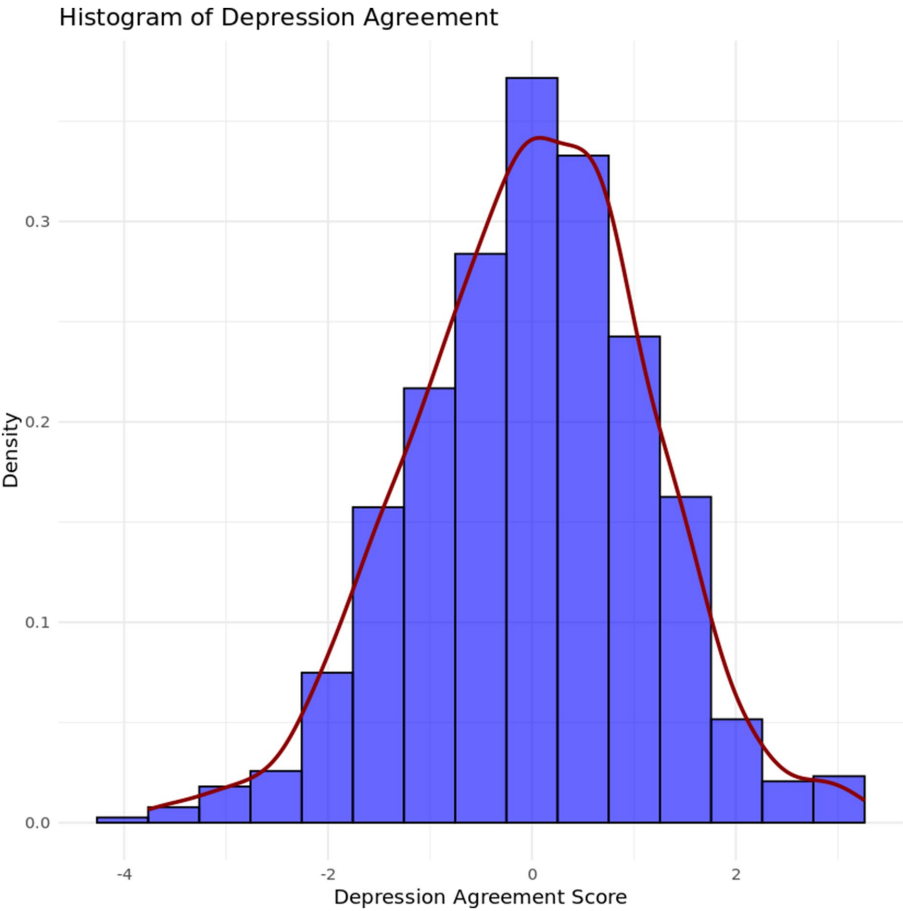


FIGURE 3  
Histogram of agreement scores on depression symptoms between caregivers and clients.

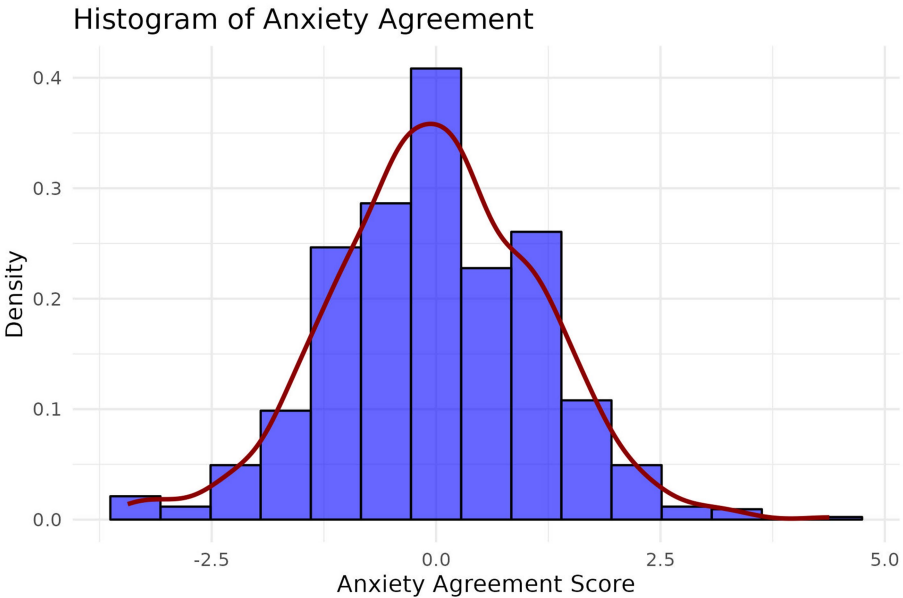


FIGURE 4  
Histogram of agreement scores on anxiety symptoms between caregivers and clients.



SE = 0.008,  $t(603) = -4.42$ ,  $p < 0.001$ ), where every additional point of family functioning predicts a 0.036 point reduction in agreement score. This linear model explains 3.1% of variance in agreement scores, meaning family functioning had a small but significant effect on depression agreement.

However, upon running a linear regression on the magnitude of disagreement and family functioning, and a logistic regression on the direction of disagreement (client-higher vs. caregiver-higher), family functioning scores were only significantly associated with the direction of disagreement  $\chi^2(1) = 6.17$ ,  $p = 0.013$ . The regression coefficient for family functioning was negative ( $\beta = -0.036$ , SE = 0.015,  $Z = -2.46$ ,  $p = 0.014$ ), indicating that as family functioning increases, the likelihood of caregivers rating their loved-one's depression symptoms higher than the client themselves decreased.

Conversely, family functioning was not significantly associated with overall anxiety symptoms agreement ( $F(1, 591) = 0.43$ ,  $p = 0.511$ ), magnitude of symptom agreement ( $F(1, 591) = 0.49$ ,  $p = 0.484$ ), or direction of symptom agreement ( $\chi^2(1) = 0.96$ ,  $p = 0.328$ ).

### 3.2.4 Caregiver well-being at intake and symptom agreement

To explore whether caregiver well-being was associated with agreement between caregiver- and child-reported symptoms, two linear regressions were conducted using the WHO-5 Well-Being Index as a predictor. Higher caregiver well-being was very weakly associated with lower agreement on depression symptoms, ( $b = -0.004$ ,  $p = 0.03$ , 95% CI  $[-0.008, -0.000]$ ), though the model only explained 1% of the variance ( $R^2 = 0.01$ ). No significant association was found between caregiver well-being and anxiety agreement,  $b = -0.002$ ,  $p = 0.34$ . These findings suggest that caregiver well-being may play a limited role in shaping agreement on symptom severity, specifically for depression, however it is not a strong influence.

### 3.2.5 Symptom agreement and treatment outcomes

Of the 763 caregiver-youth dyads included at intake, 221 (28.9%) also completed a discharge survey and were therefore included in the outcomes analyses. This analytic sample provides a meaningful window into treatment change, but the reduced proportion of available discharge data should be considered when interpreting results.

For both depression and anxiety, we fit a linear regression on the change in standardized differences in self-reported depression/anxiety on standardized measures of caregiver and self-reported depression/anxiety at intake. Figure 5 shows coefficients with 95% confidence intervals constructed from heteroskedasticity-robust standard errors from these regressions. In the left panel, we see that only the coefficient on self-reported anxiety is statistically distinguishable from zero ( $\beta = -0.636$ , 95% CI  $[-0.755, -0.517]$ ), indicating that individuals with higher levels of self-reported anxiety saw greater reductions in their self-reported anxiety levels post-treatment. For depression, both caregiver-reported depression levels ( $\beta = -0.137$ , 95% CI  $[-0.260, -0.013]$ ) and self-reported depression levels ( $\beta = -0.535$ , 95% CI  $[-0.662, -0.408]$ ) were statistically significant, though the predictive utility of the caregiver-reported level is smaller in magnitude than the self-reported level. We observe, similarly to anxiety, higher levels of self-reported depression are associated with

higher reductions in self-reported depression post-treatment, and the magnitude of this effect is similar across both constructs. In both cases, despite there being some degree of disagreement between caregiver-reported and self-reported levels, the additional variance in the caregiver-reported levels explains little additional variance in the change in the self-reported outcome.

#### 3.2.5.1 Depression

One concern with the previous results is a potential regression-to-the mean effect. Figure 6 shows changes in self-reported depression plotted against self-reported depression at intake. In general, we see that self-reported depression decreases for most participants, regardless of their position on the scale, though the limits of the original scale impose limits on the observable variation in the outcome.

To investigate heterogeneity as a function of self-reported and caregiver-reported agreement, we look at a regression of average depression level at discharge on symptom agreement at intake. Results indicated a statistically significant relationship between depression symptom agreement at intake and depression symptoms at discharge ( $F(1, 219) = 27.15$ ,  $p < 0.001$ ), and the model predicted 11.03% of the variance in depression symptoms at discharge. Additionally, the regression coefficient for depression agreement was significant ( $B = -1.64$ , SE = 0.32,  $t(219) = -5.21$ ,  $p < 0.001$ ) indicating that the higher a caregiver rates their child's depression symptoms in relation to the client's self-assessment, the less severe depression symptoms were at discharge, as seen in Figure 7.

While higher caregiver rated scores at intake were associated with lower client-reported depression symptoms at discharge, agreement level at intake was not significantly associated with likelihood of routine discharge. A logistic regression was run to determine whether overall depression agreement score, absolute depression agreement score (regardless of direction of disagreement), or the direction of agreement score (regardless of magnitude) were significant predictors of routine discharge, and found that none of the independent variables significantly predicted discharge type ( $\chi^2(3, N = 709) = 1.47$ ,  $p = 0.689$ ). Additionally, just caregiver ratings of client depression at intake did not have a significant association with likelihood of routine discharge ( $t(759.37) = 1.84$ ,  $p = 0.066$ ), indicating that once a client is in treatment, caregiver perception of symptoms does not seem to have an impact on treatment engagement.

Based on prior tests indicating that age, gender, and level of family functioning has an impact on depression symptom agreement, a final multiple linear regression was run to determine if age, gender, or family functioning have a mediating effect on depression symptoms at discharge. The overall model was statistically significant  $F(5, 188) = 4.79$ ,  $p = 0.00038$ , however it only accounted for ~11.3% of variance in depression scores. Depression symptom agreement at intake was the strongest predictor of depression scores at discharge, with caregivers scoring depression symptoms as more severe predicting lower post-treatment depression symptoms ( $\beta = -1.57$ ,  $p < 0.001$ ). Age, gender, and family functioning were not significant moderators in this model.

#### 3.2.5.2 Anxiety

As with depression, we first look for a regression-to-the mean effect. Figure 8 shows changes in self-reported anxiety plotted against self-reported anxiety at intake. In general, we again see that

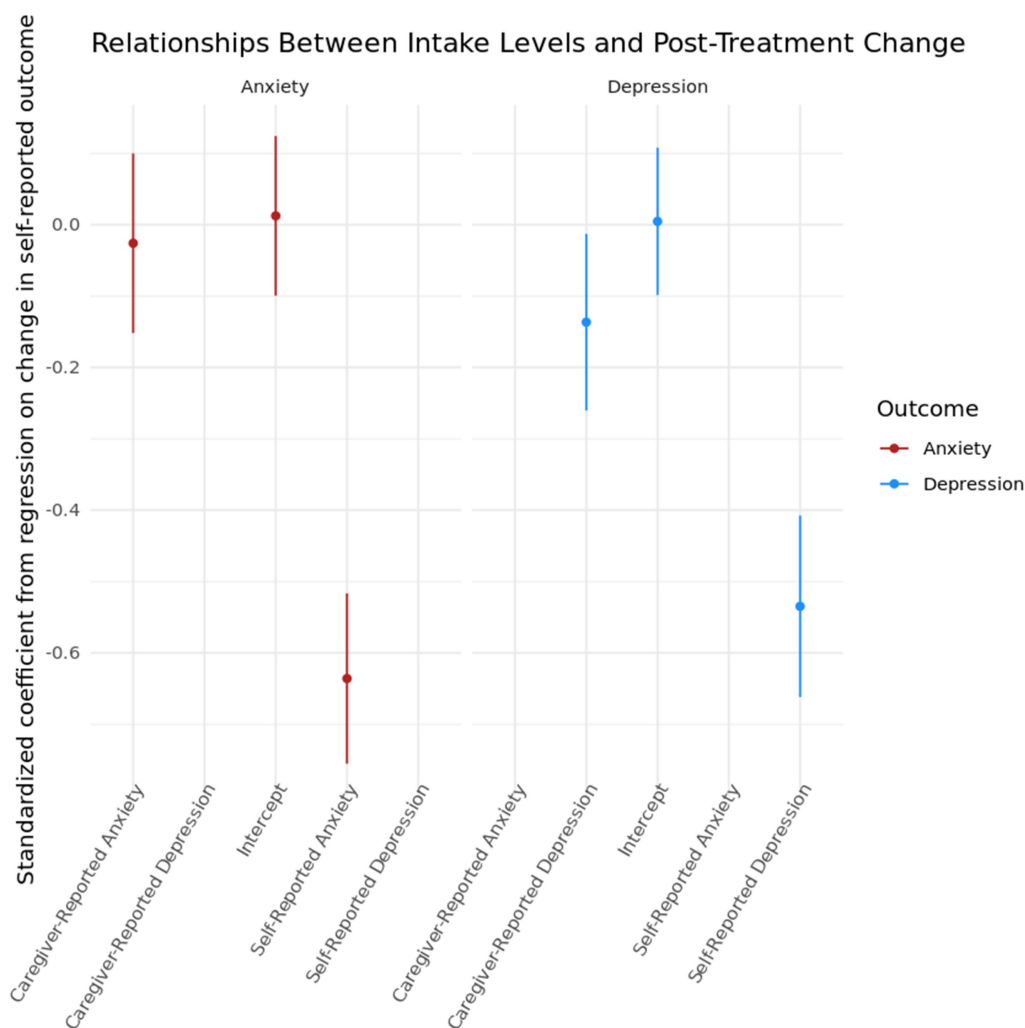


FIGURE 5

Coefficients from regressions of change in anxiety and depression on baseline levels of anxiety and depression. Anxiety and depression models were fit separately and are shown with 95% confidence intervals using heteroskedasticity-robust standard error corrections.

self-reported anxiety decreases for most participants, regardless of their starting position on the scale. The finite range of values achievable on the original scale still impose limits on the observable variation in the outcome.

The first analysis for anxiety symptoms investigated the relationship between level of agreement on anxiety symptoms at intake and anxiety score at discharge. A linear regression showed a significant relationship between anxiety agreement at intake and symptom acuity at discharge  $b = -1.14$ ,  $t(212) = -4.06$ ,  $p < 0.001$ , indicating that for every 1-point increase in anxiety agreement score at intake, GAD-7 scores at discharge fell by  $-1.14$  points, as seen in Figure 9.

Similar to the depression agreement scores, a logistic regression was run to evaluate the effect of anxiety agreement scores on likelihood of completing treatment. Independent variables were overall anxiety agreement, absolute magnitude of anxiety agreement (regardless of the direction), and the direction of anxiety agreement—whether caregivers ranked symptoms higher, or the client did (regardless of magnitude). Results of the logistic regression indicated

that none of the independent variables had a significant association with treatment completion ( $\chi^2(3, N = 647) = 1.61$ ,  $p = 0.658$ ). Additionally, caregiver ratings of anxiety symptoms at intake on their own did not have a significant association with routine discharge rates, indicating that the conclusion drawn for the depression scores remains consistent: once a client is in treatment, their agreement levels do not seem to have an impact on treatment completion.

Based on prior tests indicating that age and gender were both significantly associated with level of anxiety symptom agreement between clients and their caregivers, a multiple linear regression was run to control for both age and gender. The overall model was statistically significant ( $F(4,199) = 4.71$ ,  $p = 0.001$ ), and results indicate that when controlling for these demographics, anxiety symptom agreement at intake was the only significant predictor of anxiety symptoms at discharge ( $\beta = -1.08$ ,  $p < 0.001$ ). However, this model only accounted for  $\sim 8.7\%$  of the variance in scores at discharge ( $R^2 = 0.087$ , adjusted  $R^2 = 0.068$ ), making it a weak predictor of anxiety symptoms at discharge.

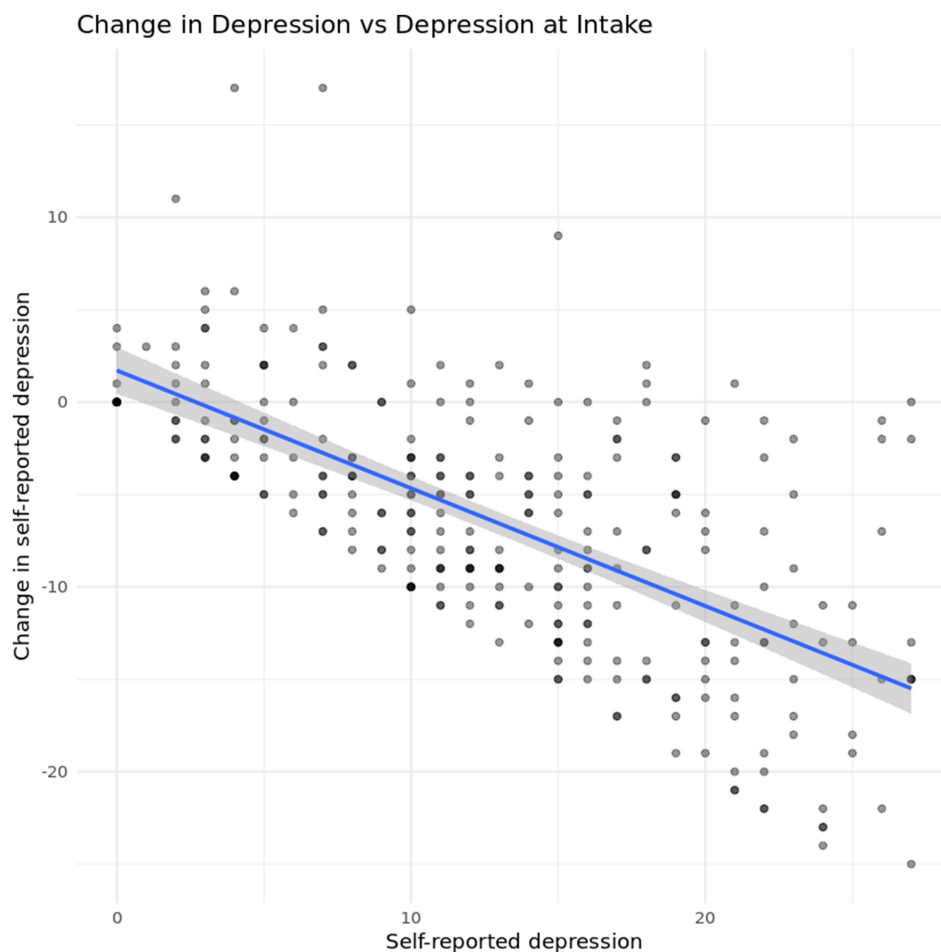


FIGURE 6  
Change in depression vs. self-reported depression at intake.

## 4 Discussion

The aim of this study was to examine client self-reported symptoms of depression and anxiety within a virtual intensive outpatient program in comparison to caregiver-reported assessments, exploring potential moderating variables and evaluating the relationship between agreement and treatment outcomes. Overall, we found modest agreement levels between clients and their caregivers when evaluating depression and anxiety symptoms. This pattern is consistent with prior research indicating only modest caregiver–child concordance on internalizing symptoms, likely reflecting differences in symptom observability and self-insight (De Los Reyes and Ohannessian, 2016; Korelitz and Garber, 2016). However, these average agreement levels concealed meaningful subgroup differences that have implications for clinical practice and future research.

In addition to the primary analyses, we conducted several exploratory regressions to gauge which demographic or family variables might warrant inclusion as covariates in outcome models. Because of the number of comparisons and lack of a preregistered confirmatory plan, these results were not corrected for multiplicity and should be interpreted as hypothesis-generating rather than confirmatory. Notably, gender emerged as a key moderator of agreement. Non-binary youth showed the largest level of discrepancy

with their caregivers. Specifically, non-binary participants were more likely to rate their symptoms as more severe than their caregivers did. This finding echoes existing literature that highlights the unique mental health burdens faced by gender-diverse youth (Eisenberg et al., 2017). Additionally, age was significantly associated with level of disagreement with older clients reporting fewer symptoms relative to their caregivers' ratings, for both anxiety and depression. However, age explained only 1% of the variance and is therefore unlikely to be meaningful. These exploratory findings should be interpreted as small effects that may reflect only statistical power rather than substantive influence.

While race and neurodivergent identity were investigated as potential moderators of agreement levels, neither category showed any significant differences. While no significant associates were found for race or neurodivergent identity, these null results should be interpreted with caution. The sample was relatively homogeneous with respect to race/ethnicity and missing data was substantial. Similarly the neurodivergent identity categories were broad, which may have obscured more specific patterns.

Results also indicated that caregiver ratings of family functioning were significantly associated with disagreement levels on depression symptoms. Specifically, in families reporting higher functioning, caregivers were less likely to overestimate client symptom severity

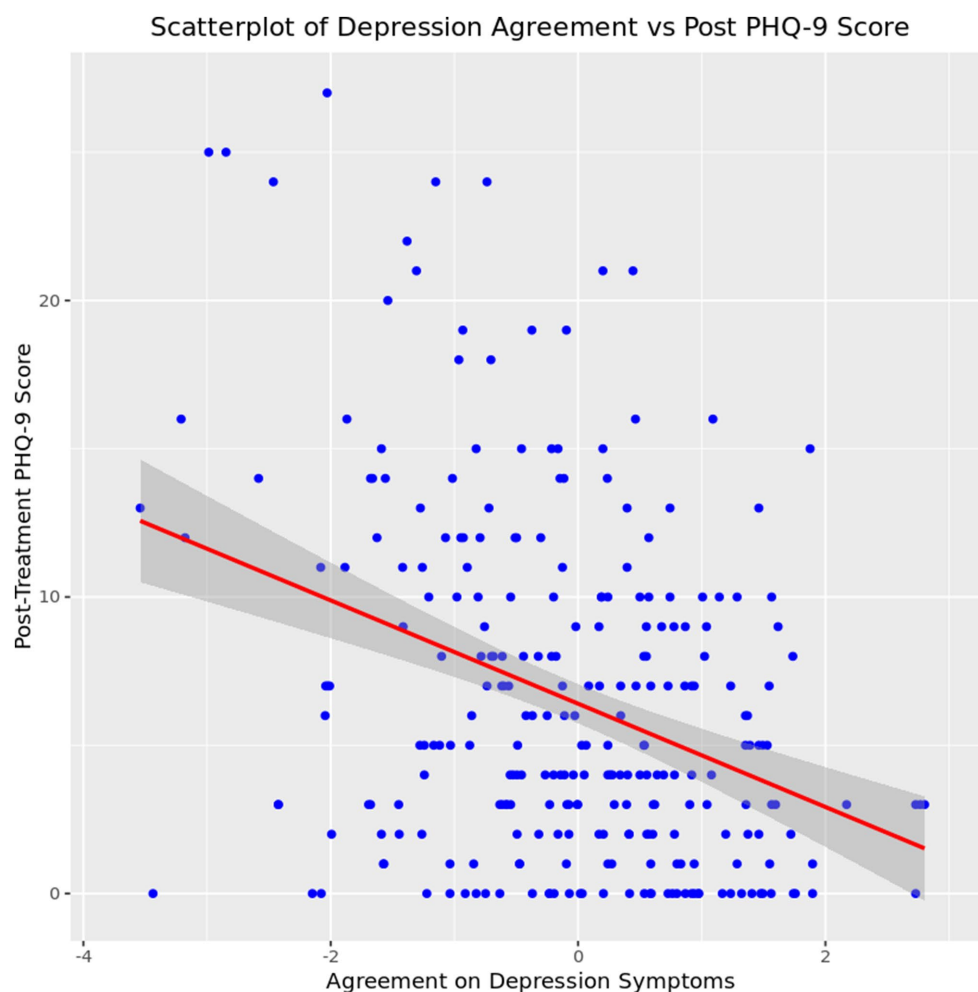


FIGURE 7  
Relationship between caregiver-youth agreement on depression symptoms at intake and youth depression severity at discharge.

compared to the client's own report. This may suggest that higher levels of family functioning are associated with greater agreement between caregiver and client perceptions of symptom severity. This aligns with family systems theory, which emphasizes that higher-functioning families often display stronger communication and shared understanding of stressors (Erdem and Safi, 2018). Agreement in symptom perception may therefore reflect healthier relational dynamics rather than only observational accuracy. However, this relationship did not hold for anxiety symptoms, suggesting that internalizing symptoms may vary in how observable or interpretable they are within family systems.

Contrary to prior literature that found that caregiver distress can bias caregival proxy ratings of child mental health (Ehrlich et al., 2011), no significant relationship between caregiver well-being and level of disagreement was found within this sample. This may be due to the fact that there was not much deviation in well-being scores between caregivers, and therefore the spread was not comprehensive enough to identify a relationship. Alternatively, caregiver well-being, while important to consider in other aspects of treatment, may not sway the perception of their loved-ones symptoms. This is an area that warrants further exploration. Given the limited variance in caregiver well-being scores, it is

possible that this study lacked the sensitivity to detect meaningful effects.

Finally, we examined the relationship between caregiver-client agreement at intake and treatment outcomes. In general, clients with higher self-reported depression and anxiety had better outcomes (operationalized as larger reductions in self-reported depression and anxiety), and after controlling for self-reported depression and anxiety at intake, caregiver-reported levels of depression and anxiety explained little additional variance in the outcome. Prior research has similarly found that client self-reports often show stronger associations with treatment response than parent ratings, particularly for internalizing symptoms (Hawley and Weisz, 2003; Ferdinand et al., 2006).

Interestingly, dyads in which caregivers scored client symptoms as more severe were associated with lower severity symptoms at discharge as measured by the PHQ-9 and GAD-7 for both anxiety and depression. At first glance, this pattern may appear counterintuitive, yet it aligns with several possible mechanisms. One plausible explanation is caregiver engagement and advocacy: When caregivers perceive distress as high, they may be more motivated to advocate for services, encourage adherence, and partner with providers, thereby indirectly facilitating client improvement. Previous work suggests that greater caregiver concern or involvement can enhance treatment

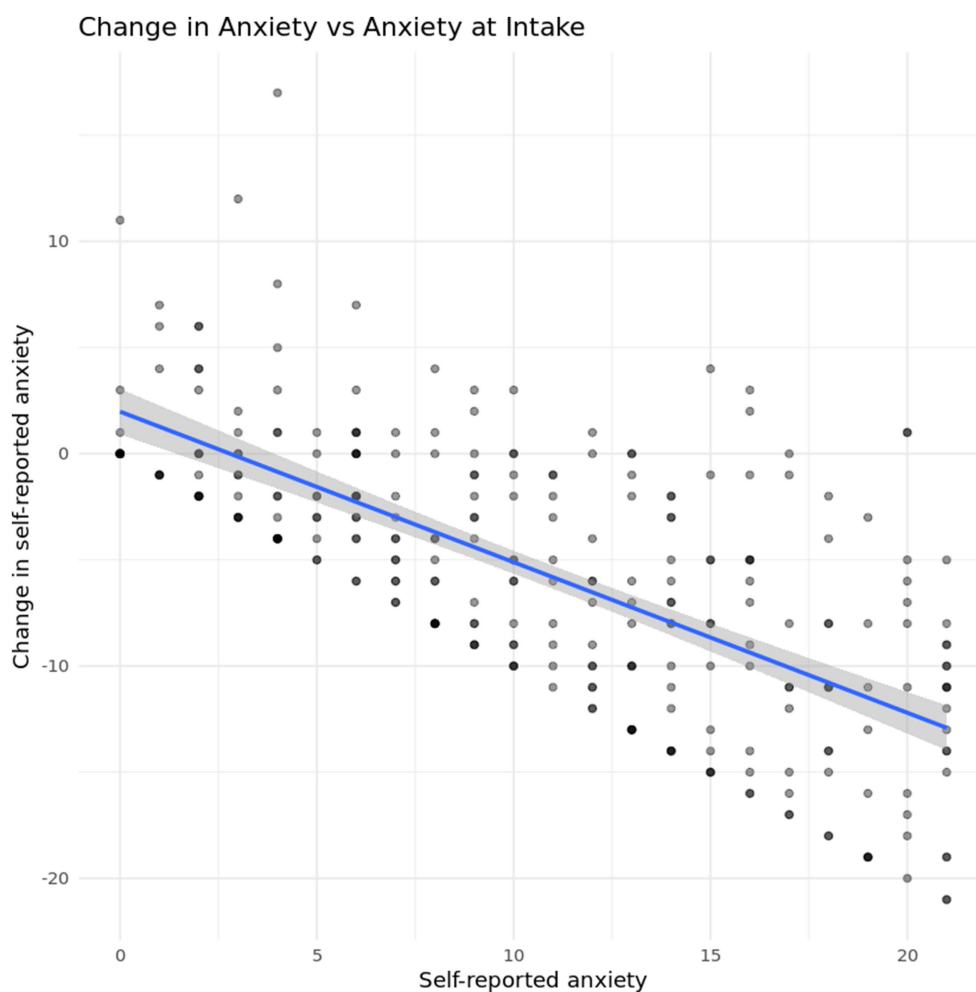


FIGURE 8  
Change in depression vs. self-reported depression at intake.

adherence and engagement, leading to better clinical improvement (Hawley and Weisz, 2005). From a family systems perspective, this could reflect a dynamic in which caregivers' acknowledgement of distress validates the client's experience, increasing motivation to engage in treatment (Erdem and Safi, 2018). Additionally, caregiver perception of high distress may allow the client to feel safer expressing and exploring emotion due to the expectation of emotion already being present (De Los Reyes et al., 2013; Hawley and Weisz, 2003). This exploration of distress may then result in better understanding of symptoms and willingness to change.

Another interpretation is that heightened caregiver concern may shape treatment processes themselves—for example, prompting earlier recognition of problems, increasing caregiver participation in family sessions, or leading to more consistent use of the caregiver portal. Although such variables were not directly measured in the current study, future work should examine whether caregiver-rated severity functions as a proxy indicator for family engagement.

Despite the improved outcomes when caregivers scored symptom severity as higher, agreement level itself was not significantly associated with the likelihood of routine discharge for either depression or anxiety. This null result is noteworthy as it suggests that while caregiver perceptions may influence the severity of symptoms

and the trajectory of symptom change, they do not appear to influence whether clients complete treatment in a virtual intensive outpatient program (IOP) setting. One possible explanation is that the structural supports inherent to IOP reduce the impact of informant agreement in predicting treatment continuation. Alternatively, agreement may be more critical at intake, influencing early engagement and symptom recognition, but less influential once clients are admitted into care.

These findings underscore the importance of assessing both client and caregiver symptom reports at intake. Specifically, caregiver perceptions of higher severity were associated with greater symptom improvement, suggesting that engaging caregivers early and validating their concerns may promote treatment buy-in and adherence. Prior work has emphasized that explicitly addressing informant discrepancies in session can improve therapeutic alliance and goal consensus (De Los Reyes, 2011; Hawley and Weisz, 2005). However, discrepancies, particularly among non-binary youth, highlight the need for clinicians to facilitate structured conversations about symptom experience to foster mutual understanding and goal alignment. The counterintuitive nature of this finding highlights the complex role of caregiver perception. It may be that caregivers who recognize and emphasize their client's distress are more likely to



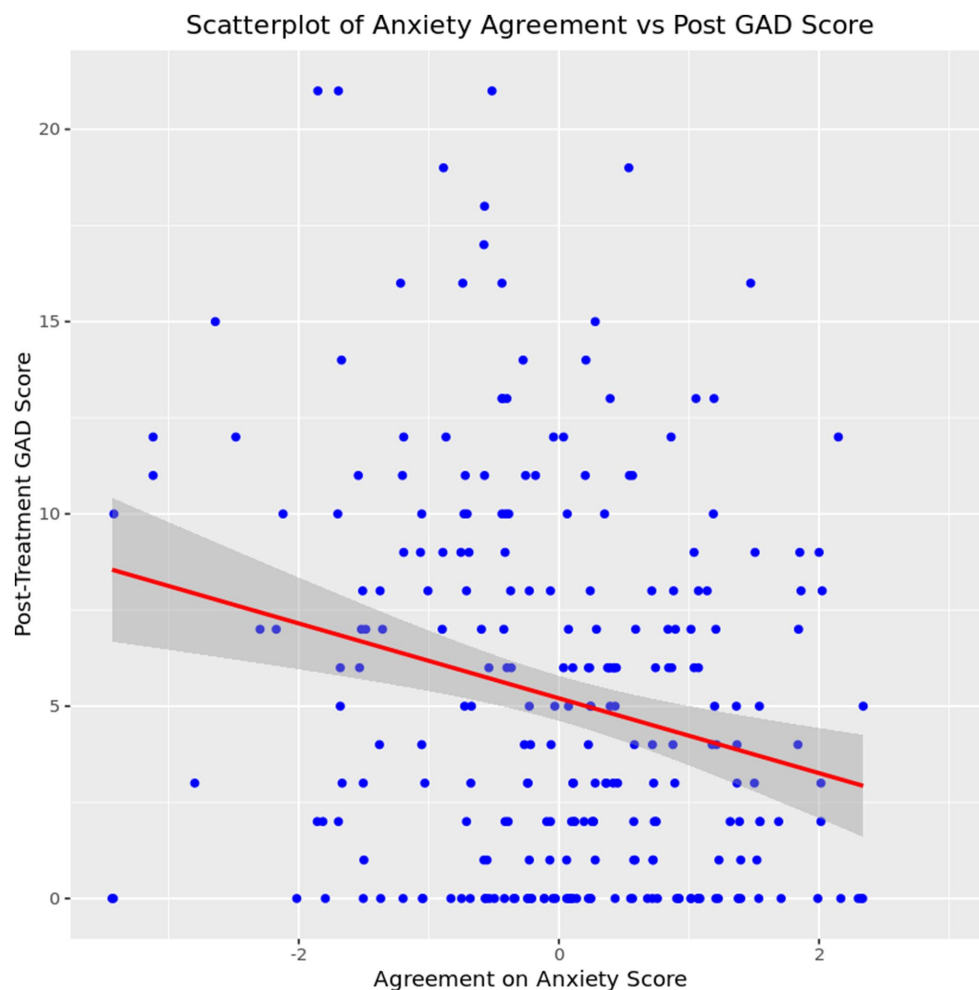


FIGURE 9  
Relationship between caregiver-youth agreement on anxiety symptoms at intake and youth anxiety severity at discharge.

advocate for services, encourage adherence, and partner with providers, thereby indirectly facilitating improvement.

## 5 Limitations and future research

Due to the amount of missing and skewed demographic data (which leaned heavily toward white, neurotypical, female participants) this dataset has limited generalizability, and conclusions should not be applied to all individuals with mental health concerns. Furthermore, several demographic and outcome variables were collapsed into broader categories to reduce small cell sizes and ensure statistical stability. For example, gender identities beyond male and female were combined into an “Other/Non-binary” category, and racial identities outside of White and Black were combined into “Other.” Similarly, discharge reasons were collapsed into “routine” versus “non-routine” to address small subgroups such as administrative discharge or referral to higher levels of care. While this approach allowed analyses to proceed without risk of overfitting or disclosure, it also limited the granularity with which subgroup differences and discharge patterns could be interpreted. Future work with larger and more diverse samples should retain finer categories and apply analytic methods

designed to handle sparse cells (e.g., penalized regression or Bayesian models) to better capture heterogeneity across groups. Future research should focus on gathering a representative sample across demographics to assess how these results may differ depending on individual identity.

This study may also have been swayed by selection bias, where families willing to participate in caregiver surveys may differ systematically compared to those who did not fill out the survey. Only a small proportion of families completed the caregiver survey, and among those, fewer than half had discharge data available, raising the possibility of non-random attrition influencing outcome estimates. This was partially demonstrated by the skewed distribution in demographics that did not align with overall population demographics within the treatment center. Future research should find ways to require the caregiver survey, or otherwise randomly select individuals to complete the caregiver survey to reduce the risk of selection bias.

Caregiver and youth surveys were collected on slightly different schedules and with different recall windows (7-day vs. 14-day). These decisions allowed analyses to proceed without risk of overfitting or disclosure but limited the granularity with which subgroup differences, discharge patterns, and agreement estimates could be interpreted. Future work should aim to retain finer categories, synchronize data

collection more tightly, and use analytic methods designed to handle sparse cells (e.g., penalized regression or Bayesian models) to better capture heterogeneity across groups. It is also important to note that agreement estimates may be influenced by the use of different instruments and informant types. Clients and caregivers completed distinct measures, and while z-standardization facilitated comparability, differences in response format, interpretation, and construct coverage may constrain the accuracy of cross-informant agreement.

Additionally, this study was limited to individuals who were already admitted to high acuity mental healthcare, which restricts the generalizability of this outcome. The question that needs to be answered may not be how agreement and disagreement impact treatment outcomes, but how they impact access to care in the first place. Further research should examine client and caregiver symptom agreement in the general population and track individuals longitudinally to determine who ends up in treatment and who does not.

Finally, as several authors are employed by the treatment program studied, there is a potential for bias in the interpretation. However, the outcomes reported were based on client self-report rather than clinician or staff ratings, reducing the risk that author employment influenced results.

## Data availability statement

For the sake of patient confidentiality, the dataset analyzed for this study is not publicly available. Please contact the authors to request access to the dataset.

## Ethics statement

The studies involving humans were approved by North Star Non-Profit Research Ethics Review Board. The studies were conducted in accordance with the local legislation and institutional requirements. Written informed consent for participation was not required from the participants or the participants' legal guardians/next of kin because the study was approved as secondary research under Common Rule 45 CFR 46.104(d)(4)(ii). Written informed consent was not obtained from the individual(s) for the publication of any potentially identifiable images or data included in this article because this study was approved as secondary research under Common Rule 45 CFR 46.104(d)(4)(ii).

## References

- Angley, M., Divney, A., Magriples, U., and Kershaw, T. (2015). Social support, family functioning and parenting competence in adolescent parents. *Matern. Child Health J.* 19, 67–73. doi: 10.1007/s10995-014-1496-x
- Bajeux, E., Klemanski, D. H., Husky, M., Leray, E., Chan Chee, C., Shojaei, T., et al. (2018). Factors associated with parent-child discrepancies in reports of mental health disorders in young children. *Child Psychiatry Hum. Dev.* 49, 1003–1010. doi: 10.1007/s10578-018-0815-7
- Bland, J. M., and Altman, D. G. (1986). Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 327, 307–310. doi: 10.1016/S0140-6736(86)90837-8
- Bosworth, T. (2019). Early treatment of anxiety: left untreated, anxiety disorders often become more complicated over time. *Pediatric News* 53:1. Available at: [https://link.gale.com/apps/doc/A583381186/AONE?u=nysl\\_oweb&sid=googleScholar&xid=60c03f30](https://link.gale.com/apps/doc/A583381186/AONE?u=nysl_oweb&sid=googleScholar&xid=60c03f30)
- Broderick, J. E., DeWitt, E. M., Rothrock, N., Crane, P. K., and Forrest, C. B. (2013). Advances in patient-reported outcomes: the NIH PROMIS® measures. *Generat. Evid. Methods Improve Patient Outcomes* 1:1015. doi: 10.13063/2327-9214.1015
- Center on the Developing Child (2016). From best practices to breakthrough impacts: a science-based approach to building a more promising future for young children and families. Cambridge: Harvard University.
- Centers for Disease Control and Prevention. (2025). Data and statistics on children's mental health. Available online at: <https://www.cdc.gov/children-mental-health/data-research/index.html> (Accessed April 2025).
- Costantini, S. (2025). How do mental health treatment delays impact long-term mortality? *Am. Econ. Rev.* 115, 1672–1707. doi: 10.1257/aer.20240226

## Author contributions

EK: Data curation, Methodology, Visualization, Conceptualization, Writing – review & editing, Project administration, Formal analysis, Writing – original draft. SH: Writing – original draft. AB: Writing – original draft. KK: Writing – review & editing, Validation, Formal analysis. JK: Writing – review & editing, Supervision. HH: Writing – review & editing. IZ: Writing – review & editing. SS: Writing – review & editing. CF: Supervision, Resources, Writing – review & editing.

## Funding

The author(s) declare that no financial support was received for the research and/or publication of this article.

## Conflict of interest

CF is the co-founder and Chief Clinical Officer of Charlie Health. EK, SS, IZ, HH, and JK are employees of Charlie Health. SH and AB are clinical practicum students with Charlie Health. KK is an independent contractor.

## Generative AI statement

The authors declare that no Gen AI was used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

- Dang, J., King, K. M., and Inzlicht, M. (2020). Why are self-report and Behavioral measures weakly correlated? *Trends Cogn. Sci.* 24, 267–269. doi: 10.1016/j.tics.2020.01.007
- De Los Reyes, A. (2011). Introduction to the special section: more than measurement error: discovering meaning behind informant discrepancies in clinical assessments of children and adolescents. *J. Clin. Child Adolesc. Psychol.* 40, 1–9. doi: 10.1080/15374416.2011.533405
- De Los Reyes, A., and Kazdin, A. E. (2005). Informant discrepancies in the assessment of childhood psychopathology: a critical review, theoretical framework, and recommendations for further study. *Psychol. Bull.* 131, 483–509. doi: 10.1037/0033-2909.131.4.483
- De Los Reyes, A., and Ohannessian, C. M. (2016). Introduction to the special issue: discrepancies in adolescent–parent perceptions of the family and adolescent adjustment. *J. Youth Adolesc.* 45, 1957–1972. doi: 10.1007/s10964-016-0533-z
- De Los Reyes, A., Talbott, E., Power, T., Michel, J., Cook, C. R., Racz, S. J., et al. (2021). The needs-to-goals gap: how informant discrepancies in youth mental health assessments impact service delivery. *Clin. Psychol. Rev.* 92:102114. doi: 10.1016/j.cpr.2021.102114
- De Los Reyes, A., Thomas, S. A., Goodman, K. L., and Kundey, S. M. (2013). Principles underlying the use of multiple informants' reports. *Annu. Rev. Clin. Psychol.* 9, 123–149. doi: 10.1146/annurev-clinpsy-050212-185617
- Ehrlich, K. B., Cassidy, J., and Dykas, M. J. (2011). Reporter discrepancies among parents, adolescents, and peers: adolescent attachment and informant depressive symptoms as explanatory factors. *Child Dev.* 82, 999–1012. doi: 10.1111/j.1467-8624.2010.01530.x
- Eisenberg, M. E., Gower, A. L., McMorris, B. J., Rider, G. N., Shea, G., and Coleman, E. (2017). Risk and protective factors in the lives of transgender/gender nonconforming adolescents. *J. Adolesc. Health* 61, 521–526. doi: 10.1016/j.jadohealth.2017.04.014
- Erdem, G., and Safi, O. A. (2018). The cultural lens approach to Bowen family systems theory: contributions of family change theory. *J. Fam. Theory Rev.* 10, 469–483. doi: 10.1111/jftr.12258
- Ferdinand, R. F., Van Der Ende, J., and Verhulst, F. C. (2006). Prognostic value of parent–adolescent disagreement in a referred sample. *Eur. Child Adolesc. Psychiatry* 15, 156–162. doi: 10.1007/s00787-005-0518-z
- Gustavson, K., Knudsen, A. K., Nesvåg, R., Knudsen, G. P., Vollset, S. E., and Reichborn-Kjennerud, T. (2018). Prevalence and stability of mental disorders among young adults: findings from a longitudinal study. *BMC psychiatry*, 18:65. doi: 10.1186/s12888-018-1647-5
- Harrell, F. E. Jr. (2023). Rms: regression modeling strategies (version) [computer software]. Available online at: <https://CRAN.R-project.org/package=rms> (Accessed April 2025).
- Hawley, K. M., and Weisz, J. R. (2003). Child, parent, and therapist (dis)agreement on target problems in outpatient therapy: the therapist's dilemma and its implications. *J. Consult. Clin. Psychol.* 71, 62–70. doi: 10.1037/0022-006X.71.1.62
- Hawley, K. M., and Weisz, J. R. (2005). Youth versus parent working alliance in usual clinical care: distinctive associations with retention, satisfaction, and treatment outcome. *J. Clin. Child Adolesc. Psychol.* 34, 117–128. doi: 10.1207/s15374424jccp3401\_11
- Holden, E., and Kobayashi-Wood, H. (2025). Adverse experiences of women with undiagnosed ADHD and the invaluable role of diagnosis. *Sci. Rep.* 15:20945. doi: 10.1038/s41598-025-04782-y
- Hughes, K., Bellis, M. A., Hardcastle, K. A., Sethi, D., Butchart, A., Mikton, C., et al. (2017). The effect of multiple adverse childhood experiences on health: a systematic review and meta-analysis. *Lancet Public Health* 2, e356–e366. doi: 10.1016/S2468-2667(17)30118-4
- Huynh, G., Masood, S., Mohsin, H., and Daniyan, A. (2024). The impact of late ADHD diagnosis on mental health outcomes in females. *Soc. Sci. Human. Open* 10:100977. doi: 10.1016/j.ssaho.2024.100977
- Irwin, D. E., Gross, H. E., Stucky, B. D., Thissen, D., DeWitt, E. M., Lai, J. S., et al. (2012). Development of six PROMIS pediatric proxy-report item banks. *Health Qual. Life Outcomes* 10, 1–13. doi: 10.1186/1477-7525-10-22
- Ismail, W. S. W., Jaafar, N. R. N., and Baharudin, A. (2017). Disagreement between parents, teachers and adolescents in reporting ADHD symptoms: what have we learned? *Med. Res. Arch.* 5. Available at: <https://esmed.org/MRA/mra/article/view/1177>
- Johnson, J. G., Harris, E. S., Spitzer, R. L., and Williams, J. B. (2002). The patient health questionnaire for adolescents: validation of an instrument for the assessment of mental disorders among adolescent primary care patients. *J. Adolesc. Health* 30, 196–204. doi: 10.1016/S1054-139X(01)00333-0
- Korelitz, K. E., and Garber, J. (2016). Congruence of parents' and children's perceptions of parenting: a Meta-analysis. *J. Youth Adolescence* 45, 1973–1995. doi: 10.1007/s10964-016-0524-0
- Lin, I. I. (1989). A concordance correlation coefficient to evaluate reproducibility. *Biometrics* 45, 255–268. doi: 10.2307/2532051
- Lopez-Espejo, M. A., Nuñez, A. C., Ruz, M., et al. (2023). Caregiver-child agreement in health-related quality of life of autistic children and adolescents. *J. Autism Dev. Disord.* 53, 977–986. doi: 10.1007/s10803-022-05689-7
- McDonald, C. A., Lopata, C., Donnelly, J. P., Thomeer, M. L., Rodgers, J. D., and Jordan, A. K. (2016). Informant discrepancies in externalizing and internalizing symptoms and adaptive skills of high-functioning children with autism spectrum disorder. *School Psychol. Quar.* 31, 467–477. doi: 10.1037/spq0000150
- McGinnis, E. W., Copeland, W., Shanahan, L., and Egger, H. L. (2022). Parental perception of mental health needs in young children. *Child Adolesc. Mental Health* 27, 328–334. doi: 10.1111/camh.12515
- McGorry, P. D., and Mei, C. (2018). Tackling the youth mental health crisis across adolescence and young adulthood. *BMJ* 362. doi: 10.1136/bmj.k3704
- Moens, M. A., Weeland, J., Van der Giessen, D., Chhangur, R. R., and Overbeek, G. (2018). In the eye of the beholder? Parent-observer discrepancies in parenting and child disruptive behavior assessments. *J. Abnorm. Child Psychol.* 46, 1147–1159. doi: 10.1007/s10802-017-0381-7
- Muggeo, V. M. R. (2022). Segmented: regression models with breakpoints/change-points estimation (version 2.1.4) [computer software] Available online at: <https://CRAN.R-project.org/package=segmented>
- Noel, J. K., Strong, S. L., Sammartino, C., Gately, K. A., and Rosenthal, S. R. (2025). Effects of delaying healthcare on mental health of college students. *Coll. Stud. Aff. J.* 43, 82–101. doi: 10.1353/csj.2025.a965918
- Olson, D. H., Portner, J., and Lavee, Y. (2013). “Family adaptability and cohesion evaluation scales II” in Handbook of measurements for marriage and family therapy. Eds. Robert S. and Norman F. (London: Routledge).
- Panther-Brick, C., Grimon, M.-P., and Eggerman, M. (2014). Caregiver—child mental health: a prospective study in conflict and refugee settings. *J. Child Psychol. Psychiatry* 55, 313–327. doi: 10.1111/jcpp.12167
- Patel, V., Flisher, A. J., Hetrick, S., and McGorry, P. (2007). Mental health of young people: a global public-health challenge. *Lancet* 369, 1302–1313. doi: 10.1016/S0140-6736(07)60368-7
- R Core Team. (2013). R: a language and environment for statistical computing (version 3.X). R foundation for statistical computing. Available online at: <https://www.R-project.org/>
- Schnyder, N., Lawrence, D., Panczak, R., Sawyer, M. G., Whiteford, H. A., Burgess, P. M., et al. (2020). Perceived need and barriers to adolescent mental health care: agreement between adolescents and their parents. *Epidemiol. Psychiatr. Sci.* 29:e60. doi: 10.1017/S2045796019000568
- Slowikowski, K. (2023). Ggrepel: automatically position non-overlapping text labels with 'ggplot2' (version 0.9.6) [computer software] Available online at: <https://CRAN.R-project.org/package=ggrepel>
- Spitzer, R. L., Kroenke, K., Williams, J. B., and Löwe, B. (2006). A brief measure for assessing generalized anxiety disorder: the GAD-7. *Arch. Intern. Med.* 166, 1092–1097. doi: 10.1001/archinte.166.10.1092
- Tarren-Sweeney, M. (2019). Do adolescents in care systematically under-report their mental health difficulties in population studies? A narrative review. *Dev. Child Welfare* 1, 251–272. doi: 10.1177/2516103219829483
- Topp, C. W., Østergaard, S. D., Søndergaard, S., and Bech, P. (2015). The WHO-5 Well-Being Index: A systematic review of the literature. *Psychotherapy and Psychosomatics*, 84, 167–176. doi: 10.1159/000376585
- Vu, M., Duhig, A. M., Tibrewal, A., Campbell, C. M., Gaur, A., Salomon, C., et al. (2023). Increased delay from initial concern to diagnosis of autism spectrum disorder and associated health care resource utilization and cost among children aged younger than 6 years in the United States. *J. Manag. Care Spec. Pharm.* 29, 378–390. doi: 10.18553/jmcp.2023.29.4.378
- Wickham, H. (2023). ggplot2: create elegant data visualisations using the grammar of graphics (version 3.5.2) [computer software]. Available online at: <https://CRAN.R-project.org/package=ggplot2>
- Wickham, H., François, R., Henry, L., and Müller, K. (2023). Dplyr: a grammar of data manipulation (version 1.1.4) [computer software]. Available online at: <https://CRAN.R-project.org/package=dplyr> (Accessed April 2025).
- Windarwati, H. D., Lestari, R., Wicaksono, S. A., Kusumawati, M. W., Ati, N. A. L., Ilmy, S. K., et al. (2022). Relationship between stress, anxiety, and depression with suicidal ideation in adolescents. *Jurnal Ners* 17, 36–41. doi: 10.20473/jn.v17i1.31216
- World Health Organization. (2024). Adolescent mental health. Available online at: <https://www.who.int/news-room/fact-sheets/detail/adolescent-mental-health> (Accessed April 2025).