

Investigating Variables as Risk Factors for Mental Health Conditions Among Autistic  
Individuals With and Without Genetic Etiologies

Leilemah Aslamy

A thesis

submitted in partial fulfillment of the  
requirements for the degree of

Master of Public Health

University of Washington

2024

Committee:

Nora Henrikson

Emily Neuhaus

Program Authorized to Offer Degree:

Public Health Genetics

©Copyright 2024

Leilemah Aslamy

University of Washington

**Abstract**

Investigating Variables as Risk Factors for Mental Health Conditions Among Autistic  
Individuals With and Without Genetic Etiologies

Leilemah Aslamy

Chair of the Supervisory Committee:

Nora Henrikson

Department of Health Systems and Population Health

This thesis compares mental health outcomes among autistic individuals with and without genetic events. Literature shows that autism spectrum disorder (ASD) is not only highly heritable but is often diagnosed with many co-occurring psychiatric medical conditions, resulting in an additive effect on the functional impairment of an individual. Copy number variants such as 16p11.2 and 22q11.2 are linked to increased ASD susceptibility and other mental health conditions. Exploring the genetic basis of ASD and co-occurring psychiatric health outcomes can strengthen our knowledge of this complex condition and potentially better clinical outcomes for those with ASD. This study includes data from the Simons Foundation Powering Autism Research (SPARK) Study DNA biobank. It uses logistic regression to investigate how variables like sex assigned at birth, age, and age of ASD diagnosis affect the relationship between ADHD, depression, and anxiety and the presence of an ASD related-genetic event.

## I. Introduction

*What is Autism? Why genetic analysis?*

Autism spectrum disorder (ASD) is a complex neurological condition that affects communication and social interaction and is often characterized by repetitive patterns of behavior and interests (Wagner et al., 2019). First outlined by American child psychiatrist, Leo Kanner, autism was primarily known as ‘infantile autism’ because he observed many children displaying similar characteristics (Shishido et al., 2013). During that time, children who presented autistic-like symptoms were formally diagnosed as having ‘childhood’ schizophrenia, and autism was frequently diagnosed by grouping separate psychological and neurodevelopmental symptoms (Shishido et al., 2013). Since then, the DSM-5 has reflected autism as part of a spectrum instead of a conglomerate of separate disorders (Shishido et al., 2013).

Gauging risk susceptibility to genes like BRCA, found in breast cancer, or investigating one’s ancestral origins are just some of the ways illustrating the increase in the adaptation of genetic data and reasoning within everyday life. (Mehrgou & Akouchekian, 2016). Scholars and researchers are motivated to explore the genetic etiology of various complex diseases to advance therapeutics and diagnostics since the boom of genetic information that came along with the Human Genome Project in the early 2000s. Genetic analysis and information shows up in healthcare practices and everyday life more and more each day. Goals that are potentially achieved through research of genomics are, earlier diagnosis and identification of health

conditions, more accurate prognosis predictions, and guiding therapeutic and clinical interventions (Johannessen et al., 2021).

Studies have shown that over 100 genes and genetic variants are associated with the likelihood of ASD (Satterstrom et al., 2020). Considering its genetic basis is critical to understanding ASD's origins and manifestations. Indeed, research is needed to investigate the genetic and epigenetic contributors to ASD to increase our knowledge of this highly complex condition, which could potentially lead to improved clinical outcomes for people with autism (Wagner et al., 2019). The use of genetic testing in autism diagnoses can affirm behavioral assessments and be helpful to parents as they navigate therapies (Little & Gunter, 2021). When left undiagnosed, ASD can significantly impact functioning later in life, often causing undue stress and harm to the autistic individual and their families (Little & Gunter, 2021).

The heritability of ASD finds an empirical basis in twin studies, highlighting the importance of investigating its genetic etiology (Wagner et al., 2019). One of the earliest papers published on the genetic study of autism in twins was released in 1977 (Folstein & Rutter, 1977). In this study, 21 twin pairs were genetically analyzed and it was found that in monozygotic (MZ) twin pairs there was a concordance rate of 36% for autism, compared to a rate of 0% concordance in the dizygotic (DZ) pairs (Folstein & Rutter, 1977). In addition, when tested for cognitive abnormalities, the concordance rate in MZ pairs was 82% compared to 10% in the DZ pairs (Folstein & Rutter, 1977). These higher concordance rates for cognitive abnormalities and autism in MZ twins compared to DZ twins, suggests a strong presence of genetic factors contributing to the development of them (Folstein & Rutter, 1977). Recent studies have estimated the heritability rate of ASD to be around 83%-90% (Tuncay et al., 2023). Although this heritability is high, extreme genetic and phenotypic heterogeneity add to the disorder's

complexity and pose a barrier to gene discovery (Betancur, 2011). Additionally, complexity rises as recent studies have shown that rare, *de novo*, or *spontaneously occurring* variants can contribute to ASD (Geschwind & State, 2015).

### *Mental Health and ASD*

In individuals with ASD, it is widely known that they may experience many co-occurring psychiatric medical conditions on top of their primary diagnosis. Some of the most commonly co-occurring medical conditions are attention-deficit/hyperactivity disorder, depression, anxiety, and schizophrenia (Lai, 2023). Existing literature reveals that co-occurring mood disorders are more frequently seen in populations with ASD than amongst the general population (Oakley et al., 2021). In a meta-analysis study comprised of 96 independent studies, those who were diagnosed with autism had a pooled prevalence of 28% for attention-deficit/hyperactivity disorder (ADHD), 20% for overall anxiety disorders, 11% for depression, and 4% for schizophrenic spectrum disorders (Lai, 2019). This data shows the significant prevalence of various mental health conditions that can exist within autistic populations, which is important to address in order to achieve a more personalized treatment methods for individuals.

The literature overwhelmingly agrees that autistic individuals are more likely to have psychiatric diagnoses than those without autism (Hossain et al., 2020). Additionally, many of these psychiatric conditions can have an additive effect when accompanied by an ASD diagnosis. A study found that autistic children who also have ADHD are three to four times as likely to experience depression, anxiety, and behavioral problems (Casseus et al., 2023). However, as seen in the general population, those with ASD are shown to be more likely to develop mood and psychiatric disorders as they increase in age (Lai, 2023). There are mental health differences

within different demographics of autistic individuals. Data suggests that autistic individuals assigned female at birth are more likely to experience psychiatric diagnoses by young adulthood than their male counterparts (Martini et al., 2022).

### *Common CNV's related to Autism*

In individuals with ASD, approximately ten percent are thought to have large chromosomal rearrangements, indicating large pieces of chromosomes are moved around, duplicated or completely missing (Shishido et al., 2013). Biologically, chromosomal rearrangement increases the number of mutations and abnormal chromosomes, causing in an increased risk of complications in DNA synthesis and protein production, therefore resulting in various gene expressions and phenotypes (Harewood & Fraser, 2014). Using comparative genomic hybridization techniques, many studies have pointed to the association between common copy number variants (CNV) and autism within people with professional ASD diagnoses (Pinto et al., 2010). Copy number variants (CNV) refers to parts of the genome in which certain DNA sections are repeated, with the number of repeats varying between individuals. The human genome has particular 'hot spots' where many CNVs occur, some of which can be related to autistic and other psychiatric traits, such as areas like 16p11.2 and 22q11.2 (Shishido et al., 2013). Genetically, there are several different mechanisms by which CNV can form including mediation by non-allelic homologous recombination, non-homologous end joining, break point induced replication, and transposon activity (Liu et al., 2012)

In most cases, common and rare CNV are not direct causes of autism but rather a possible indication of increased likelihood of ASD due to the highly heterogenic nature of the condition (Shishido et al., 2013). In a 2007 study investigating the familial heritability of CNV, it was

found that inherited CNV can increase ASD susceptibility (Autism Genome Project Consortium et al., 2007). Later that year, another study, the first of its kind, reported a statistically significant association between de novo CNV and autism (Sebat et al., 2007).

The CNV with the highest frequency in autistic populations is 16p11.2, with 1 in 100 autistic individuals being carriers of this CNV. Compared to a control non-ASD population, 16p11.2 is found at a rate of 1 in 1000 (Shishido et al., 2013). This region of DNA is known to be around 600 kb long and includes 29 different genes (Shishido et al., 2013). Along with the implications for autism, the 16p11.2 has also been found to be associated with other psychiatric conditions like obsessive compulsive disorder, bipolar disorder, schizophrenia, and obesity.

Another CNV commonly known in involvement with ASD is 22q11.2. This CNV is also commonly known as 22q11.2 deletion syndrome or DiGeorge syndrome (Shishido et al., 2013). At least 20 percent of those with this syndrome meet the diagnostic criteria for ASD. Interestingly, some studies suggest that duplications of 22q11.2 are associated more frequently with ASD than deletions (Mollon et al., 2023).

### *Autism and Social Determinants*

Analysis of social determinants of health can help advance the understanding of systemic barriers that majorly affect an individual's life. These social determinants can include factors like ethnicity, age, gender, employment, housing security, built environment, education level, discrimination, and environmental hazards. Compounded by their disabled identity, many folks with ASD experiences are intersectional identities. For example, one study showed that autistic individuals who also identify as a racial/ethnic minority and live in a low-income household are less likely to receive and engage with postsecondary education, independent living, and

health-care transition services compared to their higher-income autistic peers (Eilenberg et al., 2019). Additionally, significant data shows that autistic individuals might be over-represented within unhoused populations (Lai, 2023). Currently, the US healthcare system fails autistic individuals by the scarcity of care services in low-income settings. Sex assigned at birth is also a social determinant that affects those with ASD differently. In 2019, a study investigating gender differences in adults with ASD showed that there are significant differences between males and females when it comes to symptomatology and camouflaging (Schuck et al., 2019). Genetic evidence of gender differences in individuals with ASD suggests that there is a protective effect among those assigned female at birth (Zhang et al., 2020). Age of ASD diagnosis is also known to be positively related to increased mental health outcomes among autistic folk, indicating the significant effect that mental health has on autism-related symptoms among those who are diagnosed later in life (Mandy, 2022).

Research aimed at investigating the genetic and epigenetic contributors to ASD can increase our knowledge of this highly complex condition, which could lead to improved treatment options and clinical outcomes for people with autism (Wagner et al., 2019).

Autistic individuals experience disproportionately high rates of psychiatric and mental health conditions like ADHD, anxiety, and depression compared to their non-autistic counterparts. However, the prevalence and severity of these co-occurring diagnoses vary dramatically within autistic populations. Recent literature suggests this variability might relate to different biological pathways, including the absence or presence of known autism-related CNV. Within ASD populations, there is a need for research on how genetic subtypes may confer variations in susceptibility for co-occurring mental health conditions. A better understanding of how underlying genetic archetypes intersect with co-occurring conditions is needed to increase

the uptake and distribution of screening and care services for this vulnerable population. This project seeks to compare risk factors for and rates of psychiatric illnesses between autistic individuals with and without genetic etiologies. Ultimately, this project hopes to reveal the role of autism-related CNV for co-occurring mental health conditions and help inform future specialized treatment tailored to those with autism-related CNV/SNV.

*Specific Aims:*

Aim 1: Investigate whether the rates of mental health conditions, like anxiety, depression, and ADHD significantly differ between autistic individuals with and without autism-related CNVs/SNVs.

Aim 2: Examine how sex assigned at birth, age, and age of ASD diagnosis may predict rates of co-occurring mental health outcomes within each group.

## **II. Methods**

The population for this thesis project is people diagnosed with autism spectrum disorder (ASD) and their family members in the United States who have participated in and consented to the SPARK study. The SPARK study is an ongoing genetic autism research study in the United States, aimed at improving lives of autistic individuals through genetics by identifying causes of the condition which will hopefully inform new support services, therapies, and treatment.

Eligibility for this study involves being a proband, the first person in the family to be enrolled in the SPARK study, with an autism diagnosis and English speaking. There are 32 SPARK clinical sites all over the United States. Usually, each site is hosted in a larger research center connected

to universities and other research institutions. There are SPARK clinics in locations such as, Stanford University, Yale University, Boston's Children's Hospital, University of California Los Angeles, and University of Washington.

According to demographic data from SPARK, there is currently phenotypic data for 328,973 individuals, with more than 100,000 of them being diagnosed with ASD (*SFARI | SPARK*, 2018). The total number of female probands are 34,874 and the total number of male probands are 97,264. The total number of children under 18 with ASD is 109,161.

There are approximately 116,693 individuals for which there is SPARK genomic data available (*SFARI | SPARK*, 2018). Included in this data is 106,744 individuals (44,304 with ASD) with whole exome sequencing and 12,519 individuals (3,575 with ASD) with whole genome sequencing (*SFARI | SPARK*, 2018).

The data for this project came from the centralized SPARK phenotypic and genetic datasets, which includes genomic data on autistic individuals and their family members. The inclusion criteria for this dataset were people with ASD who also had mental health history data easily available. This study used a selected raw dataset including 79,423 individuals with ASD comprising 21,449 females and 57,974 males. Of those, 10,422 reported a diagnosis of depression, 19,258 reported a diagnosis of anxiety and 29,755 reporting diagnosis of ADHD. For the purposes of this report, individuals with dysthymia were included in the count for reported diagnosis of depression. Individuals with anxiety disorders such as panic, phobia, agoraphobia, or generalized anxiety disorder (GAD) are included in the count for reported anxiety diagnosis, however, those with social anxiety diagnosis were excluded from this count. The reported count for ADHD diagnoses also includes diagnoses of reported attention-deficit disorder (ADD). Additionally, 701 individuals indicated an identified autism related genetic event . Those with an

identified genetic event were separately grouped, as were those without identified genetic events, based on the presence or absence of autism-related CNVs or SNVs. By grouping those with known CNVs and SNVs together, there is more statistical power between the genetic event group and those without the genetic event.

The quantitative statistical analysis of this data set is a multivariable logistic regression. The outcomes are a professional diagnosis of ADHD, anxiety, and depression, which are binary in nature. Binary dependent variables must be investigated using a logistic regression model, as a linear relationship with independent variables is not possible. Logistic regression models produce odds ratios which are used to show the difference in risk between groups. This paper will include a multivariable logistic regression analysis of the dependent variables of a professional diagnosis of ADHD, anxiety, and depression and how they affect the independent variables of genetic event, sex assigned at birth, age, and age of ASD diagnosis. An example of the general form of the logistic regression model was as follows:

*ADHD Model example :*

$$\begin{aligned} (\text{ADHD}) \sim & \beta_0 \text{ ASDgene} + \beta_1 \text{ sex} + \beta_2 \text{ age} + \beta_3 \text{ age\_of\_autism\_dx} \\ & 0/1. \quad 0/1. \quad \text{male/female} \quad \text{cont.} \quad \text{cont.} \end{aligned}$$

The multivariable logistic regression statistical analyses were performed using R version 4.2.1 and RStudio. The dataset from the .sav file was read into R and multivariable logistic regression models for each mental health outcome were fitted using the 'glm' function from the 'stats' package, including the binomial family specified for binary outcome variables. Model assumptions were evaluated using the 'vif' function from the 'car' package. Model evaluations like p-values and 95% confidence intervals were estimated using the 'exp'/'confint' functions. Calibration plots were created to evaluate the model using the 'ggplot' function from the

'ggplot2' package. To further evaluate the model, brier scores were calculated and AUC-ROC curves were created using the 'roc' function in the 'pROC' package.

Calibration plots, ROC curves/AUC values, and brier scores were assessed as part of these models' evaluation. Calibration plots show the relationship between observed probabilities and predicted probabilities from our model. Slopes from these plots can give insight into the nature of miscalibration within each model. To measure the accuracy of probabilistic prediction within each model, brier scores and scaled brier scores were calculated, where a brier score of zero indicates perfect accuracy, and a scaled score closer to the max value of one indicates perfect level of accuracy. Additionally, Variance Inflation Factors (VIFs) were extracted from the models to check assumptions of multicollinearity, referring to when two or more predictor variables are significantly highly correlated. High multicollinearity is not ideal as it can unveil unreliable estimates of the multivariable logistic regression coefficients. VIF values show how much of the variance in the estimated regression coefficients is influenced by the correlation between predictor variables. VIF values equal to one indicate no multicollinearity, and values greater than 5-10 are considered high indicators of multicollinearity.

## **I. Results**

After cleaning the data, getting rid of any NA or blank values, the remaining dataset consisted of 79,127 individuals with ASD of which 21,344 were female and 57,783 male. The mean age at study enrollment was around 11.5 years, and the mean age at ASD diagnosis was around 6.5 years. Included in the sample was 10,403 individuals reporting diagnosis of depression, 19,208 individuals reporting diagnosis of anxiety and 29,670 individuals reporting

diagnosis of ADHD. There were a total of 698 individuals with genetic events identified (CNVs+SNVs). Table 1 quantifies these counts.

Multivariable logistic regression results and analysis revealed that the presence of a genetic event was significantly associated with decreased odds of having depression (OR = 0.51987, CI: 0.3829-0.6914), but not significantly associated with ADHD (OR = 1.00332 CI: 0.8573-1.1718) or anxiety (OR = 0.8358 CI:0.6906-1.0056). Males show statistically significant higher odds of having ADHD compared to females (OR = 1.2898,  $p = 6.427e-47$ ) but show a significant association with lower odds of having depression and anxiety compared to females (OR = 0.5370,  $p = 1.067e-133$ ; OR = 0.6302,  $p = 1.297e-123$ ). In all three models, for ADHD, depression, and anxiety, older age at registration and at ASD diagnosis were significantly associated with higher odds of each outcome. Table 2 shows complete results, including odds ratios, 95% confidence intervals, and p-values for all predictors are presented.

Calibration plots, ROC curves/AUC values, and brier scores were assessed, and the full list of plots and values are in Appendix A, B, and C, respectively. The ADHD model showed more deviance from the perfect calibration diagonal line, than the depression and anxiety models. All models showed positive slope values which might suggest model predictions that are under confident or too narrow. ROC curves and AUC were reported for each model and, again, the ADHD model showed to have the least statistically significant value compared to the depression and anxiety models. Interestingly, the AUC for depression was the highest at 0.8571, indicating good discriminative ability in distinguishing between autistic individuals with and without ASD related genetic events. The range of brier scores for all three models ranges from 0.0913-0.226, and the scaled brier scores ranges from 0.0336-0.201, suggesting relatively acceptable accuracy of probabilistic predictions. Full list of VIF values are in Appendix C. The

three models showed a range of VIF values for all predictor variables ranging from 1.0012-2.1978, the highest being age of ASD diagnosis (VIF = 2.1978) from the ADHD model indicating no major concerns for multicollinearity.

A second set of multivariate logistic regressions were run and odds ratios results and analyses are presented in Table 3. These are the results of stratifying the data by genetic event group, to explore potential differences/similarities between effects of the other predictor variables, like sex assigned at birth, age at registration, and age of ASD diagnosis on the outcome of ADHD, depression and anxiety among individuals with and without ASD related genetic events. Results from these six sub-regression models show that there are some differences in OR between the two groups based on the predictors, but much of the difference is shown to not be statistically significant. Notably, in both subgroups, older age at ASD diagnosis is significantly associated with higher odds of ADHD (Genetic event: OR = 1.0095, 95% CI: 1.0060-1.0133,  $p = 2.6982e-07$ ) (No genetic event: OR = 1.0016, 95% CI: 1.0013-1.0018,  $p = 3.9045e-39$ ) and depression (Genetic event: OR = 1.0063, 95% CI: 1.0025-1.0103,  $p = 0.0015$ ) (No genetic event: OR = 1.0030, 95% CI: 1.0027-1.0033,  $p = 2.1983e-90$ ). Additionally, both subgroups showed a significant association between older age at registration and higher odds of depression (Genetic event: OR = 1.0895 per year increase, 95% CI: 1.0577-1.1228,  $p = 1.5698e-08$ ) (No genetic event: OR = 1.0744, 95% CI: 1.0713-1.0775,  $p < 0.0001$ ) and anxiety (Genetic event: OR = 1.0025, 95% CI: 0.9957-1.0097,  $p = 0.4721$ ) (No genetic event: OR = 1.0025, 95% CI: 1.0023-1.0028,  $p = 1.0445e-76$ ).

In parallel to the first set of model evaluations, for all six sub-models, calibration plots, ROC curves/AUC values, and brier scores were assessed and the full list of plots and values are in Appendix A, B, and C, respectively. Much like the unstratified models, both ADHD models

showed more deviance from the perfect calibration diagonal line, than all depression and anxiety models. The depression without genetic event model showed the least deviance from the perfect calibration line out of all stratified models. All models showed positive slope values which might suggest model predictions that are under confident or too narrow. ROC curves and AUC were reported for each model and, again, the ADHD model showed to have the least statistically significant value compared to the depression and anxiety models. The AUC for both depression models were the highest at 0.873 for the group with genetic events and 0.8567 for the group without identified genetic events, indicating good discriminative ability in distinguishing between autistic individuals with and without ASD related genetic events within this outcome. The range of brier scores for all six models ranges from 0.0913-0.227, and the scaled brier scores ranges from 0.0301-0.201, suggesting relatively acceptable accuracy of probabilistic predictions. Additionally, the six models showed a range of VIF values for all predictor variables ranging from 1.0051-2.213, the highest being age of ASD diagnosis (VIF = 2.1978) from the ADHD without genetic event model indicating no major concerns for multicollinearity.

**Table 1. Demographic and Clinical Characteristics of the Study Sample**

<b>Characteristic</b>	<b>n (%)</b>
Total Sample	<b>79,127</b> (100%)
<b>Sex assigned at birth</b>	
Female	<b>21,344</b> (27.0%)
Male	<b>57,783</b> (73.0%)
<b>Mean age at study enrollment</b>	<b>11.5 years</b>
<b>Mean age of ASD diagnosis</b>	<b>6.5 years</b>
<b>Genetic events</b>	

(CNVs+SNVs)	<b>698</b> (0.9%)
<b>Mental health conditions</b>	
Depression	<b>10,403</b> (13.1%)

Table 1. Note. CNVs = copy number variations; SNVs = single nucleotide variations; ADHD = Attention Deficit Hyperactivity Disorder.

**Table 2. Odds Ratios for Mental Health Outcomes in Individuals with Autism Spectrum Disorder**

Variable	ADHD OR (95% CI, p-value)	Depression OR (95% CI, p-value)	Anxiety OR (95% CI, p-value)
<b>Genetic Events</b>			
With Genetic Events	<b>1.003</b> (0.857-1.171, p = 0.966)	<b>0.519</b> (0.382-0.691, p < 0.001)	<b>0.835</b> (0.690- 1.005, p = 0.061)
Without Genetic Events	Reference	Reference	Reference
<b>Sex Assigned at Birth</b>			
Male	<b>1.289</b> (1.245- 1.3353567, p < 0.001)	<b>0.53704842</b> (0.511- 0.564, p < 0.001)	<b>0.630</b> (0.606- 0.654, p < 0.001)
Female	Reference	Reference	Reference
<b>Age at enrollment (years)</b>	<b>1.025</b> (1.023- 1.028, p < 0.001)	<b>1.074</b> (1.071- 1.077, p < 0.001)	<b>1.052</b> (1.050- 1.055, p < 0.001)

<b>Age at ASD Diagnosis (months)</b>	<b>1.002</b> (1.001-1.002, p < 0.001)	<b>1.003</b> (1.002-1.003, p < 0.001)	<b>1.002</b> (1.002-1.003, p < 0.001)
--------------------------------------	---	---	---

*Table 2 Note. OR = Odds Ratio; CI = Confidence Interval; ADHD = Attention Deficit Hyperactivity Disorder; ASD = Autism Spectrum Disorder.*

**Table 3. Odds Ratios for Mental Health Outcomes in Individuals with Autism Spectrum Disorder, Stratified by Genetic Event Group**

<b>Variable</b>	<b>ADHD OR (95% CI, p-value)</b>	<b>Depression OR (95% CI, p-value)</b>	<b>Anxiety OR (95% CI, p-value)</b>
<b>With Genetic Events</b>			
Male	<b>1.754</b> (1.238-2.507, p = 0.002)	<b>0.822</b> (0.434-1.597, p = 0.555)	<b>0.937</b> (0.634-1.398, p = 0.750)
Female	Reference	Reference	Reference
Age at enrollment (years)	<b>1.001</b> (0.982-1.020, p = 0.910)	<b>1.089</b> (1.057-1.122, p < 0.001)	<b>1.052</b> (1.031-1.073, p < 0.001)
Age at ASD Diagnosis (months)	<b>1.009</b> (1.005-1.013, p < 0.001)	<b>1.006</b> (1.002-1.010, p = 0.001)	<b>1.002</b> (0.999-1.005, p = 0.102)
<b>Without Genetic Event</b>			
Male	<b>1.285</b>	<b>0.537</b>	<b>0.627</b>

	(1.241-1.331, p < 0.001)	(0.511-0.564, p < 0.001)	(0.604-0.652, p < 0.001)
Female	Reference	Reference	Reference
Age (years)	<b>1.026</b> (1.024- 1.028, p < 0.001)	<b>1.074</b> (1.071-1.077, p < 0.001)	<b>1.052</b> (1.050-1.055, p < 0.001)
Age at ASD Diagnosis (months)	<b>1.001</b> (1.001-1.002, p < 0.001)	<b>1.003</b> (1.002-1.003, p < 0.001)	<b>1.002</b> (1.002-1.003, p < 0.001)

*Table 3 Note. OR = Odds Ratio; CI = Confidence Interval; ADHD = Attention Deficit Hyperactivity Disorder; ASD = Autism Spectrum Disorder.*

### III. Discussion

This thesis project aimed to understand the differences in mental health conditions among autistic individuals with and without autism-related CNVs and SNVs. Another primary goal of this project was to show how demographic factors like sex, age, and age of ASD diagnosis predict co-occurring psychiatric health conditions in autistic individuals with differing genetic makeup. Results from this study found that while the presence of a genetic event does not significantly affect the risk of ADHD or anxiety, it is significantly associated with a lower risk of

depression. Age plays a particular key role in predicting likelihood of co-occurring mental health conditions in autistic individuals for all possible outcomes. Differences between rates of mental health conditions were investigated between autistic individuals with and without autism-related CNVs and SNVs, and results showed that individuals with a genetic event had significantly lower odds of having depression compared to those without a genetic event. However, there were no significant differences in the odds of having ADHD or anxiety between the two groups.

The impact of demographic factors such as sex and age on the rates of co-occurring mental health conditions were explored through odds ratios. Males were shown to have higher odds of ADHD but lower odds of depression and anxiety compared to females. Older age at registration and older age at ASD diagnosis were both associated with higher odds of ADHD, depression and anxiety. The final results of this project involved stratifying all three outcome models by genetic event group. In both subgroups, older age at ASD diagnosis and older age at registration were associated with higher odds of ADHD depression and anxiety. The effects were consistent across groups, indicating that age at enrollment and age ASD diagnosis are significant predictors of mental health outcomes regardless of genetic event status.

There are limitations to this project based on the analysis and data collection methods. Some argue that evidence for ASD-specific genes based on CNV data is insufficient (Mollon et al., 2023). However, in line with evidence from many twin and GWAS studies, the pleiotropic nature of CNV on neurodevelopmental disorders exemplifies the significant genetic correlation between ADHD and ASD traits and diagnoses (Mollon et al., 2023). Another possible limitation of this study is that it relies on the use of existing data. Thus, this project relies on the variables included in the methodology of initial data collection. Additionally, the SPARK study is a voluntary genetic autism research study. Therefore, the participant enrollment might be skewed

to reflect those of higher socioeconomic status or of a specific race and ethnicity who have better access to healthcare services systemically. Moreover, stratification by genetic event group might have reduced the sample size potentially affecting the statistical power.

Some findings from this study are consistent with that of previous studies. Our results showed that males are at higher odds of ADHD compared to females. Literature shows that ADHD is not only diagnosed more often in males compared to females but is also treated more often in males (Skogli, 2013). This may be in part due to the fact that females are far less likely to be diagnosed with ADHD and less likely to receive treatment and medications for ADHD (Martin, 2024). Furthermore, our results showed age at ASD diagnosis as a significant predictor for mental health outcomes across genetic event status. Similarly, literature shows that autistic individuals who are diagnosed later in life often have higher levels of mental health outcomes and social difficulties, and these problems only increase with age (Mandy, 2022).

This project can have various implications for future research and in clinical practice. The results of this study underline the importance of accounting for genetic factors and demographic variables in the clinical assessment and management of co-occurring psychiatric conditions among autistic folk. Results from this study also highlight the need for increased genetic counseling and education for families with autism, in particular, highlighting the potential influence genetic factors have on psychiatric health outcomes. Counseling on these topics can help families make informed decisions about further genetic testing and interventions as well as helping families understand the complex interplay between genetics, autism, and other co-occurring psychiatric conditions. Finally, this project adds to the growing literature on genetic influences on mental health outcomes among those with ASD. Future longitudinal studies are needed, within the field of ASD, to elucidate the relationships between genetic, demographic and

mental health outcome factors. Qualitative studies like focus groups or interviews, could be used to understand the lived experiences of autistic folk and their personal accounts of mental health. Potential next steps with this dataset would be to run additional multivariable logistic regressions based on other variables like race, area deprivation index, and language level.

This thesis project provides a look into the complex interplay between genetic factors, demographic factors, and mental health outcomes among those with ASD. Findings from this study show the significant role of ASD-related genetic events in affecting the odds of depression, as well as the role that age and sex have on the likelihood of ADHD, depression, and anxiety. This study lays the groundwork for a nuanced understanding of the various factors that affect mental health outcomes in the autistic community.

#### IV. References

##### References

- Autism Genome Project Consortium, Szatmari, P., & Paterson, A. (2007). Mapping autism risk loci using genetic linkage and chromosomal rearrangements. *Nature Genetics*, 39(3), 319–328. <https://doi.org/10.1038/ng1985>
- Betancur, C. (2011). Etiological heterogeneity in autism spectrum disorders: More than 100 genetic and genomic disorders and still counting. *Brain Research*, 1380, 42–77. <https://doi.org/10.1016/j.brainres.2010.11.078>
- Casseus, M., Kim, W. J., & Horton, D. B. (2023). Prevalence and treatment of mental, behavioral, and developmental disorders in children with co-occurring autism spectrum disorder and attention-deficit/hyperactivity disorder: A population-based study. *Autism Research*. <https://doi.org/10.1002/aur.2894>
- Eilenberg, J. S., Paff, M., Harrison, A. J., & Long, K. A. (2019). Disparities Based on Race, Ethnicity, and Socioeconomic Status Over the Transition to Adulthood Among Adolescents and Young Adults on the Autism Spectrum: a Systematic Review. *Current Psychiatry Reports*, 21(5). <https://doi.org/10.1007/s11920-019-1016-1>
- Folstein, S., & Rutter, M. (1977). Infantile Autism: a Genetic Study of 21 Twin Pairs. *Journal of Child Psychology and Psychiatry*, 18(4), 297–321. <https://doi.org/10.1111/j.1469-7610.1977.tb00443.x>
- Geschwind, D. H., & State, M. W. (2015). Gene hunting in autism spectrum disorder: on the path to precision medicine. *The Lancet Neurology*, 14(11), 1109–1120. [https://doi.org/10.1016/s1474-4422\(15\)00044-7](https://doi.org/10.1016/s1474-4422(15)00044-7)

Harewood, L., & Fraser, P. (2014). The impact of chromosomal rearrangements on regulation of gene expression. *Human Molecular Genetics*, 23(R1), R76–R82.

<https://doi.org/10.1093/hmg/ddu278>

Hossain, M. M., Khan, N., Sultana, A., Ma, P., McKyer, E. Lisako J., Ahmed, H. U., & Purohit, N. (2020). Prevalence of comorbid psychiatric disorders among people with autism spectrum disorder: An umbrella review of systematic reviews and meta-analyses.

*Psychiatry Research*, 287, 112922. <https://doi.org/10.1016/j.psychres.2020.112922>

Johannessen, J., Nærland, T., Hope, S., Torske, T., Kaale, A., Wirgenes, K. V., Malt, E., Djurovic, S., Rietschel, M., & Andreassen, O. A. (2022). Attitudes among parents of persons with autism spectrum disorder towards information about genetic risk and future health. *European Journal of Human Genetics*, 30(10), 1138–1146.

<https://doi.org/10.1038/s41431-021-00966-y>

Lai, M. (2023). Mental health challenges faced by autistic people. *Nature Human Behaviour*. <https://doi.org/10.1038/s41562-023-01718-2>

Lai, M.-C., Kassee, C., Besney, R., Bonato, S., Hull, L., Mandy, W., Szatmari, P., & Ameis, S. H. (2019). Prevalence of co-occurring mental health diagnoses in the autism population: a systematic review and meta-analysis. *The Lancet Psychiatry*, 6(10),

819–829. [https://doi.org/10.1016/s2215-0366\(19\)30289-5](https://doi.org/10.1016/s2215-0366(19)30289-5)

Little, I. D., & Gunter, C. (2021). Mini-Review: Genetic Literacy and Engagement With Genetic Testing for Autism Spectrum Disorder. *Frontiers in Genetics*, 12.

<https://doi.org/10.3389/fgene.2021.693158>

Liu, P., Claudia M.B. Carvalho, Hastings, P. J., & Lupski, J. R. (2012). Mechanisms for recurrent and complex human genomic rearrangements. *Current Opinion in Genetics & Development*, 22(3), 211–220. <https://doi.org/10.1016/j.gde.2012.02.012>

Luo, R., Sanders, Stephan J., Tian, Y., Voineagu, I., Huang, N., Chu, Su H., Klei, L., Cai, C., Ou, J., Lowe, Jennifer K., Hurles, Matthew E., Devlin, B., State, Matthew W., & Geschwind, Daniel H. (2012). Genome-wide Transcriptome Profiling Reveals the Functional Impact of Rare De Novo and Recurrent CNVs in Autism Spectrum Disorders. *The American Journal of Human Genetics*, 91(1), 38–55. <https://doi.org/10.1016/j.ajhg.2012.05.011>

Maenner, M. J., Shaw, K. A., Bakian, A. V., Bilder, D. A., Durkin, M. S., Esler, A., Furnier, S. M., Hallas, L., Hall-Lande, J., Hudson, A., Hughes, M. M., Patrick, M., Pierce, K., Poynter, J. N., Salinas, A., Shenouda, J., Vehorn, A., Warren, Z., Constantino, J. N., & DiRienzo, M. (2021). Prevalence and Characteristics of Autism Spectrum Disorder Among Children Aged 8 Years — Autism and Developmental Disabilities Monitoring Network, 11 Sites, United States, 2018. *MMWR. Surveillance Summaries*, 70(11), 1–16. <https://doi.org/10.15585/mmwr.ss7011a1>

Mandy, W., Midouhas, E., Hosozawa, M., Cable, N., Sacker, A., & Flouri, E. (2022). Mental health and social difficulties of late-diagnosed autistic children, across childhood and adolescence. *Journal of Child Psychology and Psychiatry*, 63(11). <https://doi.org/10.1111/jcpp.13587>

Martin, J. (2024). Why are females less likely to be diagnosed with ADHD in childhood than males? *The Lancet Psychiatry*, 11(4). [https://doi.org/10.1016/s2215-0366\(24\)00010-5](https://doi.org/10.1016/s2215-0366(24)00010-5)

Martini, M. I., Kuja-Halkola, R., Butwicka, A., Du Rietz, E., D'Onofrio, B. M., Happé, F., Kanina, A., Larsson, H., Lundström, S., Martin, J., Rosenqvist, M. A., Lichtenstein, P., & Taylor, M. J. (2022). Sex differences in mental health problems and psychiatric hospitalization in autistic young adults. *JAMA Psychiatry*, *79*(12).

<https://doi.org/10.1001/jamapsychiatry.2022.3475>

Mehrgou, A., & Akouchekian, M. (2016). The Importance of BRCA1 and BRCA2 Genes Mutations in Breast Cancer Development. *Medical Journal of the Islamic Republic of Iran*, *30*(1), 369. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4972064/>

Mollon, J., Almasy, L., Jacquemont, S., & Glahn, D. C. (2023). The contribution of copy number variants to psychiatric symptoms and cognitive ability. *Molecular Psychiatry*, *28*(4), 1480–1493. <https://doi.org/10.1038/s41380-023-01978-4>

Oakley, B., Loth, E., & Murphy, D. G. (2021). Autism and mood disorders. *International Review of Psychiatry*, *33*(3), 280–299. <https://doi.org/10.1080/09540261.2021.1872506>

Pinto, D., Pagnamenta, A. T., Klei, L., Anney, R., Merico, D., Regan, R., Conroy, J., Magalhaes, T. R., Correia, C., Abrahams, B. S., Almeida, J., Bacchelli, E., Bader, G. D., Bailey, A. J., Baird, G., Battaglia, A., Berney, T., Bolshakova, N., Bölte, S., & Bolton, P. F. (2010). Functional impact of global rare copy number variation in autism spectrum disorders. *Nature*, *466*(7304), 368–372. <https://doi.org/10.1038/nature09146>

Satterstrom, F. K., Kosmicki, J. A., Wang, J., Breen, M. S., De Rubeis, S., An, J.-Y., Peng, M., Collins, R., Grove, J., Klei, L., Stevens, C., Reichert, J., Mulhern, M. S., Artomov, M., Gerges, S., Sheppard, B., Xu, X., Bhaduri, A., Norman, U., & Brand, H. (2020). Large-Scale Exome Sequencing Study Implicates Both Developmental and

Functional Changes in the Neurobiology of Autism. *Cell*, 180(3), 568-584.e23.

<https://doi.org/10.1016/j.cell.2019.12.036>

Schuck, R. K., Flores, R. E., & Fung, L. K. (2019). Brief Report: Sex/Gender Differences in Symptomology and Camouflaging in Adults with Autism Spectrum Disorder. *Journal of Autism and Developmental Disorders*, 49(6), 2597–2604.

<https://doi.org/10.1007/s10803-019-03998-y>

Sebat, J., Lakshmi, B., Malhotra, D., Troge, J., Lese-Martin, C., Walsh, T., Yamrom, B., Yoon, S., Krasnitz, A., Kendall, J., Leotta, A., Pai, D., Zhang, R., Lee, Y.-H. ., Hicks, J., Spence, S. J., Lee, A. T., Puura, K., Lehtimaki, T., & Ledbetter, D. (2007). Strong Association of De Novo Copy Number Mutations with Autism. *Science*, 316(5823), 445–449.

<https://doi.org/10.1126/science.1138659>

SFARI | SPARK. (2018, February 7). SFARI; Simons Foundation Powering Autism Research. <https://www.sfari.org/resource/spark/>

Shishido, E., Aleksic, B., & Ozaki, N. (2013). Copy-number variation in the pathogenesis of autism spectrum disorder. *Psychiatry and Clinical Neurosciences*, 68(2), 85–95.

<https://doi.org/10.1111/pcn.12128>

Skogli, E. W., Teicher, M. H., Andersen, P. N., Hovik, K. T., & Øie, M. (2013). ADHD in girls and boys – gender differences in co-existing symptoms and executive function measures. *BMC Psychiatry*, 13(1). <https://doi.org/10.1186/1471-244x-13-298>

Tuncay, I. O., DeVries, D., Gogate, A., Kaur, K., Kumar, A., Xing, C., Goodspeed, K., Seyoum-Tesfa, L., & Chahrour, M. H. (2023). The genetics of autism spectrum disorder in an East African familial cohort. *Cell Genomics*, 3(7), 100322.

<https://doi.org/10.1016/j.xgen.2023.100322>

Wagner, K. E., McCormick, J. B., Barns, S., Carney, M., Middleton, F. A., & Hicks, S. D. (2019). Parent Perspectives Towards Genetic and Epigenetic Testing for Autism Spectrum Disorder. *Journal of Autism and Developmental Disorders*, 50(9), 3114–3125. <https://doi.org/10.1007/s10803-019-03990-6>

Zhang, Y., Li, N., Li, C., Zhang, Z., Teng, H., Wang, Y., Zhao, T., Shi, L., Zhang, K., Xia, K., Li, J., & Sun, Z. (2020). Genetic evidence of gender difference in autism spectrum disorder supports the female-protective effect. *Translational Psychiatry*, 10(1). <https://doi.org/10.1038/s41398-020-0699-8>

## V. Appendices

### Appendix A - Calibration Plots

Figure 1 - Calibration plot for ADHD Model:

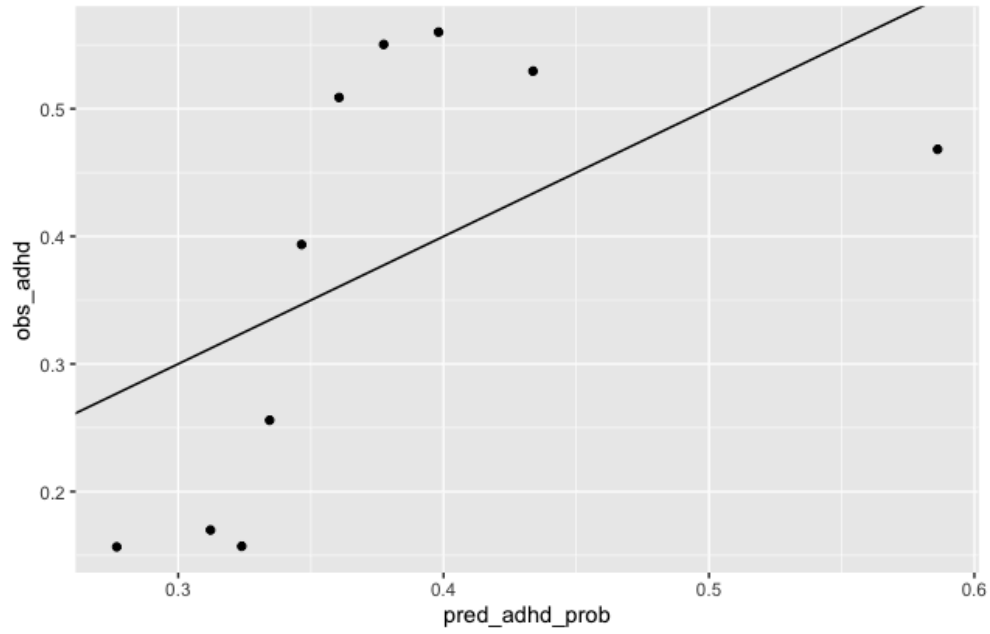


Figure 2 - Calibration plot for Depression Model:

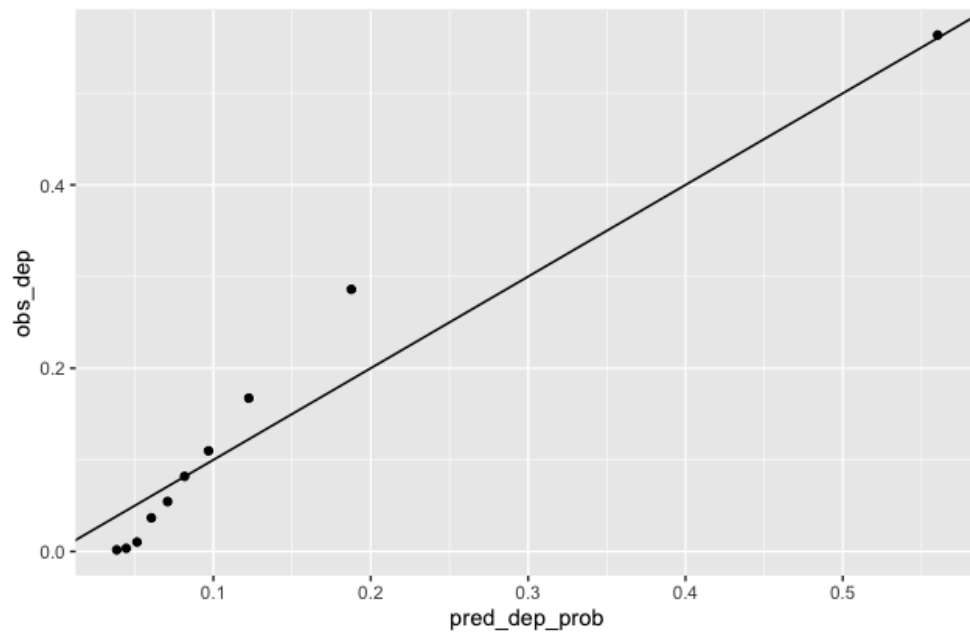


Figure 3 - Calibration plot for Anxiety Model:

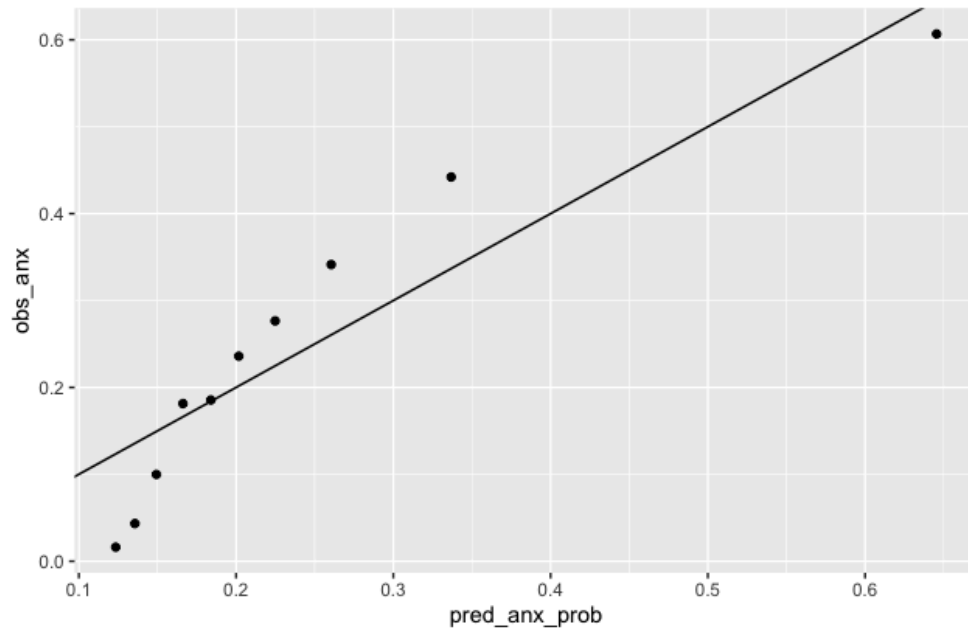


Figure 4 - Calibration plot for ADHD With Genetic Event Model:

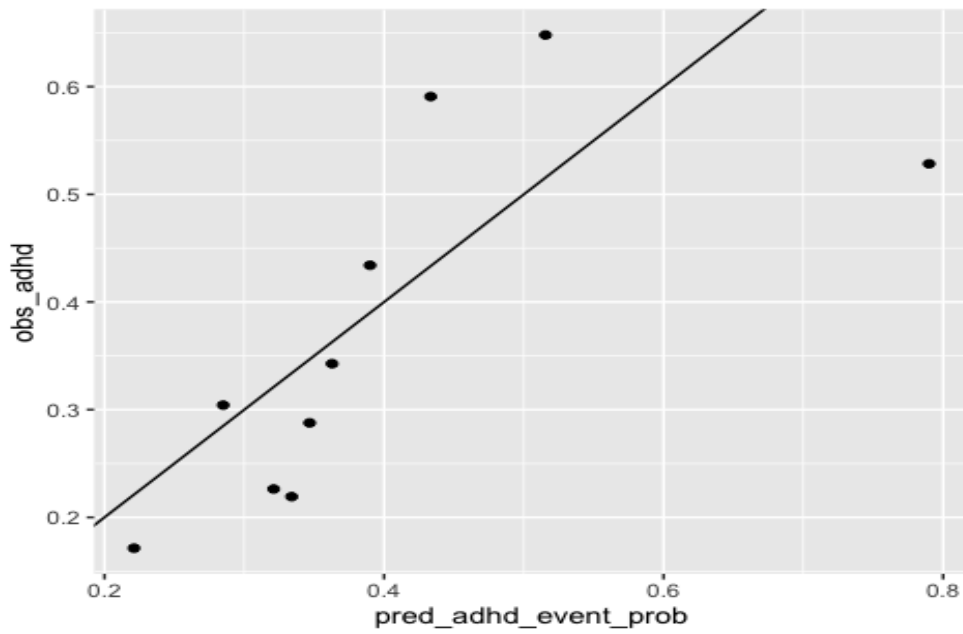


Figure 5 - Calibration plot for ADHD Without Genetic Event Model:

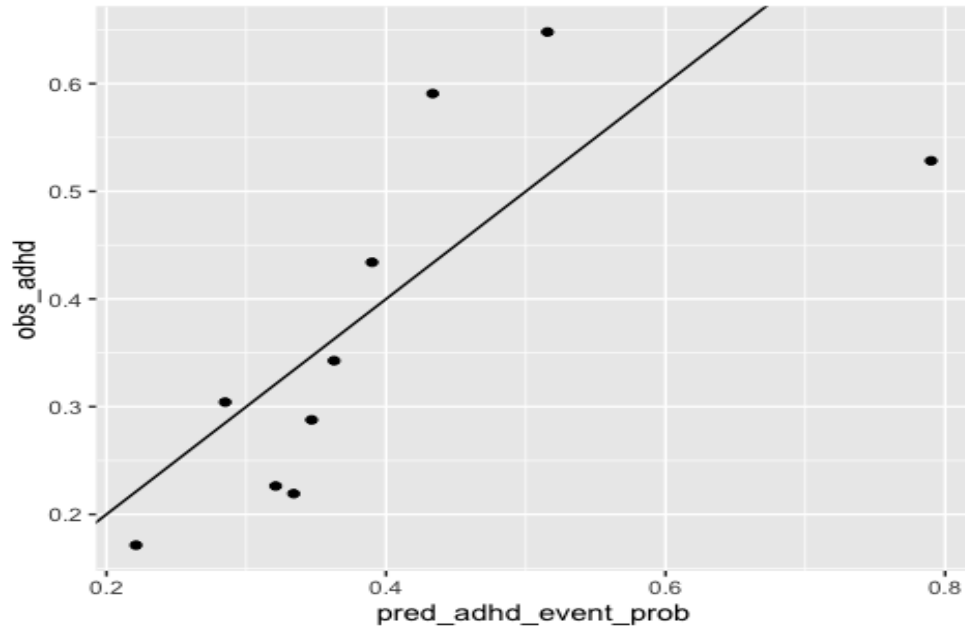


Figure 6 - Calibration plot for Depression With Genetic Event Model:

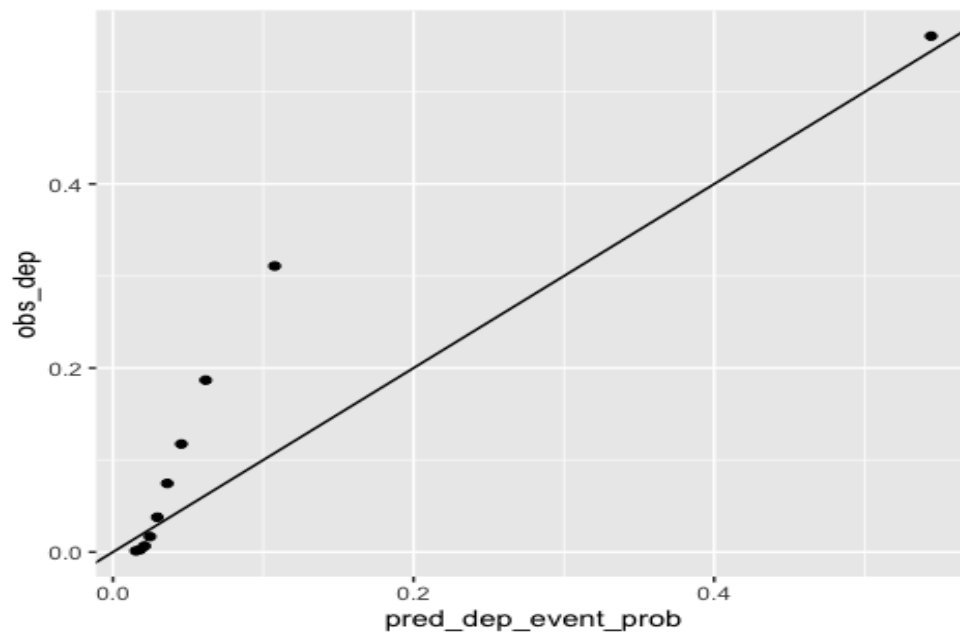


Figure 7 - Calibration plot for Depression Without Genetic Event Model:

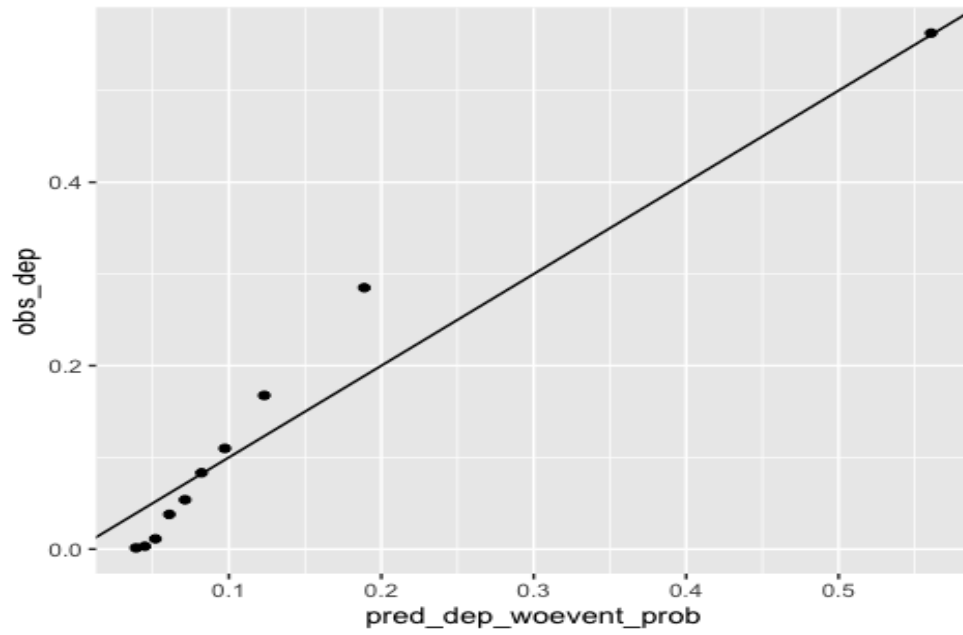


Figure 8 - Calibration plot for Anxiety With Genetic Event Model:

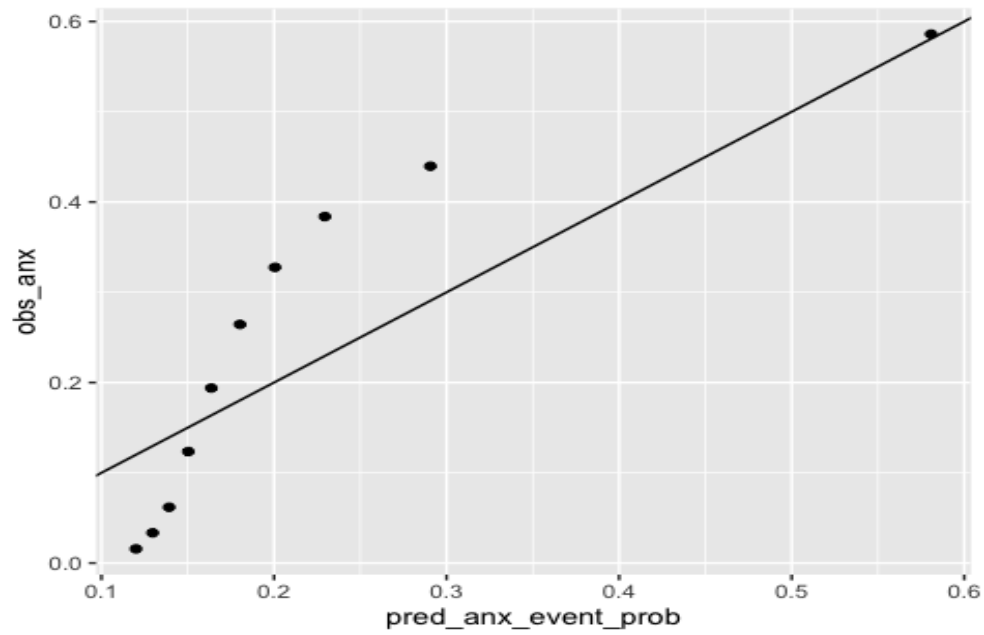
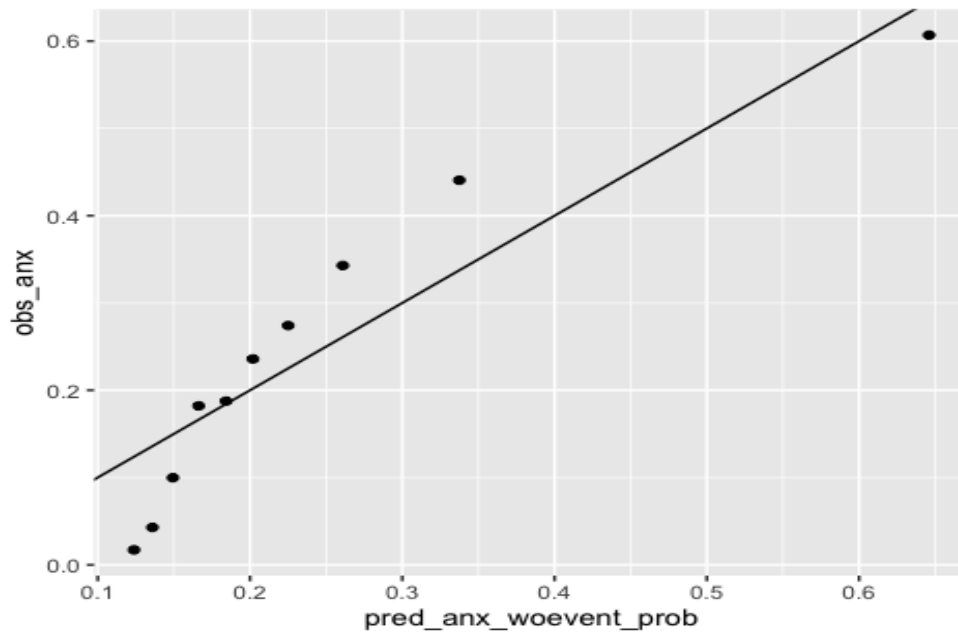


Figure 9 - Calibration plot for Anxiety Without Genetic Event Model:



## Appendix B - AUC-ROC Curves

Figure 10 - ROC for ADHD Model: AUC=0.6798

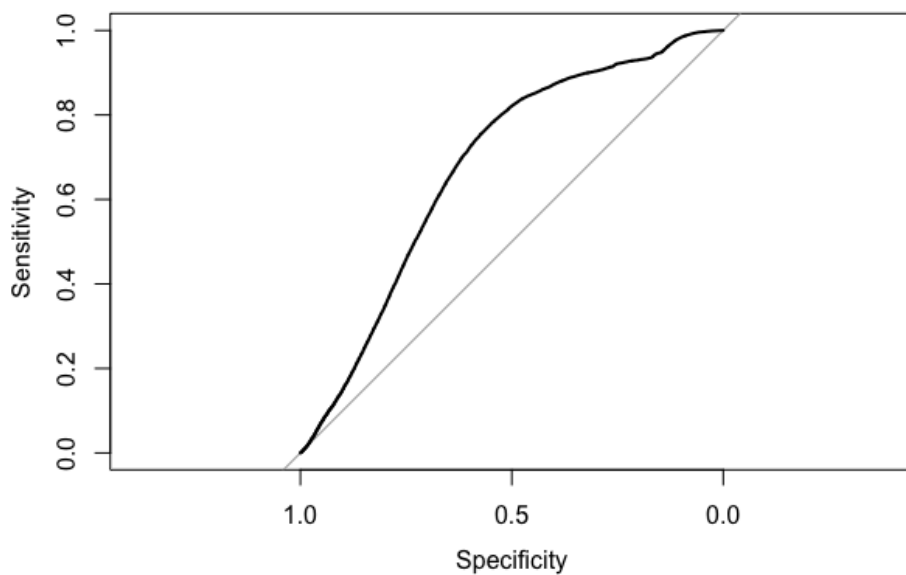


Figure 11 - ROC for Depression Model: AUC=0.8571

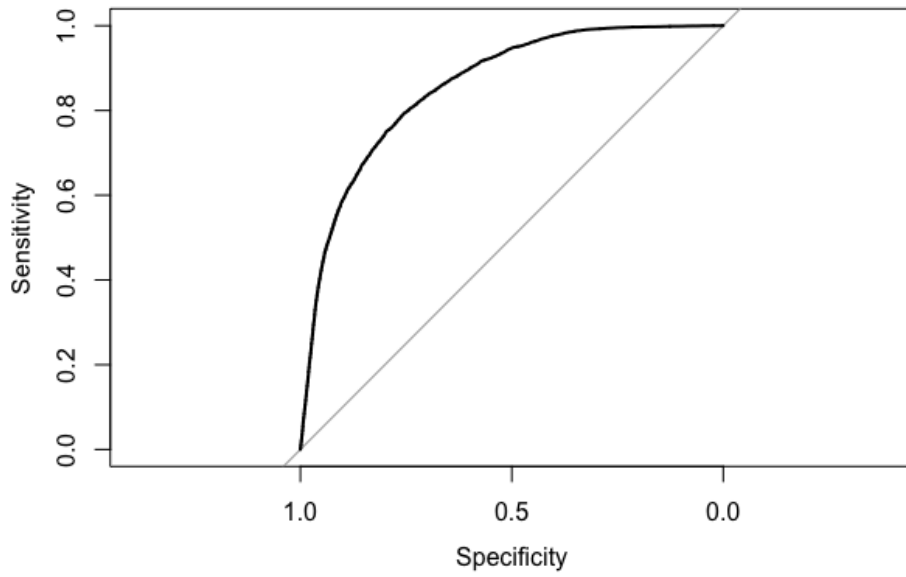


Figure 12 - ROC for Anxiety Model: AUC=0.7652

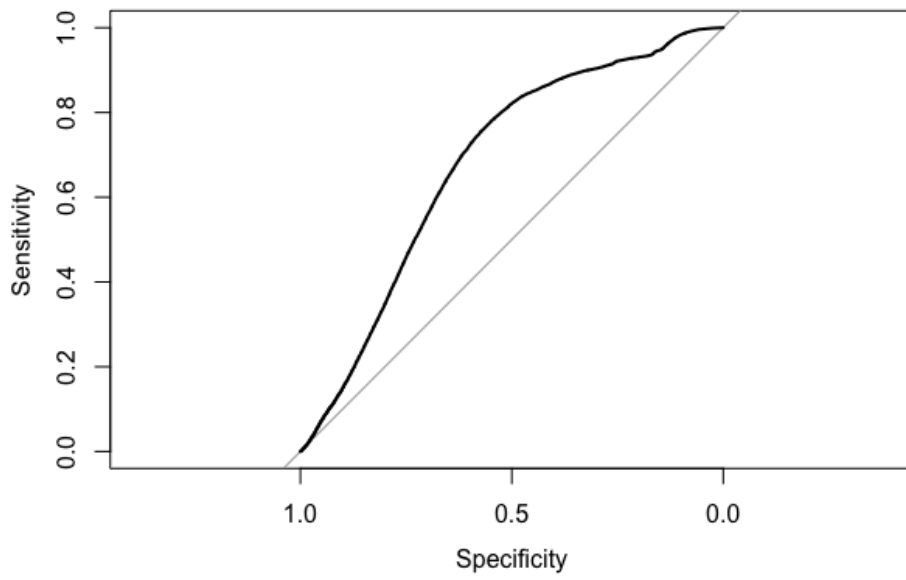


Figure 13 - ROC for ADHD With Genetic Event Model: AUC=0.6752

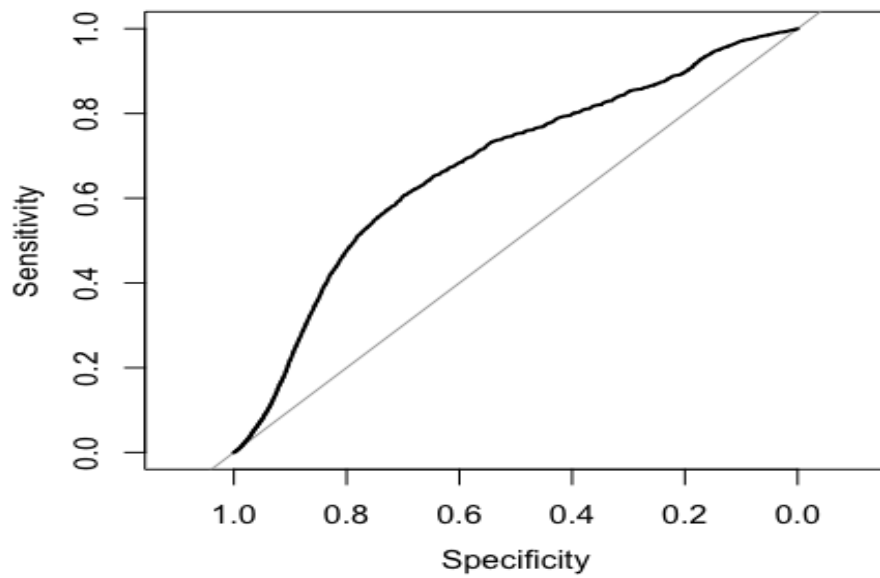


Figure 14 - ROC for ADHD Without Genetic Event Model: AUC = 0.6798

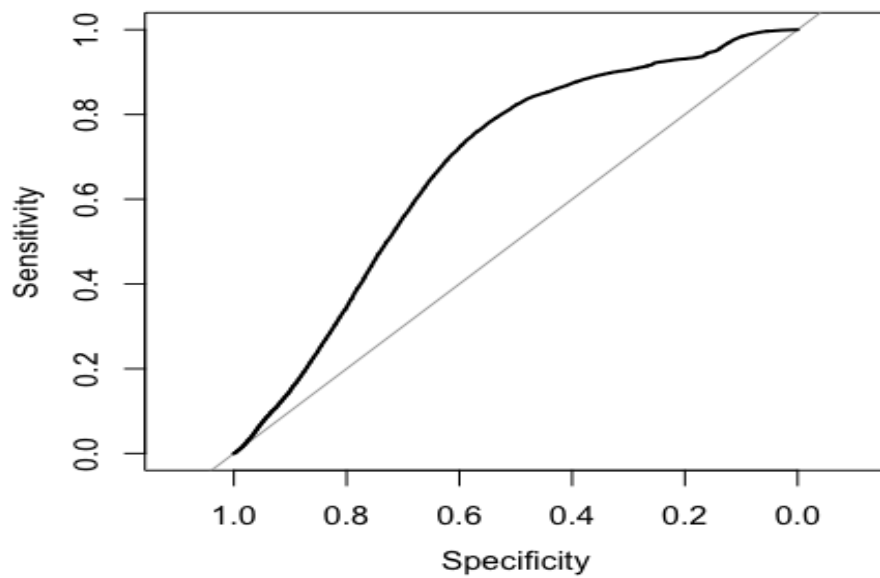


Figure 15 - ROC for Depression With Genetic Event Model: AUC = 0.873

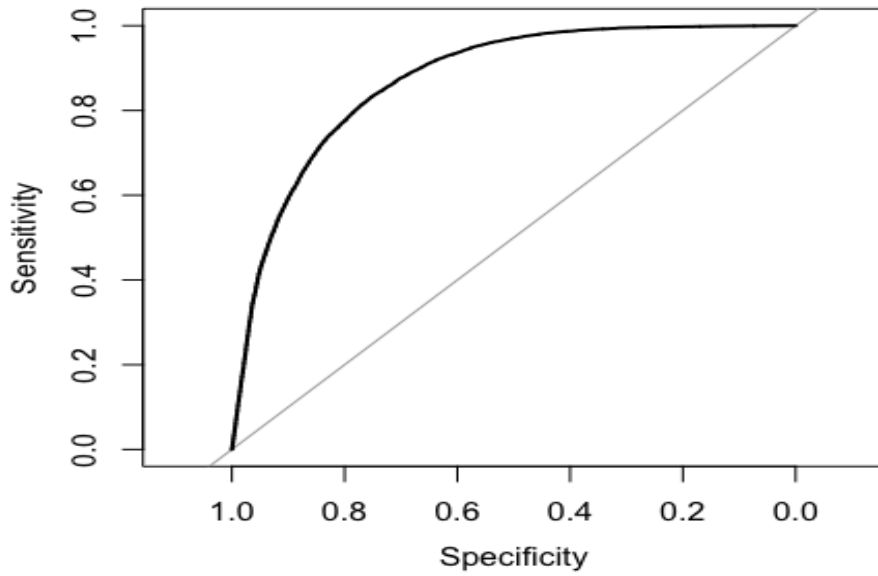


Figure 16 - ROC for Depression Without Event Model: AUC = 0.8567

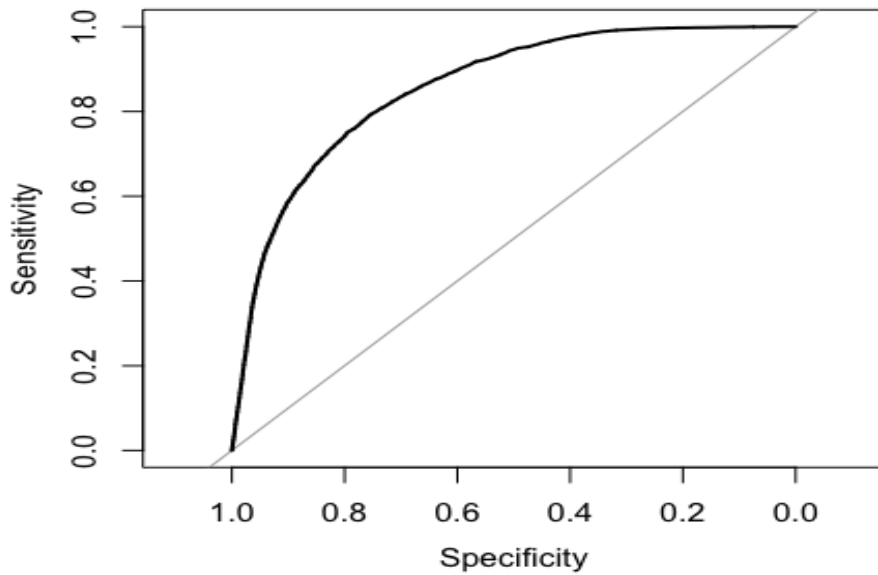


Figure 17 - ROC for Anxiety With Event Model: AUC = 0.7815

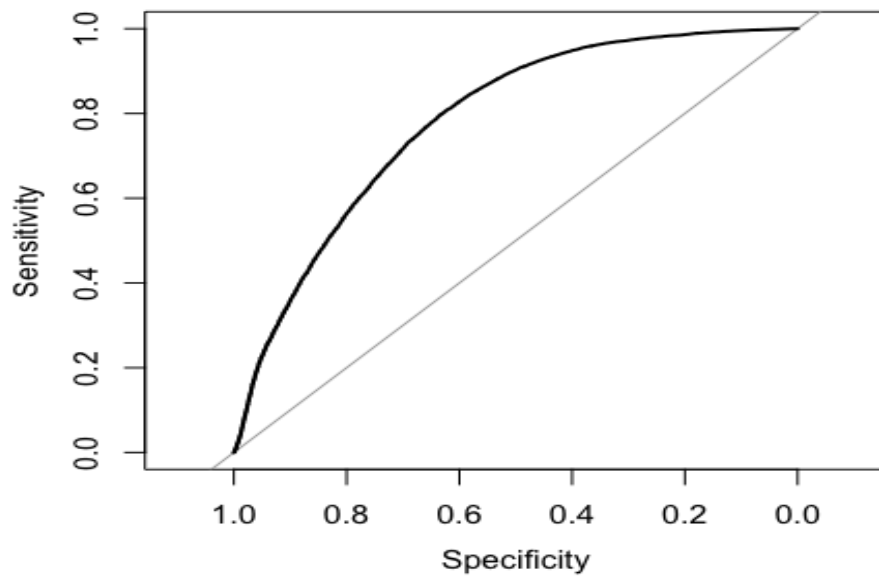
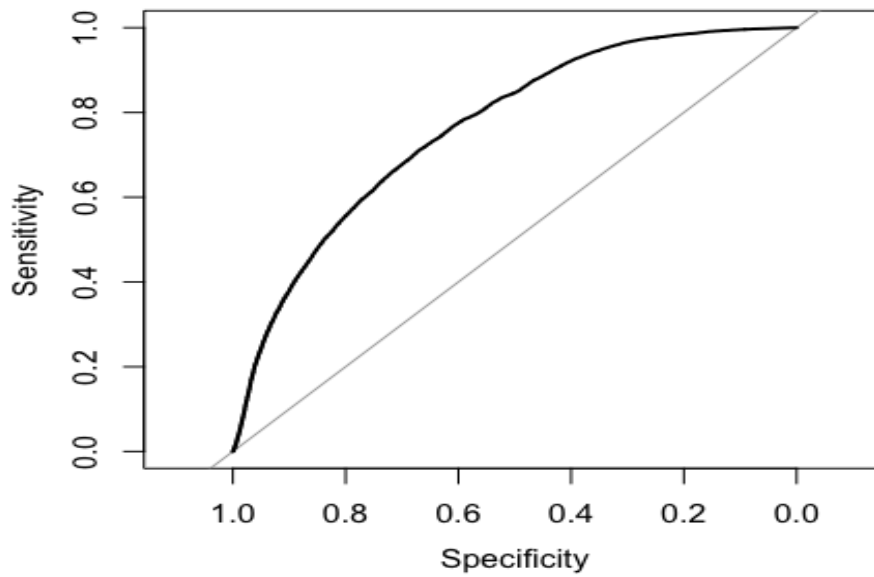


Figure 18 - ROC for Anxiety Without Genetic Event Model: AUC = 0.7649



## Appendix C - Brier Scores and VIF Values

Table 4 - Full list of Brier scores and VIF values

Value	ADHD	Depression	Anxiety	ADHD +Gene	ADHD -Gene	Dep +Gene	Dep -Gene	Anx + Gene	Anx - Gene
<b>Brier Scores</b>									
Brier	0.226	0.0913	0.160	0.227	0.226	0.0971	0.0913	0.164	0.160
Scaled Brier	0.0336	0.201	0.127	0.0301	0.0336	0.150	0.201	0.110	0.127
<b>VIF values</b>									
Genetic_Event_TRUE	1.001290	1.001838	1.001646						
Sex	1.077770	1.057191	1.049955	1.019844	1.078363	1.008977	1.057508	1.005145	1.050253
Age	2.146399	1.928241	1.882578	1.201705	2.161393	1.173037	1.936624	1.258284	1.891361
Age @ASD Diagnosis	2.197873	1.979967	1.928705	1.223407	2.213020	1.175892	1.989237	1.263000	1.938121