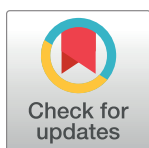


## RESEARCH ARTICLE

# Influences of race and clinical variables on psychiatric genetic research participation: Results from a schizophrenia sample

Rose Mary Xavier<sup>1\*</sup>, Yuktha Shanavas<sup>2</sup>, Brian M. Britt<sup>3</sup>, Wales T. George<sup>4</sup>

**1** School of Nursing, University of North Carolina Chapel Hill, Chapel Hill, North Carolina, United States of America, **2** College of Medicine, SUNY Upstate Medical University, Syracuse, New York, United States of America, **3** College of Arts and Sciences, University of North Carolina Chapel Hill, Chapel Hill, North Carolina, United States of America, **4** Mental and Behavioral Health Services, Veterans Affairs Medical Center, Durham, North Carolina, United States of America

\* [rxavier@email.unc.edu](mailto:rxavier@email.unc.edu)

## Abstract

Advances in genetics has led to a better understanding of both genetic and environmental contributions to psychiatric mental health disorders. But psychiatric genetics research is predominantly Eurocentric, and individuals of non-European ancestry continue to be significantly underrepresented in research studies with potential to worsen existing mental health disparities. The objective of this study was to examine factors associated with genetic study participation in a schizophrenia sample. The study sample was extracted from the Clinical Antipsychotics Trial of Intervention Effectiveness (CATIE) schizophrenia study which enrolled 1493 patients with chronic schizophrenia between the ages of 18–65 years and incorporated an optional genetic sub-study. Using a logistic regression model (N = 1249), we examined sociodemographic and clinical variables that were independently associated with the outcome i.e., participation in the genetic sub-study. The genetic sub-study had a lower proportion of Black (30% in genetic vs 40% in CATIE overall) and other race (4% vs 6%) participants. Increased severity of psychopathology symptoms (odds ratio [OR] = 0.78,  $p = 0.004$ ) decreased the odds whereas better reasoning scores (OR = 1.16,  $p = 0.036$ ) increased the odds of genetic study participation. Compared to Black participants, White participants were significantly more likely to participate in the genetic sub-study (OR = 1.43,  $p = 0.009$ ). Clinical factors in addition to race significantly impact genetic study participation of individuals with chronic schizophrenia. Our findings highlight the need for future research that examines the interactive effects of race and clinical factors such as symptom severity on psychiatrically ill individuals' choice to participate in genetics studies and to identify targeted strategies to increase equitable representation in psychiatric genetics research.

## OPEN ACCESS

**Citation:** Xavier RM, Shanavas Y, Britt BM, George WT (2023) Influences of race and clinical variables on psychiatric genetic research participation: Results from a schizophrenia sample. PLoS ONE 18(4): e0284356. <https://doi.org/10.1371/journal.pone.0284356>

**Editor:** Rikinkumar S. Patel, Duke University Medical Center: Duke University Hospital, UNITED STATES

**Received:** January 29, 2023

**Accepted:** March 29, 2023

**Published:** April 12, 2023

**Copyright:** © 2023 Xavier et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** All relevant data are located at [https://nda.nih.gov/edit\\_collection.html%3Fid=2081](https://nda.nih.gov/edit_collection.html%3Fid=2081).

**Funding:** The authors received no specific funding for this work.

**Competing interests:** The authors have declared that no competing interests exist.

## Introduction

Genetic studies have overwhelmingly focused on individuals of European ancestry [1–3]. Despite concerted efforts to increase representation, individuals of African, Latin American, Hispanic, Indigenous and East Asian ancestry continue to be severely underrepresented in

genetic studies [1]. For genetic studies of psychiatric and mental health disorders, this underrepresentation is quite pronounced [4, 5]. For example, the largest genome wide association study of schizophrenia (N = 320,404) comprised of 74.3% European samples 17.5% Asian, 5.5% African American and 2.5% Latino samples [6] whereas for bipolar disorder (N = 41,917) 100% of the study sample were European [7]. Scientific insights gained from one population does not carry over to another as differences in allele frequencies and genetic background confound the association of genes with the disorder or phenotype under study [8]. Subsequently, findings from predominantly European samples have poor predictive value in individuals of non-European ancestry [4]. Underrepresentation of individuals of non-European ancestry in psychiatric genetic studies thus has the potential to worsen existing mental health disparities [4]. Understanding factors that determine research participation is important to facilitate improved participation in psychiatric genetic studies and is critically important for underrepresented patient participants.

Patients' research priorities often don't align well with those of researchers and funding agencies [9]. What patients' perceive as important can influence their research participation levels. Studies that have specifically examined factors affecting psychiatric and mental health research participation have identified fear of invasive procedures [10], the type of research (e.g., genetic studies, neuroimaging etc.), a lack of comprehension regarding research topic and study procedures, fear of treatment side effects, study safety and social stigma surrounding mental health illnesses as major factors that influence participation [11, 12]. The specific psychiatric diagnosis of the participant is also associated with their research participation. For example, compared to patients with depression, patients with schizophrenia report a lower approval of psychiatric research and a decreased readiness to participate in research using questionnaires and those requiring blood draws [10]. For both diagnostic groups, there is a lower level of willingness to participate in studies that are medication trials or that use neuroimaging techniques [10]. Patients with schizophrenia generally do not positively regard research if the methods were perceived as dangerous and involuntary [13].

Studies that have examined factors that impact research participation in psychosis and schizophrenia have primarily focused on the motivations and incentives that increase research participation. These studies show that the main factors incentivizing a schizophrenia patient's participation included altruism (i.e., helping others and science) [13, 14] and monetary rewards [10]; how participants were informed and referred to research studies also influence research participation among patients with psychosis. For example, clear communication and a positive patient-clinician relationship were both key factors that increased research participation [15]. When the patient-clinician relationship is generally positive, patients were more likely to be referred to psychosis research studies [15].

Sociodemographic characteristics can influence who are exposed to opportunities for research participation as well as who are likely to participate in research studies. For example, gender and minority status influence how and where an individual access psychiatric care which in turn can influence recruitment into research studies [16]. Black and other persons of color experience multiple obstacles to active engagement in research including time constraints, lack of health insurance, lack of media access, and transportation issues to and from research sites [11, 17]. Black participants' also have a higher level of mistrust regarding participation in genetic research compared to White participants regardless of their socioeconomic status and education levels [11, 18–20]. Black participants report concerns about the safety behind procedures as well as confidentiality about mental illness in families [11, 16]. But there is a dearth of knowledge regarding the social, demographic, and clinical factors associated with genetic research participation in the severely mentally ill such as those with schizophrenia and related psychosis.

The primary purpose of this study was to examine differences in sociodemographic and clinical characteristics of schizophrenia patients who chose to participate in a genetic study compared to those who chose not to. Specifically, we examined sociodemographic factors (such as age, gender, race) and clinical factors including psychopathology symptoms, insight and neurocognitive function known to influence participation in psychiatric research in general. As a secondary aim, we also investigated differences in longitudinal trajectories of symptoms and clinical outcomes between the participants who donated the DNA sample for the genetic study compared to those who did not.

## Methods

We obtained our study sample from the Clinical Antipsychotics Trial of Intervention Effectiveness (CATIE) schizophrenia study conducted between 2001 and 2004. Funded by the National Institute of Mental Health (NIMH), CATIE was a multiphase randomized controlled trial ( $N = 1493$ ) that compared the effectiveness of four atypical antipsychotic medications—quetiapine, ziprasidone, olanzapine, and risperidone—against a typical antipsychotic, perphenazine, in patients with chronic schizophrenia between the ages of 18–65 years [21]. CATIE excluded those with treatment resistant schizophrenia. CATIE also incorporated an optional genetic sub-study that aimed to identify genetic risk variants associated with schizophrenia; ~51% ( $N = 738$ ) of the CATIE sample participated in the genetic study. We accessed deidentified CATIE data from the NIMH Data Archive (NDA, <https://nda.nih.gov/>) after obtaining relevant data access permissions from the NDA and exempt ethical approval from Duke University IRB.

## Measures

*Genetic study participation* was measured based on if the participant donated a sample for the genetic sub-study—a dichotomous variable coded as ‘1’ for ‘Yes’ and ‘0’ for ‘No’.

*Psychopathology symptoms* were rated using the positive and negative syndrome scale (PANSS) [22] which has a total of 30 items with each item scored between 1 (absent) to 7 (extreme); the higher the scores the more severe the symptoms with total scores of <58 indicating mild illness, 58–75 as moderately ill, 76–95 markedly ill and 96–116 as severely ill. PANSS has 3 subscales—7 items assessing positive symptoms, 7 negative symptoms and 16 items focused on general psychopathology symptoms [23]. PANSS is a reliable and extensively validated scale [24] with good construct validity, and predictive validity [22]. We have previously reported a Cronbach’s  $\alpha$  of 0.87 in our study sample [25].

*Insight* was measured by the insight and treatment attitudes questionnaire (ITAQ), a rater administered 11-item scale with good construct validity and reliability (Cronbach’s  $\alpha \sim 0.95$ ) [25], measures the participant’s attitude towards treatment i.e., treatment insight and their awareness of illness i.e., illness insight [26, 27]. ITAQ total scores range from 0 (indicating no insight) to 22 (indicating full awareness of illness), with higher scores indicating better insight. In this study we used the ITAQ total scores for analysis.

*Symptom severity* was assessed from both the patient and the clinician perspectives using the *Clinical global impression of symptom severity* (CGIS) [28]—the patient version measures the participant’s impression of their symptom severity. The clinician version is rated by clinicians to measure the global symptoms severity. Both items are scored from 1 to 7, with 1 being “Normal, not ill” and 7 being “Very severely ill”.

Severity of *depressive symptoms* in the CATIE study was measured by the *Calgary depression rating scale* (CDRS), a valid and reliable measure for depression in schizophrenia patients

[29]. CDRS measures nine mood symptoms which are scored from 1–4: less overall score indicates fewer depression symptoms.

*Neurocognition* measure was a composite score derived from working memory, verbal memory, processing speed, reasoning, and vigilance domain scores. Details on validity and reliability including interrater reliability and how the neurocognitive composite scores were calculated are published elsewhere [30].

*Health summary* was measured by the SF-12, a 12-item participant reported survey of mental and physical health. We used physical and mental health summary variables in our study; these variable scores range from 0 to 100. Higher scores indicate better general physical and mental health with scores <50 used for physical summary to determine a physical condition and a score of <42 indicating depression.

Patient symptoms and functioning over the preceding 4 weeks are measured using the *Quality of Life Scale* (QLS) [31], a measure with high sensitivity to subtle changes and effects of treatment [32]. It consists of 21 items, each rated on a 7-point primarily relying on clinician judgment.

The *MacArthur Competence Assessment Tool-Clinical Research* (MacCAT-CR) [33] was used to measure the capacity of the potential research participant to make competent decisions about research participation. The MacCAT-CR has good construct validity and reliability with an intraclass correlation ranging between 0.84–0.98 [34]. We used MacCAT 4 subscales in our analysis. (1) *Understanding*—the sum of the 13-item subset of the MacCAT-CR which pertains to a participant's ability to understand provided information about a research project and its protocols. Scores can range from 0–26 with  $\leq 15$  suggests poor understanding. (2) *Appreciation*—the sum of the 3 items, can range between 0–6 and pertains to a participant's ability to evaluate the consequences of study participation on themselves. (3) *Choice* is determined from the participant's response to the item that pertains to a patient's ability to clearly communicate a decision about their willingness to participate or not to participate in a research study. Scores can range between 0–2. (4) *Reasoning* is the sum of the 4 items with scores that can range between 0–8 and pertains to the patient's ability to reason about their participation in a study by comparing alternatives and their consequences on themselves.

*Medication adherence* in CATIE was measured by the clinicians' judgement based on data collected from patient report at all assessment visits, monthly pill counts and any other available information [21]. This information was coded on a scale of 1 to 4, with 1 being adherent always/most of the time, 2 for usually, 3 for sometimes and 4 for never/almost never. This was reverse coded for analysis.

## Statistical analysis

All analysis were implemented in R statistical software. We performed basic quality control including examining for duplicate entries on the data obtained from the NDA. We tested for assumptions of normality by examining data distributions of all variables included in our analysis. Variables in our analysis were all near normally distributed. Since there were missing data, we tested the missing mechanism using the Little's test implemented in the R package *naniar* [35]. The data did not meet the assumption of missing completely at random ( $\chi^2 = 450$ ,  $df = 292$ ,  $p < 0.001$ ) and subsequently we performed multiple imputation to estimate missing data using the R package *MissForest* [36]. We conducted analysis on both the datasets—first on the original data with missing values and a second on the imputed dataset.

Primary aim: To examine if socio-demographic (age, gender, race) and clinical variables (psychopathology symptoms, insight, depression, neurocognition, quality of life, capacity for decision making and antipsychotic medication) that were independently associated with

genetic study participation, we estimated a logistic regression model implemented in the R package *rms* [37] with blood sample donation status for genetic sub-study participation as the outcome variable using baseline/study entry data.

Secondary aim: To investigate if there were longitudinal changes in clinical variables between the overall and genetic sub-study groups, we estimated linear mixed models implemented in the R package *lme4* [38], assuming participant level random effects, with relevant clinical variables (psychopathology symptoms, insight, depression, neurocognition, general health, quality of life, capacity for decision making, illness severity) as the outcome variable and blood sample donation as the dependent variable adjusting for the effects of age, gender, race, and medication adherence across the study period.

## Results

Our overall study sample were 74% male and included 60% White participants, 35% Black participants and 5% were of other races—0.5% American Indian/Alaska Native, 2.3% Asian, 0.3% Hawaiian or Pacific Islander, 1.8% of mixed race, 0.1% unknown or not reported. Since the sample size for groups other than White and Black participants were small, we collapsed them into an ‘Other’ group for all analysis. Descriptive summaries of other variables are provided in Table 1. 52.4% of the overall sample participated in the genetic sub-study. Differences between groups that participated in the genetic study compared to those who chose not to are provided in Table 2. The genetic sub-study had a lower proportion of Black (30% in genetic sub-study vs 40% in CATIE overall) and Other (4% vs 6%) participants.

For baseline data, the proportion of missing data for all variables of interest except neurocognition were less than 2%; neurocognition was missing for ~9% of the sample. We fitted a logistic regression model using maximum likelihood estimation to examine which sociodemographic and clinical variables at baseline were independently associated with genetic study participation. 95% confidence intervals and p-values were computed using the Wald approximation. Since the estimates for both the complete and imputed data were similar, we report and interpret our results from the complete data. Please see Table 3 for full results of the model. Our results show that increased severity of psychopathology symptoms assessed by PANSS scores significantly decreased the odds of whether a participant gave a DNA sample (odds ratio [OR] = 0.77,  $p = <0.001$ ). A participant’s increased level of reasoning competency increased the odds of them giving a sample (OR = 1.16,  $p = 0.036$ ). Compared to Black

**Table 1. Descriptive summary of variables at baseline.**

Variables	Mean (SD)	Range	Interquartile range
Age in years	40.5 (11.1)	18–67	17
PANSS	75.9 (17.7)	31–140	23
Insight	18.1 (5.1)	0–22	7
Depression	4.5 (4.4)	0–22	6
Neurocognition (standardized)	0 (1)	-3.3–2.9	1.3
General physical health	48.2 (10.1)	13–68	14.5
General mental health	40.9 (11.7)	8–67	19
Quality of Life	2.7 (3.1)	0.2–5.9	1.5
Competency			
Understanding	23.4 (3.1)	0–26	4
Appreciation	4.5 (1.3)	0–6	2
Reasoning	5.7 (2)	0–8	3
Choice	2 (0.1)	0–2	0

<https://doi.org/10.1371/journal.pone.0284356.t001>

Table 2. Socio-demographic characteristics by genetic sub-study participation.

Variable	Participants who gave a DNA sample		Patients who did not give a DNA sample		Analysis		
	Mean	SD	Mean	SD	t-test	df	p-value
Age(years)	40.9	11.0	40.1	11.2	-1.43	1441	0.153
Patient's education(years)	11.6	3.5	11.4	3.7	-0.75	1411	0.456
Duration since first prescribed antipsychotic medication (years)	16.3	11.3	14.3	10.6	0.51	1415	0.611
Duration since first diagnosed (years)	14.0	10.9	16.2	10.8	-0.18	1416	0.860
	Frequencies (proportion)				$\chi^2$	df	p-value
Sex					0.61	1	0.437
Male	.73		.75				
Female	.27		.25				
Race					<b>23.1</b>	<b>2</b>	<b>1.0e-05</b>
Black	.30		.40				
Other	.04		.06				
White	.66		.53				
Employment					1.28	3	0.526
Full-time	.06		.07				
Part-time	.09		.08				
Did not work	.85		.85				
Marital Status					4.0	4	0.406
Divorced	.22		.19				
Married	.10		.12				
Never Married	.59		.60				
Separated	.06		.06				
Widowed	.02		.03				
Medications switched at the study beginning					<b>15.1</b>	<b>2</b>	<b>5.2e-04</b>
Newly treated	.24		.33				
Same Medication	.15		.13				
Switched Medication	.60		.53				
Antipsychotic medication					<b>13.6</b>	<b>3</b>	<b>3.6e-04</b>
None	.26		.35				
Second Generation	.57		.51				
First Generation	.07		.07				
First + Second generation	.10		.08				

Bolded values are statistically significant at  $p < 0.05$

<https://doi.org/10.1371/journal.pone.0284356.t002>

participants, White participants were more likely to give a blood sample for genetic study participation (OR = 1.43,  $p = 0.003$ ). Please see Fig 1 for the predicted probabilities of genetic study participation for varying levels of psychopathology symptoms by race. Compared to participants who were not on any medications at baseline, those who on antipsychotic medications were more likely to provide a blood sample for genetic sub-study participation. But this association was only statistically significant for second generation antipsychotic medications (OR = 1.41,  $p = 0.013$ ). Other variables included in the model were not statistically significant.

Missing data proportions for longitudinal data are provided in the S1 Table. In CATIE, all structured assessments including the measures used in this study were done at baseline, 6, 12 and 18 months and at the end of the study. We included these timepoints for our longitudinal analyses as data was consistently collected on most participants. We fitted linear mixed models

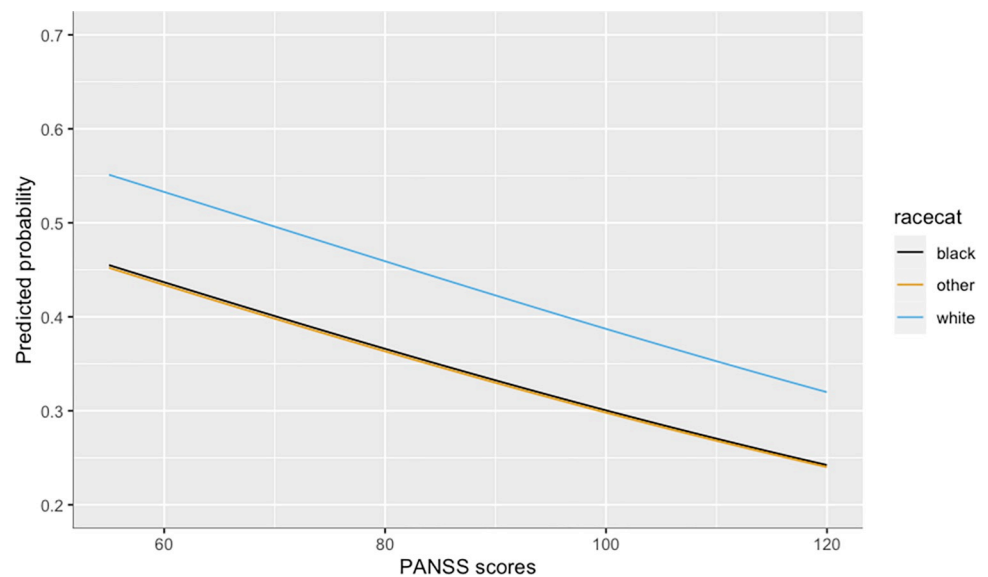
Table 3. Logistic regression results for baseline data (N = 1249).

	Variable	Complete data		
		Standardized Coefficient (log odds)	95% CIs	p-value
Sociodemographic variables	Age	0.06	-0.07, 0.19	0.350
	Gender (ref = female)	-0.04	-0.31, 0.24	0.787
	Education in years	0.01	-0.03, 0.04	0.743
	Race (ref = Black)			
	<b>White</b>	<b>0.39</b>	<b>0.13, 0.64</b>	<b>0.003</b>
	Other	-0.01	-0.57, 0.55	0.967
Clinical variables	Antipsychotic medication at baseline (ref = no med)			
	<b>2<sup>nd</sup> generation</b>	<b>0.34</b>	<b>0.07, 0.61</b>	<b>0.013</b>
	1 <sup>st</sup> generation	0.16	-0.33, 0.65	0.522
	1 <sup>st</sup> + 2 <sup>nd</sup> generation	0.35	-0.08, 0.78	0.110
	Insight	0.05	-0.07, 0.17	0.418
	Depression	0.09	-0.03, 0.22	0.147
	<b>PANSS total</b>	<b>-0.26</b>	<b>-0.40, -0.12</b>	<b>&lt;0.001</b>
	Neurocognition	-0.07	-0.21, 0.07	0.308
	Quality of life	-0.10	-0.25, 0.04	0.162
	MacArthur Competencies			
	Understanding	-0.12	-0.25, 0.009	0.068
	Appreciation	0.05	-0.09, 0.19	0.510
	<b>Reasoning</b>	<b>0.15</b>	<b>0.01, 0.29</b>	<b>0.036</b>
	Choice	0.07	-0.05, 0.19	0.246

Bolded values are statistically significant at  $p < 0.05$

<https://doi.org/10.1371/journal.pone.0284356.t003>

using restricted maximum likelihood estimation on both the original dataset and imputed dataset to examine longitudinal changes in 16 clinical variables by genetic study participation (Table 4). Estimations from original data were very similar to imputed data estimations, so we



**Fig 1. Predicted probabilities of genetic study participation from the logit model.** Y-axis indicates probabilities and x-axis indicates psychopathology symptom scores measured by PANSS. At any level of psychopathology symptoms, Black and Other participants were less likely to participate in the genetic study compared to White participants.

<https://doi.org/10.1371/journal.pone.0284356.g001>

Table 4. Longitudinal changes in clinical variables by genetic sub-study participation.

Outcome Variable	N	Beta Coefficient for sample provided (ref = No)	p-value <sup>a</sup>
<b>Insight</b>	1309	<b>0.61</b>	<b>0.014<sup>b</sup></b>
Illness severity, patient rating	1316	-0.05	0.435
<b>Illness severity, clinician rating</b>	1316	<b>-0.20</b>	<b>8.19e-06</b>
Depression*	1319	0.20	0.306
<b>PANSS</b>	1320		
<b>Total</b>		<b>-4.63</b>	<b>1.57e-07</b>
<b>Positive</b>		<b>-1.40</b>	<b>4.89e-07</b>
<b>Negative</b>		<b>-0.98</b>	<b>0.001</b>
<b>General</b>		<b>-2.26</b>	<b>6.8e-07</b>
Neurocognition	1141	-0.05	0.340
General physical Health	1307	-0.37	0.410
General mental Health	1307	-0.09	0.858
Quality of life	1305	0.02	0.659
<b>MacArthur Competency Subscales</b>	1230		
Understanding		-0.21	0.148
Appreciation		-0.00	0.998
Reasoning		0.04	0.652
Choice		0.01	0.221

Note: Genetic study participation indexed as sample provided is the explanatory variable of interest for all outcome variables. Gender, age, race, medication adherence and medication adherence x time (interaction) were included as covariates in all models; these estimates are not provided in the table.

\*Imputed model was significant for depression with beta 0.44,  $p = 0.016$

<sup>a</sup> Uncorrected p-values reported

<sup>b</sup> not significant at the Bonferroni corrected p-value threshold of  $<0.003$  for 16 linear mixed models reported in this table.

<https://doi.org/10.1371/journal.pone.0284356.t004>

report and interpret results only from the original data. Compared to others, participants in the genetic sub-study longitudinally had significantly better insight ( $\beta = 0.63$ ,  $p < 0.05$ ), were rated by clinician as less severely ill ( $\beta = -0.24$ ,  $p < 0.001$ ) and had less severe psychopathology symptoms ( $\beta = -5.43$ ,  $p < 0.001$ ) across the study period.

## Discussion

Our study examined sociodemographic and clinical characteristics of genetic study participation using a relatively large sample of individuals with chronic schizophrenia. We report that among individuals with chronic schizophrenia, clinical factors (severity of psychopathology symptoms, reasoning abilities and illness severity in general) and sociodemographic factors such as self-reported race significantly influence genetic study participation. Our finding that individuals with severe symptoms and poorer reasoning scores (which are strongly correlated with cognitive dysfunction in schizophrenia; please refer to [S1 Fig](#)) were less likely to participate presents a challenge for investigations into the genetic contributors to schizophrenia. Individuals with the most severe illnesses can be highly informative samples not just to understand illness etiology, but also to develop tailored interventions that improve clinical outcomes for patient subgroups with severe illnesses. But participation in research is imperative to ensure benefits from research. Our study findings highlight the need to actively engage and recruit severely ill individuals. It is also important to consider factors that support retention as psychiatric genetic discoveries progress towards clinical applications—discovery studies generally require a single contact whereas clinical genetic studies often need to be longitudinal. But



there is a critical gap in our knowledge as to the why there are differences in initial participation as well as retention rates in individuals with schizophrenia. Future studies should be designed to facilitate this type of inquiry and must engage individuals with lived experience not just as research participants but as research partners and investigators. Approaches such as community based participatory research has demonstrated value in psychiatric services research and should be considered in psychiatric and mental health genetics research.

Psychiatric genetics research in general (not just schizophrenia genetics research) has a significant underrepresentation of individuals of non-European ancestry [4, 6, 39]. Despite concerted efforts (such as through policy initiatives by the NIH), this underrepresentation continues to widen [2, 39, 40]. Research findings from the general population show that there is a historically justified lack of trust among Black individuals when it comes to genetics research participation [11, 20]. The general mistrust in biomedical research stems not just from historical events e.g., the Tuskegee syphilis study, but is strengthened by discriminatory experiences Black Americans continue to face within the current health care system [41]. There are no studies that have examined genetics research participation specifically in severely psychiatrically ill groups of non-European ancestry. A study that examined racial differences in participation for a genetics of nicotine dependence study showed that willing to participate was not a significant barrier once the racial minority groups were reached by targeted recruitment and engagement [42]. It is unknown if such targeted recruitment and engagement would suffice to improve participation of individuals who have a severe mental illness or those with more severe symptoms in general. Studies of race and ethnic differences in utility of genetic findings and ethical concerns of individuals with severe mental illnesses are scant; most of them predominantly including White individuals [43–45]. Our study findings demonstrate the gap in schizophrenia genetic studies, but extant literature shows that this is also the case for other mental illnesses such as bipolar disorder [39]. Race is not a valid genetic variable, but it is a valid and important social variable that influences important mental health outcomes and as such deserves due consideration in the design stage itself. This is important not just to ensure that research benefits are equitably distributed, but is a necessity to accurately quantify genetic risk [5].

Our study findings show a statistically significant association of genetic study participant by the type of antipsychotic medication at baseline—those who were on second generation antipsychotic medications alone were more likely to participate. Second generation medications are the first line treatment for schizophrenia and patients who are on second generation medications tend to be more treatment responsive or are likely to have fewer years of treatment since initial diagnosis [46]. Our study design and data does not allow for further inquiry into this but suggest the need for further investigation. The CATIE sample though a nationally representative is a US based sample; further studies are warranted to examine if these findings are similar in other populations. Our study findings are also limited in that, the design allows us to only examine initial participation and not retention, which is critically important for clinical genetics studies. Nevertheless, our study highlights the need to understand the intersectional influences of clinical and sociodemographic factors in an individual's decision to participate in a genetic study. The genetic study was optional in CATIE; this allowed us to examine factors influencing participation for schizophrenia patients, but such studies involving diagnosed individuals are critically warranted in other psychiatric disorders including major depressive, bipolar disorders and eating disorders among others.

## Supporting information

**S1 Table. N of complete data for relevant variables across time points.**  
(DOCX)

**S1 Fig. Correlogram of relevant variables from baseline.** Measures for variables are included in parenthesis. CDRS–Calgary Depression Rating Scale; CGIS–Clinical Global Impression of Symptom Severity; ITAQ–Illness and Treatment Attitude Questionnaire; MacCAT-CR–MacArthur Competence Assessment Tool- Clinical Research; PANSS–Positive and Negative Syndrome Scale; SF-12–12-item Short Form Survey; QLS–Heinrich-Carpenter Quality of Life Scale.  
(DOCX)

## Author Contributions

**Conceptualization:** Rose Mary Xavier.

**Data curation:** Rose Mary Xavier.

**Investigation:** Rose Mary Xavier.

**Methodology:** Rose Mary Xavier.

**Project administration:** Rose Mary Xavier.

**Resources:** Rose Mary Xavier.

**Software:** Rose Mary Xavier.

**Supervision:** Rose Mary Xavier.

**Visualization:** Rose Mary Xavier.

**Writing – original draft:** Rose Mary Xavier, Yuktha Shanavas, Brian M. Britt, Wales T. George.

**Writing – review & editing:** Rose Mary Xavier, Wales T. George.

## References

1. Popejoy AB, Fullerton SM. Genomics is failing on diversity [Internet]. 2016 [cited 2021 May 7]. Available from: <http://www.nature.com/news/genomics-is-failing-on-diversity-1.20759>
2. Peterson RE, Kuchenbaecker K, Walters RK, Chen C-Y, Popejoy AB, Periyasamy S, et al. Genome-wide Association Studies in Ancestrally Diverse Populations: Opportunities, Methods, Pitfalls, and Recommendations. *Cell*. 2019 Oct 17; 179(3):589–603. <https://doi.org/10.1016/j.cell.2019.08.051> PMID: 31607513
3. Atkinson EG, Bianchi SB, Ye GY, Martínez-Magaña JJ, Tietz GE, Montalvo-Ortiz JL, et al. Cross-ancestry genomic research: time to close the gap. *Neuropsychopharmacology* [Internet]. 2022 Jun 23; Available from: <https://doi.org/10.1038/s41386-022-01365-7> PMID: 35739257
4. Martin AR, Kanai M, Kamatani Y, Okada Y, Neale BM, Daly MJ. Clinical use of current polygenic risk scores may exacerbate health disparities. *Nat Genet*. 2019 Apr; 51(4):584–91. <https://doi.org/10.1038/s41588-019-0379-x> PMID: 30926966
5. Martin AR, Daly MJ, Robinson EB, Hyman SE, Neale BM. Predicting Polygenic Risk of Psychiatric Disorders. *Biol Psychiatry*. 2019 Jul 15; 86(2):97–109. <https://doi.org/10.1016/j.biopsych.2018.12.015> PMID: 30737014
6. Trubetskoy V, Pardiñas AF, Qi T, Panagiotaropoulou G, Awasthi S, Bigdeli TB, et al. Mapping genomic loci implicates genes and synaptic biology in schizophrenia. *Nature*. 2022 Apr; 604(7906):502–8. <https://doi.org/10.1038/s41586-022-04434-5> PMID: 35396580
7. Mullins N, Forstner AJ, O’Connell KS, Coombes B, Coleman JRI, Qiao Z, et al. Genome-wide association study of more than 40,000 bipolar disorder cases provides new insights into the underlying biology. *Nat Genet*. 2021 Jun; 53(6):817–29. <https://doi.org/10.1038/s41588-021-00857-4> PMID: 34002096
8. Freedman ML, Reich D, Penney KL, McDonald GJ, Mignault AA, Patterson N, et al. Assessing the impact of population stratification on genetic association studies. *Nat Genet*. 2004 Apr; 36(4):388–93. <https://doi.org/10.1038/ng1333> PMID: 15052270

9. Crowe S, Fenton M, Hall M, Cowan K, Chalmers I. Patients', clinicians' and the research communities' priorities for treatment research: there is an important mismatch. *Res Involv Engagem* [Internet]. 2015 Dec; 1(1). Available from: <http://dx.doi.org/10.1186/s40900-015-0003-x>
10. Schäfer I, Burns T, Fleischhacker WW, Galderisi S, Rybakowski JK, Libiger J, et al. Attitudes of patients with schizophrenia and depression to psychiatric research: a study in seven European countries. *Soc Psychiatry Psychiatr Epidemiol*. 2011 Feb; 46(2):159–65. <https://doi.org/10.1007/s00127-010-0181-7> PMID: 20119828
11. Murphy E, Thompson A. An exploration of attitudes among black Americans towards psychiatric genetic research. *Psychiatry*. 2009 Summer; 72(2):177–94. <https://doi.org/10.1521/psyc.2009.72.2.177> PMID: 19614555
12. Roberts LW, Kim JP. Receptiveness to participation in genetic research: A pilot study comparing views of people with depression, diabetes, or no illness. *J Psychiatr Res*. 2017 Nov; 94:156–62. <https://doi.org/10.1016/j.jpsychires.2017.07.002> PMID: 28719815
13. Weiss Roberts L, Warner TD, Brody JL. Perspectives of patients with schizo—phrenia and psychiatrists regarding ethically important aspects of research participation. *Am J Psychiatry*. 2000; 157:67–74. <https://doi.org/10.1176/ajp.157.1.67> PMID: 10618015
14. Zullino D, Conus P, Borgeat F, Bonsack C. Readiness to participate in psychiatric research. *Can J Psychiatry*. 2003 Aug; 48(7):480–4. <https://doi.org/10.1177/070674370304800709> PMID: 12971019
15. Bucci S, Butcher I, Hartley S, Neil ST, Mulligan J, Haddock G. Barriers and facilitators to recruitment in mental health services: care coordinators' expectations and experience of referring to a psychosis research trial. *Psychol Psychother*. 2015 Sep; 88(3):335–50. <https://doi.org/10.1111/papt.12042> PMID: 25257960
16. Brown G, Marshall M, Bower P, Woodham A, Waheed W. Barriers to recruiting ethnic minorities to mental health research: a systematic review. *Int J Methods Psychiatr Res*. 2014 Mar; 23(1):36–48. <https://doi.org/10.1002/mpr.1434> PMID: 24474683
17. Hussain-Gambles M, Atkin K, Leese B. Why ethnic minority groups are under-represented in clinical trials: a review of the literature. *Health Soc Care Community*. 2004 Sep; 12(5):382–8. <https://doi.org/10.1111/j.1365-2524.2004.00507.x> PMID: 15373816
18. Furr LA. Perceptions of genetics research as harmful to society: differences among samples of African-Americans and European-Americans. *Genet Test*. 2002 Spring; 6(1):25–30. <https://doi.org/10.1089/109065702760093889> PMID: 12180073
19. Laskey SL, Williams J, Pierre-Louis J, O'Riordan M, Matthews A, Robin NH. Attitudes of African American premedical students toward genetic testing and screening. *Genet Med*. 2003 Jan; 5(1):49–54. <https://doi.org/10.1097/00125817-200301000-00008> PMID: 12544476
20. McQuillan GM, Porter KS, Agelli M, Kington R. Consent for genetic research in a general population: the NHANES experience. *Genet Med*. 2003 Jan; 5(1):35–42. <https://doi.org/10.1097/00125817-200301000-00006> PMID: 12544474
21. Stroup TS, McEvoy JP, Swartz MS, Byerly MJ, Glick ID, Canive JM, et al. The National Institute of Mental Health Clinical Antipsychotic Trials of Intervention Effectiveness (CATIE) Project: Schizophrenia Trial Design and Protocol Development. *Schizophr Bull*. 2003 Jan 1; 29(1):15–31. <https://doi.org/10.1093/oxfordjournals.schbul.a006986> PMID: 12908658
22. Kay SR, Opler LA, Lindenmayer JP. The Positive and Negative Syndrome Scale (PANSS): rationale and standardisation. *Br J Psychiatry Suppl*. 1989 Nov;(7):59–67. PMID: 2619982
23. Leucht S, Kane JM, Kissling W, Hamann J, Etschel E, Engel RR. What does the PANSS mean? *Schizophr Res*. 2005 Nov 15; 79(2–3):231–8. <https://doi.org/10.1016/j.schres.2005.04.008> PMID: 15982856
24. Opler MGA, Yavorsky C, Daniel DG. Positive and Negative Syndrome Scale (PANSS) Training: Challenges, Solutions, and Future Directions. *Innov Clin Neurosci*. 2017 Dec 1; 14(11–12):77–81. PMID: 29410941
25. Xavier RM, Pan W, Dungan JR, Keefe RSE, Vorderstrasse A. Unraveling interrelationships among psychopathology symptoms, cognitive domains and insight dimensions in chronic schizophrenia. *Schizophr Res*. 2018; 193:83–90. <https://doi.org/10.1016/j.schres.2017.07.002> PMID: 28693755
26. Mcevoy JP, Apperson LJ, Applebaum P. Insight in schizophrenia, its relationship to acute psychopathology. *J Nerv Ment Dis*. 1989; 177:43–7.
27. Zhou Y, Rosenheck R, Mohamed S, Zhang J, Chang Q, Ou Y, et al. Insight in inpatients with schizophrenia: relationship to symptoms and neuropsychological functioning. *Schizophr Res*. 2015 Feb; 161(2–3):376–81. <https://doi.org/10.1016/j.schres.2014.12.009> PMID: 25533592
28. Busner J, Targum SD. The clinical global impressions scale: applying a research tool in clinical practice. *Psychiatry (Edgmont)*. 2007; 4:28–37. PMID: 20526405

29. Addington D, Addington J, Schissel B. A depression rating scale for schizophrenics. *Schizophr Res*. 1990 Jul; 3(4):247–51. [https://doi.org/10.1016/0920-9964\(90\)90005-r](https://doi.org/10.1016/0920-9964(90)90005-r) PMID: 2278986
30. Keefe RSE, Bilder RM, Harvey PD, Davis SM, Palmer BW, Gold JM, et al. Baseline neurocognitive deficits in the CATIE schizophrenia trial. *Neuropsychopharmacology*. 2006 Sep; 31(9):2033–46. <https://doi.org/10.1038/sj.npp.1301072> PMID: 16641947
31. Heinrichs DW, Hanlon TE, Carpenter WT Jr. The Quality of Life Scale: an instrument for rating the schizophrenic deficit syndrome. *Schizophr Bull*. 1984; 10(3):388–98. <https://doi.org/10.1093/schbul/10.3.388> PMID: 6474101
32. Cramer JA, Rosenheck R, Xu W, Thomas J, Henderson W, Charney DS. Quality of life in schizophrenia: a comparison of instruments. Department of Veterans Affairs Cooperative Study Group on Clozapine in Refractory Schizophrenia. *Schizophr Bull*. 2000; 26(3):659–66. <https://doi.org/10.1093/oxfordjournals.schbul.a033484> PMID: 10993404
33. acArthur Competence Assessment Tool-Clinical Research (MacCAT-CR). Competence Assessment Tool-Clinical Research (MacCAT-CR). 1996;
34. Carpenter WT Jr, Gold JM, Lahti AC, Queern CA, Conley RR, Bartko JJ, et al. Decisional capacity for informed consent in schizophrenia research. *Arch Gen Psychiatry*. 2000 Jun; 57(6):533–8. <https://doi.org/10.1001/archpsyc.57.6.533> PMID: 10839330
35. Tierney NJ, Cook DH. Expanding tidy data principles to facilitate missing data exploration, visualization and assessment of imputations [Internet]. arXiv [stat.CO]. 2018. Available from: <http://arxiv.org/abs/1809.02264>
36. Stekhoven D. missForest: missForest is a nonparametric, mixed-type imputation method for basically any type of data for the statistical software R [Internet]. Github; [cited 2022 Jul 28]. Available from: <https://github.com/stekhoven/missForest>
37. rms: Regression Modeling Strategies [Internet]. Comprehensive R Archive Network (CRAN). [cited 2023 Mar 22]. Available from: <https://cran.r-project.org/web/packages/rms/>
38. Linear Mixed-Effects Models using “Eigen” and S4 [R package lme4 version 1.1–32]. 2023 Mar 14 [cited 2023 Mar 22]; Available from: <https://cran.r-project.org/web/packages/lme4/index.html>
39. Akinhanmi MO, Biernacka JM, Strakowski SM, McElroy SL, Balls Berry JE, Merikangas KR, et al. Racial disparities in bipolar disorder treatment and research: a call to action. *Bipolar Disord*. 2018 Sep; 20(6):506–14. <https://doi.org/10.1111/bdi.12638> PMID: 29527766
40. Fatumo S, Chikowore T, Choudhury A, Ayub M, Martin AR, Kuchenbaecker K. A roadmap to increase diversity in genomic studies. *Nat Med* [Internet]. 2022 Feb 10; Available from: <https://doi.org/10.1038/s41591-021-01672-4> PMID: 35145307
41. Scharff DP, Mathews KJ, Jackson P, Hoffsuemmer J, Martin E, Edwards D. More than Tuskegee: understanding mistrust about research participation. *J Health Care Poor Underserved*. 2010 Aug; 21(3):879–97. <https://doi.org/10.1353/hpu.0.0323> PMID: 20693733
42. Hartz SM, Johnson EO, Saccone NL, Hatsukami D, Breslau N, Bierut LJ. Inclusion of African Americans in genetic studies: what is the barrier? *Am J Epidemiol*. 2011 Aug 1; 174(3):336–44. <https://doi.org/10.1093/aje/kwr084> PMID: 21633120
43. Moldovan R, McGhee KA, Coviello D, Hamang A, Inglis A, Ingvoldstad Malmgren C, et al. Psychiatric genetic counseling: A mapping exercise. *Am J Med Genet B Neuropsychiatr Genet*. 2019 Dec; 180(8):523–32. <https://doi.org/10.1002/ajmg.b.32735> PMID: 31222934
44. Costain G, Esplen MJ, Toner B, Scherer SW, Meschino WS, Hodgkinson KA, et al. Evaluating genetic counseling for individuals with schizophrenia in the molecular age. *Schizophr Bull*. 2014 Jan; 40(1):78–87. <https://doi.org/10.1093/schbul/sbs138> PMID: 23236078
45. Hippman C, Lohn Z, Ringrose A, Inglis A, Cheek J, Austin JC. “Nothing is absolute in life”: understanding uncertainty in the context of psychiatric genetic counseling from the perspective of those with serious mental illness. *J Genet Couns*. 2013 Oct; 22(5):625–32. <https://doi.org/10.1007/s10897-013-9594-2> PMID: 23604904
46. Domino ME, Swartz MS. Who are the new users of antipsychotic medications? *Psychiatr Serv*. 2008 May; 59(5):507–14. <https://doi.org/10.1176/ps.2008.59.5.507> PMID: 18451006