**Background**

Multi-center studies are increasingly necessary to advance psychiatric neuroscience and evaluate novel interventions for patients with schizophrenia. However, generalizing results even from large, well-powered cohorts to usual-care clinical contexts remains challenging. Of the many factors that contribute to this generalizability barrier include race, gender, and ethnicity disparities in schizophrenia research. Here, we provide a novel way of quantifying the disparity within a large multi-site study and describe the extent of additional sampling necessary to achieve racial, gender, and ethnic parity in schizophrenia research.

**Methods**

The analysis cohort included 841 healthy control subjects (HCS) and 1060 subjects with schizophrenia (SZ) between ages 18 and 65 from four of the five sites in the Consortium on the Genetics of Schizophrenia 2 (COGS2) study: Los Angeles, New York, Philadelphia, and Seattle. We generated a Diversity Index using methodology from National Equity Atlas and previously published entropy index calculation to characterize the diversity within this sample. We created 24 categories based on race (White, African-American, American Indian/Alaska Native, Asian, Pacific Islander/Native Hawaiian, and Mixed/Other), ethnicity (Hispanic/Latino or non-Hispanic/Latino), and gender (Male/Female) based on demographic classification of COGS2 and generated Diversity Index values for both HCS and SZ subjects for each city. These proportions and Diversity Indexes were compared to those generated from American Community Survey (ACS) census data for each of the above cities from 2010-2014, limited to individuals between ages 18 and 65. We carried out a simulation algorithm which 1) randomly resampled COGS2 HCS and SZ cohorts, separately, to identify how many resamples would be required to approximate the ACS Diversity Index of COGS2 study sites, and 2) randomly resampled the SZ cohort to identify how many resamples would be required to approximate the HCS Diversity Index in COGS2. All simulations were repeated 1,000 times and each resample was set to 50% of the COGS2 recruitment cohort.

**Results**

Across COGS2, non-Hispanic African-Americans males, non-Hispanic Caucasian males, non-Hispanic Native Hawaiian/Pacific Islander males, and non-Hispanic Mixed/Other males were significantly overrepresented (Holm correction, p < 0.05) in HCS and SZ cohorts compared to ACS demographic data, while Hispanic/Latino Caucasian females and Hispanic/Latino Mixed/Other were significantly underrepresented (Holm correction, p < 0.05). In HCS, non-Hispanic Caucasian females and non-Hispanic Native Hawaiian/Pacific Islander females were significantly overrepresented (Holm correction, p < 0.05), while non-Hispanic Asian-American males were significantly underrepresented (Holm correction, p < 0.05). In SZ, non-Hispanic African-American females and Hispanic/Latino African-American males were significantly overrepresented (Holm correction, p < 0.05), while non-Hispanic Asian-American females and non-Hispanic Caucasian females were significantly underrepresented compared to ACS data (Holm correction, p < 0.05).

On average the HCS cohort required 22.7 resamples (Standard Deviation: 6.27) to approximate background ACS race, ethnicity and gender proportions of the cities in which they were conducted, while SZ required 46.8 resamples (Standard Deviation: 11.0). On average, 17.6 resamples were required (Standard Deviation: 6.69) for the SZ to approximate the HCS race, ethnicity and gender proportions in COGS2.

**Conclusion**

Data implies that current recruitment and ascertainment strategies would require the cohort of COGS2 to be nearly 12.5 times larger for HCS and 24.5 times larger for SZ to achieve racial, ethnic, and gender parity with the cities in which COGS2 was conducted. The SZ cohort in COGS2 would need to be 10 times larger to approximate racial, ethnic and gender parity with the recruited COGS2 HCS cohort. This methodology provides 1) a novel way to quantify and compare diversity of recruited cohorts between and across large SZ studies, 2) a useful metric to periodically assess diversity in longitudinal studies while they are ongoing, and 3) an objective measure which can be used by funders and policy makers to help allocate additional resources in order enhance representation in multisite studies in SZ.