Gender, Age and Geographical Representation over the Past 50 Years of Schizophrenia Research

Luz María Alliende[a,b], Leticia Czepielewski[c,d], David Aceituno[a, e], Carmen Paz Castañeda[f], Camila Diaz[g], Barbara Iruretagoyena[a], Carlos Mena[a,h], Cristian Mena[f,i], Juan Pablo Ramirez-Mahaluf[a], Ángeles Tepper[a], Javiera Vasquez[a], Lais Fonseca[j,k], Vivianne Machado[j], Camilo E. Hernández[l], Cristian Vargas-Upegui[l], Gladys Gomez-Cruz[m], Luis F. Kobayashi-Romero[m], Tomas Moncada-Habib[m], Sara Evans-Lacko[n], Rodrigo Bressan[j], Clarissa Gama[o,p], Carlos Lopez-Jaramillo[l], Camilo de la Fuente-Sandoval[m], Alfonso Gonzalez-Valderrama[f,i], Juan Undurraga[f,q], Ary Gadelha[j,k], Nicolás A. Crossley[a,r,*], for the ANDES Network.

a) Department of Psychiatry, School of Medicine, Pontificia Universidad Católica de Chile, Chile.
b) Department of Psychiatry and Behavioral Neuroscience, The University of Chicago, USA.
c) Laboratory of Molecular Psychiatry, Centro de Pesquisa Clínica, Hospital de Clínicas de Porto Alegre, Porto Alegre, Brazil.
d) Programa de Pós-Graduação em Psicologia, Instituto de Psicología.Universidade Federal do Rio Grande do Sul (UFRGS), Porto Alegre, Brazil.
e) Psychiatry and Mental Health Service, Complejo Asistencial Dr. Sótero del Río, Puente Alto, Chile.
f) Early Intervention Program, Instituto Psiquiátrico José Horwitz Barak, Chile.
g) Pharmacovigilance Program, Instituto Psiquiátrico José Horwitz Barak, Chile.
h) Institute of Cognitive Neuroscience, University College London, London, UK.
i) School of Medicine, Universidad Finis Terrae, Santiago, Chile
j) Interdisciplinary Laboratory in Clinical Neuroscience (LiNC), Department of Psychiatry, Federal University of Sao Paulo, Brazil

k) Schizophrenia Program (PROESQ), Department of Psychiatry, Federal University of Sao Paulo, Brazil.

l) Department of Psychiatry, Faculty of Medicine, University of Antioquia, Colombia.

m) Laboratory of Experimental Psychiatry, Instituto Nacional de Neurología y Neurocirugía, Mexico.

n) Care Policy and Evaluation Centre, London School of Economics and Political Science, UK.

o) Laboratory of Molecular Psychiatry, Hospital de Clínicas de Porto Alegre (HCPA), Porto Alegre, RS, Brazil.

p) Graduate Program in Psychiatry and Behavioral Sciences, Departamento de Psiquiatria e Medicina Legal, Universidade Federal do Rio Grande do Sul (UFRGS), Porto Alegre, RS, Brazil.

q) Department of Neurology and Psychiatry. Faculty of Medicine, Clínica Alemana Universidad del Desarrollo. Santiago, Chile.

r) Department of Psychosis Studies, Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, UK.

* = corresponding author, Nicolas Crossley, Escuela de Medicina UC, Diagonal Paraguay 362, Santiago, Chile. ncrossley@uc.cl.

Highlights

- We here examined changes in geographic location, gender and age of participants, in schizophrenia studies from a large database of over 3,000 studies published in the last 50 years.

- We found that studies from low- and middle-income countries have been increasing since the 1990s, although are still under-represented when considering the size of the world's population they represent.

- Women have been historically under-represented in schizophrenia studies. This is less of a problem in low- and middle-income countries, particularly China.

- Women's under-representation in studies has been improving significantly, and we could expect upper-income countries to reach representativeness within the next decade.

- There has not been any significant change in time in the mean age of participants.

Abstract

Previous studies have suggested that subjects participating in schizophrenia research are not representative of the demographics of the global population of people with schizophrenia, particularly in terms of gender and geographical location. We here explored if this has evolved throughout the decades, examining changes in geographical location, gender and age of participants in studies of schizophrenia published in the last 50 years. We examined this using a meta-analytical approach on an existing database including over 3,000 studies collated for another project. We found that the proportion of studies and participants from low-and-middle income countries has significantly increased over time, with considerable input from studies from China. However, it is still low when compared to the global population they represent. Women have been historically under-represented in studies, and still are in high-income countries. However, a significantly higher proportion of female participants have been included in studies over time. The age of participants included has not changed significantly over time. Overall, there have been improvements in the geographical and gender representation of people with schizophrenia. However, there is still a long way to go so research can be representative of the global population of people with schizophrenia, particularly in geographical terms.

Keywords: schizophrenia, diversity, gender, age, global diversity, low-and-middle-income countries, high-income countries

1. Introduction

Over the past 50 years the landscape of schizophrenia research has changed considerably. Similar to trends in other research areas, the number of studies focused on schizophrenia have close to doubled every 9 years (Lammer et al, 2018; Noorden, 2014). However, it remains unclear whether this increase in the number of publications has involved changes in the demographic composition of study participants' age and gender and their country of origin. This is particularly important since subjects participating in schizophrenia studies might not be representative of the demographics of the global population of people with schizophrenia (Paketci, 2020; Patel & Sumathipala, 2001). Furthermore, exposure to social determinants in mental health, such as the effects of poverty, urbanicity, and community violence, distributes unevenly between geographical regions, gender, and race (Crossley, 2018; Burkard, 2021). This lack of diversity limits research generalizability and highlights the unequal distribution of medical advances. There are ongoing calls towards greater representation and diversity in study participant's gender and background for all disciplines (Nature Methods, 2021).

Echoes of this calls towards greater diversity in research participants have reached schizophrenia research. There have been efforts and calls to breach gaps in representation in schizophrenia research participants, who are overwhelmingly from western, educated, industrialized, rich, and democratic societies (Burkhard et al, 2021). In terms of gender representation, there is a reported gap in the epidemiological occurrence of schizophrenia between females and males and the proportion of male and female participants included in non-epidemiological studies, with men being overrepresented in studies beyond what would be representative of epidemiological data (Longenecker et al, 2010). Significant efforts have been put in place to breach the historical gender divide in participants in health research- including institutional measures such as placing National Institute of Health (NIH) guidelines for gender inclusion for grant approval (NIH, 2021). This determines that women should be included in research studies in a way that can ensure that research findings can be generalized to the entire population being studied. Similarly, a considerable gap between psychiatric research from Euro-American countries and the rest of the world has also been found, with considerably less research being published from developing nations (Patel & Sumathipala, 2001). Additionally, driven by early intervention approaches, a recent focus has been placed on the study of the early stages of

schizophrenia and first episode patients (McGorry, Killackey & Yung, 2008). Some people have suggested that the development of early intervention services might be at the expense of other psychosis services (Castle & Singh, 2015). One could also hypothesize that this is also mirrored in research studies, with more recent ones including younger subjects, with less representation of older subjects.

The present work analyzed how these representation gaps, specifically those related to gender, geographical location, and age, have changed in the past 50 years in schizophrenia research. We used a large existing database of more than 3,000 schizophrenia studies published in the last 50 years, collated for a systematic review of educational attainment in schizophrenia. This database included studies from different fields, ranging from neuroimaging to genetics. We hypothesized that representation from low- and middle-income countries, women, and younger participants, would have increased over time.

2. Methods

2.1 Search and study selection

The present study was conducted using data from an ongoing metanalysis on educational achievement in people with schizophrenia. Its protocol is registered in the International prospective register of systematic reviews (PROSPERO) under the identification number (CRD42020220546).

That project explored educational attainment in subjects with schizophrenia, focusing on years of education as the main outcome. It included a wide range of studies, as long as they reported: a) data from subjects deemed to be suffering from schizophrenia based on a validated diagnostic criteria (such as DSM or ICD), b) included adults (mean age over 18 years old), c) reported their educational attainment expressed as years of education (and not in any other form such as high-school or completed A-levels) with a measure of its variance, and d) included the average age of the participants. Studies in English, Spanish, Portuguese, German and French were considered. There were no restrictions on the type of study included.

The original search was performed on the 24th November 2020, and was conducted using the following search terms “Schizophrenia [mesh terms] OR Schizophren* AND "educational status"[mesh] OR "years of education" OR schooling” in PUBMED and PsycINFO. It was limited to publications between 1970 and 2020. The initial hits were subdivided among 22 of the co-authors, and studies considered eligible, and their data extraction, were reviewed by another author. Duplicates of the same data published in two different articles were excluded.

The PRISMA flow diagram is presented in Figure 1. From 32,593 initial possible hits, 3,278 publications reporting on 3,321 samples (as some publications reported on more than one) were considered to fulfill selection criteria and included.

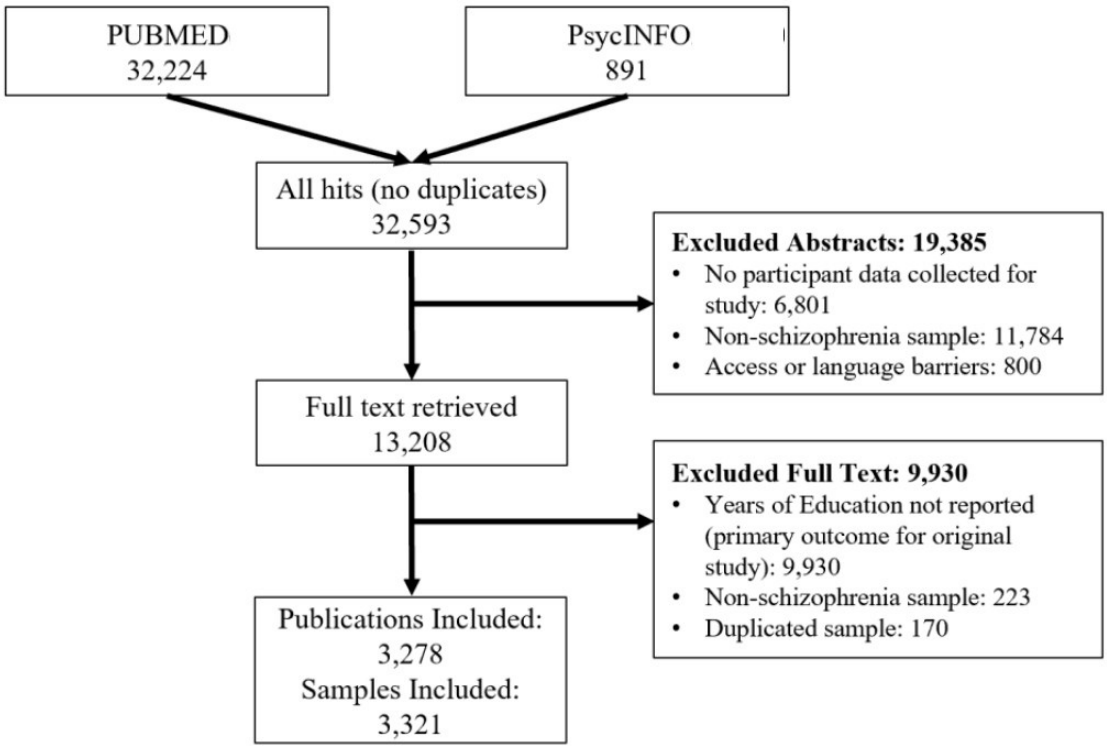

Figure 1. Abstract and full text article selection process.

2.2 Variable Operationalization

Countries of publication were classified as either part of high-income countries according to the World Bank´s 2022 income classification or as part of any other income bracket (upper-middle income, lower-middle-income, and low income). Country of publication was also classified by their geographical region (East Asia and Pacific, Europe and Central Asia, Latin America and the Caribbean, Middle East and North Africa, North America, South Asia, Sub-Saharan Africa). Gender in each sample was determined as the raw proportion of female participants out of the total number of participants. The age variables for each sample were summarized as the mean age and standard deviation of age of all schizophrenia participants.

2.3 Analyses

To analyze the representation of people with schizophrenia from different income regions, the total number of studies per year (from 1971 to 2020) for high income and low-and-middle-

income country categories were compared using chi-squared. The total number of participants included in studies for each year was also compared between those two groups using an independent sample t-test. To examine changes with time of those two variables (numer of studies per year, and total number of participants per year), we used a general linear model incorporating World Bank Classification (upper or low-middle-income), publication year, and their interaction as regressors.

To determine whether gender representation had changed over time, we performed a meta-analysis and meta-regression using a random-effect model, weighted according to the inverse of the variance of the studies included, and fitted using a maximum-likelihood estimator with R's *metafor* package (Viechtbauer, 2010). We built two models. First, to determine the trajectory of gender representation over time, a model was constructed using year of publishing as a moderator. A second model was constructed including publication year, the World-Bank classification of the country of the study (upper or low-and-middle-income), as well as their interaction.

The same approach was used to analyze the effect of year of publication on the age of the participants. This involved a first simple meta-analytical model including only year of publication as a predictor, followed by a second model also including the income classification of the country of publication and their interaction.

Lastly, given the large number of studies from low-and-middle-income countries originating from China (63.44%), analyses were replicated without studies from this country to corroborate whether any observed effects were driven by this country only.

3. Results

This study included a database of 3,321studies including a total of 317,385 participants. We now describe the characteristics and changes across time of these studies.

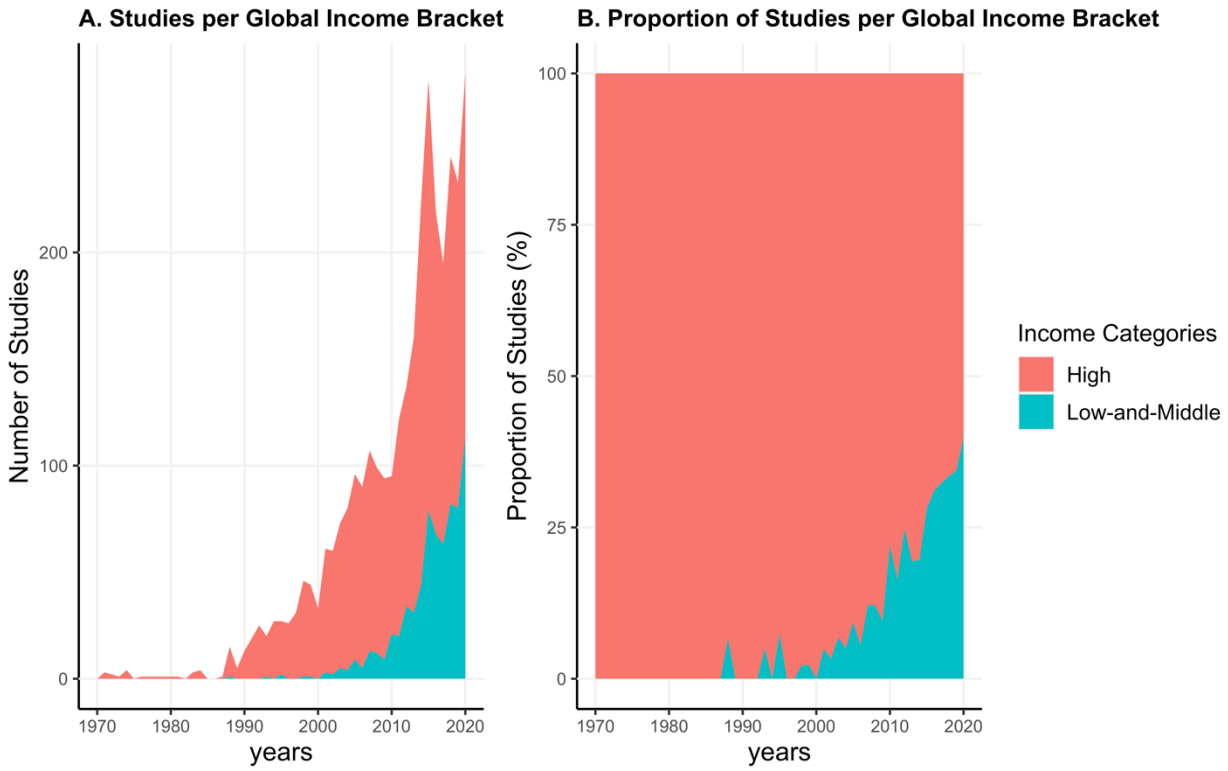

Figure 2. A. Proportion of all studies from countries classified as high-income or upper-middle income, lower-middle-income, or low-income countries (low-and-middle-income countries). B. Proportion of all studies from each of the World Bank´s geographical regions.

3.1 Geographical representation in studies

3,300 studies included information about the country where the study was performed. There was a significantly larger number of publications from high-income countries (N=2,607, 79% of all studies included; $\chi^2(1) = 1110.1$, $p < 10^{-15}$). Most participants were also from the group of high-income countries (*N*= 228,708 (72.66%): $\chi^2(1) = 64635$, $p < 10^{-15}$) in this dataset. To put this into context, the total world's population living in high-income countries as of 2020

corresponds only to 15.73% of the world population (The World Bank, 2020). The number of studies from each income bracket and geographical area over time can be seen in Figure 2.

There was a significant effect of time on the proportion of studies from middle-and-lower income countries over the total number of studies published each year, with an increase in the proportion of 0.1 percentage points per year (t-value=8.57, $p<10^{-10}$; Figure 2). However, this was clearly a non-linear relationship, with the proportion of published studies from low- and middle-income countries being almost zero until the 1990's, and then representing between 27.8% to 39.3% in the last 5 years (2015-2020). The same pattern was found for total number of participants from low- and middle- income countries over the total number of patients with schizophrenia included each year. A linear regression was also conducted using the proportion of participants (instead of studies) per year from low-and-middle-income countries as a variable of interest. There was a significant effect of time on the proportion of participants from low-and-middle-income countries, with a 1 percentage point increase per year (t=8.57, p<0.001).

Subsequent analyses showed that the increase in studies and participants from low-and-middle-income countries was largely driven by China. When studies from China were excluded, the proportion of studies from lower income countries increased by only 0.23 percentage points each year. The increase over time in participants from low-and-middle-income countries remained significant when removing studies from China (t=7.62, p<0.001).

3.2 Gender Representation

Out of the whole sample 34.83% of participants were women. When running our simpler model including year of publication only, there was a significant effect of time on gender proportion (z=10.73, $p <10^{-26}$, with significant heterogeneity I^2=98.34%). In 2014 - the median publication year for all studies- 35.49% of all participants were female, and each additional year showed a 0.41 percentage point increase. To put this in context, a meta-analysis from 2003 (Aleman, Kahn & Selten) exploring the risk of schizophrenia and sex, suggested that the proportion of women with schizophrenia was 41.32% (95% CI=39.06% to 42.48%). If ongoing trends continue, the lower end of this study's confidence interval would be reached in about 9 years.

When the income bracket of the study's publication was included as an additional moderator in the model, as well as its interaction with year of publication, the effect of time remained significant but slightly smaller than in the simpler model (β=0.003, z=8.03, p $<10^{-15}$). The income bracket for the country of publication also showed a significant effect on the proportion of females in studies. Low-and-middle-income countries showed a significantly higher proportion of women in their studies. In 2014, low-and-middle-income studies had on average 39.51% female participants, while high-income countries had an average of 34.1% female participants ((z=6.92, p $<10^{-11}$). The interaction between publication year and income bracket was not significant (p =0.84), suggesting women's participation in research was growing at a similar pace across the world. However, the difference between regions meant that representativeness (according to the Aleman meta-analysis) was achieved in studies from low-and-middle-income countries more than 5 years ago, while it will take around 10 years more to achieve that in upper-income countries. The results of the model can be seen in Figure 3.

The effect on time on the proportion of female participants remained when studies from China where excluded (β=0.003, z=8.12, p $<10^{-15}$). When country of publication's income bracket was included in the model as well as its interaction with year of publication, the effect of country of publication's income remained significant (β=0.03, z=2.51, p =0.012). The effect of time using this more comprehensive model remained significant (β=0.003, z=8.05, p $<10^{-15}$), but also the interaction between time and country of publication's income bracket (β=-0.005, z=-2.28, p=0.023). This meant that low- and middle-income countries (excluding China) presented a decrease in participation of women in research.

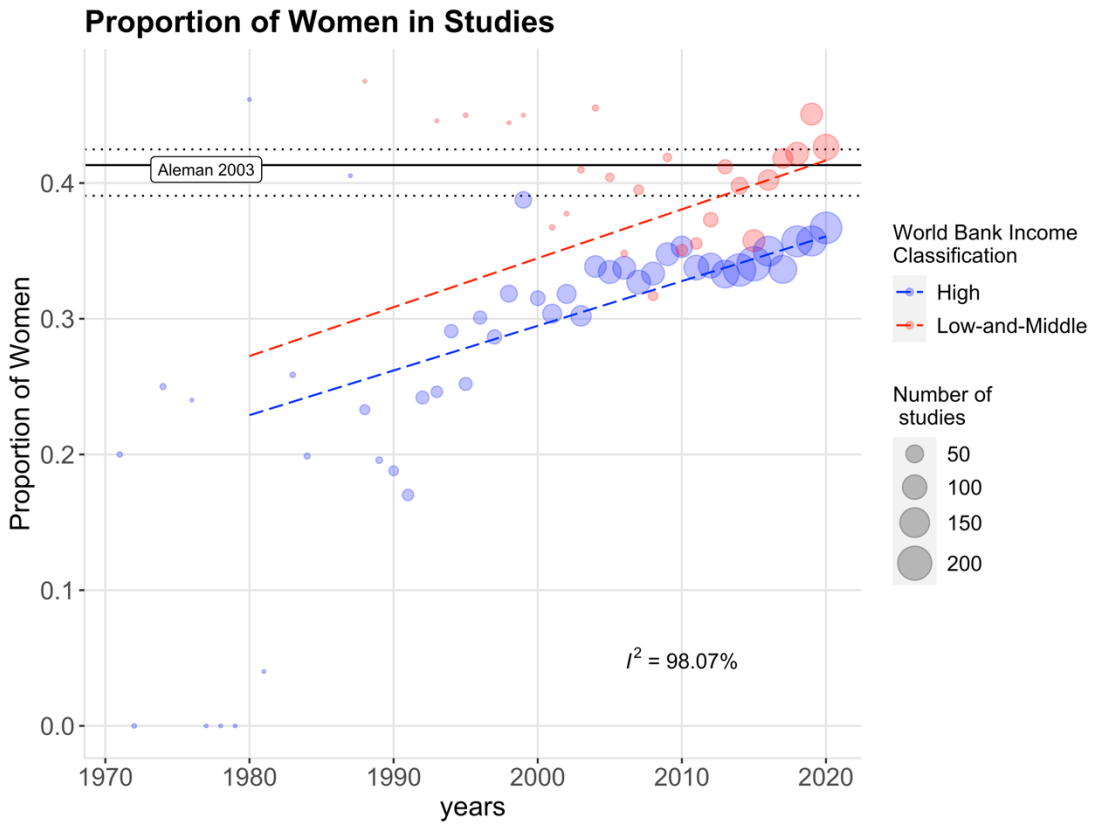

Figure 3. Annual proportion of women in schizophrenia in research papers from 1971-2020 according to the country of publication's World Bank income classification. The middle horizontal line reports the expected proportion of women (41.32%) from the global population of people with schizophrenia (with the 95% confidence interval in dashed lines), estimated in the Aleman, Kahn and Selten's 2003 metanalysis' on epidemiological studies. Dashed lines in red and blue represent the meta-analytic model including all studies, alongside the residual variance of the model (I^2).

Another way of looking at gender bias is to look at proportion of studies including single-gender participants, either male or female. Out of all studies, 217 reported no female participants while only reported 17 no male participants. 84.79% of the studies with no female participants where from high-income countries and 88.24% of studies with no male participants were from high-income countries. See Figure 4 for a visualization of the proportion of studies over the years that have no female and no male participants.

Figure 4. Percentage of overall studies with no male or no female participants over time.

3.3 Age

Participants across studies had a mean age of 37.08 years (sd= 9.05). The meta-analytical model that only included time as a predictor on participants age was found to be significant (decrease of 0.09 on average age per year, z=-4.64, p<0.00001, I^2=99.17%). However, when including the income bracket for the study's country of publication as a moderator for the model, as well as its interaction with the paper's year of publication, the effect of time did not remain significant (β=-0.03, z=-1.55, p=0.12, model with an I^2=99.13%). In this new model the country of publication's income bracket was significant (β=-4.33, z=-10.79, p<10^{-26}), and the interaction between country´s income bracket and year of publication was not significant (β=0.04, z=0.46, p=0.64). In other words, the initial analysis showing a significant effect of time was a result driven by the younger age of participants in low- and middle-income countries, and the increase of studies in time from these countries reported previously. Figure 5 shows results from this meta-analytic model.

Results were unchanged when studies from China were excluded. There was an initial effect of time on participants age when no other variables were included (β=-0.04, z=-2.16, p=0.03), which was no longer significant after including the country's income bracket that was

driving most of the effect (β=-4.09, z=-7.01, $p<10^{-11}$), with no significant interaction between income bracket and time.

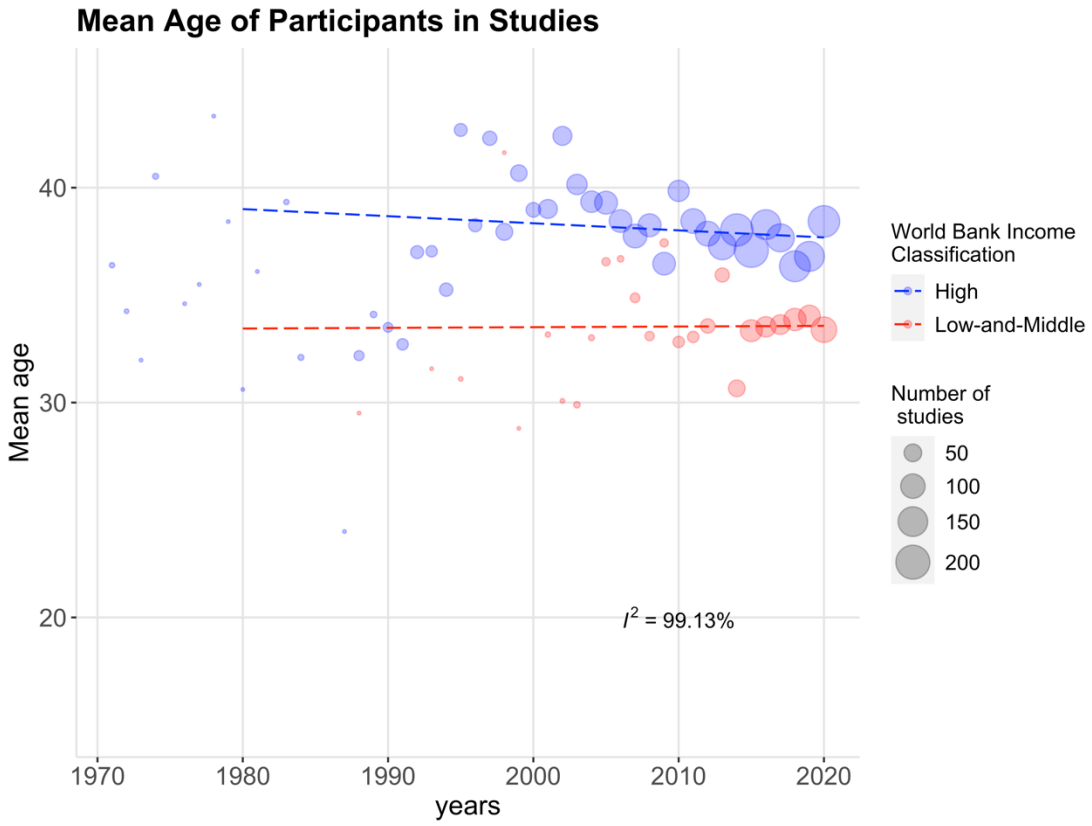

Figure 3. Average age of participants per year in schizophrenia research papers between 1971-2020 according to the country of publication's World Bank income classification. Dashed lines in red and blue represent the meta-analytic model including all studies, alongside the residual variance of the model (I^2). As can be seen, participants in low- and middle-income countries are younger, which has not changed much through the decades.

4. Discussion

Research studies in schizophrenia should include diverse participants representative of the global population of people with schizophrenia in terms of gender, age and country of origin to ensure that research findings are relevant for diverse populations. By analyzing a large database of research studies published in the last 50 years, we found that inequality is still present in terms of geography, gender and age.

In terms of geographical region, considering that research studies cost money, it is unsurprising to find that upper-income countries are more represented in the literature. However, the magnitude is nevertheless striking, with overall only 21% of participants belonging to regions of the world that correspond to 84% of the population. We did find an increase in this representativeness, which is in line with previous reports (Large et al., 2010), that has mostly been driven by China.

Our results also confirmed an existing gender bias, with women being under-represented in the total corpus of research. Our data also show that studies from low and middle-income countries recruit a proportion of women that is currently within the one expected in the global population of patients. Our sub-group analysis excluding China showed that this was mostly driven by this country. This is not the case for upper-income countries, where there is still a significant under-representation of women in research participants. Encouragingly, as our data show, there have been gains in the last years across countries, and within a decade we might expect that the gender proportion of studies in high-income studies will also be representative of the population of people with schizophrenia.

In terms of age representativeness, although we initially saw an effect of time, this was mostly driven by younger participants from low- and middle-income countries, and the increase in their participation over time. This is expected considering their demographic composition (The World Bank, 2020). Potential concerns about a neglecting effect on older population due to the interest in the early stages of the disorder, particularly the critical period (Birchwood, Todd & Jackson, 1998), seem unsubstantiated.

This study is limited by the nature of the database used. The original search terms and inclusion criteria were focused on educational achievement in schizophrenia. This in itself should

not be related to gender, geographical areas, or age. However, there might still be a potential for bias in the way different countries, with different educational systems, report education in their studies. For example, under-estimating the literature from certain countries in which other measures of educational attainment other than completed years of education are used (for example, GCSEs in the UK or High school in the USA). However, the database used included more than 3,000 studies that allowed us, with the limitation described above, to provide a picture of the evolution of representativeness. Further research should consider creating a new database with more general search terms to avoid these types of biases.

In conclusion, we here show that there are historical and current gaps in the representation of women and developing nations in SZ research, with no changes in the age of participants. These gaps have improved over the past 50 years but there remains to be considerable underrepresentation of both groups. Further efforts must be made to increase the representativeness of the global population in research.

5. Funding acknowledgements

This collaboration was made possible thanks to the Programa Iberoamericano de Ciencia y Tecnología para el Desarrollo CYTED Redes to ANDES (218RT0547).

Camilo de la Fuente-Sandoval was supported by Consejo Nacional de Ciencia y Tecnología - Sistema Nacional de Investigadores, Mexico, (CONACyT - SNI); and National Institutes of Health Grant Nos. R21 MH117434 and R01 MH110270. Nicolas Crossley and Juan Undurraga were supported by ANID-PIA-ACT192064, ANID-FONDECYT 1180358, 1200601. Juan Pablo Ramirez-Mahaluf was supported by ANID-FONDECYT post doctorate funding 3190311.

Conflict of Interest

Nicolas Crossley has received personal fees from Janssen, outside the submitted work. Ary Gadelha has been a consultant and/or advisor to or has received honoraria from: Aché, Daiichi-Sankyo, Torrent, Bayer, Cristalia, Daiichi-Sankyo and Janssen. None of the other authors has potential conflicts of interest to be disclosed.

References

Accounting for diversity. (2021). Nature methods, 18(6), 577. https://doi.org/10.1038/s41592-021-01193-5

Aleman, A., Kahn, R. S., & Selten, J. P. (2003). Sex differences in the risk of schizophrenia: evidence from meta-analysis. *Archives of general psychiatry*, 60(6), 565–571. https://doi.org/10.1001/archpsyc.60.6.565

Birchwood, M., Todd, P., & Jackson, C. (1998). Early intervention in psychosis. The critical period hypothesis. *The British journal of psychiatry*. Supplement, 172(33), 53–59.

Burkhard, C., Cicek, S., Barzilay, R., Radhakrishnan, R., & Guloksuz, S. (2021). Need for Ethnic and Population Diversity in Psychosis Research. *Schizophrenia bulletin*, 47(4), 889–895. https://doi.org/10.1093/schbul/sbab048

Castle, D., & Singh, S. (2015). Early intervention in psychosis: Still the 'best buy'? British Journal of Psychiatry, 207(4), 288-292. doi:10.1192/bjp.bp.114.160630

Crossley, N. A., Alliende, L. M., Ossandon, T., Castañeda, C. P., González-Valderrama, A., Undurraga, J., Castro, M., Guinjoan, S., Díaz-Zuluaga, A. M., Pineda-Zapata, J. A., López-Jaramillo, C., Reyes-Madrigal, F., León-Ortíz, P., de la Fuente-Sandoval, C., Czepielewski, L. S., Gama, C. S., Zugman, A., Gadelha, A., Jackowski, A., & Bressan, R. (2019). Imaging Social and Environmental Factors as Modulators of Brain Dysfunction: Time to Focus on Developing Non-Western Societies. Biological psychiatry. *Cognitive neuroscience and neuroimaging*, 4(1), 8–15. https://doi.org/10.1016/j.bpsc.2018.09.005

Lammer, E., Klingelhöfer, D., Bendels, M., Ohlendorf, D., Spallek, M., & Groneberg, D. A. (2018). Development of the global schizophrenia research under epidemiological and socio-economic influences. *Schizophrenia research*, *199*, 458–460. https://doi.org/10.1016/j.schres.2018.04.001

Large, M., Nielssen, O., Farooq, S., & Glozier, N. (2010). Increasing rates of psychiatric publication from low- and middle-income countries. *The International journal of social psychiatry*, 56(5), 497–506. https://doi.org/10.1177/0020764009102412

Longenecker, J., Genderson, J., Dickinson, D., Malley, J., Elvevåg, B., Weinberger, D. R., & Gold, J. (2010). Where have all the women gone?: participant gender in epidemiological and non-epidemiological research of schizophrenia. *Schizophrenia research*, 119(1-3), 240–245. https://doi.org/10.1016/j.schres.2010.03.023

McGorry, P. D., Killackey, E., & Yung, A. (2008). Early intervention in psychosis: concepts, evidence and future directions. *World psychiatry : official journal of the World Psychiatric Association (WPA)*, *7*(3), 148–156. https://doi.org/10.1002/j.2051-5545.2008.tb00182.x

National Institute of Health (March 23, 2021) Inclusion of Women and Minorities as Participants in Research Involving Human Subjects. NIH Central Resource for grants and Funding Information. https://grants.nih.gov/policy/inclusion/women-and-minorities.htm. Accessed September 2021.

Noorden, R.V., (2014). Natureblogs. Newsblog. URL: http://blogs.nature.com/news/2014/05/global-scientific-output-doubles-every-nine-years.html, Accessed September 2021.

Paketci S. (2020). Racial and ethnic inequalities in genetic studies of schizophrenia. *European neuropsychopharmacology : the journal of the European College of Neuropsychopharmacology*, *37*, 102. https://doi.org/10.1016/j.euroneuro.2020.07.006

Patel, V., & Sumathipala, A. (2001). International representation in psychiatric literature: survey of six leading journals. *The British journal of psychiatry : the journal of mental science*, *178*, 406–409. https://doi.org/10.1192/bjp.178.5.406

Viechtbauer, W. (2010). Conducting Meta-Analyses in R with the metafor Package. *Journal of Statistical Software*, *36*(3), 1–48. https://doi.org/10.18637/jss.v036.i03

The World Bank. (2020). Data Bank. World Bank Open Data. https://data.worldbank.org. Accessed: 9/29/2021