Schizophrenia Research 204 (2019) 404–406

Contents lists available at ScienceDirect

Schizophrenia Research

journal homepage: www.elsevier.com/locate/schres

Letter to the Editor

Implications of an admixed Brazilian population in schizophrenia polygenic risk score

The Psychiatric Genomics Consortium (PGC) has implemented a tool, called Polygenic Risk Score (PRS), which compiles data from hundreds to millions of common variants into a single measure, making it a valuable instrument to investigate the genetic risk of complex diseases, like schizophrenia (SCZ) (Dudbridge, 2013; Purcell et al., 2009). To calculate the SCZ-PRS, a reference sample must be defined, but most subjects in currently available samples are Caucasian, and doubts remain about the reliability and efficiency of SCZ-PRS in admixed samples (Birnbaum and Weinberger, 2017; Vassos et al., 2017). We tested if SCZ-PRS would differentiate patients with schizophrenia and healthy controls in a Brazilian sample, producing a higher mean PRS in patients. In addition, we investigated if interracial admixture influences the results.

We genotyped 177 patients with SCZ (from the Schizophrenia Program of Universidade Federal de São Paulo - UNIFESP) and 242 controls (HumanOmniExpress BeadChips – Illumina, USA, which interrogate up to 730,525 genetic markers). This study was approved by the Research Ethics Committee of the Universidade Federal de São Paulo/CEP – UNIFESP and all participants agreed their participation on the study signing an informed consent. The genomic imputation was performed by the Sanger Imputation Server using Eagle2 for phasing and the 1000 Genomes sample (version 3) as a reference panel. To generate the PRS we used PRSice software using the default options. The SCZ sample from PGC was set as a reference sample and our genomic imputed and non-imputed samples as targets.

Initially, we calculated the PRS for all the 419 individuals without imputation (649,401 SNPs), which could explain 6.8% of the variance with p-value threshold of 0.5 (Supplementary Table 1, Supplementary Fig. 1A). The same was done in our genomic imputed samples (10,704,150 SNPs) and the PRS tool could explain 5.2% of variance, with a p-value threshold of 0.001 (Table S1, Fig. S1B). Then, we explored the PRS distribution of the non-genomic imputation samples, along with the PRS distribution of five ethnicities from 1000 Genomes Project and observed that the Caucasian group showed a lower PRS mean when compared with the other groups, followed by Mexicans and Asians (Supplementary Fig. 2). The African group had the highest PRS mean while the Brazilian population showed an intermediate value with an overlap among the other ancestries.

Subsequently, we calculated PRS for only individuals who self-declared as Caucasian (red dots in the PC graph – Supplementary Fig. 3), excluding from the analysis those who self-declared as other descent. We found an explained variance of 9.2% without genomic imputation, with a 0.5 threshold (Supplementary Table 1; Supplementary Fig. 1C). In our genomic imputed samples, the PRS tool could explain 6.7% of variance, with a p-value threshold of 0.001 (Supplementary Fig. 1D).

Even considering only self-declared Caucasian individuals (BR_CEU), our sample continues to be highly heterogeneous. Thus, we used ENIGMA protocol to plot the principal components of our sample along with the 1000 Genomes sample to observe how individuals are distributed according to their ancestry (ENIGMA, 2013) and divided them into four different groups, arbitrarily, based on the PC graph: Caucasians (PC-selected), Latin 1, Latin 2, and Asians (see Supplementary Table 2). The explained variance and the p-value threshold of each group are described in Table S1 and the PRS distribution of the non-imputed samples can be seen in Fig. 1.

Lastly, we compared the PRS p-threshold mean between cases and controls for all analyses using a *t*-test in R Studio for each analysis (with and without genomic imputation). We found a statistical difference in the PC-selected Caucasians and Latin 2 groups, both with and without genomic imputation (p-values: Caucasians $= 8.4 \times 10^{-8}$ and 3.6×10^{-7}, respectively; p-values Latin 2 $= 2.7 \times 10^{-5}$ and 8.5×10^{-5}, respectively) (Supplementary Table 3).

We were able to observe the influence of an admixed population from our results. The larger GWAS and genomic imputation tools available for SCZ have been designed based on European, North America, and Asian population samples (Martin et al., 2017). But even in a small and admixed sample, we could to differentiate cases and controls with the PRS tool, even though the discriminative power of PRS differed significantly according to ancestry and to the use of imputed and non-imputed data.

Although producing similar results to those of PGC, that found an explained variance around 20%, our results may have been less effective because PGC have a very homogenous cohort from European and East Asian ancestry (Schizophrenia Working Group of the Psychiatric Genomics Consortium, 2014), while our cohort has a very mixed ancestry pattern.

Furthermore, we believe that our imputation (considering all individuals, self-declared Caucasians, Latin 1, and Latin 2 groups) may introduce more errors than expected for this approach (Huang et al., 2009), mainly in the phasing stage, since these group's results showed lower variance explained with imputed samples. We believe the reason of these results is due to the lack of representative of the Brazilian admixed pattern in the reference panel.

The best models to differentiate patients with schizophrenia and controls in our cohort was selecting only PC-selected Caucasians and PC-selected Asians, because we could diminish the heterogeneity of the sample. The PC-selected Asians is the most homogenous group in our cohort and we found a better result with our genomic imputed samples, corroborating the idea that PRS tool show better results with more homogeneous samples and for ancestries widely explored in the majority of studies.

In conclusion, new approaches should be developed to apply PRS with equal explained variance from those homogenous populations. Furthermore, we need to identify the differences between patients with schizophrenia and controls beyond PRS results, as environmental factors also contribute to the disease's onset (Rethelyi et al., 2013).

https://doi.org/10.1016/j.schres.2018.07.026

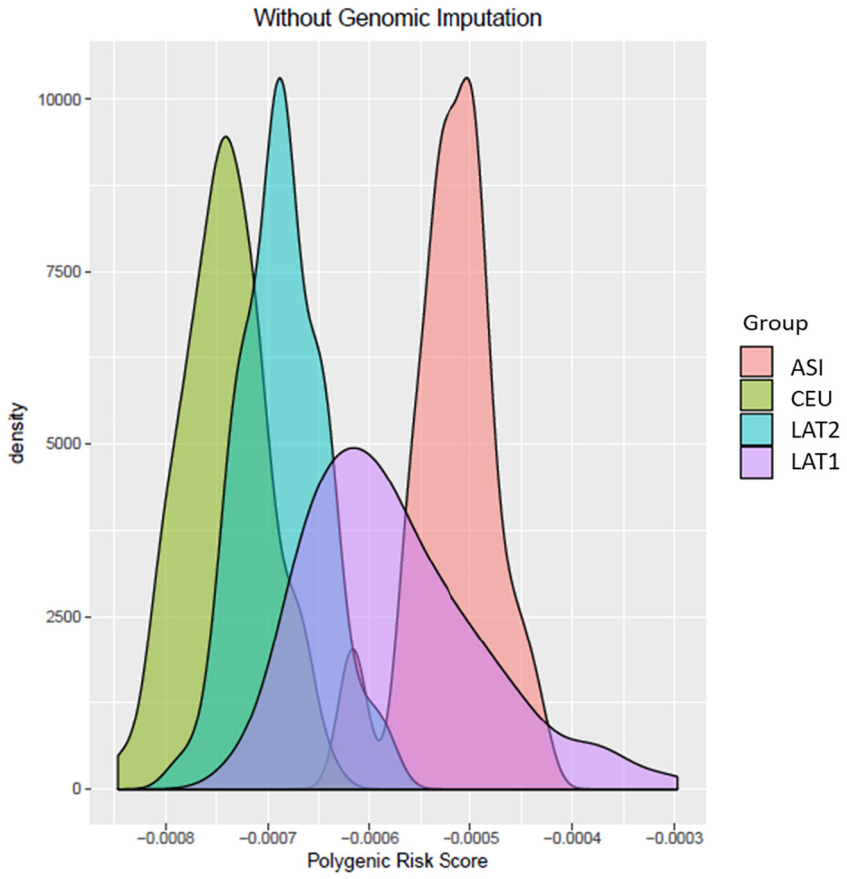

Fig. 1. PRS distribution of the non-imputed samples for the PC-selected Caucasian, Latin 1, Latin 2 and PC-selected Asians.

Conflict of interest

The authors declare that they have no conflict of interest.

Role of funding source

This study was supported by Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP - 2017/20059-0; 2010/08968-6; 2011/50740-5; 2014/50830-1; 2014/07280-1; 2016/04983-7).

Contributions

The contributions of each author to the paper are as follows: F. Talarico, M. Santoro, V. Ota, and S. Belangero designed the current study. A. Gadelha collected clinical data. F. Talarico and M. Santoro designed and implemented the statistical analysis and interpreted the results. F. Talarico wrote the first draft of the manuscript. All authors contributed to and have approved the final manuscript.

Acknowledgements

The authors would like to thank the patients, their families, the psychiatrists and nurses for their participation in this study and FAPESP for funding.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.schres.2018.07.026.

References

Birnbaum, R., Weinberger, D.R., 2017. Genetic insights into the neurodevelopmental origins of schizophrenia. Nat. Rev. Neurosci. 18, 727–740.
Dudbridge, F., 2013. Power and predictive accuracy of polygenic risk scores. PLoS Genet. 9, e1003348.
ENIGMA, C, 2013. ENIGMA 1KGP_p3v5 Cookbook.
Huang, L., Li, Y., Singleton, A.B., Hardy, J.A., Abecasis, G., Rosenberg, N.A., Scheet, P., 2009. Genotype-imputation accuracy across worldwide human populations. Am. J. Hum. Genet. 84, 235–250.
Martin, A.R., Gignoux, C.R., Walters, R.K., Wojcik, G.L., Neale, B.M., Gravel, S., Daly, M.J., Bustamante, C.D., Kenny, E.E., 2017. Human demographic history impacts genetic risk prediction across diverse populations. Am. J. Hum. Genet. 100, 635–649.
Purcell, S.M., Wray, N.R., Stone, J.L., Visscher, P.M., O'Donovan, M.C., Sullivan, P.F., Sklar, P., 2009. Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. Nature 460, 748–752.
Rethelyi, J.M., Benkovits, J., Bitter, I., 2013. Genes and environments in schizophrenia: the different pieces of a manifold puzzle. Neurosci. Biobehav. Rev. 37, 2424–2437.
Schizophrenia Working Group of the Psychiatric Genomics Consortium, 2014. Biological insights from 108 schizophrenia-associated genetic loci. Nature 511, 421–427.
Vassos, E., Di Forti, M., Coleman, J., Iyegbe, C., Prata, D., Euesden, J., O'Reilly, P., Curtis, C., Kolliakou, A., Patel, H., Newhouse, S., Traylor, M., Ajnakina, O., Mondelli, V., Marques, T.R., Gardner-Sood, P., Aitchison, K.J., Powell, J., Atakan, Z., Greenwood, K.E., Smith, S., Ismail, K., Pariante, C., Gaughran, F., Dazzan, P., Markus, H.S., David, A.S., Lewis, C.M., Murray, R.M., Breen, G., 2017. An examination of polygenic score risk prediction in individuals with first-episode psychosis. Biol. Psychiatry 81, 470–477.

Fernanda Talarico[1]
Departamento de Morfologia e Genética da Universidade Federal de São Paulo (UNIFESP), São Paulo, Brazil
Laboratório Interdisciplinar de Neurociências Clínicas (LiNC) do Departamento de Psiquiatria da UNIFESP, São Paulo, Brazil

Marcos Santoro[1]
Laboratório Interdisciplinar de Neurociências Clínicas (LiNC) do Departamento de Psiquiatria da UNIFESP, São Paulo, Brazil
Departamento de Psiquiatria da UNIFESP, São Paulo, Brazil

Vanessa K. Ota
Departamento de Morfologia e Genética da Universidade Federal de São Paulo (UNIFESP), São Paulo, Brazil

[1] Both authors contributed equally to this work.

Laboratório Interdisciplinar de Neurociências Clínicas (LiNC) do Departamento de Psiquiatria da UNIFESP, São Paulo, Brazil

Ary Gadelha
Laboratório Interdisciplinar de Neurociências Clínicas (LiNC) do Departamento de Psiquiatria da UNIFESP, São Paulo, Brazil
Departamento de Psiquiatria da UNIFESP, São Paulo, Brazil

Renata Pellegrino
Hakon Hakonarson
The Children's Hospital of Philadelphia, Philadelphia, United States of America

Rodrigo Bressan
Laboratório Interdisciplinar de Neurociências Clínicas (LiNC) do Departamento de Psiquiatria da UNIFESP, São Paulo, Brazil
Departamento de Psiquiatria da UNIFESP, São Paulo, Brazil

Jair J. Mari
Laboratório Interdisciplinar de Neurociências Clínicas (LiNC) do Departamento de Psiquiatria da UNIFESP, São Paulo, Brazil

Sintia Belangero
Departamento de Morfologia e Genética da Universidade Federal de São Paulo (UNIFESP), São Paulo, Brazil
Laboratório Interdisciplinar de Neurociências Clínicas (LiNC) do Departamento de Psiquiatria da UNIFESP, São Paulo, Brazil
Departamento de Psiquiatria da UNIFESP, São Paulo, Brazil
Corresponding author at: Genetics Division, Rua Botucatu, 720, Leitão da Cunha, 1st floor, 04023-900 São Paulo, Brazil.
E-mail address: sinogueira@gmail.com.

29 May 2018