



Guest editorial: Special issue on machine learning in schizophrenia

M. Mallar Chakravarty ^{a, b, c, *}

^a Computational Brain Anatomy (CoBra Laboratory), Douglas Institute, Montreal, Quebec, Canada

^b Department of Psychiatry, McGill University, Montreal, Quebec, Canada

^c Biological and Biomedical Engineering, McGill University, Montreal, Quebec, Canada

ARTICLE INFO

Article history:

Received 18 October 2019

Accepted 20 October 2019

Available online 9 November 2019

Keywords:

Schizophrenia

Machine learning

Artificial intelligence

Precision medicine

Personalized medicine

Precision psychiatry

Machine learning methods have been employed in various domains including natural language processing, the financial industry, and image recognition (LeCun et al., 2015; Bengio et al., 2013). Once considered a computational advance only of use to computer scientists and those possessing some realm of technical sophistication, increases in the availability of computing resources, the increased emphasis on multi-/trans-disciplinary research, and the increased availability of open source software packages (e.g.: Scikit-learn (<http://scikit-learn.org/stable/>) and TensorFlow (<https://www.tensorflow.org/>) libraries) has increased the use of machine learning methodologies in biomedical research (Abraham et al., 2014). While the move towards using machine learning in daily clinical practice will likely involve a phenomenal amount of data to adequately train and test generalizable models which properly parse through phenotypic heterogeneity (Bzdok; Yeo, 2017), to date there are no clear rules and guidelines with respect to how to generate new models. Importantly, there is still much to do with respect to understanding how to leverage these sophisticated methods in order to robustly investigate pathophysiological processes or use them in an effort to enhance practices related to precision or personalized medicine (Dwyer et al., 2018). Like many new technologies, it is clear that investigators are excited about the

promise of these techniques, but it is unclear how they can be implemented in heterogeneous disorders to provide believable predictions at the single-subject/patient level or as a means of elucidating complex genotype-phenotype relationships (although, as we will see below, the early results are promising).

The focus of this special issue is the use of machine learning methodologies as it pertains to schizophrenia and related psychosis. All of the studies included in this special issue have leveraged neuroimaging data (discussed further below). This is not altogether surprising, as there is a rich history of the use of machine learning techniques with large complex data types (such as magnetic resonance imaging data (Klöppel et al., 2012)). However, the current challenges facing machine learning applications with respect to service delivery and understanding phenotypic variation across neural, behavioural, and clinic indices remain unaddressed. In a recent review examining machine learning techniques for case control diagnosis, Kambeitz and colleagues demonstrate that there are major discrepancies with respect to training and testing methodologies (Kambeitz et al., 2015). They further report that most machine learning methods achieve sensitivities and specificity on the order of ~80%. It is unclear as to what would be the acceptable performance standards should be before integration of these methods into clinical trials, single subject-level predictions, or treatment planning.

One of the major obstacles is data availability. If we take Alzheimer's disease as an example, the availability of large-scale neuroimaging and clinical databases (such as the Alzheimer's disease neuroimaging initiative (Weiner et al., 2017)) allows individuals to demonstrate the feasibility of using large, well-structured datasets of neural phenotypes alongside clinical stages. However, to date, the availability of such data to accelerate schizophrenia research are considerably lacking. Although many recent collaborative large-scale initiatives have been established, it may be a challenge to integrate all of these databases to establish methodologies for prognosis (as has been done in Alzheimer's (Bhagwat et al., 2018); however there are definitely encouraging signs to this end in recent studies (Antonucci et al., 2019)). This lack of data availability gives rise to concerns to the generalizability (the applicability of a trained machine learning model across different datasets without suffering a loss in performance) of the models being developed. Naturally, these challenges highlight the significant opportunities that exist in the context of schizophrenia research and many of these opportunities are highlighted in this special issue.

* GH-2112, Cerebral Imaging Centre, Douglas Institute, 6875 Boulevard LaSalle, Montréal, QC, H4H 1R3, Canada.

E-mail address: mallar@cobralab.ca.

Interestingly there are clear trends in this special issue with respect to the contributions of the authors. The special issue begins with a critical appraisal of previous machine learning-based classification studies seeking to identify schizophrenia patients at a single subject level using neuroimaging data and demonstrates how seemingly insignificant design choices, such as image processing techniques and choice of machine learning model, may have a significant impact on results (Winterburn et al., 2017). Further in the realm of using magnetic resonance imaging to identify patients and controls, one contribution also examines the ability to effectively classify drug-naïve schizophrenia patients from normal controls using supervised techniques (Xiao et al., 2017). One paper examines the difficult problem of future prognosis in individuals deemed to be clinically at high-risk for conversion towards schizophrenia (Zarogianni et al., 2017). An interesting problem for machine learning methods also addressed here, is the problem of differential diagnosis in cases where individuals share clinical features (Palaniyappan et al., 2018). Another important theme in this special issue is heterogeneity in schizophrenia. A practical guide on how to deal with this topic in the context of machine learning is provided by Schnack (Schnack et al., 2017). The remaining articles covering heterogeneity do so by examining neuroanatomical heterogeneity using semi-supervised techniques (Honnorat et al., 2017), how symptom-level heterogeneity can be leveraged to better classify subgroups of schizophrenia patients using neuroanatomy (Talpaluru et al., 2019), and using clinical and biological data to derive subtypes in the Bipolar-Schizophrenia Network for Intermediate Phenotypes (BSNIP) (Mothi et al., 2018). Finally the special issue ends with a critical appraisal of machine learning in the context of examining heterogeneity in schizophrenia and serves as an excellent guide for future studies (Tandon et al., 2019).

In conclusion, as availability of computing software and technology improves alongside an improved ethic towards data sharing and open science, it is clear that progress will be made in the use of machine learning in the context of schizophrenia research, and ultimately in clinical applications. However, there are still significant quandaries that are left to be addressed, including: ethical issues with respect to data privacy and access, clinical responsibility, how to best deal with site-specific differences and potential differences in clinical assessment criteria, the most effective method for adopting machine learning based models in clinical settings and how to test and validate these models, and how to integrate machine learning methods into clinical training and the optimal methodology for integrating these methodologies alongside trained clinicians. Not all challenges that I mention here are technical, but they require deeper consideration beyond leveraging technical innovations. Embracing machine learning in clinical practice will require that we also embrace the challenges that access to technology provides us beyond the obvious technical ones.

Role of funding source

None.

Declaration of competing interest

None.

Acknowledgements

MMC would like to thank all the authors, Dr. Henry Nasrallah for his valuable advice in preparing this special issue, and Ms. Vanessa Valiquette for help with administrative support. Dr. Chakravarty is a Fonds de Recherche Santé Québec Junior 2 Scholar. He also receives support from the Canadian Institutes of Health Research, National Sciences and Engineering Research Council of Canada, and the McGill University Healthy Brains for Healthy Lives initiative (a Canada First Research Excellence Fund Program).

References

- Abraham, A., Pedregosa, F., Eickenberg, M., Gervais, P., Mueller, A., Kossaifi, J., Gramfort, A., Thirion, B., Varoquaux, G., 2014. Machine learning for neuroimaging with scikit-learn. *Front. Neuroinf.* 8, 14.
- Antonucci, L.A., Penzel, N., Pergola, G., Kambeitz-Illankovic, L., Dwyer, D., Kambeitz, J., Haas, S.S., Passiatore, R., Fazio, L., Caforio, G., Falkai, P., Blasi, G., Bertolino, A., Koutsouleris, N., 2019. Multivariate classification of schizophrenia and its familial risk based on load-dependent attentional control brain functional connectivity. *Neuropsychopharmacology*. <https://doi.org/10.1038/s41386-019-0532-3>.
- Bhagwat, N., Viviano, J., Voineskos, A.N., Chakravarty, M.M., 2018. Modeling and prediction of clinical symptom trajectories in Alzheimer's disease using longitudinal data. *PLoS Comput. Biol.* 14, e1006376.
- Bengio, Y., Courville, A., Vincent, P., 2013. Representation learning: a review and new perspectives. *IEEE Trans. Pattern Anal. Mach. Intell.* 35 (8), 1798–1828.
- Bzdok, D., Yeo, B.T.T., 2017. Inference in the age of big data: future perspectives on neuroscience. *Neuroimage* 155, 549–564.
- Dwyer, D.B., Falkai, P., Koutsouleris, N., 2018. Machine learning approaches for clinical psychology and psychiatry. *Annu. Rev. Clin. Psychol.* 14, 91–118.
- Honnorat, N., Dong, A., Meisenzahl-Lechner, E., Koutsouleris, N., Davatzikos, C., 2017. Neuroanatomical heterogeneity of schizophrenia revealed by semi-supervised machine learning methods. *Schizophr. Res.* <https://doi.org/10.1016/j.schres.2017.12.008>.
- Klöppel, S., Abdulkadir, A., Jack Jr., C.R., Koutsouleris, N., Mourão-Miranda, J., Vemuri, P., 2012. Diagnostic neuroimaging across diseases. *Neuroimage* 61, 457–463.
- Kambeitz, J., Kambeitz-Illankovic, L., Leucht, S., Wood, S., Davatzikos, C., Malchow, B., Falkai, P., Koutsouleris, N., 2015. Detecting neuroimaging biomarkers for schizophrenia: a meta-analysis of multivariate pattern recognition studies. *Neuropsychopharmacology* 40, 1742–1751.
- LeCun, Y., Bengio, Y., Hinton, G., 2015. Deep learning. *Nature* 521 (7553), 436–444.
- Mothi, S.S., Sudarshan, M., Tandon, N., Tamminga, C., Pearlson, G., Sweeney, J., Clementz, B., Keshavan, M.S., 2018. Machine learning improved classification of psychoses using clinical and biological stratification: update from the bipolar-schizophrenia network for intermediate phenotypes (B-SNIP). *Schizophr. Res.* <https://doi.org/10.1016/j.schres.2018.04.037>.
- Palaniyappan, L., Deshpande, G., Lanka, P., Rangaprakash, D., Iwabuchi, S., Francis, S., Liddle, P.F., 2018. Effective connectivity within a triple network brain system discriminates schizophrenia spectrum disorders from psychotic bipolar disorder at the single-subject level. *Schizophr. Res.* <https://doi.org/10.1016/j.schres.2018.01.006>.
- Schnack, H.G., 2017. Improving individual predictions: machine learning approaches for detecting and attacking heterogeneity in schizophrenia (and other psychiatric diseases). *Schizophr. Res.* <https://doi.org/10.1016/j.schres.2017.10.023>.
- Talpaluru, A., Bhagwat, N., Devenyi, G.A., Lepage, M., Mallar Chakravarty, M., 2019. Identifying schizophrenia subgroups using clustering and supervised learning. *Schizophr. Res.* <https://doi.org/10.1016/j.schres.2019.05.044>.
- Tandon, N., Tandon, R., 2019. Using machine learning to explain the heterogeneity of schizophrenia. Realizing the promise and avoiding the hype. *Schizophr. Res.* <https://doi.org/10.1016/j.schres.2019.08.032>.
- Weiner, M., 2017. ADNI MRI data. *J. Neuro. Sci.* 381, 13–14.
- Winterburn, J.L., Voineskos, A.N., Devenyi, G.A., Plitman, E., de la Fuente-Sandoval, C., Bhagwat, N., Graff, A., Knight, J., Chakravarty, M.M., 2017. Can we accurately classify schizophrenia patients from healthy controls using magnetic resonance imaging and machine learning? A multi-method and multi-dataset study. *Schizophr. Res.* <https://doi.org/10.1016/j.schres.2017.11.038>.
- Xiao, Y., Yan, Z., Zhao, Y., Tao, B., Sun, H., Li, F., Yao, L., Zhang, W., Chandan, S., Liu, J., Gong, Q., Sweeney, J.A., Lui, S., 2017. Support vector machine-based classification of first episode drug-naïve schizophrenia patients and healthy controls using structural MRI. *Schizophr. Res.* <https://doi.org/10.1016/j.schres.2017.11.037>.
- Zarogianni, E., Storkey, A.J., Borgwardt, S., Smieskova, R., Studerus, E., Riecher-Rössler, A., Lawrie, S.M., 2017. Individualized prediction of psychosis in subjects with an at-risk mental state. *Schizophr. Res.* <https://doi.org/10.1016/j.schres.2017.08.061>.