THE ENIGMA TOOLBOX: MULTISCALE NEURAL CONTEXTUALIZATION OF MULTISITE NEUROIMAGING DATASETS

Sara Larivière[1], Casey Paquola[2,3], Bo-yong Park[4,5], Jessica Royer[6], Yezhou Wang[7], Oualid Benkarim[8], Reinder Vos de Wael[9], Sofie L. Valk[10,11], Sophia I. Thomopoulos[12], Matthias Kirschner[13], ENIGMA Consortium[14], Lindsay B. Lewis[15,16], Alan C. Evans[17,18], Sanjay M. Sisodiya[19,20], Carrie R. McDonald[21], Paul M. Thompson[22], Boris C. Bernhardt[23]

[1]*Multimodal Imaging and Connectome Analysis Laboratory, McConnell Brain Imaging Centre, Montreal Neurological Institute and Hospital, McGill University, Montreal, Quebec, Canada*

[2]*Multimodal Imaging and Connectome Analysis Laboratory, McConnell Brain Imaging Centre, Montreal Neurological Institute and Hospital, McGill University, Montreal, Quebec, Canada*

[3]*Institute of Neuroscience and Medicine (INM-1), Research Center, Jülich, D-52425 Jülich, Germany*

[4]*Multimodal Imaging and Connectome Analysis Laboratory, McConnell Brain Imaging Centre, Montreal Neurological Institute and Hospital, McGill University, Montreal, Quebec, Canada*

[5]Department of Data Science, Inha University, Incheon, South Korea

[6]*Multimodal Imaging and Connectome Analysis Laboratory, McConnell Brain Imaging Centre, Montreal Neurological Institute and Hospital, McGill University, Montreal, Quebec, Canada*

[7]*Multimodal Imaging and Connectome Analysis Laboratory, McConnell Brain Imaging Centre, Montreal Neurological Institute and Hospital, McGill University, Montreal, Quebec, Canada*

[8]*Multimodal Imaging and Connectome Analysis Laboratory, McConnell Brain Imaging Centre, Montreal Neurological Institute and Hospital, McGill University, Montreal, Quebec, Canada*

[9]*Multimodal Imaging and Connectome Analysis Laboratory, McConnell Brain Imaging Centre, Montreal Neurological Institute and Hospital, McGill University, Montreal, Quebec, Canada*

[10]*Otto Hahn Research Group for Cognitive Neurogenetics, Max Planck Institute for Cognitive and Brain Sciences, Leipzig, Germany*

[11]*INM-7, FZ Jülich, Jülich, Germany*

[12]*Imaging Genetics Center, Mark & Mary Stevens Neuroimaging and Informatics Institute, Keck School of Medicine, University of Southern California, Marina del Rey, California, United States*

[13]*Department of Psychiatry, Psychotherapy and Psychosomatics, Psychiatric Hospital, University of Zurich, Switzerland*

[14]*http://enigma.ini.usc.edu*

[15]*McConnell Brain Imaging Centre, Montreal Neurological Institute and Hospital, McGill University, Montreal, Quebec, Canada*

[16]*McGill Centre for Integrative Neuroscience, McGill University, Montreal, Quebec, Canada*

[17]*McConnell Brain Imaging Centre, Montreal Neurological Institute and Hospital, McGill University, Montreal, Quebec, Canada*

[18]*McGill Centre for Integrative Neuroscience, McGill University, Montreal, Quebec, Canada*

[19]*Department of Clinical and Experimental Epilepsy, UCL Queen Square Institute of Neurology, London, United Kingdom*

[20]*Chalfont Centre for Epilepsy, Bucks, United Kingdom*

[21]*Department of Psychiatry, Center for Multimodal Imaging and Genetics, University of California San Diego, La Jolla, California, United States*

[22]*Imaging Genetics Center, Mark & Mary Stevens Neuroimaging and Informatics Institute, Keck School of Medicine, University of Southern California, Marina del Rey, California, United States*

[23]*Multimodal Imaging and Connectome Analysis Laboratory, McConnell Brain Imaging Centre, Montreal Neurological Institute and Hospital, McGill University, Montreal, Quebec, Canada*

CORRESPONDENCE TO:
Boris C. Bernhardt
E-mail: boris.bernhardt@mcgill.ca

Sara Larivière
E-mail: sara.lariviere@mail.mcgill.ca

Among big data neuroscience initiatives, the ENIGMA (Enhancing NeuroImaging Genetics through Meta-Analysis) Consortium—a worldwide alliance of over 2,000 scientists diversified into over 50 Working Groups—has yielded some of the largest studies of the healthy and diseased human brain. Through harmonized procedures, and by sharing site-specific brain metrics (*e.g.*, cortical thickness) or aggregated statistical maps, ENIGMA has set the stage for large-scale analyses comparing findings across different topics or disorders [1, 2]. In parallel, increasingly available resources offer opportunities to contextualize findings across multiscale brain organization. Examples include the Allen Human Brain Atlas[3] (AHBA; microarray-derived *postmortem* gene expression), the BigBrain Project[4] (3D *postmortem* human brain histology), and the Human Connectome Project[5] (HCP; high-definition *in vivo* functional and structural connectomics). Here we introduce the ENIGMA Toolbox, an open ecosystem for integration and visualization of multisite ENIGMA results and their multiscale neural contextualization. Our Toolbox relies on an efficient codebase for exploring and analyzing big data, aiming to facilitate and homogenize follow-up analyses of ENIGMA, or other, MRI datasets around the globe.

To advance and simplify cross-disorder analysis and multiscale neural contextualization of neuroimaging findings, the ENIGMA Toolbox offers the ability to access over 100 ENIGMA-derived statistical maps (**FIG. 1A**), to visualize and manipulate cortical and subcortical surface data and generate publication-ready figures (**FIG. 1B**), and to contextualize neuroimaging findings at the microscale (*postmortem* gene expression and cytoarchitecture; **FIG. 1C**) and macroscale (structural and functional connectome properties; **FIG. 1D**). To increase generalizability and usability, our Toolbox is compatible with most neuroimaging data and supports the mapping between parcellation maps and vertexwise surface space. To ensure cross-software compatibility, Toolbox users can also export data results to a range of file formats.

To deepen our understanding of the molecular and cellular underpinnings of healthy and diseased brain organization, our Toolbox provides microscale neural contextualization workflows to cross-reference neuroimaging findings against the AHBA brain-wide gene expression data. Toolbox users can import microarray expression datasets, visualize brain maps of gene expression levels, relate gene expression to properties of brain organization, and identify genes that are spatially correlated with a given brain map. The cell body-stained BigBrain data[4] and the von Economo and Koskinas cytoarchitectural atlas[6] are also accessible within our Toolbox. These digitized brain maps are invaluable for linking brain microstructure to functional localization and enable users to speculate on the underlying cytoarchitectural composition of, for instance, structurally abnormal areas in specific diseases. When combined, transcriptomic and cytoarchitectonic decoding can embed neuroimaging findings in a rich neurobiological context and yield insights into the etiology of several brain disorders.

At the macroscopic level, network connectivity offers a vantage point to quantify brain reorganization in diseases that are increasingly being conceptualized as network disorders. Our Toolbox provides tools to relate surface maps to normative connectome properties derived from functional and structural HCP data. Building on prior neurodegenerative[7, 8], psychiatric[9], and epilepsy[10] research, Toolbox users can build hub susceptibility models to assess the vulnerability of highly connected network hubs to disease-related effects and disease epicenter models to identify regions whose connectivity profiles herald disease-related effects. Combined, these two network models can advance our understanding of how connectome architecture relates to morphological abnormalities across a range of disorders. Network models can be further enriched

with microstructural properties to inject multiscale information into cortical and subcortical morphometric findings.

By bridging the gap between pre-established data processing protocols and analytic workflows, we hope that the ENIGMA Toolbox facilitates neuroscientific contextualization of results and cross-consortia initiatives. We are eager for researchers and clinicians to test hypotheses beyond traditional case-control comparisons. We hope that our platform will lead to novel and harmonized analyses in global neuroimaging initiatives.

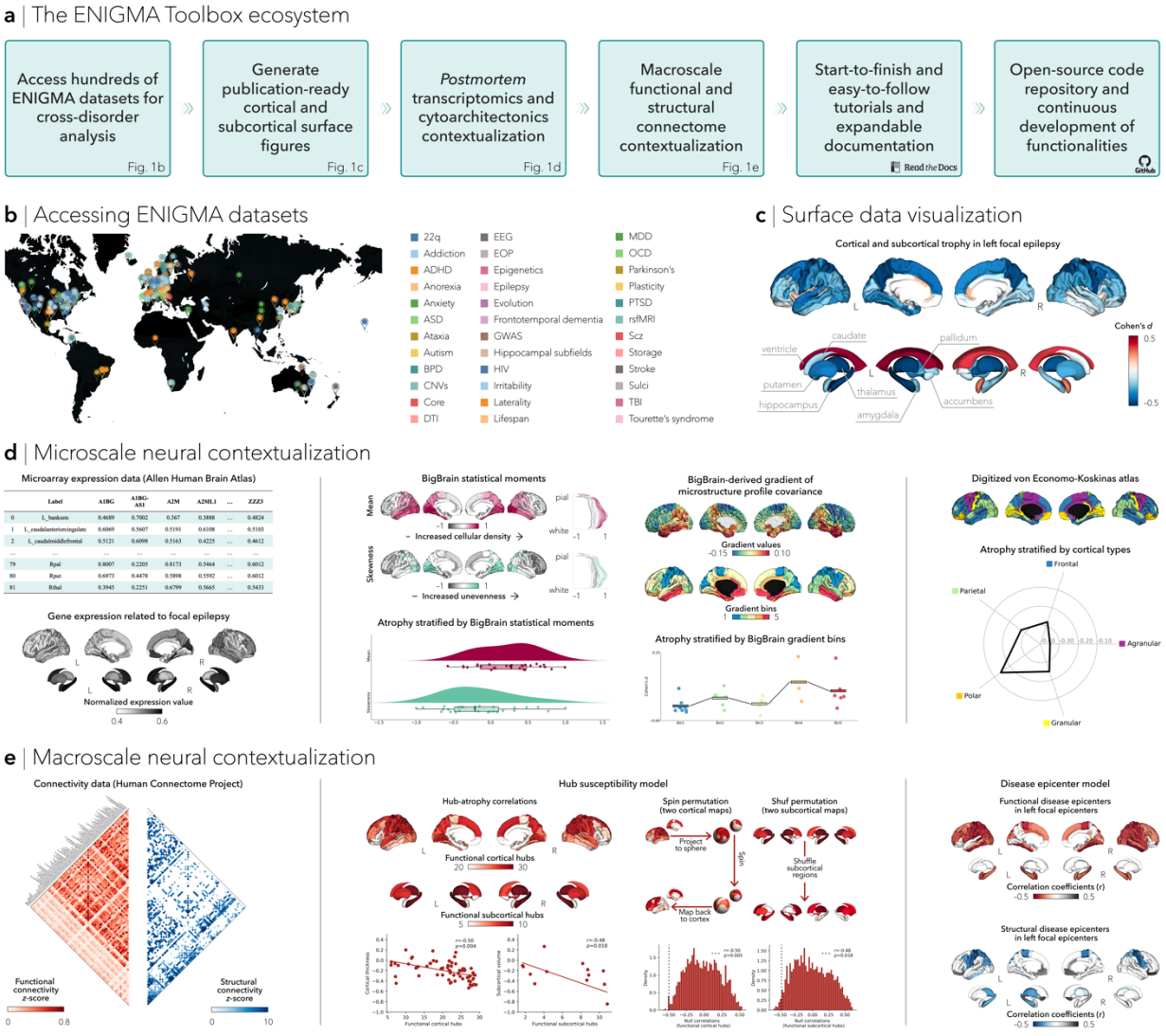

FIGURE 1. **Overview of the ENIGMA Toolbox**. (**a**) World map of a subset of Working Groups. 100+ summary statistics from published studies are accessible within the ENIGMA Toolbox. (**b**) Surface visualization tools are provided to project cortical and subcortical data results to the surface. As an example, we displayed gray matter atrophy in left focal epilepsy. (**c**) To contextualize neuroimaging data with respect to microscale brain organization, Toolbox users can fetch disease-related gene expression data (here, we displayed the average expression levels of focal epilepsy genes (left)), stratify atrophy (or other effect maps) according to BigBrain statistical moments (middle left) and gradient (middle right), and stratify atrophy patterns according to cytoarchitectonic classes (right). (**d**) To contextualize neuroimaging data with respect to macroscale brain organization, Toolbox users can load preprocessed functional and structural connectivity data (left), build hub susceptibility models to assess relationships between hub regions and atrophy patterns (middle left) and assess statistical significance of two surface maps using spin

permutation testing (middle right), and identify cortical and subcortical disease epicenters that herald patterns of atrophy.

CODE AVAILABILITY

Our Toolbox is available in Python and Matlab and complemented with expandable online documentation (http://enigma-toolbox.readthedocs.io). Derivative data (*e.g.*, summary statistics from published ENIGMA studies) and codes are openly accessible under the terms of the BSD-3-Clause license at http://github.com/MICA-MNI/ENIGMA. Requests to work on a project with subject-level data can be proposed to the Working Group via the chairs (http://enigma.ini.usc.edu/). Users seeking help are encouraged to subscribe and post their questions to the ENIGMA Toolbox mailing list at https://groups.google.com/g/enigma-toolbox.

AUTHOR CONTRIBUTIONS

Core developers: S.L., B.C.B.
Toolbox beta testing: B.-Y.P., J.R., C.P., S.L.V., Y.W, M.K.
Writing: S.L., B.C.B.; revised and approved by other listed co-authors.

ACKNOWLEDGMENTS

Many scientists around the world contributed to ENIGMA but did not take part in the writing of this report. A full list of contributors to ENIGMA is available here: http://enigma.ini.usc.edu/about-2/consortium/members/.

The authors would like to express their gratitude to the open science initiatives that made this work possible: (*i*) The ENIGMA Consortium (core funding for ENIGMA was provided by the NIH Big Data to Knowledge (BD2K) program under consortium grant U54 EB020403 to P.M.T.), (*ii*) The Allen Human Brain Atlas, (*iii*) BigBrain/HIBALL, and (*iv*) The Human Connectome Project (Principal Investigators: David Van Essen and Kamil Ugurbil; 1U54MH091657) funded by the 16 NIH Institutes and Centers that support the NIH Blueprint for Neuroscience Research; and by the McDonnell Center for Systems Neuroscience at Washington University.

S.L. acknowledges funding from Fonds de la Recherche du Québec – Santé (FRQ-S) and the Canadian Institutes of Health Research (CIHR). C.P. was funded through a postdoctoral FRQ-S fellowship. O.B. was funded by a Healthy Brains for Healthy Lives (HBHL) postdoctoral fellowship. B.-y.P. was funded by the National Research Foundation of Korea (NRF-2020R1A6A3A03037088), Molson Neuro-Engineering fellowship from the Montreal Neurological Institute and Hospital, and FRQ-S. J.R. was supported by CIHR. R.V.d.W. was funded by studentships from the Savoy Foundation for Epilepsy and the Richard and Ann Sievers award. S.L.V. was supported by the Otto Hahn award of the Max Planck Society. M.K. acknowledges funding from the Swiss National Science Foundation (P2SKP3_178175). S.M.S. was supported by the Epilepsy Society, UK. Part of this work was undertaken at University College London Hospitals, which received a proportion of funding from the NIHR Biomedical Research Centres funding scheme. C.R.M. acknowledges funding from the National Institutes of Health (NINDS R01NS065838 and R21 NS107739). B.C.B. acknowledges research funding from the SickKids Foundation (NI17-039), the Natural Sciences and Engineering Research Council of Canada (NSERC; Discovery-1304413), CIHR (FDN-154298), Azrieli Center for Autism Research

(ACAR), an MNI-Cambridge collaboration grant, salary support from FRQ-S (Chercheur-Boursier), BrainCanada, the Helmholtz BigBrain Analytics and Learning Lab (Hiball), and the Canada Research Chairs (CRC) Program.

ETHICS DECLARATION

The authors declare no competing interests. P.M.T. received partial grant support from Biogen, Inc., and consulting payments from Kairos Venture Capital, Inc., for work unrelated to ENIGMA and this manuscript.

Editor's note: this article has been peer reviewed.

REFERENCES

1. Patel, Y. et al. *Jama Psychiatry* (2020).
2. Boedhoe, P.S. et al. *American Journal of Psychiatry* **177**, 834-843 (2020).
3. Hawrylycz, M.J. et al. *Nature* **489**, 391-399 (2012).
4. Amunts, K. et al. *Science* **340**, 1472-1475 (2013).
5. Van Essen, D.C. et al. *Neuroimage* **62**, 2222-2231 (2012).
6. von Economo, C.F. & Koskinas, G.N., Vol. (J. Springer, 1925).
7. Zhou, J., Gennatas, E.D., Kramer, J.H., Miller, B.L. & Seeley, W.W. *Neuron* **73**, 1216-1227 (2012).
8. Zheng, Y.-Q. et al. *PLoS biology* **17**, e3000495 (2019).
9. Gollo, L.L. et al. *Nature neuroscience* **21**, 1107-1116 (2018).
10. Larivière, S. et al. *Sci Adv* **6** (2020).