

Continuous Large-scale Neuroimaging Analyses

Neda Jahanshad, PhD

Assoc Prof Neurology & Biomedical Engineering

Director of USC LoBeS: Laboratory of Brain eScience

Assoc Director of the ENIGMA Consortium: <http://enigma.usc.edu>

Mark and Mary Stevens Neuroimaging and Informatics Institute, Keck School of Medicine

Information Sciences Institute Affiliate

University of Southern California

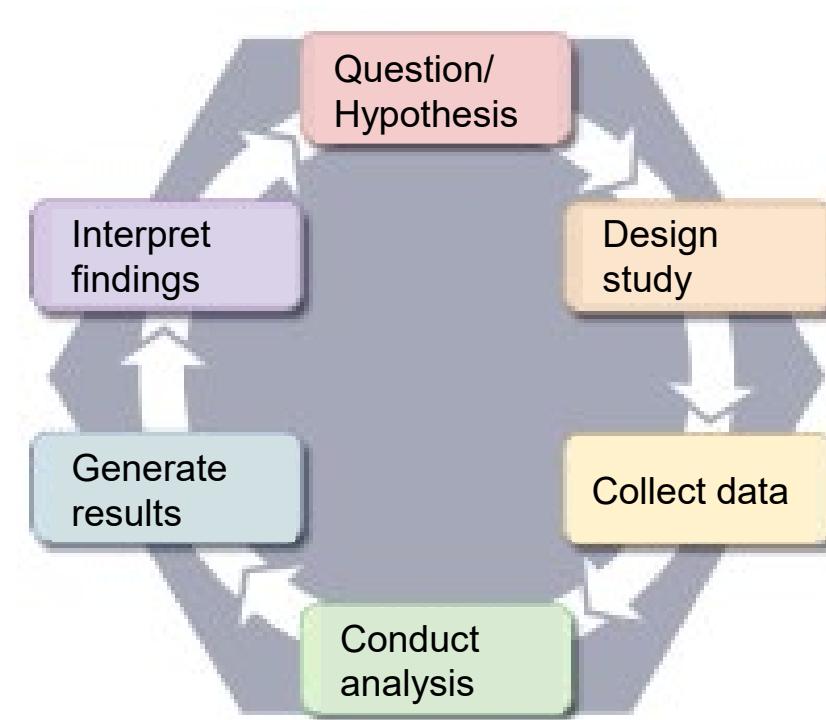
neda.jahanshad@usc.edu



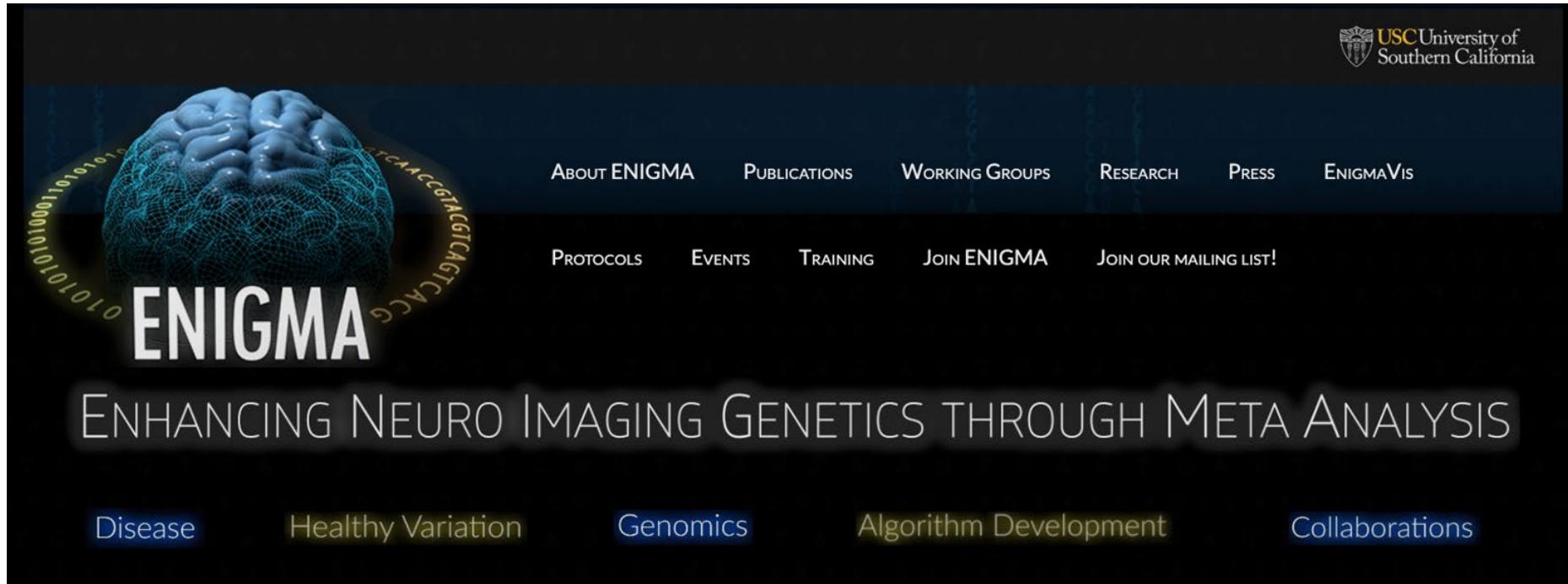
USC Mark and Mary Stevens
Neuroimaging and Informatics Institute



Continuous Analysis in Research



Context: The ENIGMA Consortium



The screenshot shows the ENIGMA consortium's website. At the top right is the USC University of Southern California logo. Below it is a navigation bar with links: ABOUT ENIGMA, PUBLICATIONS, WORKING GROUPS, RESEARCH, PRESS, and ENIGMAVis. Underneath this is a secondary row with links: PROTOCOLS, EVENTS, TRAINING, JOIN ENIGMA, and JOIN OUR MAILING LIST!. The main title "ENIGMA" is prominently displayed in large white letters on the left, with a blue brain graphic behind it. A circular DNA sequence graphic is also visible. Below the title is the subtitle "ENHANCING NEURO IMAGING GENETICS THROUGH META ANALYSIS". At the bottom, there are five tabs: Disease, Healthy Variation, Genomics (which is highlighted in blue), Algorithm Development, and Collaborations.



Background: **What** is ENIGMA? **Why** and **How** did it start?



What is ENIGMA?

A dynamic and collaborative investigator driven neuroimaging consortium that is organically expanding in scale and scope.



Motivation: Reproducibility, Reliability and Generalizability

Why Most Published Research Findings Are False

John P. A. Ioannidis

PLoS Medicine | www.plosmedicine.org

0696

August 2005 | Volume 2 | Issue 8 | e124

Power failure: why small sample size undermines the reliability of neuroscience

Katherine S. Button^{1,2}, John P. A. Ioannidis³, Claire Mokrysz¹, Brian A. Nosek⁴, Jonathan Flint⁵, Emma S. J. Robinson⁶ and Marcus R. Munafò¹

NATURE REVIEWS | NEUROSCIENCE

VOLUME 14 | MAY 2013 | 365

MOTHERBOARD
TECH BY VICE

Why Two Decades of Brain Research Could Be Seriously Flawed

A new paper casts doubt on 40,000 studies.

By Kate Lunau

Jul 5 2016, 10:10pm



Nature: Brain Imaging Studies Are Most Likely False

Small MRI studies inflate effect sizes, leaving the brain imaging research literature cluttered with false positives.

By Peter Simons - March 21, 2022

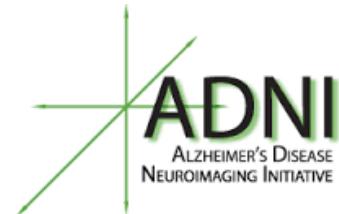


4013

ENIGMA

Large-scale multi-site datasets

Biobanks and coordinated large scale studies are becoming more and more common.
These are *tremendous* resources!



PARKINSON'S
PROGRESSION
MARKERS
INITIATIVE

Play a Part in Parkinson's Research

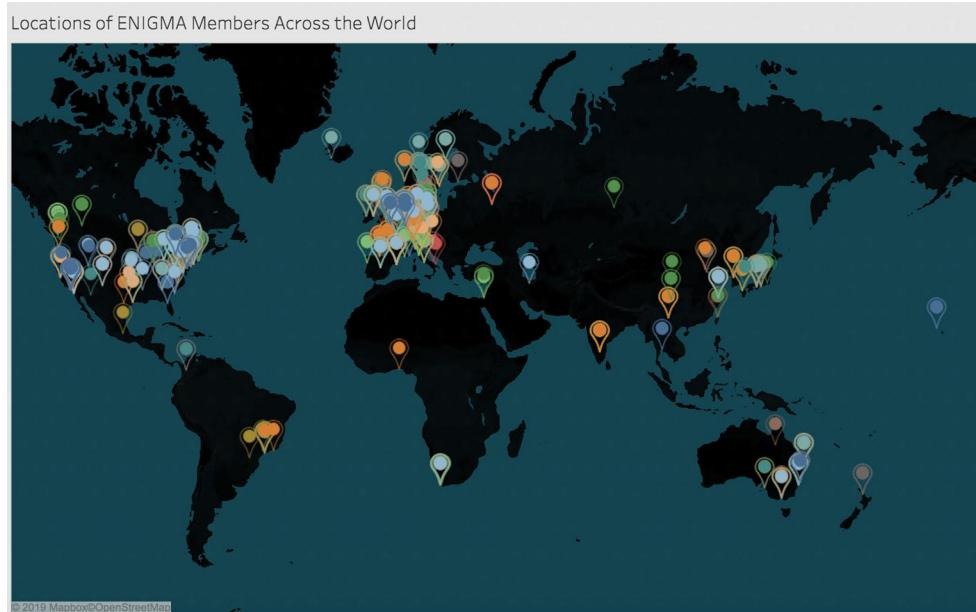


HUMAN
Connectome
PROJECT

And many others...

What about all the data from individual uncoordinated studies?

Many studies have been conducted that have collected neuroimaging data over the years, often having completed data collection years ago, and also completed the primary study. Many pilot studies that were on their own underpowered, data reuse!



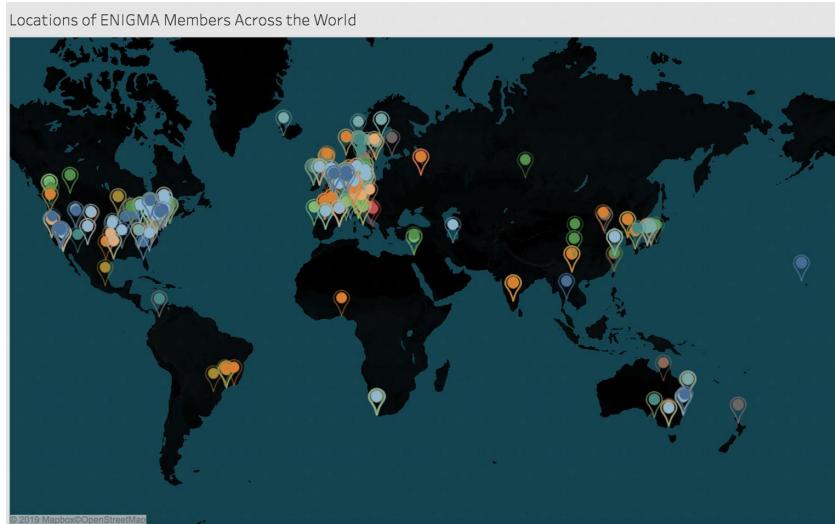
Can still be used to help answer questions about the brain and its disorders

Study specific rare or uncommon conditions around the world

Help answer questions about reliability and generalizability



Can we work with data from private and public sources together?



“Data” from individual sites can be combined with publicly available resources for larger scale initiatives, including numerous projects within (and across) ENIGMA.

Not all data can be openly shared, but collaboration is still possible



PARKINSON'S
PROGRESSION
MARKERS
INITIATIVE

Play a Part in Parkinson's Research

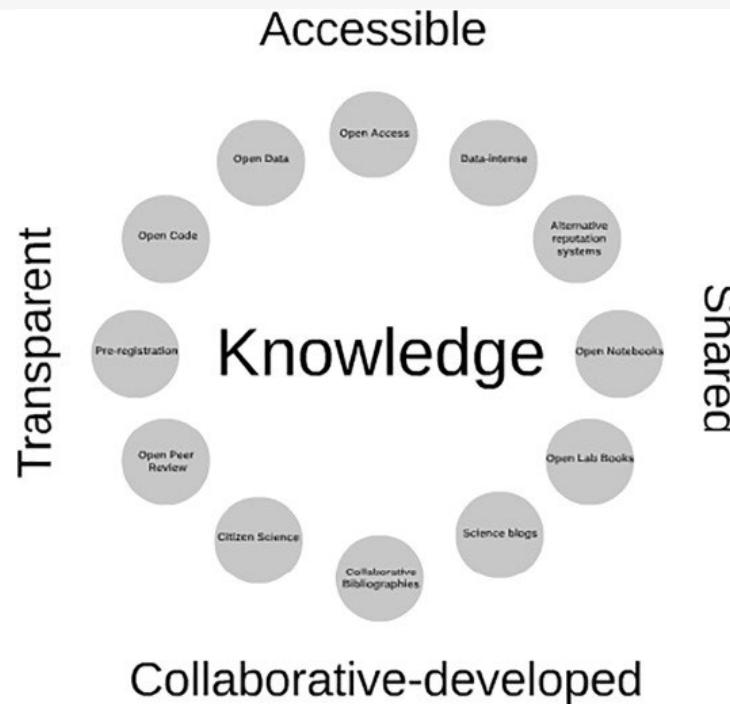


Adolescent Brain Cognitive Development®
Teen Brains. Today's Science. Brighter Future.



Not all data can be fully shared, but Open Science is still possible

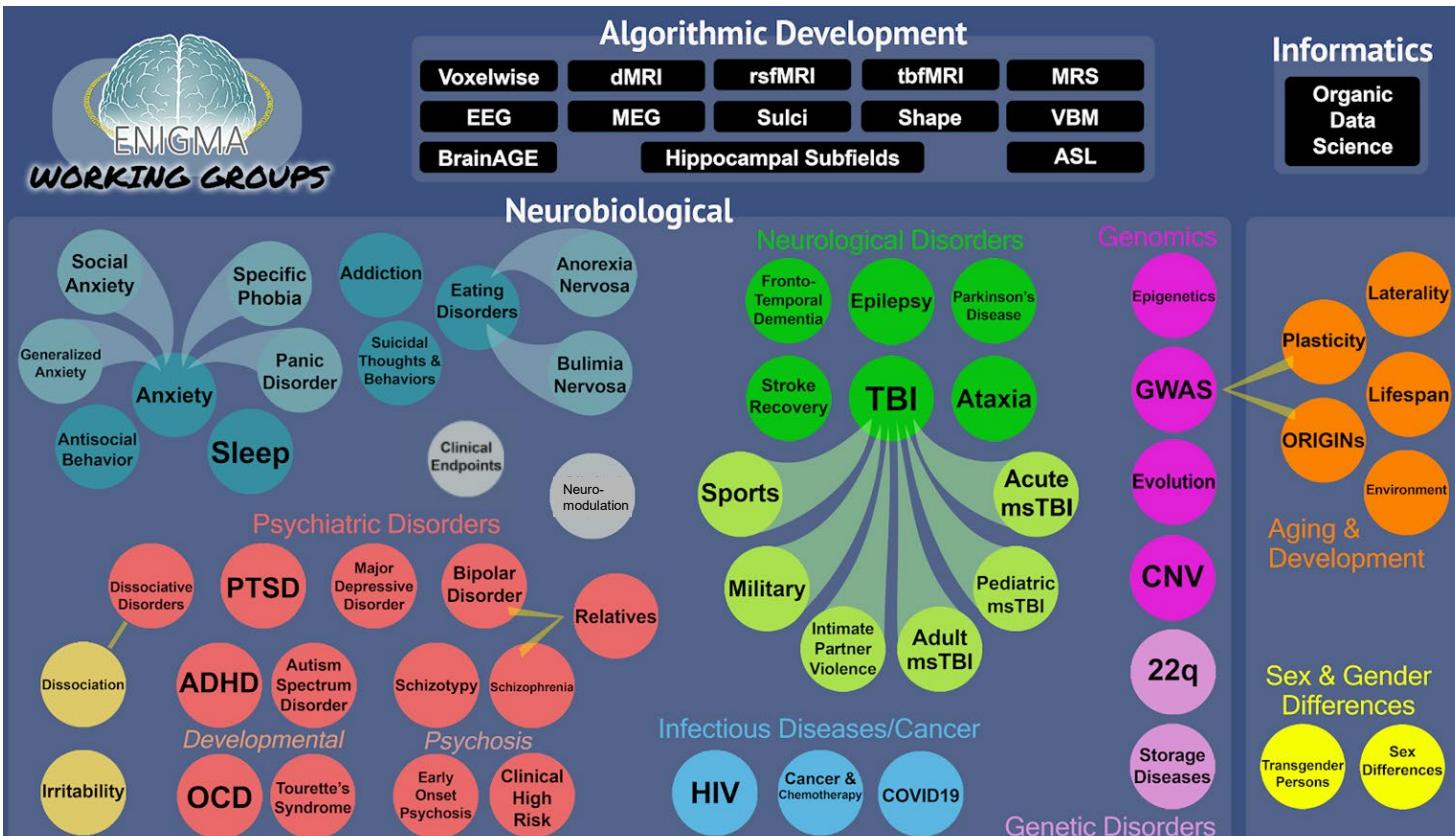
Open Science is transparent and accessible knowledge that is shared and developed through collaborative networks



Vicente-Saez and Martinez-Fuentes,

Journal of Business Research 2018

ENIGMA Organization - Working Groups



ENIGMA contains numerous working groups (WG).

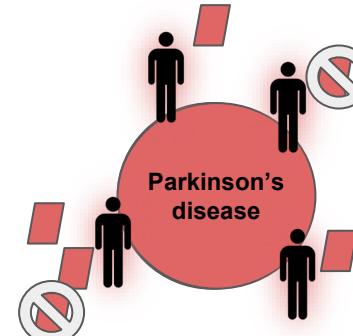
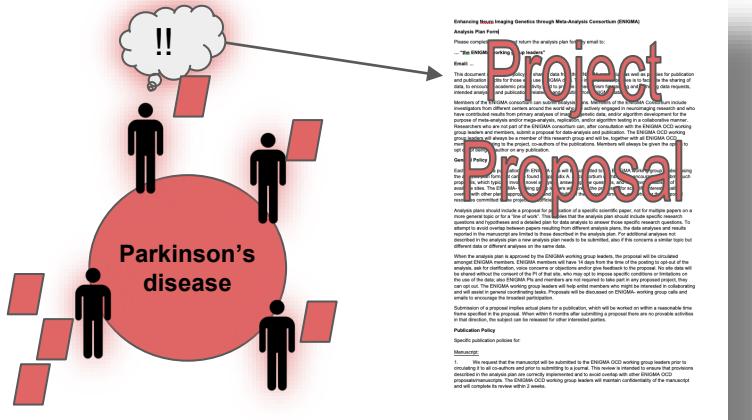
Computational: Imaging / data analytics / informatics

Neurobiological: Healthy variation / genetics / specific conditions or disorders / groups of disorders

ENIGMA Organization - Data (?)

ENIGMA is not a data repository. Data from participating sites / cohorts have a variety of different sharing and access rules.

However, any working group member can propose projects. Projects will need to be approved by WG chairs and presented to members, who can then opt in or out of the project. The project lead can then access data from groups that have opted-in through cohort-specific data access requirements. Projects are collaborative processes and updates and discussions occur on regular group-wide communications.



ENIGMA Organization - Working Groups Projects

Initial core WG projects aim to grow the network, establish collaborations, understand individual groups' strengths and limitations, and ask very straightforward and broad questions of the data from ALL participating sites/cohorts.

For example: assess the effect of the condition/trait/disease of interest on each of the available brain imaging traits (subcortical volumes, cortical thickness/area).



Any interested person (usually with relevant data/tool) can join at anytime.

There is no requirement to share individual participant level data to join ENIGMA.

Analytic plans are discussed, agreed upon, documented and distributed. Developed and tested analysis protocols and workflows are distributed to individual cohorts representatives.

This initial project helps determine the scope and feasibility of future projects.



Data Structures can help formulate projects



LOBES

Parkinson Disease Working Group

Covariates

DEMOGRAPHICS	COGNITIVE	NEUROPSYCHIATRIC
<input type="checkbox"/> MDS-UPDRSTOTAL <input checked="" type="checkbox"/> SEX <input type="checkbox"/> CITYDWELLING <input type="checkbox"/> SCOPA-SLEEP <input type="checkbox"/> SUBJECTSES <input type="checkbox"/> SIDEONSET <input type="checkbox"/> OTHERMOTOREVALUATION <input type="checkbox"/> LEDDSCORE	<input type="checkbox"/> WORDLISTLEARNINGTASK <input type="checkbox"/> DIGITSPANBACKWARD <input type="checkbox"/> DIGITSPANFORWARD <input type="checkbox"/> SENTENCESCONSTRUCTION <input type="checkbox"/> REVS15WORDTEST <input type="checkbox"/> SWCT <input checked="" type="checkbox"/> MoCA <input type="checkbox"/> MMSE	<input type="checkbox"/> SASDQ <input type="checkbox"/> RBDIQ <input type="checkbox"/> RBDSQ <input type="checkbox"/> STAXI <input type="checkbox"/> BAI <input type="checkbox"/> TAS-20 <input type="checkbox"/> IDS <input type="checkbox"/> SHAPS

List of Cohorts

Name
STANFORD
ARMENIA
VUMC2
SHANGAI
UPENN

Rows per Page: 5 ▾ 1-5 of 5 < >

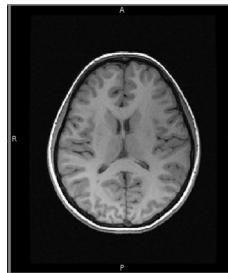
Project proposal can also act as a pre-registered analysis plan



How does it work?

Challenges of pooling diverse data

Imaging acquisitions are all different



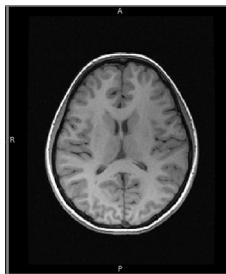
Study designs are often unique:

- Different populations
- Different inclusion/ exclusion criteria
- Different assessments/ batteries and instruments



Challenges of pooling diverse data

Imaging acquisitions are different



that's ok!

Study designs are often unique:

- Different populations
- Different inclusion/ exclusion criteria
- Different assessments/ batteries and instruments

A biomarker has limited clinical utility if it is only evident when measured with instruments that are not widely available.



Motivation: Reproducibility, Reliability, and **Generalizability**

With large diverse, heterogeneous datasets, we can readily assess generalizability.

Will the findings hold in a different population? With a different method: acquisition / analytical ?

If not, why? Biological? Methodological?



Image Processing / Feature Extraction Workflows in ENIGMA

ENIGMA's general approach to **pooling diverse data** has largely revolved around analytical standardization.

Detailed and documented analytic plans are central to the workings of the consortium.

- Imaging metrics
- Clinical / biological metrics / covariates
- Quality assurance
- Statistical workflow
- Test / retest / evaluate



Feature Robustness, Detailed Documentation, Containerization

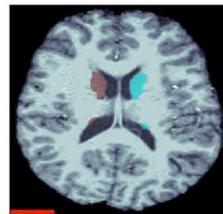
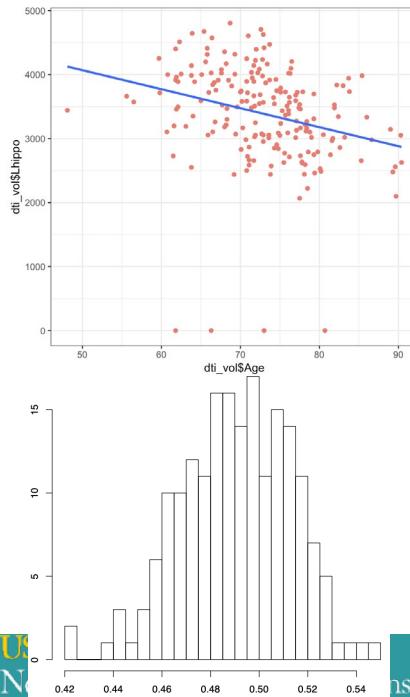
Imaging protocols

- 1) Does it work for multiple imaging acquisitions
- 2) Is it reliable across acquisitions? -- multiple test retest datasets
- 3) Is it straightforward to implement -- often build on existing commonly used tools
- 4) Are features biologically meaningful?

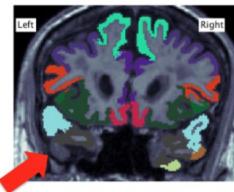
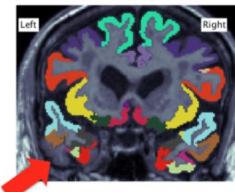
Analytical Harmonization: Quality Assurance

Specifying a single image processing workflow will help minimize methodological variability across multisite analyses. Even with containerization of workflows, some datasets may behave unexpectedly. Certain workflows may extract more accurate features for some datasets than others. Feasibility?

The documentation of quality assurance procedures is critical!

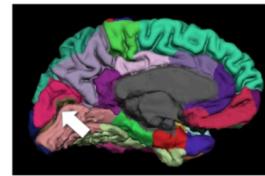


Internal QC: Moderate
Temporal Pole Underestimation

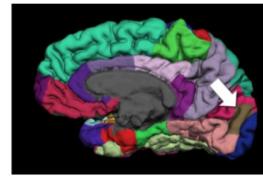


Pericalcarine Overestimations (<5% cases)

GOOD

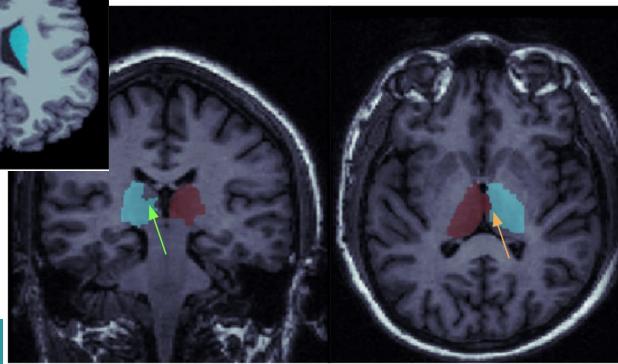
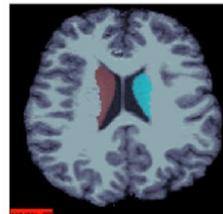


BAD



- Segmentation confined to calcarine sulcus

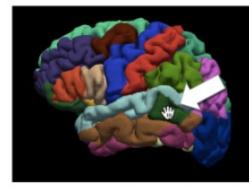
- Segmentation overestimates pericalcarine region
- Also failed regions (lesions)



Bad examples BanksSTS
(List affected regions in QC sheet)



Fail BanksSTS + STG

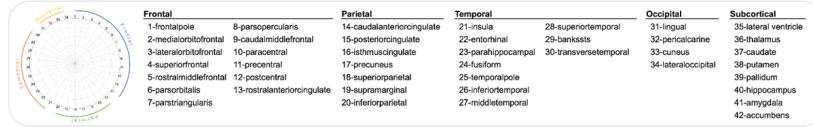


Fail BanksSTS + STG + MTG

New Updates? Test Reliability Reproducibility Compatibility

Numerous factors can contribute to differences in the derived imaging metrics (i.e., volume) even those from the same scan.

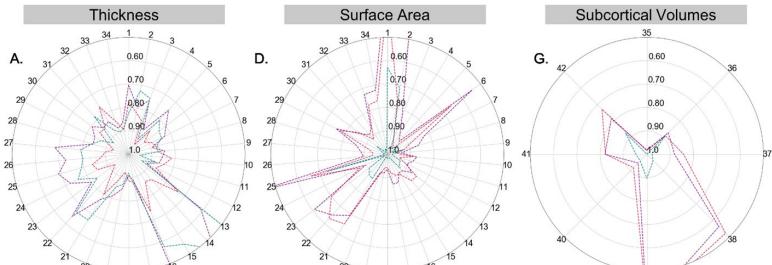
Can we quantify the degree to which these choices matter, and specifically for what features?



- - - v5.3 vs. v6.0
- - - v5.3 vs. v7.1
- - - v6.0 vs. v7.1



Bilateral



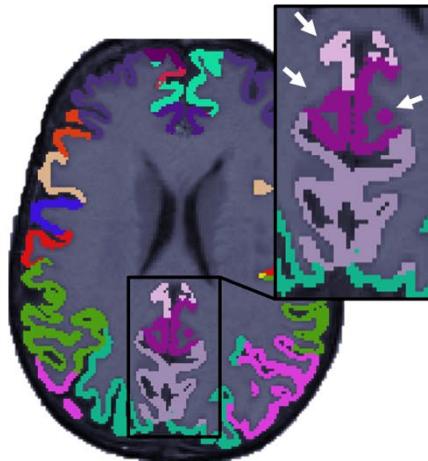
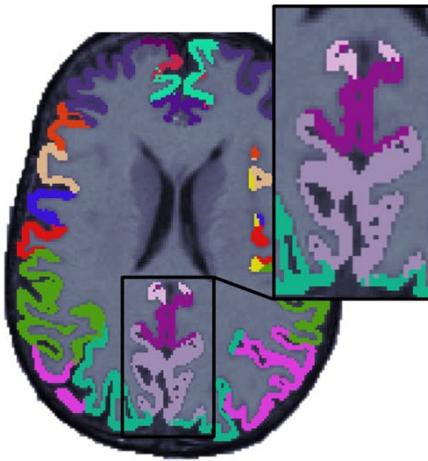
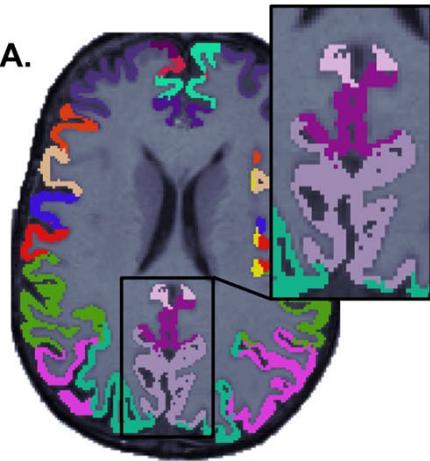
Elizabeth Haddad

FreeSurfer v5.3

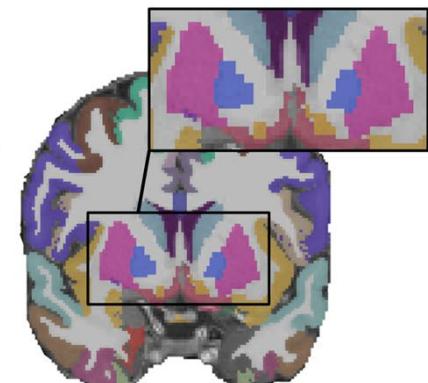
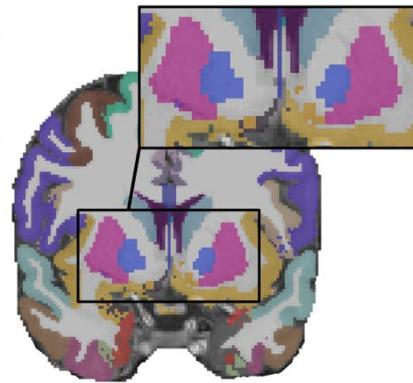
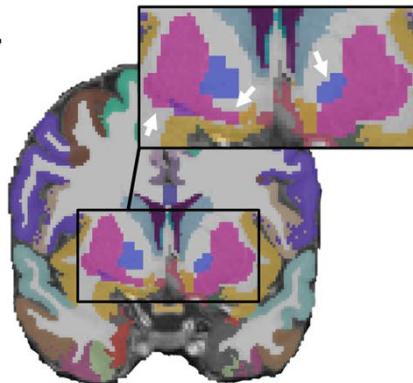
FreeSurfer v6.0

FreeSurfer v7.1

A.



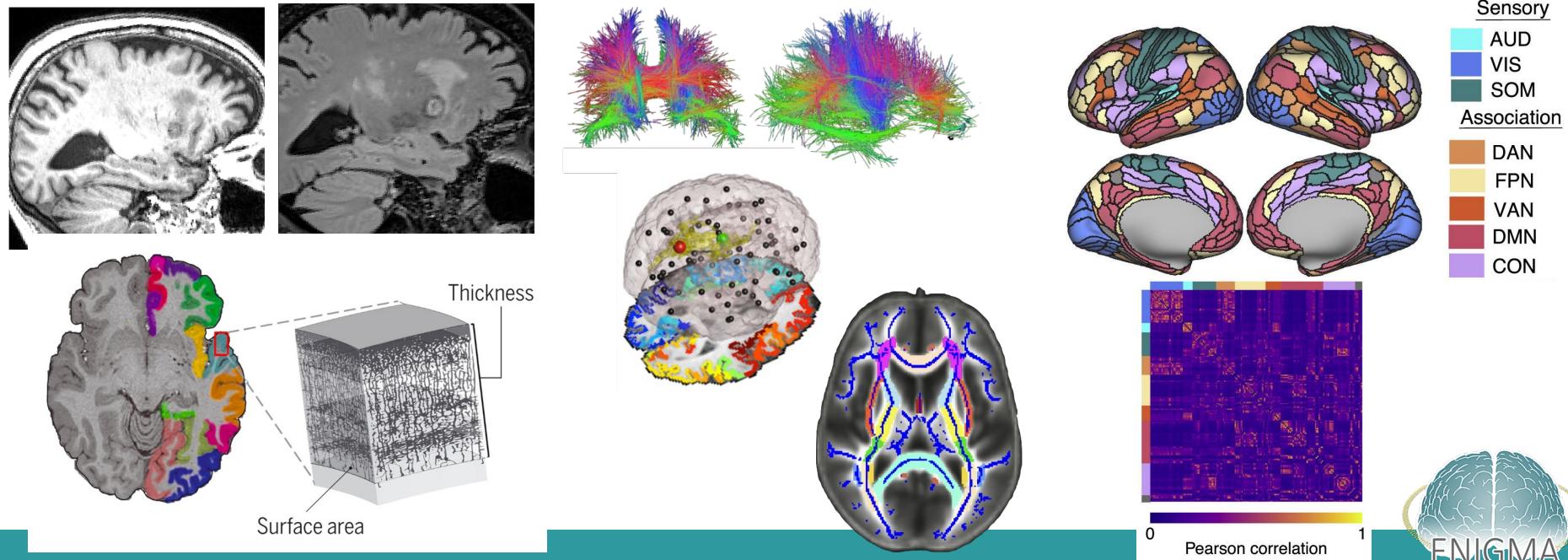
B.



Different MRI modalities can reveal different structural and functional properties of biological tissue

T2-FLAIR, susceptibility, diffusion, function - resting state, task

Continuous development of new protocols



Continuously evolving through active multidisciplinary research

Standardization efforts are all accelerating the pace of multi-site, multi-study analyses

- Standardization of data formats (BIDS; Gorgolewski 2016) makes it more straightforward to develop new software that relies on particular framework / data
- Building on fMRIprep (Esteban 2019) → HALFPIPE (Waller 2023)
- TBSS (Smith 2006) → ENIGMA-DTI (Jahanshad 2013)

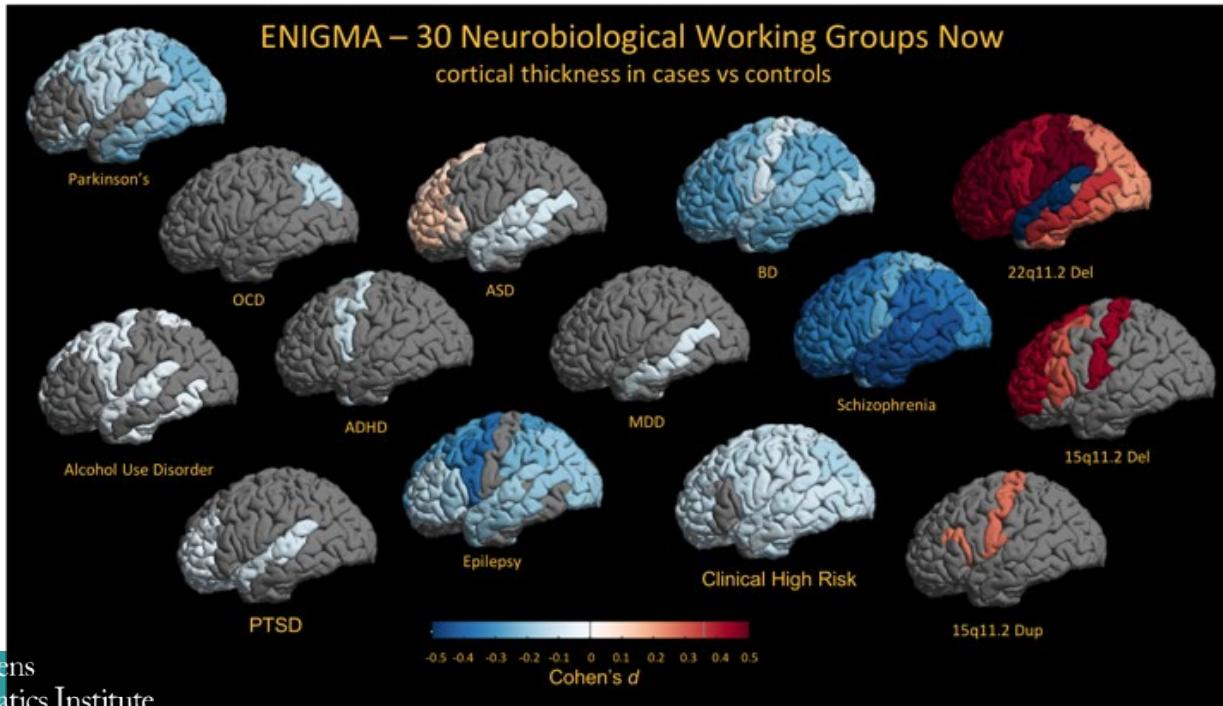
Projects and protocols routinely being re-evaluated and tested

- Automated volume extractions (old vs new versions) vs DL based extractions vs voxelwise analyses (VBM)



Standardized Extractions, Detailed Documentation, Containerization

Documented protocols allow for projects across working groups to use the same image analysis approaches, and comparisons can be made across diagnostic groups

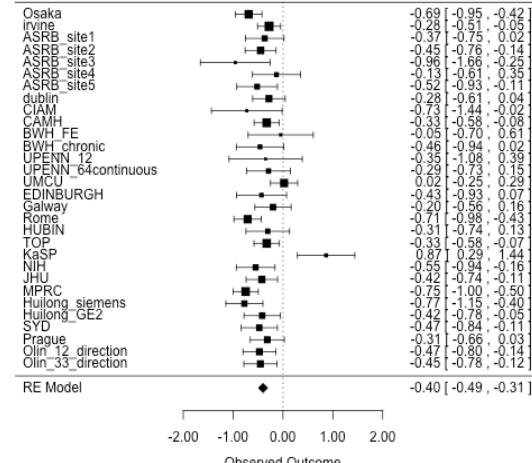


Meta and “Mega” analysis projects within ENIGMA

Meta-analysis projects require only summary statistics from each cohort. Overall effect sizes are a weighted average of individual cohorts (inverse variance).

Coordinated meta-analysis are different from literature based meta-analysis as they are not subject to publication bias and include results for all tested features for all cohorts

- + Easy to see heterogeneity/outliers in effects - site data/ analytics / processing effects
- + Distributed computation
- Analyses (generally*) limited to associations with common traits -- available across population (genetics), good distribution within each sample (most case/control)
- Can take lots of time / troubleshooting



Genetic Projects within ENIGMA have all been Meta Analysis



RESEARCH ARTICLE SUMMARY

CORTICAL GENETICS

The genetic architecture of the human cerebral cortex

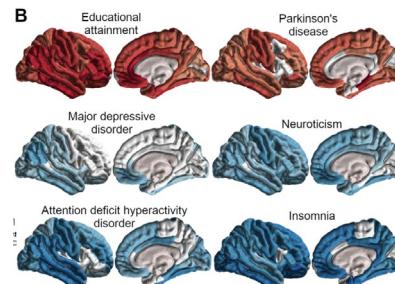
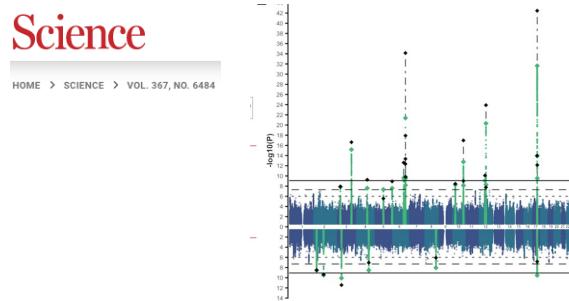
Katrina L. Grabsky*† and Neda Jahanshad*† et al.

Genetics Protocols

ENIGMA-git / ENIGMA

ENIGMA Consortium GWAS Protocols GWAS of the Cortex (ENIGMA3)

Version 1.1 – July 14, 2015
Written by Derek Hibar and Neda Jahanshad
Please address any questions our Google group. Also, please check our FAQ to see if your question is answered.
We have tutorial videos available for running our scripts (run0, run1, run2) on our YouTube channel as well.



Meta Analysis within Schizophrenia Working Group

Meta-Analysis > *Biol Psychiatry*. 2018 Nov 1;84(9):644-654.

doi: 10.1016/j.biopsych.2018.04.023. Epub 2018 May 14.

Cortical Brain Abnormalities in 4474 Individuals With Schizophrenia and 5098 Control Subjects via the Enhancing Neuro Imaging Genetics Through Meta Analysis (ENIGMA) Consortium

> *Mol Psychiatry*. 2018 May;23(5):1261-1269. doi: 10.1038/mp.2017.170. Epub 2017 Oct 17.

Widespread white matter microstructural differences in schizophrenia across 4322 individuals: results from the ENIGMA Schizophrenia DTI Working Group

RESEARCH ARTICLE | MEDICAL SCIENCES |

Large-scale analysis of structural brain asymmetries in schizophrenia via the ENIGMA consortium

Dick Schijven, Merel C. Postema, Masaki Fukunaga

, +145, and Clyde Francks

Authors Info & Affiliations

Edited by Marcus Raichle, Washington University in St Louis School of Medicine, St. Louis, MO; received August 24, 2022; accepted February 3, 2023

March 28, 2023 | 120 (14) e2213880120 | <https://doi.org/10.1073/pnas.2213880120>

nature > molecular psychiatry > articles > article

Article | [Open Access](#) | Published: 09 December 2022

Brain ageing in schizophrenia: evidence from 26 international cohorts via the ENIGMA Schizophrenia consortium

nature > molecular psychiatry > original article > article

[Open Access](#) | Published: 02 June 2015

Subcortical brain volume abnormalities in 2028 individuals with schizophrenia and 2540 healthy controls via the ENIGMA consortium

Articles

 Full Access

The Relationship Between White Matter Microstructure and General Cognitive Ability in Patients With Schizophrenia and Healthy Participants in the ENIGMA Consortium

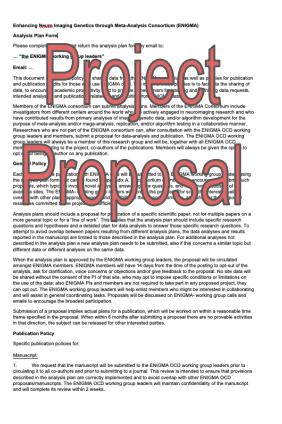


Meta and “Mega” analysis projects within ENIGMA

“Mega” analysis in this case refers to the analysis of data centrally* with all subject level data, as opposed to meta-analysis of statistical results.

Often done using features already extracted from imaging protocols, rarely using raw images

- + Less back and forth with participating sites
- + Can assess rarer effects, data pooling for amassing sufficient sample sizes for a trait not well distributed at sites, can include sites that have cases/controls only, machine/deep learning
- Site effects need to be taken into consideration, data harmonization
- (Recommend always doing site-wise analysis for quality assurance / validation)



ENIGMA Parkinson's disease working group



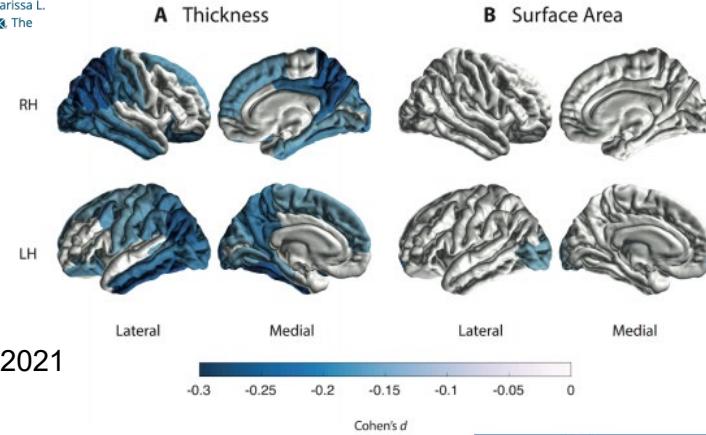
RESEARCH ARTICLE | Open Access | ⓘ

International Multicenter Analysis of Brain Structure Across Clinical Stages of Parkinson's Disease

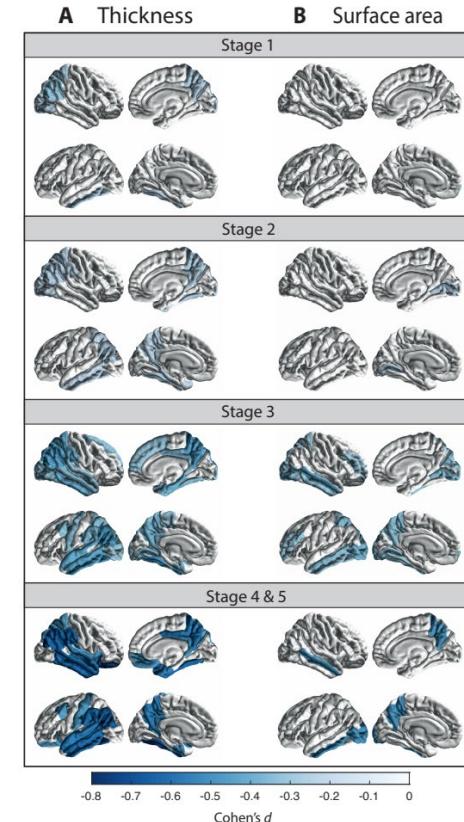
Max A. Laansma MSc, Joanna K. Bright BSc, Sarah Al-Bachari PhD, MBChB, Tim J. Anderson MD, PhD, Tyler Ard PhD, Francesca Assogna PhD, Katherine A. Baquero MSc, Henk W. Berendse MD, PhD, Jamie Blair PhD, Fernando Cendes MD, PhD, John C. Dalrymple-Alford PhD, Rob M.A. de Bie MD, PhD, Ines Debove MD, PhD, Michiel F. Dirksen MD, PhD, Jason Druzgal MD, PhD, Hedley C.A. Emsley PhD, FRCR, Gaetan Garraux PhD, Rachel P. Guimaraes PhD, Boris A. Gutman PhD, Rick C. Helmich MD, PhD, Johannes C. Klein MD, PhD, Clare E. Mackay PhD, Corey T. McMillan PhD, Tracy R. Melzer PhD, Laura M. Parkes PhD, Fabrizio Piras PhD, Toni L. Pitcher PhD, Kathleen L. Poston MD, Mario Rango MD, PhD, Leticia F. Ribeiro MD, Cristiane S. Rocha PhD, Christian Rummel PhD, Lucas S.R. Santos MD, Reinhold Schmidt MD, PhD, Petra Schwingenschuh MD, PhD, Gianfranco Spalletta MD, PhD, Letizia Squarcina PhD, Odile A. van den Heuvel MD, PhD, Chris Vriend PhD, Jian-Jie Wang PhD, Daniel Weintraub MD, PhD, Roland Wiest PhD, Clarissa L. Yasuda MD, PhD, Neda Jahanshad PhD, Paul M. Thompson PhD, Ysbrand D. van der Werf PhD, The ENIGMA-Parkinson's Study, ... See fewer authors ^

First published: 20 July 2021 | <https://doi.org/10.1002/mds.28706>

Pooling data across sites allowed for well-powered analyses not only of PD cases vs controls, but also PD at various Hoehn & Yahr stages vs controls.



Laansma, Bright et al Movement Disorders 2021



HY1	HY2	HY3	HY4&5
N _{PD} 437	N _{PD} 941	N _{PD} 263	N _{PD} 85
N _{IHC} 847	N _{IHC} 908	N _{IHC} 502	N _{IHC} 329

Many exciting developments in statistical data harmonization

README.md

ComBatHarmonization

Harmonization of multi-site imaging data with ComBat

Jean-Philippe Fortin, Drew Parker, Birkan Tunc, Takanori Watanabe, Michael Elliott, Kosha Ruparel, David R Roalf, Theodore D Satterthwaite, Ruben Gur, Raquel E Gur, Robert T Schultz, Ragini Verma, Russell T Shinohara. Harmonization Of Multi-Site Diffusion Tensor Imaging Data. *NeuroImage*, 161, 149–170, 2017

Jean-Philippe Fortin, Nicholas Cullen, Yvette I. Sheline, Warren D. Taylor, Irem Aselcioglu, Philip A. Cook, Phil Adams, Crystal Cooper, Maurizio Fazio, Patrick J. McGrath, Melvin McInnis, Mary L. Phillips, Madhukar H. Trivedi, Myrna M. Weissman, Russell T. Shinohara. Harmonization of cortical thickness measurements across scanners and sites. *NeuroImage*, 161, 104–120, 2018

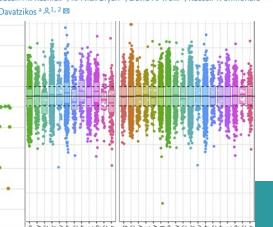
W. Evan Johnson and Cheng Li, Adjusting batch effects in microarray expression data using empirical Bayes methods. *NeuroImage*, 43, 127, 2007.



NeuroImage
Volume 208, March 2020, 116450

Harmonization of large MRI datasets for the analysis of brain imaging patterns throughout the lifespan

Raymond Ponponio^a Guray Erus^a, Mohammad Habes^{a,b} Jimit Doshi^a, Dhivya Srinivasan^a, Elizabeth Mamourian^a, Vishnu Bashyam^a, Ilyas M. Nasrallah^{a,c}, Theodore D. Satterthwaite^a, Yong Fan^a, Lenore J. Launer^c, Colin L. Masters^d, Paul Maruff^d, Chuanjun Zhou^{e,f}, Henning Volzke^b, Sterling C. Johnson^a, Jürgen Fröpp^a, Nikolas Koutsourelis^a, Daniel H. Wolf^a, Raquel Gur^{a,b}, Ruben Gur^{a,b}, John Morris^m, Marilyn S. Albert^a, Hans J. Grabe^a, Susan M. Resnick^a, R. Nick Bryan^a, David A. Wolk^b, Russell T. Shinohara^{a,b,c}, Haochang Shou^{a,c,f}, Christos Davatzikos^a



NeuroImage
Volume 220, 15 October 2020, 117129

Longitudinal ComBat: A method for harmonizing multi-site longitudinal imaging data

NEUROIMAGE
ELSEVIER
FREE Full-Text Article

Neuroimage, 2020 Oct 15; 220: 117127.
doi: [10.1016/j.neuroimage.2020.117127](https://doi.org/10.1016/j.neuroimage.2020.117127)

PMCID: PMC7573655
PMID: 32634595

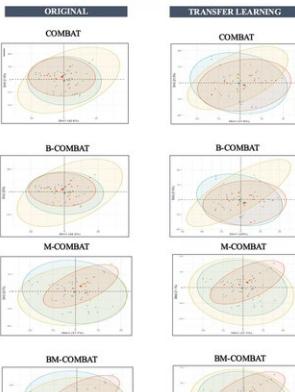
Neuroharmony: A new tool for harmonizing volumetric MRI data from unseen scanners

Rafael Garcia-Dias^a Cristina Scarpazza^{a,b} Lea Baecker^a Sandra Vieira^a Walter H.L. Pinaya^{a,c} Aiden Corvin^d Alberto Redolfi^e Barnaby Nelson^{f,g} Benedicto Crespo-Facorro^{h,i,j,k} Colm McDonaldⁱ Diana Tordesillas-Gutiérrez^{h,m} Dara Cannon^l David Mothersellⁿ Dennis Hernaus^o Derek Morris^p Esther Seilen-Suero^{h,k} Gary Donohoe^q Giovanni Frisoni^{q,t} Giulia Tronchin^t João Sato^c Machteid Marcelis^o Haren^s Oliver Gruber^{t,u} Patrick McGorry^{f,g} Paul Amminger^{f,g} René S. Kahn^{z,aa} Rosa Ayesta-Ariola^{h,k} Therese van Amelsvoort^o Calhoun^{ab,s} Wiepke Cahn^z and Andrea Mechelli^a

Accommodating site variation in neuroimaging data using normative and hierarchical Bayesian models

Johanna M. M. Bayer^a Richard Dinga^a Seyed Mostafa Kia^a Akhil R. Kottaram^a Thomas Wolfers^a Jinglei Lv^a Andrew Zalesky^a Lianne Schmaal^a Andre Marquand^a doi: <https://doi.org/10.1101/2021.02.09.430363>

This article is a preprint and has not been certified by peer review (what does this mean?).



PLOS ONE

OPEN ACCESS PEER-REVIEWED
RESEARCH ARTICLE

A transfer learning approach to facilitate ComBat-based harmonization of multicentre radiomic features in new datasets

Ronnick Daiane^a François Lucie, Ingrid Masson, Ronan Algral, Joanne Alferi, Caroline Reinhold, Olivier Pradier, Ulrike Schick, Dimitris Vavsek^a, Mathieu Hatt^a



NeuroImage
Volume 245, 15 December 2021, 118703

A multi-scanner neuroimaging data harmonization using RAVEL and ComBat

Mahbaneh Eshaghzadeh Torbati^a, Davneet S. Minhas^b, Ghasan Ahmad^c, Erin E. O'Connor^c, John Muschelli^c, Charles M. Laymon^b, Zixi Yang^b, Ann D. Cohen^b, Howard J. Aizenstein^b, William E. Klunk^b, Bradley T. Christian^f, Seong Jae Hwang^{a,b}, Ciprian M. Crainiceanu^d, Dana L. Tudorascu^{a,b,c}



Harmonization Approaches - Very active field of research

Many site adjustment / harmonization approaches are statistical harmonization techniques.

> Neuroimage. 2020 Sep;218:116956. doi: 10.1016/j.neuroimage.2020.116956. Epub 2020 May 26.

Increased power by harmonizing structural MRI site differences with the ComBat batch adjustment method in ENIGMA

Job: Front. Neuroinform., 08 January 2019
Volume 12 - 2018 |
<https://doi.org/10.3389/fninf.2018.00102>

This article is part of the Research Topic
Collaborative Efforts for Understanding the Human Brain
[View all 15 Articles >](#)

An Empirical Comparison of Meta- and Mega-Analysis With Data From the ENIGMA Obsessive-Compulsive Disorder Working Group



NeuroImage

Volume 261, 1 November 2022, 119509



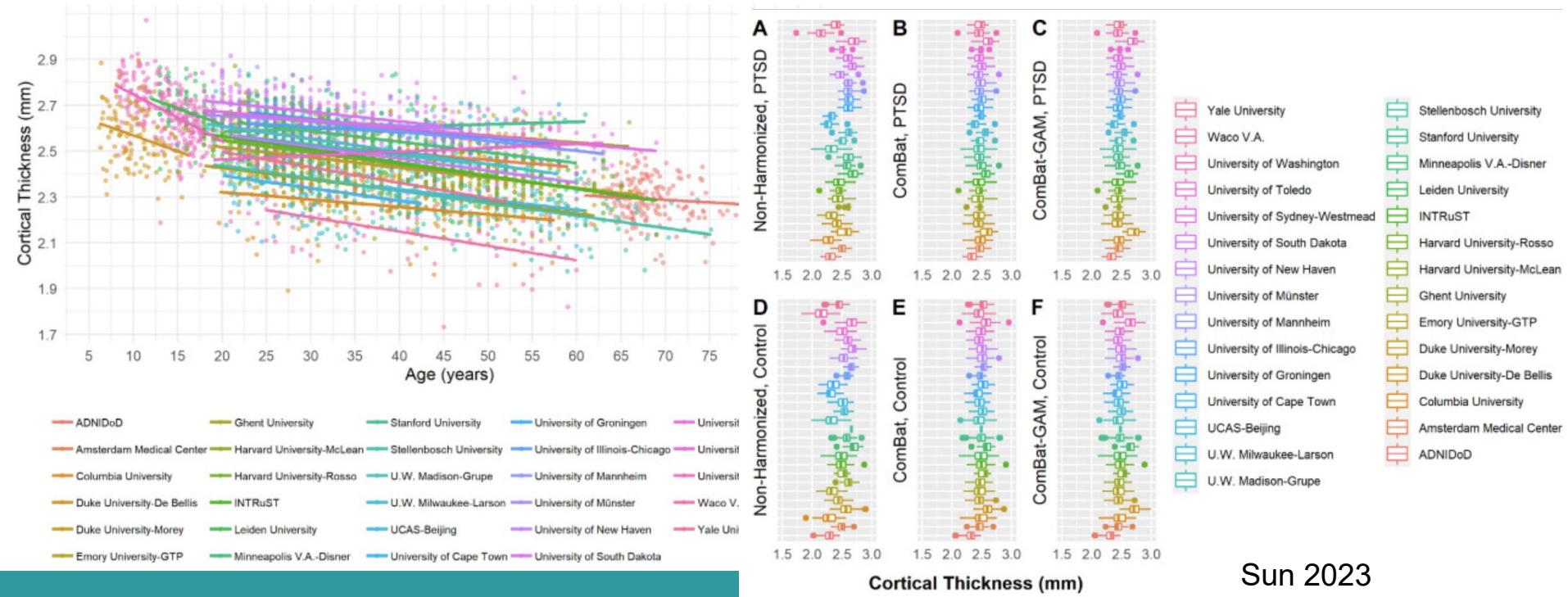
A comparison of methods to harmonize cortical thickness measurements across scanners and sites

Delin Sun ^{1 2 66}, Gopalkumar Rakesh ^{1 2}, Courtney C. Haswell ^{1 2}, Mark Logue ^{3 4 5 6},
C. Lexi Baird ^{1 2}, Erin N. O'Leary ⁷, Andrew S. Cotton ⁷, Hong Xie ⁷, Marijo Tamburrino ⁷,
Tian Chen ^{7 8}, Emily L. Dennis ^{8 9 10 11}, Neda Jahanshad ⁹, Lauren E. Salminen ⁹,
Sophia I. Thomopoulos ⁹, Faisal Rashid ⁹, Christopher R.K. Ching ⁹, Saskia B.J. Koch ^{12 13},
Jessie L. Frijling ¹², Laura Nawijn ^{12 14}, Mirjam van Zuiden ¹²...Rajendra A. Morey ^{1 2}

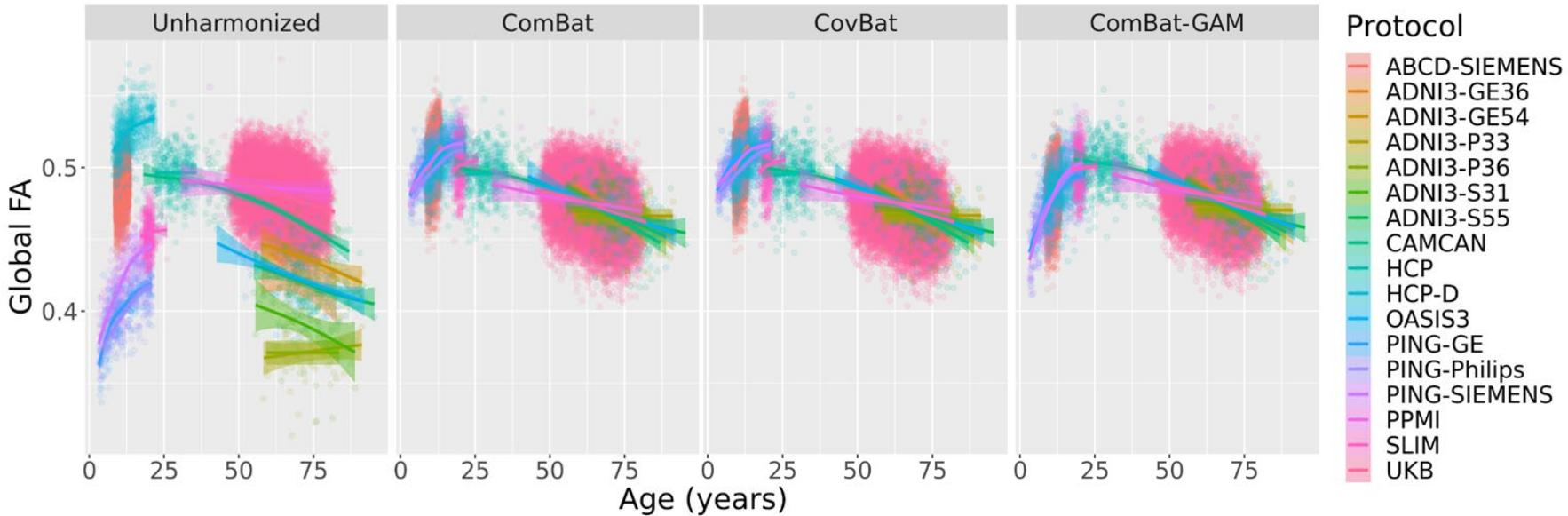


Harmonization Approaches - Very active field of research

Many harmonization approaches are statistical harmonization techniques.
COMBAT - very often used family of statistical harmonization methods



ComBat-GAMpreserves expected non-linear lifespan trajectories



Zhu et al - eHarmonize - in preparation



Many exciting developments in statistical data harmonization

README.md

ComBatHarmonization

Harmonization of multi-site imaging data with ComBat

Jean-Philippe Fortin, Drew Parker, Birkan Tunc, Takanori Watanabe, Michael Elliott, Kosha Ruparel, David R Roalf, Theodore D Satterthwaite, Ruben C Gur, Raquel E Gur, Robert T Schultz, Ragini Verma, Russell T Shinohara. Harmonization Of Multi-Site Diffusion Tensor Imaging Data. *NeuroImage*, 161, 149–170, 2017

Jean-Philippe Fortin, Nicholas Cullen, Yvette I. Sheline, Warren D. Taylor, Irem Aselcioglu, Philip A. Cook, Phil Adams, Crystal Cooper, Maurizio Fazio, Patrick J. McGrath, Melvin McInnis, Myrna M. Weissman, Russell T. Shinohara. Harmonizing volumetric MRI brain thickness measurements across sites. *NeuroImage*, 104–120, 2018

W. Evan Johnson and Cheng Li. Adapting gene expression data using empirical Bayes methods. *NeuroImage*, 46, 127, 2007.



NeuroImage

Volume 220, 15 October 2020, 117129



Longitudinal ComBat: A method for harmonizing multi-site longitudinal imaging data ★

Cook ^{a, d}, Christos Davatzikos ^e, Yvette I. Sheline ^{c, d, f}, Russell T. Shinohara ^b, Daniel H. Wolf ^b, Raquel Gur ^{b, e}, John Morris ^b, Marilyn S. Albert ^b, Hans J. Grabe ^b, Susan M. Resnick ^b, R. Nick Bryan ^b, David A. Wolk ^b, Russell T. Shinohara ^{b, c, f}, Haochang Shou ^{b, c, f}, Christos Davatzikos ^{a, b, c, d, e, f}. Longitudinal ComBat: A method for harmonizing multi-site longitudinal imaging data. *NeuroImage*, 2020 Oct 15; 220: 117127. doi: [10.1016/j.neuroimage.2020.117127](https://doi.org/10.1016/j.neuroimage.2020.117127)

PMCID: PMC7573655

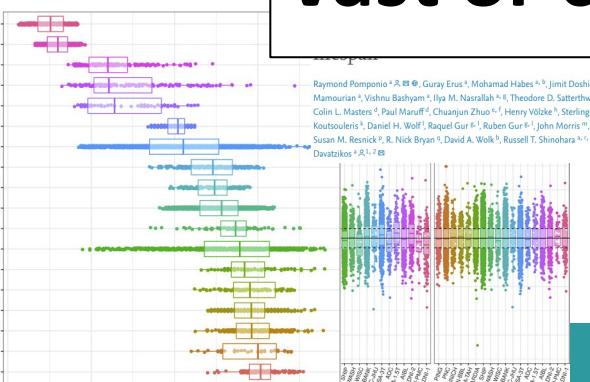
PMID: 32634595

NEUROIMAGE

ELSEVIER
FREE Full-Text Article

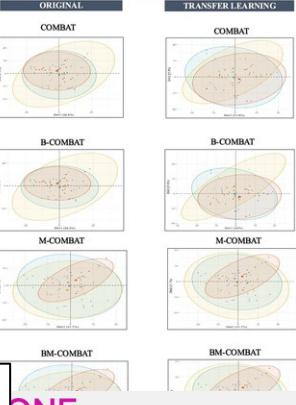
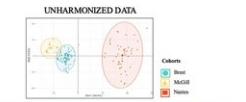
Neuroharmony: A new tool for harmonizing volumetric MRI data from multiple sites

What if site differences are too vast or confounded?



Raymond Pomponio ^{a, b, d}, Guray Erus ^a, Mohamad Habes ^{a, b}, Jimit Doshi ^a, Dhivya Srinivasan ^a, Elizabeth Mamourian ^a, Vishnu Bashyam ^a, Ilyas M. Nasrallah ^{a, b}, Theodore D. Satterthwaite ^a, Yong Fan ^a, Lenore J. Launer ^c, Colin L. Masters ^d, Paul Maruff ^d, Chuanjun Zhou ^{a, f}, Henning Völzke ^b, Sterling C. Johnson ^a, Jürgen Fripp ^a, Nikolas Koutsourelis ^a, Daniel H. Wolf ^b, Raquel Gur ^{b, e}, Ruben C. Gur ^b, John Morris ^b, Marilyn S. Albert ^b, Hans J. Grabe ^b, Susan M. Resnick ^b, R. Nick Bryan ^b, David A. Wolk ^b, Russell T. Shinohara ^{b, c, f}, Haochang Shou ^{b, c, f}, Christos Davatzikos ^{a, b, c, d, e, f}. Neuroharmony: A new tool for harmonizing volumetric MRI data from multiple sites. *NeuroImage*, 2021 Oct 15; 254: 118036. doi: [10.1016/j.neuroimage.2021.118036](https://doi.org/10.1016/j.neuroimage.2021.118036)

This article is a preprint and has not been certified by peer review (what does this mean?).



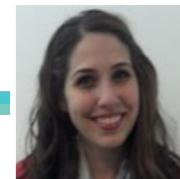
A multi-scanner neuroimaging data harmonization using RAVEL and ComBat

Mahbanch Eshaghzadeh Torbati ^a, Davneet S. Minhas ^b, Ghasan Ahmad ^c, Erin E. O'Connor ^c, John Muschelli ^c, Charles M. Laymon ^b, Zixi Yang ^b, Ann D. Cohen ^b, Howard J. Aizenstein ^b, William E. Klunk ^b, Bradley T. Christian ^b, Seong Jae Hwang ^{a, b}, Ciprian M. Crainiceanu ^d, Dana L. Tudorascu ^{a, b, c, d}.



Harmonization Needs Care - Confounded Cases

JAMA Network | Open™



Talia Nir, PhD

Dan Stein MD, PhD

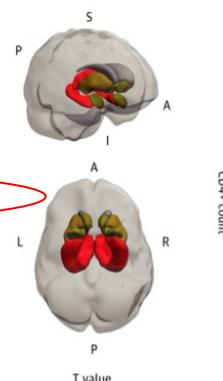
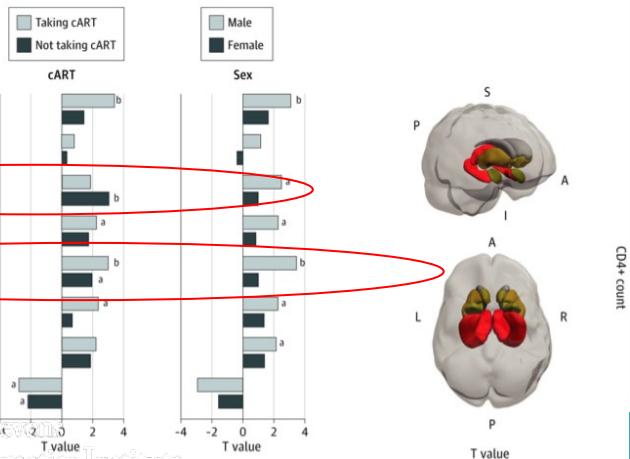
Original Investigation | Infectious Diseases

Association of Immunosuppression and Viral Load With Subcortical Brain Volume in an International Sample of People Living With HIV

Talia M. Nir, PhD; Jean-Paul Fouche, PhD; Jintananat Ananworanich, MD, PhD; Beau M. Ances, MD, PhD; Jasmina Boban, MD, PhD; Bruce J. Brew, MD; Joga R. Chaganti, MD; Linda Chang, MD, MS; Christopher R. K. Ching, PhD; Lucette A. Cysique, PhD; Thomas Ernst, PhD; Joshua Faskowitz, BA; Vikash Gupta, PhD; Jaroslaw Harezlak, PhD; Jodi M. Heaps-Woodruff, PhD; Charles H. Hinkin, PhD; Jacqueline Hoare, PhD; John A. Joska, PhD; Kalpana J. Kallianpur, PhD; Taylor Kuhn, PhD; Hei Y. Lam, BA; Meng Law, MD; Christine Lebrun-Fréney, MD, PhD; Andrew J. Levine, PhD; Lydiane Mondot, MD; Beau K. Nakamoto, MD; Bradford A. Navia, MD, PhD; Xavier Pennec, PhD; Eric C. Porges, PhD; Lauren E. Salminen, PhD; Cecilia M. Shikuma, MD; Wesley Surento, MS; April D. Thamés, PhD; Victor Valcour, MD, PhD; Matteo Vassallo, MD; Adam J. Woods, PhD; Paul M. Thompson, PhD; Ronald A. Cohen, PhD; Robert Paul, PhD; Dan J. Stein, PhD; Neda Jahanshad, PhD

CD4 continuous

A Subcortical volumes and CD4+ cell counts



Site/Cohort confounded by demographic variables

Table 1. Demographic and Clinical Information by Study and Scanning Site

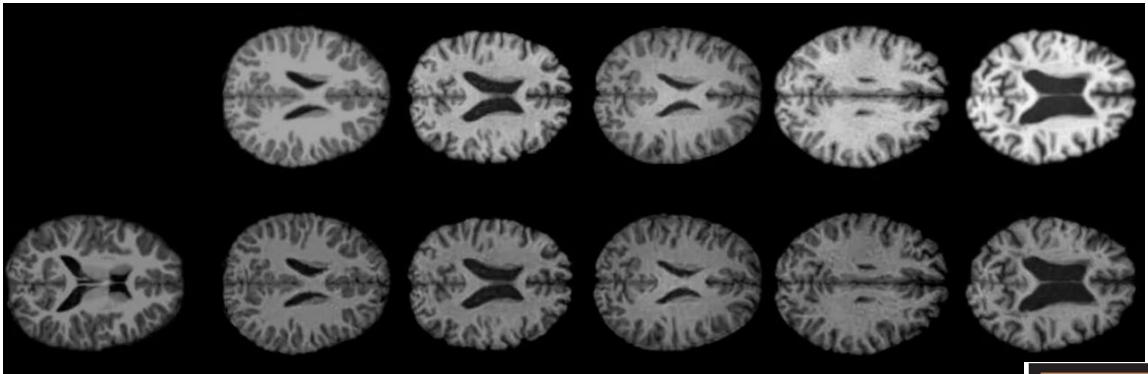
Site	Total, No.	Male, No. (%)	Age, mean (SD) [range], y	Taking cART	Detectable viral load	No. (%) of participants	
						CD4+ cell count, mean (SD), /µL	CD4+ cell count <200/µL
HIVNC (7 sites), US	218	187 (86)	48.5 (8.3) [24-71]	218 (100)	61 (28) [n = 217]	378.0 (231.9)	48 (22)
Site 1: University of California, San Diego	21	21 (100)	47.6 (5.0) [37-56]	21 (100)	13 (65) [n = 20]	313.7 (328.4)	11 (52) 7 (35) [n = 20]
Site 2: Harbor UCLA Medical Center, Los Angeles, California	52	42 (81)	46.6 (9.1) [24-70]	52 (100)	8 (15)	350.2 (173.0)	14 (27) 5 (10)
Site 3: Stanford University, Stanford, California	10	9 (90)	46.4 (10.4) [31-62]	10 (100)	0	313.5 (157.5)	2 (20) 0
Site 4: Colorado	35	34 (97)	49.2 (7.9) [31-66]	35 (100)	14 (40)	446.6 (246.0)	4 (11) 6 (17)
Site 5: Pittsburgh, Pennsylvania	19	18 (95)	49.9 (9.7) [34-71]	19 (100)	4 (21)	438.9 (298.2)	3 (16) 3 (16)
Site 6: Rochester University, Rochester, New York	41	26 (63)	48.4 (7.7) [26-62]	41 (100)	11 (27)	382.0 (207.8)	6 (15) 7 (17)
Site 7: University of California, Los Angeles	40	37 (93)	50.7 (8.4) [28-69]	40 (100)	11 (28)	370.7 (224.7)	8 (20) 11 (28)
University of Hawaii, Honolulu, US ¹⁷	175	159 (91)	47.6 (10.5) [20-74]	157 (90)	54 (33) [n = 164]	475.2 (281.4)	31 (18) 38 (23) [n = 164]
University of Hawaii, Honolulu, US ¹⁸	53	45 (85)	50.9 (8.0) [40-71]	53 (100)	6 (11)	491.1 (208.5)	4 (8) 2 (4)
University of California, San Francisco, US	50	49 (98)	63.6 (2.5) [60-69]	49 (98)	14 (29) [n = 49]	529.3 (218.5)	0 4 (8) [n = 49]
Brown University, Providence, Rhode Island, US	79	48 (61)	45.2 (9.5) [23-65]	66 (84)	22 (28)	476.2 (228.8)	7 (9) 21 (27)
University of California, Los Angeles, US ¹⁹	12	12 (100)	46.2 (8.5) [26-57]	9 (75)	6 (50)	604.1 (289.1)	1 (8) 4 (33)
University of California, Los Angeles, US ²⁰	51	46 (90)	50.9 (13.2) [24-76]	51 (100)	21 (41)	603.2 (287.7)	2 (4) 9 (18)
University of New South Wales, New South Wales, Australia ²¹	39	38 (97)	52.8 (8.2) [39-75]	39 (100)	10 (26)	612.5 (269.0)	0 6 (15)
University of New South Wales, New South Wales, Australia ²²	68	68 (100)	55.3 (6.7) [44-69]	68 (100)	1 (1)	549.5 (273.6)	5 (7) 1 (1)
SEARCH-011 Consortium, Bangkok, Thailand	61	26 (43)	34.2 (7.0) [22-56]	0	61 (100)	236.0 (139.0)	25 (41) 61 (100)
University of Cape Town, South Africa	181	25 (14)	32.4 (5.0) [22-46]	0	148 (97) [n = 152]	225.5 (149.2)	95 (53) 129 (88) [n = 147]
Nice University, Nice, France	155	122 (79)	45.4 (10.0) [22-81]	126 (76)	56 (36)	580.9 (277.6)	11 (7) 39 (25)
University of Novi Sad, Novi Sad, Serbia	61	55 (90)	44.3 (10.9) [25-66]	61 (100)	10 (16)	616.3 (346.2)	1 (2) 10 (16)

Abbreviations: cART, combination antiretroviral therapy; HIVNC, HIV Neuroimaging Consortium; SEARCH, South East Asian Research Collaboration in HIV.

More Active Research in Harmonization Methods



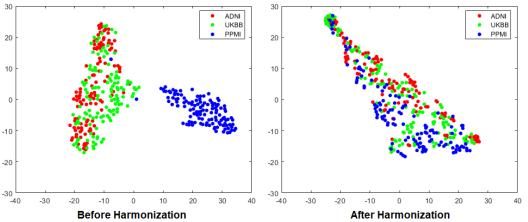
Image Harmonization with Style-Encoding GAN



Liu et al, MICCAI 2021



Mengting Liu, PhD



HUMAN BRAIN MAPPING

Open Access

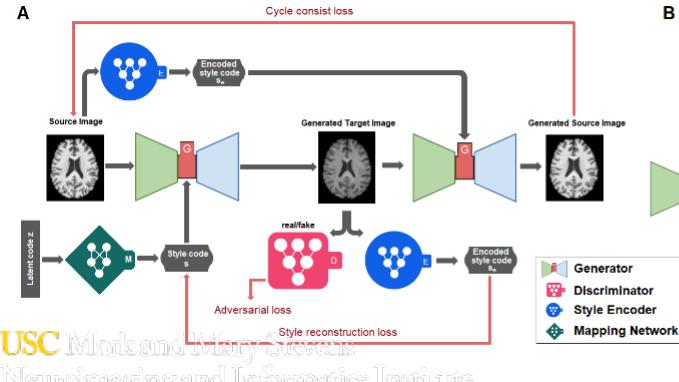
RESEARCH ARTICLE | [Open Access](#) |

Style transfer generative adversarial networks to harmonize multisite MRI to a single reference image to avoid overcorrection

Mengting Liu Alyssa H. Zhu, Piyush Maiti, Sophia I. Thomopoulos, Shruti Gadewar, Yaqiong Chai, Hosung Kim Neda Jahanshad for the Alzheimer's Disease Neuroimaging Initiative

First published: 20 July 2023 | <https://doi.org/10.1002/hbm.26422>

COV: Convolution
IN: Instance Normalization
AdIN: Adaptive Instance Normalization
FC: Fully Connected Layer



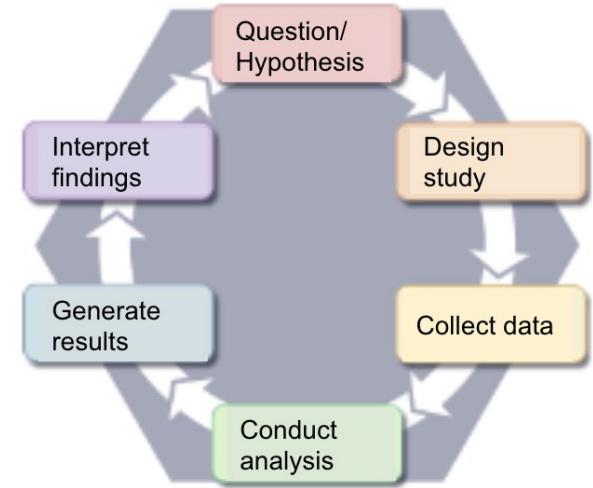
We can continuously integrate more data

More studies can be aggregated

New populations

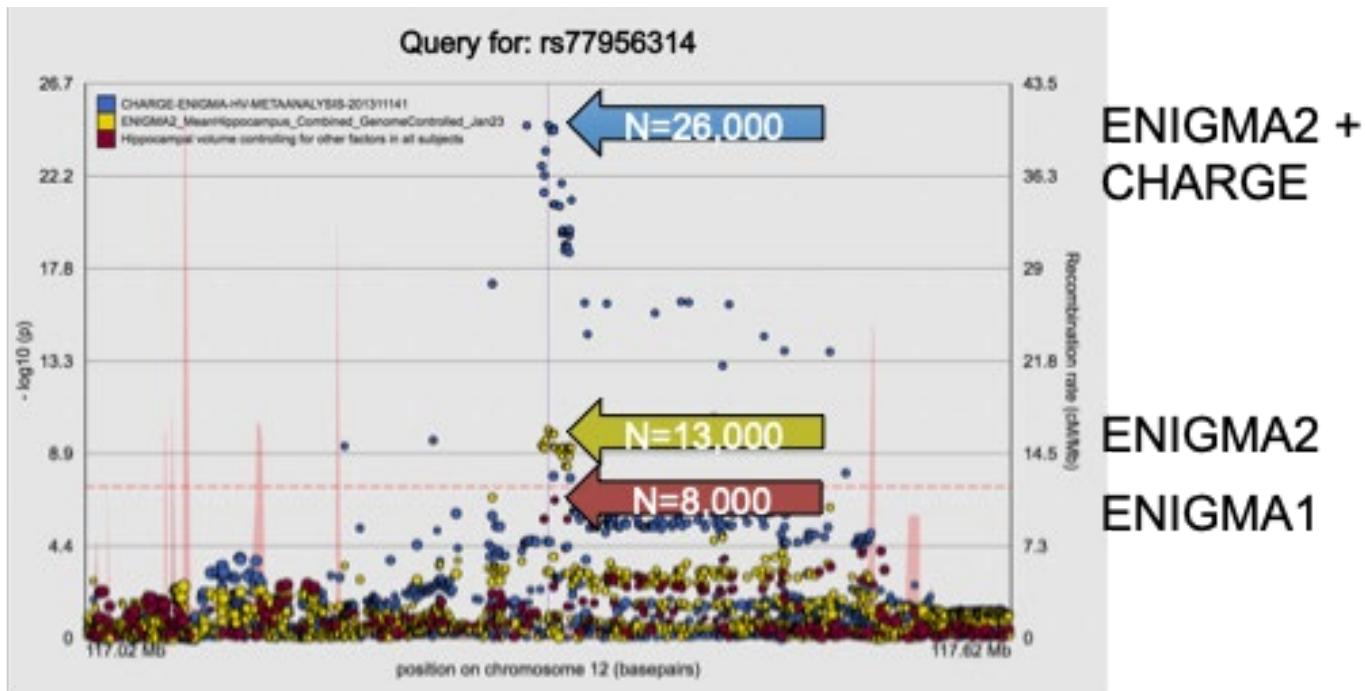
New technologies

Or nothing new! Can also be used to validate findings, refine conclusions and confidence



Can always expand on published work

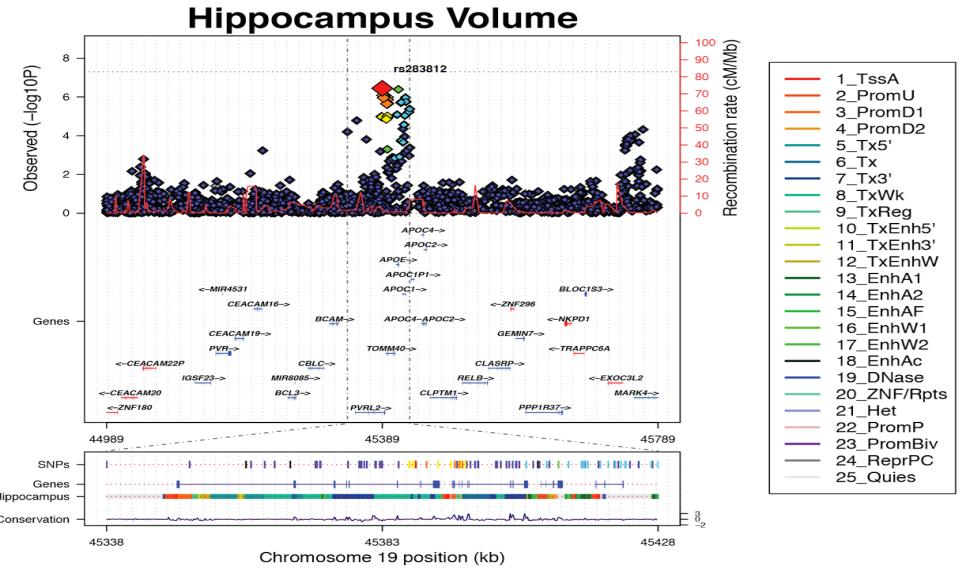
Hippocampal volume GWAS across initiatives



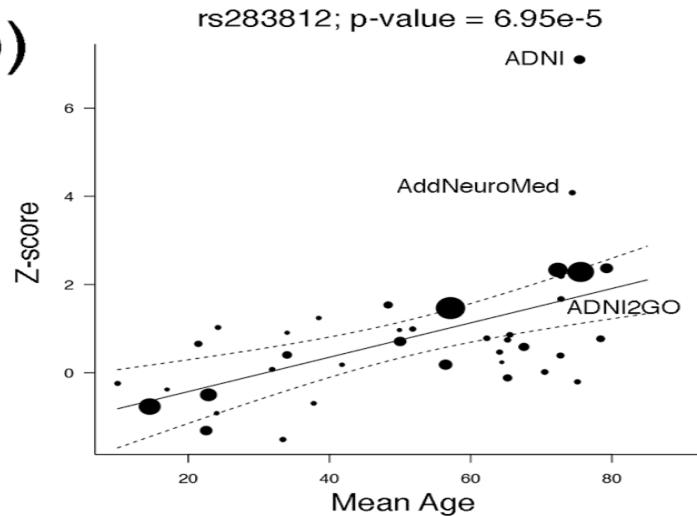
Explore Generalizability / Lack of

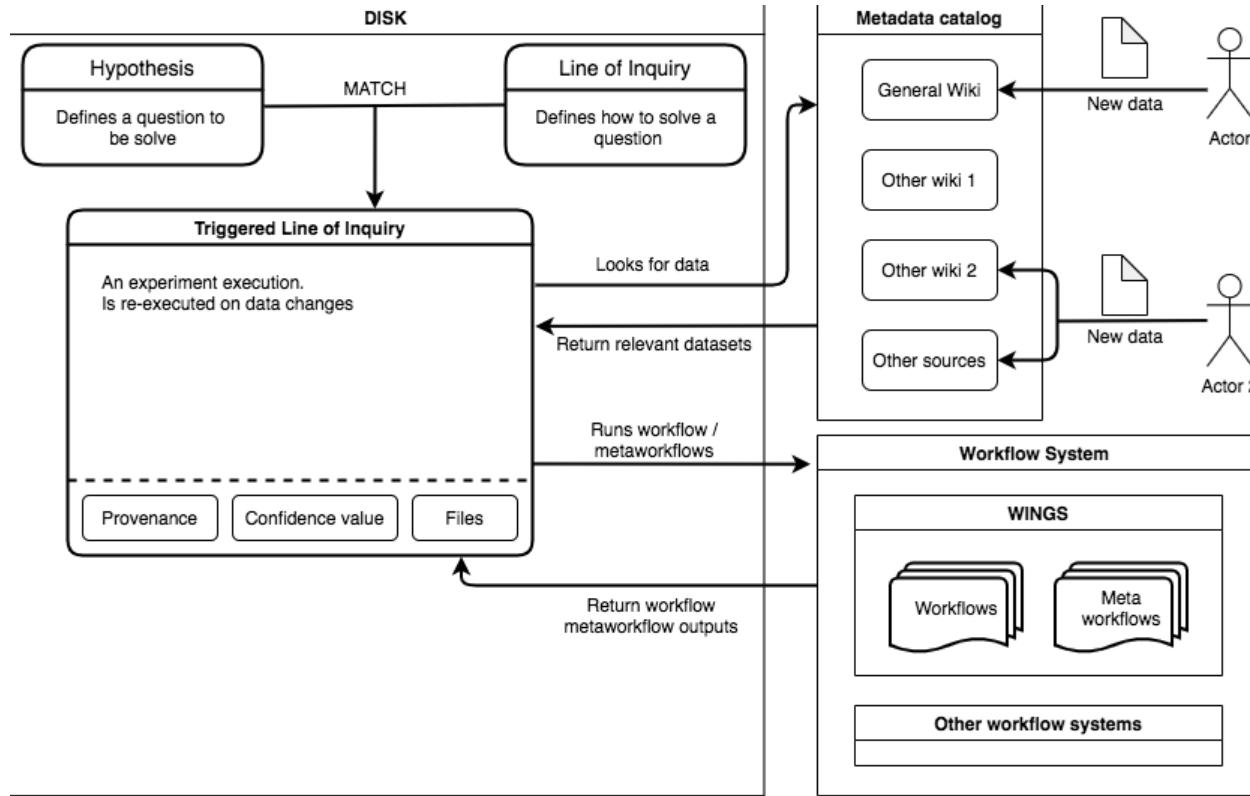
APOE – Near GWAS Significance - Why?

(a)

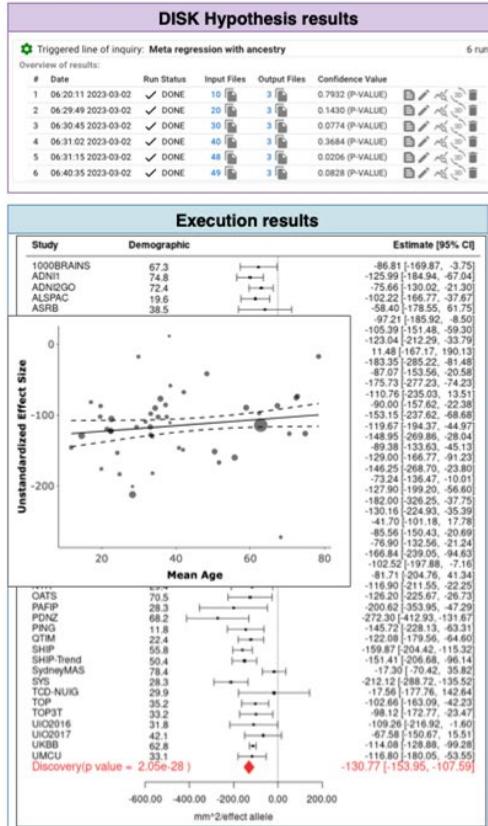


(b)





Data Standards
Documentation /
Containerization
Analytic
Provenance



DISK Hypothesis

Hypothesis or question:

The hypothesis or question to be tested:
Is the effect size of rs1080066 on SA of Precentral associated with HasAge Mean for cohorts of European ancestry

DISK Line of Inquiry

Data query:

```

Data source: Enigma_wiki Data obtained from the Enigma.ODS.wiki (Ontology) a collaborative wiki for different working groups of Enigma
Data query *
1 ?Cohorts a enigma:Category-3ACohort_-28E-29 .
2 ?Cohorts enigma:Property-3AHasEthnicGroup_-28E-29 ?EthnicGroup .

...
15 ?ProjectResults enigma:Property-3ASNP ?Genotype .
16 ?ProjectResults enigma:Property-3AHasContentUrl_-28E-29 ?CohortData .

```

Methods:

Meta-Regression

Meta regression is considered as an extended statistical model of meta analysis, regressing the effect size against variable(s) of interest to account for the systematic differences of the effect sizes being meta-analyzed

area: ?Region

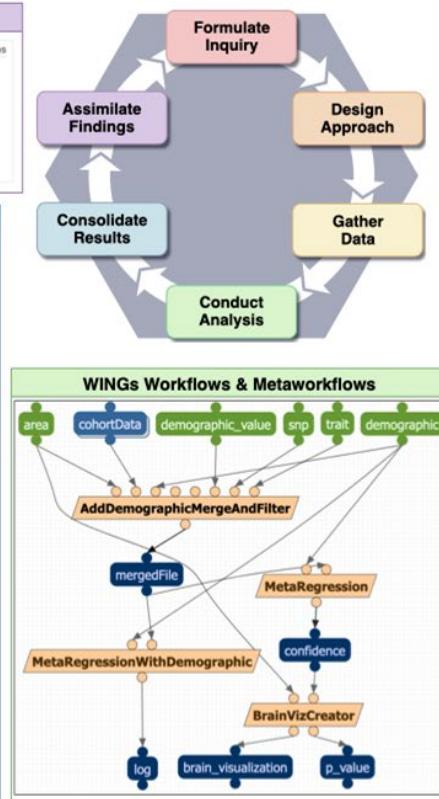
cohortData: ?CohortData

demographic: ?DemographicLabel

demographic_value: ?TraitDemographic

sep: ?Genotype

trait: ?BrainImagingTrait



ENIGMA Consortium Wikis

Main Page

Contents [hide]

- Welcome to the ENIGMA Organic Data Science wiki!
- Recent changes
- Random page
- Help
- Tools
- What links here

Recently Changed Pages

- Bodhoo Premika SW
- Bommes Jessica
- Chakravarty Mallar
- Zwers Marcel P
- Brydt Joshua D
- Ames David
- Bustillo Juan R

ABCD proj ENIGMA3 Cortical GWAS (CohortProject (E))

HasAge Mean (E)	9.96
HasBrainScanDataType (E)	T1w
HasNumberOfFemaleSex (E)	1582
HasNumberOfMaleSex (E)	2026
HasNumberOfParticipants (E)	3,608
HasProjectResults (E)	ABCD ENIGMA3 Cortical GWAS Results
IsPartOfProject (E)	Proj ENIGMA3 Cortical GWAS
	ENIGMA3 Cortical GWAS

Extra information

Incoming Properties

ABCD > HasCohortProject (E) > ABCD proj ENIGMA3 Cortical GWAS



ENIGMA Resources - continuously being updated



Protocols

ENIGMA-git

► ENIGMA sMRI Imaging

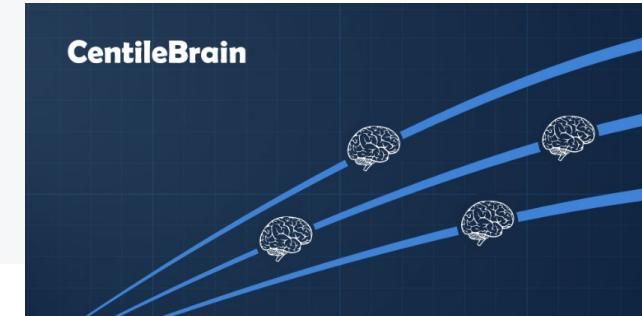
► ENIGMA DTI Imaging

► ENIGMA Genetics Protocols

► ENIGMA Functional Imaging

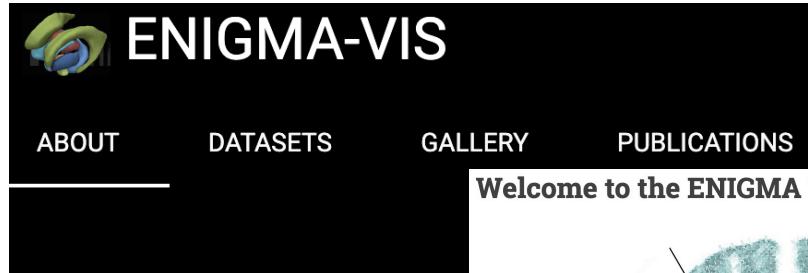
► ENIGMA Statistical Protocols

► ENIGMA Visualization tools



ENIGMA-U: Free Online Neuroscience and Neuroimaging Course

<https://biabl.org/enigma-u/>



Welcome to the ENIGMA TOOLBOX 🙌

ENIGMA
TOOLBOX



Thank you to the thousands of ENIGMA participants-- group leads, investigators and research volunteers!

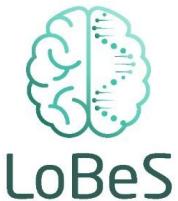
QUESTIONS?

Contact:

Neda.Jahanshad@ini.usc.edu

Join ENIGMA:

<http://enigma.ini.usc.edu>



ENIGMA and Global Neuroscience: A Decade of Large-Scale Studies of the Brain in Health and Disease across more than 40 Countries

AUTHORS

Paul Thompson, Neda Jahanshad, Christopher R. K. Ching, Lauren E. Salminen, Sophia Thomopoulos, Joanna Bright, Bernhard T. Baune, Sara Bertolini, Janita Bralten, Willem B. Bruin, Robin Bülow, Jian Chen, Yann Chye, Udo Dannowski, Carolien de Kovel, Gary Donohoe, Lisa Elyer, Stephen V. Faraone, Pauline Favre, Courtney Filippi, Thomas Frodl, Daniel Garijo, Yolanda Gil, Hans J. Grabe, Katrina L. Grasby, Tomas Hajek, Laura Han, Sean Hatton, Kevin Hilbert, Tiffany C. Ho, Laurena Hollerer, Georg Homuth, Norbert Hosten, Josselin Houenou, Iliyan Ivanov, Tianye Jia, Sinead Kelly, Marieke Klein, Jun Soo Kwon, Max A. Laansma, Jeanne Leerssen, Ulrike Lueken, Abraham Nunes, Joseph O'Neill, Nils Opel, Fabrizio Piras, Federica Piras, Merel Postema, Elena Pozzi, Natalia Shatokhina, Carles Soriano-Mas, Gianfranco Spalletta, Daqiang Sun, Alexander Teumer, Amanda K. Tilot, Leonardo Tozzi, Celia van der Merwe, Eus J. W. Van Someren, Guido van Wingen, Henry Völzke, Esther Walton, Lei Wang, Anderson M. Winkler, Katharina Wittfeld, Margaret J. Wright, Je-Yeon Yun, Guohao Zhang, Yanli Zhang-James, Bhim M. Adhikari, Ingrid Agartz, Moji Aghajani, André Aleman, Robert R. Althoff, Andre Altmann, Ole A. Andreassen, David A. Baron, Brenda L. Bartnik-Olson, Janna Marie Bas-Hoogendam, Arielle R. Baskin-Sommers, Carrie E. Bearden, Laura A. Berner, Premika S. W. Boedhoe, Rachel M. Brouwer, Jan K. Buitelaar, Karen Caeyenberghs, Charlotte A. M. Cecil, Ronald A. Cohen, James Cole, Patricia J. Conrod, Stephane De Brito, Sonja M. C. de Zwart, Emily Dennis, Sylviane Desrivières, Danai Dima, Stefan Ehrlich, Carrie Espenek, Graeme Fairchild, Simon E. Fisher, Jean-Paul Fouché, Clyde Francks, Sophia Frangou, Barbara Franke, Hugh P. Garavan, David C. Glahn, Nynke A. Groenewold, Tiril Gurholt, Boris A. Gutman, Tim Hahn, Ian Harding, Dennis Hernaus, Derrek P. Hibar, Frank Hillary, Martine Hoogman, Hilleke E. Hulshoff Pol, Maria Jalbrzikowski, George A. Karkashadze, Eduard Klapwijk, Rebecca C. Knickmeyer, Peter Kochunov, Inga Koerte, Xiang-Zhen Kong, Sook-Lei Liew, Alexander P. Lin, Mark W. Logue, Eileen Lunders, Fabio Macciardi, Scott Mackay, Andrew R. Mayer, Carrie R. McDonald, Agnes B. McMahon, Sarah E. Medland, Gemma Modinos, Rajendra A. Morey, Sven Mueller, Pratik Mukherjee, Leyla Namazova-Baranova, Talia M. Nir, Alexander Olsen, Peristera Paschou, Daniel Pine, Fabrizio Pizzagalli, Miguel E. Rentería, Jonathan D. Rohrer, Philipp G. Sämann, Lianne Schmaal, Gunter Schumann, Mark S. Shiroishi, Sanjay M. Sisodiya, Dirk J. A. Smit, Ida E. Sønderby, Dan J. Stein, Jason L. Stein, Masoud Tahmasian, David Tate, Jessica Turner, Odile van den Heuvel, Nic van der Wee, Ysbrand D. van der Werf, Theo G. M. van Erp, Neeltje van Haren, Daan van Rooij, Laura S. van Velzen, Ilya Veer, Dick J. Veltman, Julio E. Villalon-Reina, Henrik Walter, Christopher D. Whelan, Elisabeth A. Wilde, Mojtaba Zarei, Vladimir Zelman



Individual cohort studies would not be possible without international and national support:

NHMRC, DCRC, BMBF, Swedish Research Council, German Ministry of Cultural Affairs, Social Ministry of the Federal State of Mecklenburg-West Pomerania, NGFN, Siemens, ISCIII, SENY Fundació, NOW, BBMRI-NL, CBF, Hersenstichting Nederland, Alzheimer's Australia Dementia Research MJFF Foundation, Autism Speaks, NIDA, NIMH, NIBIB, NICHD, NINDS, NIA Kavli Foundation



Questions?

neda.jahanshad@usc.edu



Data-quality dependent protocols

Sometimes, may need to restrict participating sites and datasets based on image acquisition protocols

Much of the clinically focused datasets are optimized to collect information from multiple domains and may limit time in the scanner, so acquisition protocols are not always optimal for all analyses

% of ENIGMA	DTI b<=2000	HARDI b<=2000	Multishell b<=2000	DTI b>=3000	HARDI b>=3000	Multishell b>=3000
> 2.5mm	~10%	~2%				
2-2.5mm	~70%	~10%	~5%	~2%		
< 2mm						Max tracts



Analytical Harmonization: Quality Assurance

Tens of thousands of scans have been visually assessed for quality according to ENIGMA guidelines → Let's use this information!

We are working to develop tools that help automate much of the process.



Shruti Gadewar

Region and hemisphere specific automatic quality control for MRI-derived cortical segmentations



2021 IEEE 18th International Symposium on Biomedical Imaging (ISBI)
April 13-16, 2021, Nice, France

REGION SPECIFIC AUTOMATIC QUALITY ASSURANCE FOR MRI-DERIVED CORTICAL SEGMENTATIONS

Shruti Gadewar, Alyssa H. Zhu, Sophia I. Thomopoulos, Zuocheng Li,
Iyad Ba Gari, Piyush Maiti, Paul M. Thompson, Neda Jahanshad

Imaging Genetics Center, Mark and Mary Stevens Neuroimaging and Informatics Institute,
Keck School of Medicine, University of Southern California, Los Angeles, CA, USA

The screenshot shows a dark-themed web application interface. At the top right is a circular icon with a white cat-like head. Below it, the text "USC-IGC / FreeSurfer_Cortex_AutoQC" is displayed in blue, with "Public" in smaller text to its right. The main area features a 3D brain model with various colored regions representing different cortical segments. A legend on the left indicates color coding for different regions.

ENIGMA Suicidal Thoughts and Behaviors

Site		Age HC (years)	Age CC (years)	Age Attempt	% female HC	% female CC	% female Attempt	Total N HC	Total N CC	Total N Attempt
<i>A. Mood disorders only sample</i>										
DEP-ARREST-CLIN - MOODS (France)	-	22.0 (18–25)	19.0 (18–23)	-	70.0	55.6	0	10	9	
Houston BD (USA)	14.0 (8–25)	15.0 (8–25)	17.0 (11–24)	48.5	34.7	69.2	97	75	13	
Sydney Bipolar Kids and Siblings	21.0 (13–25)	22.5 (16–25)	22.0 (18–25)	50.0	62.5	66.7	64	16	9	
Sydney Brain and Mind Centre (Australia)	-	19.5 (15–25)	20.5 (15–24)	-	66.7	100.0	0	48	8	
University of Minnesota (USA)	16.0 (14–20)	16.0 (12–19)	17.5 (12–19)	66.7	78.0	62.5	12	41	8	
University of Texas- Austin - Bipolar Seed Program (USA)	21.0 (18–25)	21.0 (18–25)	21.0 (19–25)	69.2	77.8	66.7	26	18	9	
UWashington/Harvard (USA)	11.5 (8–16)	15.0 (12–16)	14.5 (9–17)	53.7	55.6	75.0	54	9	8	
Total	16.0 (8–25)	17.0 (8–25)	19.0 (9–25)	53.0	58.1	70.3	253	217	64	
<i>B. Transdiagnostic sample</i>										
DEP-ARREST-CLIN - MOODS (France)		22.0 (18–25)	19.0 (18–23)		70.0	55.6	0	10	9	
Fondazione Santa Lucia - Schizophrenia sample (Italy)		23.0 (16–25)	24.0 (20–25)		5.3	57.1	0	19	7	
Houston BD (USA)	14.0 (8–25)	15.0 (8–25)	17.0 (11–24)	48.5	34.7	69.2	97	75	13	
PAFIP1 (Spain)		22.0 (17–25)	22.0 (19–24)		25.5	14.3	0	51	7	
PAFIP2 (Spain)		21.0 (17–25)	22.5 (17–25)		24.1	50.0	0	58	10	
Sydney Bipolar Kids and Siblings (Australia)	21.0 (13–25)	22.0 (16–25)	22.0 (18–25)	50.0	60.9	66.7	64	23	9	
Sydney Brain and Mind Centre (Australia)		19.0 (13–25)	19.0 (15–24)		68.3	100	0	101	10	
University of Minnesota (USA)	16.0 (14–20)	16.0 (12–19)	17.5 (12–19)	66.7	78.0	62.5	12	41	8	
University of Texas- Austin - Bipolar Seed Program (USA)	21.0 (18–25)	21.0 (18–25)	21.0 (19–25)	69.2	77.8	66.7	26	18	9	
UWashington/Harvard (USA)	11.5 (8–16)	12.5 (8–16)	13.0 (9–17)	53.7	52.8	77.8	54	36	9	
Total	16.0 (8–25)	19.5 (8–25)	20.0 (9–25)	53.0	48.4	63.7	253	432	91	



ENIGMA

