Gene-brain-behavior continuums across neurodegenerative disorders in ONDRI

Derek Beaton, Allison Dilliot, Joel Ramirez, Chris J. M. Scott, Mario Masellis, Paula M. McLaughlin, Stephen C. Strother

JB Orange, Angela Roberts, Angela Troyer, Brian Levine, Donna Kwan, Kelly Sunderland, Malcolm A. Binns, Robert A. Hegele, John F Robinson, Emily C Evans, Sali MK Farhan, Ekaterina Rogaeva, Ming Zhang, Dennis E Bulman, Sandra E. Black, Morris Freedman, Richard H. Swartz, Demetrios Sahlas, Anthony Lang, David Grimes, Mandar Jog, Tom Steeves, Elizabeth Finger, Carmela Tartaglia, Lorne Zinman, John Turnbull, Christian Shoesmith, ONDRI Investigators

dbeaton@research.baycrest.org

@derek__beaton

INTRODUCTION

- Recent work across psychiatric and neurological disorders has sought to better understand (or even redefine) clinical and pathological aspects of disorders. Often these works focus on “bottom up” perspectives such as the Research Domain Criteria (RDoC), as opposed to a “top down” perspective (DSM-like).
- The Ontario Neurodegenerative Disease Research Initiative is a multicentre investigation of neurodegenerative and cerebrovascular disorders that has:
  - breadth of disorders: Alzheimer’s Disease/Mild Cognitive Impairment (ADMCI), vascular disease with possible cognitive impairment (VCI), Parkinson’s (PD), amyotrophic lateral sclerosis (ALS), and frontotemporal dementia (FTD),
  - depth of assessments, and
  - expertise within and across assessments and disorders
- We leverage that breadth, depth, and expertise to identify gene-brain-behavior profiles from clinical and potential biomarker measures in a data-driven way. We focus our analyses on key measures and techniques as determined by domain experts within ONDRI. This includes:
  - a comprehensive cognitive battery
  - structural neuroimaging with an emphasis on pathological measures
  - well-established pathological genetic markers (ApoE, MAPT, C9orf72)
  - multivariate statistical techniques to handle mixed data (categorical, ordinal, continuous) with *a priori* structures
- We approach the problem in two ways:
  - A DSM-inspired way that emphasizes the clinical groups through a discriminant analysis (with classification), and
  - An RDoC-inspired way that (1) reveals spectrums/continuums and (2) ignores the clinical groups, in order to (3) reveal groups from the results

METHODS

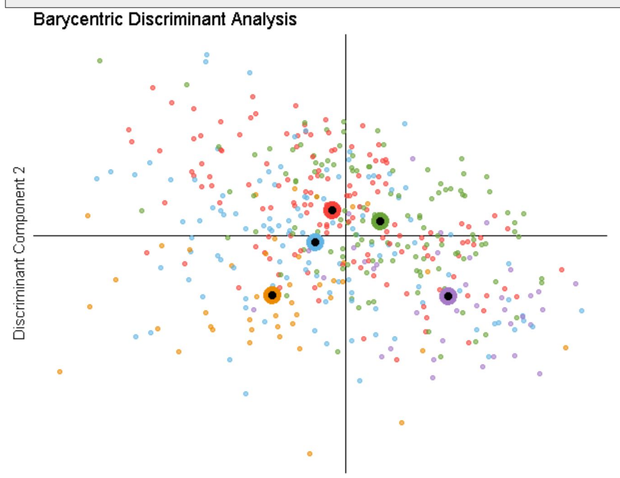
- ONDRI includes five well characterized cohorts with a range of disorders, pathology, and cognitive impairment. Some of these disorders are inherently heterogeneous, e.g.,
  - VCI have covert or overt strokes in various locations and size
  - Some individuals in FTD, PD, and ALS have overt strokes
  - AD/MCI includes MCI, probable AD with amnestic presentation, and probably AD with non-amnestic presentation
  - FTD spans multiple subtypes including behavioral variant, progressive supranuclear palsy, primary progressive aphasia, semantic dementia, corticobasal syndrome, and mixed FTD diagnoses.
- Genetic data were: *APOE* genotype (categorical), *MAPT* diplotype (categorical), and *C9orf72* G4C2 repeat expansions (ordinal). SABRE produced 28 bilateral lobar region volumes (e.g., inferior parietal, hippocampus) for: normal appearing white & grey matter (NAWM/GM); ventricular & sulcal CSF (v/sCSF); periventricular & deep white matter hyperintensities (p/dWMH); periventricular & deep lacunes; perivascular spaces (PVS); and stroke volumes. Volumes were expressed as % NAGM, % NAWM, and % of remaining tissue “aberrant” types within each region (summed to 100% per region) for each participant. The neuropsychological battery included 20 variables from 14 clinical tests across 5 domains: Attention/working memory; visuospatial; memory; language; executive function.
- We used Multiple Factor Analysis (MFA) and a discriminant version of MFA (dMFA) to accommodate the multiple data structures: we had 113 columns grouped together into 36 subtables: 3 genetics (each gene), 28 neuroimaging (each region), and 5 cognitive (each domain).
- MFA and dMFA were both supplemented with resampling procedures to identify stable components, and the measures that contribute to those components. Additionally, we used hierarchical clustering for the MFA results to reveal data-driven groups and compared those with the clinically defined groups.

Some of this research was conducted with the support of the Ontario Brain Institute, an independent non-profit corporation, funded partially by the Ontario government. The opinions, results, and conclusions are those of the authors and no endorsement by the Ontario Brain Institute is intended or should be inferred. DB is partly supported by a Canadian Institutes of Health Research grant (MOP 201403) to SCS.

This work was completed on behalf of the ONDRI. The authors would like to acknowledge the ONDRI Founding Authors: Robert Bartha, Sandra E. Black, Michael Borrie, Dale Corbett, Elizabeth Finger, Morris Freedman, Barry Greenberg, David A. Grimes, Robert A. Hegele, Chris Hudson, Anthony E. Lang, Mario Masellis, William E. McIlroy, Paula M. McLaughlin, Manuel Montero-Odasso, David G. Munoz, Douglas P. Munoz, J.B. Orange, Michael J. Strong, Stephen C. Strother, Richard H. Swartz, Sean Symons, Maria Carmela Tartaglia, Angela Troyer, and Lorne Zinman.

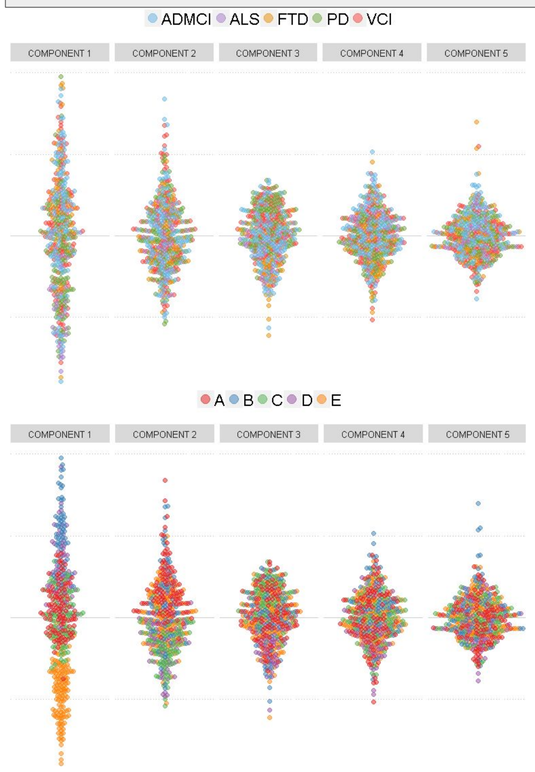
RESULTS

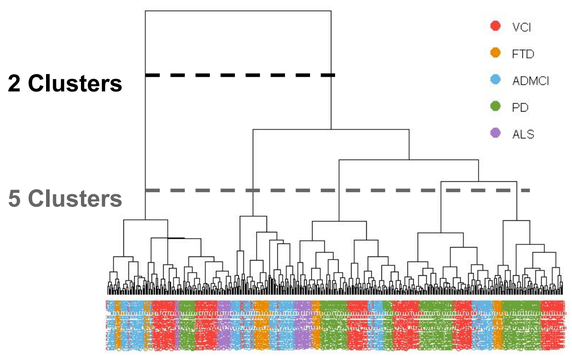
ANALYSIS OF CLINICALLY DEFINED GROUPS



Actual	ADMCI N=112	ALS N=36	FTD N=42	PD N=126	VCI N=140
Predicted					
ADMCI	**39 (34.8%)**	1	6	11	20
ALS	18	**27 (75.0%)**	3	34	27
FTD	25	1	**27 (64.3%)**	11	24
PD	18	3	3	**47 (37.3%)**	23
VCI	12	4	3	23	**46 (32.9%)**

ANALYSIS TO FIND PATHOLOGICAL SPECTRUMS AND GROUPS





		ADMCI N=112 (%)	ALS N=36 (%)	FTD N=42 (%)	PD N=126 (%)	VCI N=140 (%)
2 clusters						
	Cluster {A+B+C+D}	91 (81%)	11 (31%)	37 (88%)	80 (63%)	111 (79%)
	Cluster E	21 (19%)	25 (69%)	5 (12%)	46 (37%)	29 (21%)
5 clusters						
	Cluster A	25 (22%)	4 (11%)	6 (14%)	28 (22%)	43 (31%)
	Cluster B	29 (26%)	1 (3%)	9 (21%)	8 (6%)	16 (11%)
	Cluster C	20 (18%)	5 (14%)	2 (5%)	34 (27%)	21 (15%)
	Cluster D	17 (15%)	1 (3%)	20 (48%)	10 (8%)	21 (15%)
	Cluster E	21 (19%)	25 (69%)	5 (12%)	46 (37%)	29 (21%)

RESULTS

- **dMFA produced 4 components**. We only discuss the first two. Component 1 shows a strong dissociation driven by FTD vs. ALS. Component 2 shows a smaller dissociation of {ADMCI/VCI/PD} vs. {FTD/ALS}. Little-to-no genetic contributions in this model.
  - Fixed effects classification show that ALS and FTD—even with its many subtypes—are relatively high. The other groups are misclassified almost as often as they are correctly classified
  - PD and ALS generally show higher intact cognition and in particular, higher preservation of white matter
  - ADMCI and VCI (especially) show more diffuse aberrant tissue and lower overall performance on the cognitive battery
  - FTD has low performance on the cognitive battery with highly focused aberrant tissue: generally right and left temporal regions
- **MFA revealed 5 components and within those 5 components two configurations for clustering: a two group solution and a five group solution**.
  - MFA Component 1 shows relatively high and intact cognition vs. *MAPT* diplotypes and *APOE* risk alleles with aberrant volumes. Component 1 dissociates Cluster E (bottom, orange, highly intact) from the rest.
  - Component 2 dissociates two aberrant tissue patterns: superior frontal and parietal (in clusters C & D; lower, purple-green) vs. basal ganglia/thalamic, inferior parietal and posterior temporal (in clusters A & B; upper, red-blue).
  - Component 3 dissociates preserved aspects of language and executive function vs. higher proportions of aberrant hippocampal tissue and *APOE* risk alleles.
  - Together, Components 4 and 5 help isolate Cluster A: a group comprised of individuals with greater frontal, temporal, and hippocampal aberrant tissue with some preserved aspects of attention/working memory and executive function.
- **Both approaches show that:**
  - FTD-like and ALS-like gene-brain-behavior patterns stand out, and in particular are very different from one another
  - PD shows three types of patterns
  - ADMCI and VCI are highly heterogeneous

DISCUSSION

- The components and clusters reflect different gene-brain-behavior continuums and groups that capture the diversity of neurodegenerative disorders in the ONDRI sample, where clinically-defined cohorts do not map onto data-driven clusters (derived from clinically meaningful measures). We see: 1) ADMCI and VCI distributed across all clusters, 2) PD generally spread across 3 clusters (A, C & E), and 3) both FTD and ALS generally fall into different clusters. Instead, we see more global effects (Component 1), dissociable pathologies (Component 2), and an overall heterogeneity of disorders across components and clusters (which may reflect multi-morbid individuals).
- When focused on groups (dMFA) genetics do not play a strong role in separating these groups, whereas when focused on revealing groups (MFA) we see that ApoE an MAPT do play a role
- Similar to neuropsychiatric domains, we approach neurodegenerative disorders in a way that combines concepts from both the RDoC and the DSM to identify underlying continuums and groups that cross clinical “boundaries”
- ONDRI—as a singular heterogeneous cohort of neurodegenerative and cerebrovascular diseases—shows that particular types of spectrums exist across clinically defined disorders. In particular, such work can help us focus on which groups are clinically well-defined (e.g., FTD and ALS) as opposed to needing to further understand the underlying heterogeneous pathology of disorders (PD, ADMCI, VCI)

Farhan, S. M., Bartha, R., Black, S. E., Corbett, D., Finger, E., Freedman, M., ... & Kleinstiver, P. W. (2017). The Ontario neurodegenerative disease research initiative (ONDRI). *Canadian Journal of Neurological Sciences*, *44*(2), 196-202.

Abdi, H., Williams, L. J., & Valentin, D. (2013). Multiple factor analysis: principal component analysis for multitable and multiblock data sets. *Wiley Interdisciplinary reviews: computational statistics*, *5*(2), 149-179.

Bécue-Bertaut, M., & Pagès, J. (2008). Multiple factor analysis and clustering of a mixture of quantitative, categorical and frequency data. *Computational Statistics & Data Analysis*, *52*(6), 3255-3268.

Dilliott, A. A., Evans, E. C., Farhan, S. M., Ghani, M., Sato, C., Zhang, M., ... & Strong, M. J. (2019). Genetic Variation in the Ontario Neurodegenerative Disease Research Initiative. *Canadian Journal of Neurological Sciences*, *46*(5), 491-498.

Ramirez, J., Holmes, M. F., Scott, C. J. M., Ozzoude, M., Adamo, S., Szilagyi, G. M., Goubran, M., Gao, F., Arnott, S. R., Lawrence-Dewar, J. M., Beaton D., Strother, S. C., Munoz, D. P., Masellis, M., Swartz, R. H., & The ONDRI Investigators (2019). Ontario Neurodegenerative Disease Research Initiative (ONDRI): Structural MRI methods & outcome measures. Preprint on *bioRxiv*.

McLaughlin P. M., Sunderland K. M., Beaton D., Binns M. A., Kwan D., Levine B., Orange JB., Peltsch A. J., Roberts A. C., Strother S. C., Troyer A., & ONDRI Investigators (accepted). The Quality Assurance and Quality Control Protocol for Neuropsychological Data Collection and Curation in the Ontario Neurodegenerative Disease Research Initiative (ONDRI) Study. *Assessment*.