

# NIAGADS Update: Spring 2024

Li-San Wang, PhD University of Pennsylvania

COI: Nothing to Disclose

# **NIAGADS** at a Glance



# **Data Sharing**

113 datasets
>183,000 samples
497 Unique Requests
238 Institutional Certifications

#### Userbase



>330 articles cite U24

>4,900 verified users

#### ADSP

73 cohorts23 current grant subawards

Exomes 20,503

#### Genomes

**~350%** increase from 2021 to 2024

(16,905 in 2021 to ~60,000 genomes in 2024)

### ADSP Support

14 workgroups /
 5 consortia
 >500 members
 52 institutions

### **NIAGADS 2023 Growth**





+17% datasets

+24%

articles that cite U24

+28% verified users

### **ADSP WGS Releases**



#### **Released in 2023**

#### December

R4 WGS Annotation PHC 2023 Release

#### Coming Soon in 2024

Q2	SV joint genotype called pVCF
4-	QC'd chrX/multi-allelic pVCF

R5 WGS- ~62k genomes CRAMs, gVCFs, preview pVCF

# ADSP 2023 Deliverables – NG00067v11 Phenotype Harmonization Consortium (PHC)



Cognition	Fluid Biomarker	Neuropathology	Vascular Risk Factors
•Harmonized composite scores (PHC_EF, PHC_MEM, PHC_LAN, PHC_VSP)	<ul> <li>Amyloid ß,Tau, pTau</li> <li>Z-scores</li> <li>A/T classification</li> <li>Raw data</li> <li>Platform Information</li> </ul>	<ul> <li>Mapped AD Phenotypes (NIA-AA ABC score, Thal phasing, Braak staging, CERAD score)</li> <li>Cerebral amyloid angiopathy, lewy bodies, vascular brain injury, TDP- 43 proteinopathy and hippocampal sclerosis</li> </ul>	•Vascular Composite (Hypertension, diabetes, heart disease [myocardial infarct, congestive heart failure, etc.], BMI, height, weight, stroke, smoking, medication, blood pressure, CVR score)

Imaging – DTI	Imaging – FLAIR	Imaging – MRI	Imaging – PET
<ul> <li>PreQual Processed Data (Image-Based BIDS)</li> <li>Free Water Corrected Metrics</li> <li>Scanner Harmonization</li> <li>Multi-Site Calibration</li> <li>FA and MD Scalars</li> </ul>	•Total WMH Volume	<ul> <li>Preprocessed Imaging Collection</li> <li>MUSE Regions of Interest</li> <li>FreeSurfer Regions of Interest</li> <li>SPARE-AD Scores</li> <li>AD Signature (Schwartz)</li> </ul>	<ul> <li>Preprocessed Imaging Collection</li> <li>Global SUVRs</li> <li>Centilloid Values</li> <li>GMM Probabilities</li> <li>Amyloid Status</li> </ul>



Courtesy Tim Hohman / Vanderbilt: timothy.j.hohman@vumc.org

# ADSP 2023 Deliverables – NG00067v11 Phenotype Harmonization Consortium (PHC)

Included in ng00067.v11 Harmonization in Progress; to be included in future release Domain data not available for harmonization

Cohorts	Autopsy	Cardiovascular Risk Factors	Cognition	Fluid Biomarkers	Imaging – DTI	Imaging – FLAIR	Imaging – PET	Imaging – T1
A4	N/A		$\sim$	N/A	N/A	$\checkmark$	$\sim$	
ACT	$\checkmark$		$\checkmark$	N/A			N/A	
ADNI	N/A	$\sim$	$\checkmark$	$\checkmark$	$\sim$		$\sim$	$\sim$
ADFBS (NIA-LOAD)	$\checkmark$	$\checkmark$	$\checkmark$	N/A	N/A	N/A	N/A	N/A
EFIGA		$\sim$	$\sim$	N/A	N/A	N/A	N/A	N/A
Knight ADRC			$\checkmark$	$\checkmark$				
NACC	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$
ROS/MAP/MARS	$\checkmark$	$\checkmark$	$\checkmark$	N/A	$\checkmark$		N/A	
WHICAP			$\sim$	N/A		$\sim$		$\sim$
UPenn				$\checkmark$				
WHICAP		$\sim$	$\sim$			$\sim$	N/A	
WRAP	N/A	$\checkmark$	$\checkmark$		$\sim$		$\checkmark$	$\sim$

# ADSP 2024 Deliverables R5 WGS Release



ADSP Release	Primary Ancestry/Ethnicity	Cases	Controls	Other	Totals
GCAD Release 5	Black/African American	2,646	5,608	777	9,031
	Asian	1,273	4,083	5	5,361
	Latino/Hispanic	4,488	10,248	2,777	17,513
	Non-Hispanic White	9,124	14,459	6,067	29,650
	Total	17,531	34,398	9,626	61,555

#### **ADRC** samples in R5

#### 4,590

**EXOMES** 22% of total ADSP WES

#### 12,629

genomes 20% of total ADSP WGS

**ADRCs can get WGS/WES data from their own samples from us!** Send an email to <u>niagads@pennmedicine.upenn.edu</u> to initiate.

# **ADRC GWAS Data Release**



#### ADC 1–15 GWAS available via DSS

- QC'd PLINK files on GRCh38
- Updated ID schemas for seamless mapping of phenotypes to genotypes
- TOPMedR2 imputed files in .bgen format
- Curated phenotypes provided by the ADSP-PHC
- ADSP to NACC ID mapping file for any ADRC participants that also have WGS, WES, or ADSP-PHC harmonized phenotypes available through the ADSP dataset (ng00067)

#### ADC 16 & 17 GWAS

available to the ADRCs

- Raw array data available
- Total of 2,292 participants from 38 ADRCs

#### GWAS SNP Array Data for

29,676

ADRC Samples Passing QC and available for request.

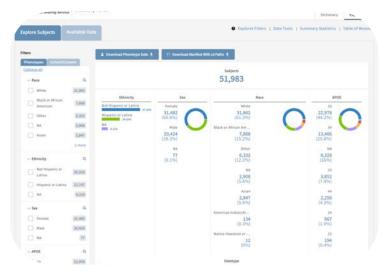
# **Supporting FAIR Principles**



Initiatives	Findable	Accessible	Interoperable	Reusable
Gen3 Platform	Q		$\leftarrow$	
<b>RAS</b> Authentication			$\leftarrow$	
Data Submission Portal	Q			
DOI Implementation	Q			
ADGC GWAS Standardization				
Outreach Initiatives	Q			

# **Supporting FAIR Principles**







're-approval et preliminary details from submitter

ssign data intake manager for assessmer

stake data manger communicates approva

tatus over email and updates submission

tatus using the portal API

status

**Experiment Details Upload & Validation** user-dashboard to track submission templates provided for metadata files (e.g., subject phenotypes, file manifests)

documented, workflow-driven forms metadata files and forms validate guiding users from study -> dataset -> automatically at submission xperiment details



**ADSP Data Discovery Platform** Integration powered by Gen3 Technology

Streamline data submission and validation using **DSS Data Submission Portal** 

**Digital Object Identifiers (DOI)** Enabled through NIH DataCite **Consortium** membership

### NIAGADS Open Access Data and Annotation Resources

unrestricted data and annotations (variant, gene, sequence) available for immediate access



**Open Access Portal** 



& advp<sup>2</sup>

GenomicsDB<sup>3</sup>

direct data downloads for GWAS summary statistics (p-values) and ADSP Variant Annotations

https://dss.niagads.org

75,207

standardized, searchable functional genomics tracks

https://tf.lisanwanglab.org/FILER/

curated AD-GWAS catalog, documenting

>18k

published associations

https://advp.niagads.org

AD-genetics knowledgebase and Genome Browser, containing

>260M

annotated variants

https://www.niagads.org/genomics



1. Kuksa et al. 2022, NAR Genomics and Bioinformatics, 4(1): lqab123. 2. Kuksa et al. 2021, Journal of Alzheimer's Disease, 86(1): 461-477. 3. Greenfest-Allen et al. 20204, Alzheimer's & Dementia, 20(2):1123-1136

# NIAGADS Open Access NIAGADS API\*

Lead: Emily Greenfest-Allen



filer

Alzheimer's Disease Variant Portal





**Open Access Portal** 

#### **GenomicsDB**

Alzheimer's Genomics Database





- Programmatic access
- Bulk, ranged, and genome wide queries
- Integration into analysis pipelines
- Better support ADSP activities and AI/ML pipelines

#### conforms to OpenAPI standards (& is FAIR)

\* under development; beta-testing version available summer 2024

# **Supporting Data Use**

#### **User Support**

NIAGADS offers several support resources launched over the last year:



*Held at least once a quarter* 

https://www.youtube.com/@niagads1814



All NIAGADS documentation in one place, new informational pages to address user questions, dedicated FAQ page



Users can now submit a ticket to work on resolving technical issues with the NIAGADS team.

#### Awareness

Expanding NIAGADS reach through increased investigator awareness will be accomplished through informational outreach sessions at AAIC and ASHG in 2024



Exhibit Hall Booth

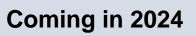
Visit us at booth 1329D at AAIC!



**Poster Presentations** 



ADRC workshop at AAIC in partnership with NACC, Sage, and NCRAD



- Monthly newsletter
- More video tutorials, user support pages, blog posts



### **AAIC 2024 Activities**



# AAIC Exhibit Hall Booth Visit us at booth 1329D at AAIC



### See you this July in Philadelphia!

