Workshop on Identifying Midlife Social Exposures that Might Modify Risk for Cognitive Impairment Associated with Early Life Disadvantage

Session 4: Social and Health Policies

Jason Fletcher

Vilas Distinguished Achievement Professor of Public Affairs, Agricultural and Applied Economics, and Population Health Sciences

University of Wisconsin-Madison

Challenges

- Addressing difficult problems, likely no easy answers
- Important disciplinary differences in what to focus on and priorities in research design and conventions
  - Disagreement about focus on measures of brain pathology or cognitive function or diagnosis
- What are we looking for: silver bullets or complexity?
- Do we have the data infrastructure to find what we are looking for?

Social science concerns

- Are the data representative of some population of interest?
- Are the outcomes usefully measured?
  - What are tradeoffs between measurement error and sample size
- Are we collecting a large enough sample to match our envisioned effect size?
- Do we have in place strong data sharing and reproducibility protocols?

Session: Health and Social Policies

- Constrained question: Identifying Midlife Social Exposures that Might Modify Risk for Cognitive Impairment Associated with Early Life Disadvantage
- Less constrained: Identifying Social/Health Policies that Might Modify Risk for Cognitive Impairment

- Not education policies (other session's focus)
  - Note that we have no broad social/health policies that target cognitive impairment, so presumably any effects are unintended, likely modestly sized

Aligning target outcomes and exposures

- Less constrained: Identifying Social/Health Policies that Might Modify Risk for Cognitive Impairment
- Even less constrained: Identifying Social/Health Policies that Might Shape Later Life Cognitive
- We have representative, modest-to-big-sized data results for:
  - Early nutrition
  - Lead and environmental contaminants
  - Resources, family SES, area-level SES
  - Life course occupations
  - Health care access
- These studies sometimes are tied in with cognitive decline / diagnosis (often with mixed, underpowered results) but not brain pathology

Questions we cannot answer well

- Less constrained: Identifying Social/Health Policies that Might Modify Risk for Cognitive Impairment
- More constrained: Identifying Social/Health Policies that Might shape old-age brain pathology
  - Mis-matched data, unrepresentative samples, but deep phenotyping

Promising opportunities / directions

- Consider complexity and small/modest effect sizes

- Consider diagnosis, mortality outcomes (not cognition) collected in big data
- Recognize that HRS-sized data and convenience/clinical samples are risky options

- Electronic Health Records (David)
- Administrative data / matched with surveys
- Linked Historical / administrative / mortality databases

Ex: Big data, Medicare Claims

Hazed and Confused: The Effect of Air Pollution on Dementia

Get access ›

Kelly C Bishop ✉, Jonathan D Ketcham, Nicolai V Kuminoff

The Review of Economic Studies, Volume 90, Issue 5, October 2023, Pages 2188–2214,
https://doi.org/10.1093/restud/rdac078
Published: 13 December 2022 **Article history** ▾

FIGURE I: DEMENTIA DIAGNOSIS BY AGE AND GENDER IN 2013

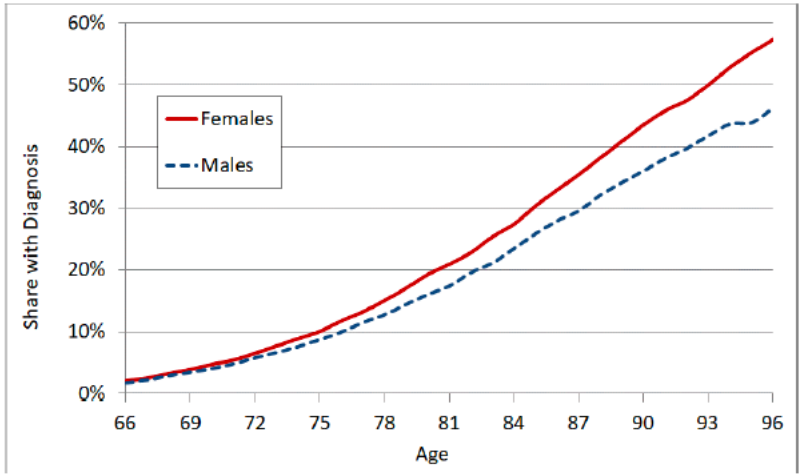


TABLE I—AVERAGE MARGINAL EFFECT OF CUMULATIVE $PM_{2.5}$ ON THE PROBABILITY OF A NEW DEMENTIA DIAGNOSIS

	(1)	(2)	(3)	(4)	(5)	(6)
(1 μg/m^3 increase in decadal $PM_{2.5}$)	0.629***	0.124	1.545***	2.283***	2.384***	2.151***
	(0.058)	(0.105)	(0.536)	(0.565)	(0.568)	(0.846)
individual & neighborhood covariates		x	x	x	x	x
$PM_{2.5}$ control function			x	x	x	x
survival control function				x	x	x
polynomial functions and interactions					x	x
heterogeneity by exposure duration						x
F-statistic on $PM_{2.5}$ instruments			496	498	498	165 to 489
number of individuals: dementia function	1,179,094	1,179,094	1,179,094	1,179,094	1,179,094	989,751 to 2,293,270
Chi-square statistic on survival instruments				3,813	3,813	1,166 to 2,274
number of individuals: survival function				2,439,904	2,439,904	2,439,904

Ex: Big Surveys

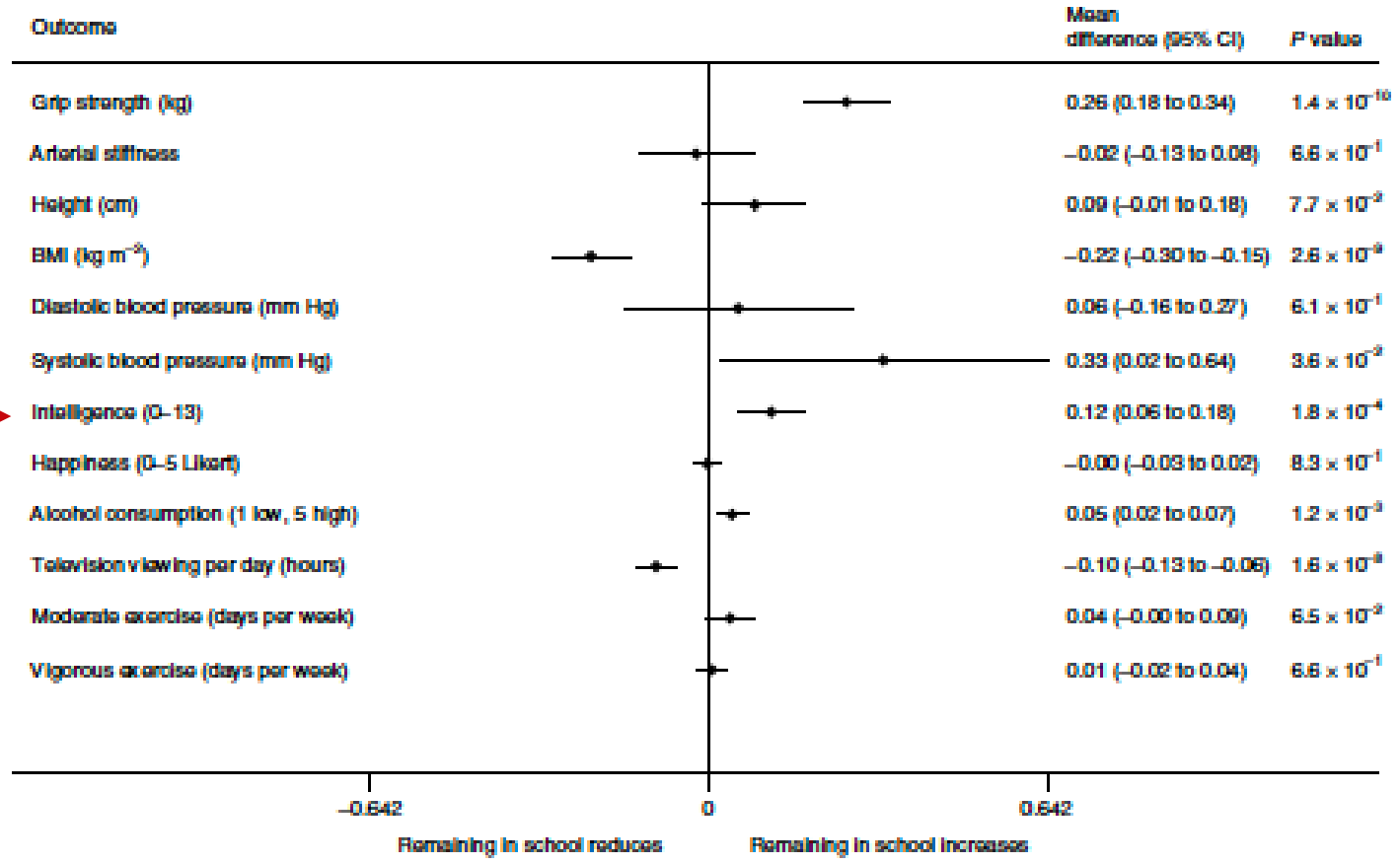
nature human behaviour

LETTERS

https://doi.org/10.1038/s41562-017-0279-y

The causal effects of education on health outcomes in the UK Biobank

Neil M. Davies[1,2]*, Matt Dickson[3], George Davey Smith[1,2], Gerard J. van den Berg[1,4] and Frank Windmeijer[1,4]



Focusing on dementia mortality outcomes

- Allows other ‘big’ data from vital statistics
- Obstacle
  - US Death Records have limited information, cannot speak to life course
- Filling in this gap
  - “Big” Surveys linked with death records
    - National Longitudinal Mortality Study (>2M, >20K AD deaths)
    - Mortality Disparities in American Communities (>3M, 20K AD deaths)
    - Diet and Health Study (400K. >1K AD deaths)
- However, restricted access

[Under Construction] Obtain a Public-Use File (Direct access to limited version of the MDAC Analysis File)

An MDAC public-use file is under construction and expected to be available in early 2023. Due to the confidential nature of the MDAC Title 13 data, the MDAC Public-use file will be limited in content.

Diet and Health Study

PLOS ONE

RESEARCH ARTICLE

Geographic variation in Alzheimer's disease mortality

Michael Topping[1,2], Jinho Kim[2,3,4]*, Jason Fletcher[1,2,5,6,7]

ELSEVIER

SSM - Population Health
Volume 15, September 2021, 100841

Association and pathways of birth in the stroke belt on old age dementia and stroke Mortality ☆, ☆☆

Michael Topping [a b], Jinho Kim [b c d], Jason Fletcher [a b e]

J Alzheimers Dis. 2023 ; 93(3): 1007–1016. doi:10.3233/JAD-230086.

Area-Level Infant Mortality Exposure in Early Life and Alzheimer's Disease Mortality: Examining Variation Based on Age, Sex, and Place of Birth

Emerging infrastructures

The IPUMS Multigenerational Longitudinal Panel: Progress and Prospects

Steven Ruggles, University of Minnesota†

Julia A. Rivera Drew, University of Minnesota

Catherine A. Fitch, University of Minnesota

J. David Hacker, University of Minnesota

Jonas Helgertz, Lund University

Matt A. Nelson, University of Minnesota

Nesile Ozder, University of Minnesota

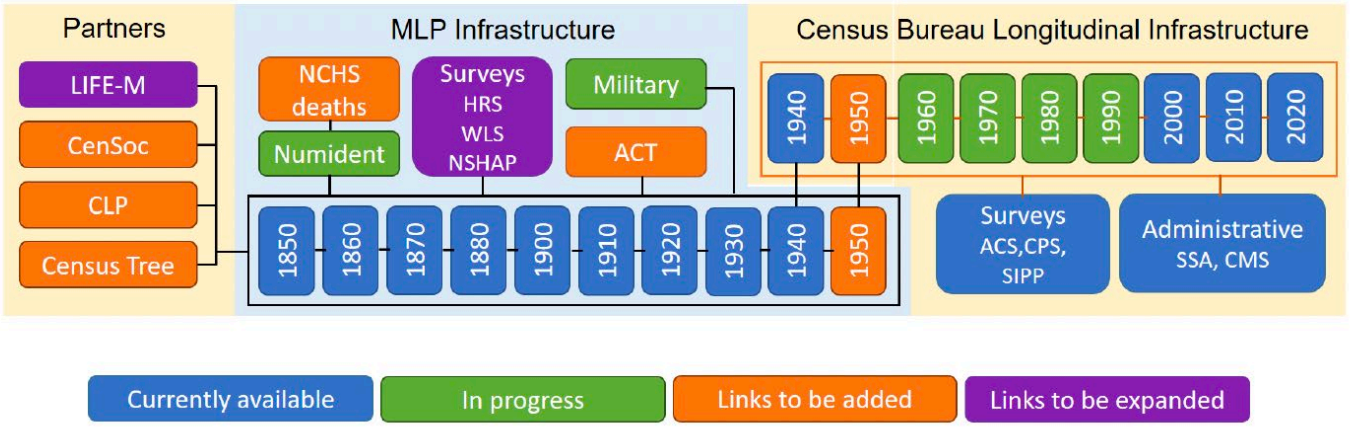
Matthew Sobek, University of Minnesota

John Robert Warren, University of Minnesota

July 2024

Working Paper No. 2024-05
DOI: https://doi.org/10.18128/IPUMS2024-05

Figure 1. Existing and planned MLP infrastructure and links to collaborating projects



Ongoing challenges: Data

- The majority of data that focuses on dementia is
  - too small, unless we are committed to silver bullets only
  - Without effort to consider representativeness, which is an unforced error
- A second tranche of data is representative, still too small, and lacking the deep phenotyping

Example: Parity and AD incidence: HRS

The Relationship Between Fertility History and Incident Dementia in the U.S. Health and Retirement Study

FREE

Alison Gemmill, PhD ✉, Jordan Weiss, PhD

The Journals of Gerontology: Series B, Volume 77, Issue 6, June 2022, Pages 1118–1131,

The 15,361 respondents who comprised the analytic sample contributed 173,282 person-years (median [interquartile range] length of follow-up: 13.0 [7.0–16.0] years) of follow-up over the study period. Over the 173,282 person-years of follow-up, 3,208 cases of all-cause dementia were observed yielding a crude incidence rate of 18.5/1,000 person-years. There were 2,030 cases of all-cause dementia among women (crude incidence rate: 19.9/1,000 person-years) and 1,178 cases were observed among men (crude incidence rate: 16.6/1,000 person-years).

Table 2. Associations Between Parity and Incident Dementia in Models Stratified by Gender, csHR, and 95% CI

Open in new tab

	Model 1		Model 2	
Characteristic	**Men *n* = 6,486**	**Women *n* = 8,875**	**Men *n* = 6,486**	**Women *n* = 8,875**
Parity				
0 (reference)	1.00	1.00	1.00	1.00
1	1.16 (0.85, 1.59)	0.99 (0.80, 1.22)	1.13 (0.82, 1.56)	0.96 (0.78, 1.18)
2	0.92 (0.70, 1.20)	1.04 (0.88, 1.24)	1.01 (0.77, 1.32)	1.12 (0.95, 1.33)
3	1.04 (0.79, 1.37)	0.98 (0.82, 1.16)	1.16 (0.89, 1.53)	0.98 (0.82, 1.17)
≥4	1.33 (1.02, 1.73)	1.31 (1.11, 1.55)	1.12 (0.85, 1.48)	1.09 (0.92, 1.28)

Discussion

In our study that uses nationally representative, population-based data from the United States, we find no strong evidence of an association between three commonly used measures of fertility history and incident dementia after adjusting for confounders and while accounting for the semicompeting risk of death. In crude

Parity and AD: UK Biobank

- Yan Zheng and Jason Fletcher. “The Association between Parity and Odds of Alzheimer’s Disease and Dementias Status” forthcoming at *Demography*
- UK Biobank (500,000 respondents, reporting on their parents)
- Parental AD: 37K cases for moms, 20K cases for dads

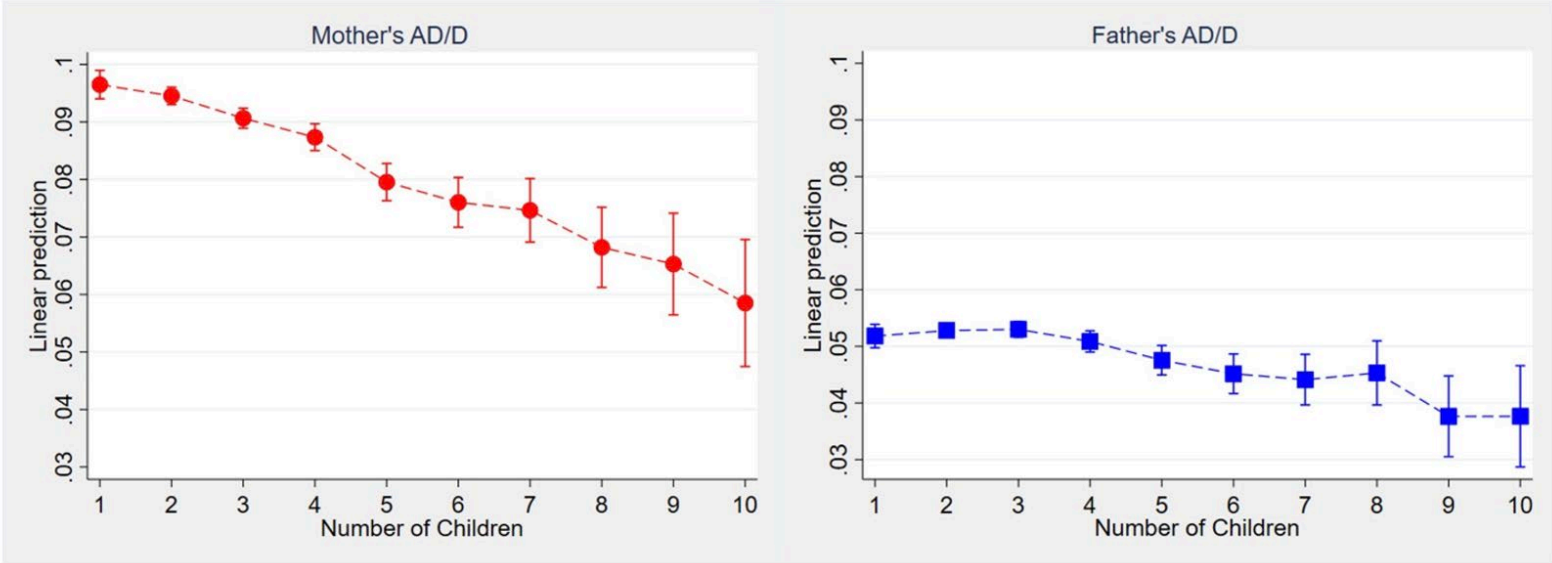


Figure 1. Association between number of children and parent’s odds of Alzheimer’s disease or dementias (AD/D). Results are linear prediction based on the results from Table 2 Model 1.

Ongoing challenges: Data

- The majority of data that focuses on dementia is
  - too small, unless we are committed to silver bullets only
  - Without effort to consider representativeness, which is an unforced error
- A second tranche of data is representative, still too small, and lacking the deep phenotyping
- Administrative data, especially Medicare is prohibitively expensive
  - Funders/stakeholders should think about what we are failing to uncover because we can’t ask the important questions
- Large Scale survey data seems to confuse size with representative (UK Biobank, All of Us)

Big, but unrepresentative samples

Reweighting the UK Biobank to reflect its underlying sampling population substantially reduces pervasive selection bias due to volunteering

Sjoerd van Alten[1,2], Benjamin W. Domingue[3], Titus Galama[1,2,4,5], and Andries T. Marees[1]

[1] *Vrije Universiteit Amsterdam, School of Business and Economics*
[2] *Tinbergen Institute*
[3] *Stanford University Graduate School of Education*
[4] *University of Southern California, Dornsife Center for Economic and Social Research and Department of Economics*
[5] *Erasmus University Rotterdam, Erasmus School of Economics*

May 13, 2022

nature human behaviour

Article

https://doi.org/10.1038/s41562-023-01579-9

Participation bias in the UK Biobank distorts genetic associations and downstream analyses

Received: 27 September 2022

Accepted: 7 March 2023

Published online: 27 April 2023

Check for updates

Tabea Schoeler[1,2], Doug Speed[1,3], Eleonora Porcu[4], Nicola Pirastu[4], Jean-Baptiste Pingault[2,5] & Zoltán Kutalik[1,2,6]

While volunteer-based studies such as the UK Biobank have become the cornerstone of genetic epidemiology, the participating individuals are rarely representative of their target population. To evaluate the impact of selective participation, here we derived UK Biobank participation probabilities on the basis of 14 variables harmonized across the UK Biobank and a representative sample. We then conducted weighted genome-wide association analyses on 19 traits. Comparing the output from weighted genome-wide association analyses ($n_{effective}$ = 94,643 to 102,215) with that from standard genome-wide association analyses (n = 263,464 to 283,749), we found that increasing representativeness led to changes in SNP effect sizes and identified novel SNP associations for 12 traits. While heritability estimates were less impacted by weighting (maximum change in h^2, 5%), we found substantial discrepancies for genetic correlations (maximum change in r_g, 0.31) and Mendelian randomization estimates (maximum change in β_{STD}, 0.15) for socio-behavioural traits. We urge the field to increase representativeness in biobank samples, especially when studying genetic correlates of behaviour, lifestyles and social outcomes.

What an outsider might see when reviewing ADRD research--conventions

- Very high priority on phenotype/outcome measurement
- Uncertainty around what to measure and how to classify ADRD status; seemingly large changes over time
- Naïve / cursory focus on social environment
- Indifference to representativeness
  - Mostly researchers do not even try to address the issue
- High levels of monopoly power

What an outsider might see when reviewing ADRD research--output

- Enormous NIH and private funding
- Lots of research output: data collection, academic papers, clinical trials
- Little progress on treatment or cure
- Instances of scientific fraud

Backing out the implicit conventions in ADRD

- Biological / genetic effects are, basically, universal
  - → representative samples are not essential
- Focus on looking for silver bullets
  - →Big samples are not essential
- Data are valuable to individual PIs and teams
  - →no reproducibility efforts, very limited data sharing infrastructure, lack of norms around research transparency
- →We **need** to recognize the risk associated with these conventions in failing to make any progress on ADRD

Potential lessons from genetic discovery

- Paradigm shift from silver bullets to incremental effects/complexity
- Candidate genes ended up being a monumental waste of funding and effort
  - Search for silver bullets, buttressed by animal model evidence
- OGOD→ infinitesimal model of genetic effects
- Bigger N dominates measurement precision
- Researcher degrees of freedom: Need for replication and reproducibility

Next steps?

- Level up the research—I have outlined many problems that have been solved in social science research
- Should there be a handoff of funding and focus, away from medical sciences?
- Entertain alternative theories/conceptual focuses
  - What if there is no silver bullet?
  - and
  - What if the incremental and complex determinants are environmental and interactive?
    - Pollution, chemicals
    - Early life course
- These issues all point to the enormous potential to leverage underused and new data, and the hurdles are surmountable

Summary

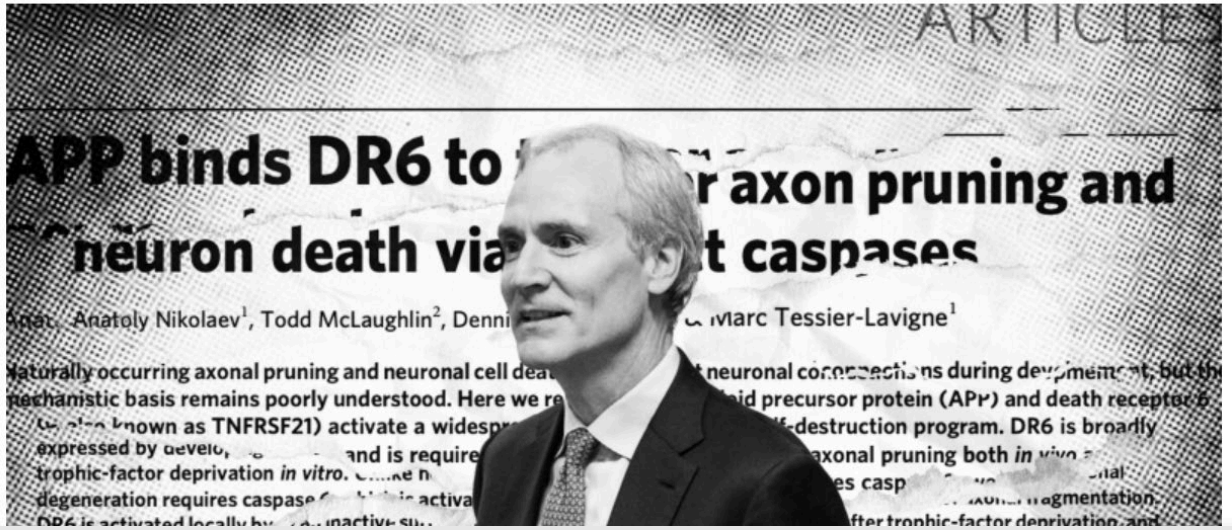
- Shockingly high level of resources focused on medical studies of ADRD, compared to other conditions
  - Likely that these investments were not “shovel ready”
  - Produced minimal advances
- Use of small convenience samples with little regard for the problems this can pose, especially when effect sizes are small compared to the size of the likely biases
- Silver bullets are assumed; this is an extraordinarily risky assumption to maintain, as produced evidence can make us dumber
- Monopolistic system of investigation

Examples: Fraud

The Stanford Daily

Blockbuster Alzheimer's paper retracted by former Stanford president after a decade of resistance

Marc Tessier-Lavigne had declined to withdraw the paper as recently as this summer, when he resigned as president



Examples: naïve to social environment

- Neighborhood Atlas / Area Deprivation Index

T Act Project	Year	Sub	Principal Investigator(s)/ Project Leader(s)	Organization	Fiscal Year	Admin IC	Funding IC	FY Total Cost by IC	Similar Projects
The Neighborhoods Study: Contextual Disadvantage and Alzheimer's Disease and Related Dementias (ADRD)									
5 R01 AG070883-03			KIND, AMY J. BENDLIN, BARBARA BRIGITTA	UNIVERSITY OF WISCONSIN-MADISON	2023	NIA	NIA	$6,354,915	View
The Neighborhoods Study: Contextual Disadvantage and Alzheimer's Disease and Related Dementias (ADRD)									
3 R01 AG070883-03S1			KIND, AMY J. BENDLIN, BARBARA BRIGITTA	UNIVERSITY OF WISCONSIN-MADISON	2023	NIA	NIA	$3,494,004	View

Rehkopf, David H., and Robert L. Phillips Jr. "The Neighborhood Atlas Area Deprivation Index And Recommendations For Area-Based Deprivation Measures: Perspective offers recommendations for improving area-based deprivation measures." *Health Affairs* 42, no. 5 (2023): 710-711.

Hannan, Edward L., Yifeng Wu, Kimberly Cozzens, and Brett Anderson. "The Neighborhood Atlas Area Deprivation Index For Measuring Socioeconomic Status: An Overemphasis On Home Value: Study examines the Neighborhood Atlas Area Deprivation Index as a tool to measure socioeconomic status." *Health Affairs* 42, no. 5 (2023): 702-709.

Petterson, Stephen. "Deciphering the Neighborhood Atlas Area Deprivation Index: the consequences of not standardizing." Health Affairs Scholar 1, no. 5 (2023): qxad063.

Examples: Non-Universal Genetic effects

Pervasive biases in proxy GWAS based on parental history of Alzheimer's disease

Yuchang Wu[1,2,†], Zhongxuan Sun[1,†], Qinwen Zheng[1], Jiacheng Miao[1], Stephen Dorn[1], Shubhabrata Mukherjee[3], Jason M. Fletcher[2,4,5], Qiongshi Lu[1,2,*]

[1] Department of Biostatistics and Medical Informatics, University of Wisconsin-Madison Madison, WI, USA
[2] Center for Demography of Health and Aging, University of Wisconsin-Madison, Madison, WI, USA
[3] Department of Medicine, University of Washington, Seattle, WA, USA
[4] Department of Population Health Sciences, University of Wisconsin-Madison, Madison, WI, USA.
[5] La Follette School of Public Affairs, University of Wisconsin-Madison, Madison, WI, USA