A simple metric to assess the influence of recruitment bias in dementia research-related autopsy cohorts:

Putting it to the test with NACC data

Pete Nelson

Kathryn Gauthreaux

DOI: 10.1002/alz.13422

RESEARCH ARTICLE

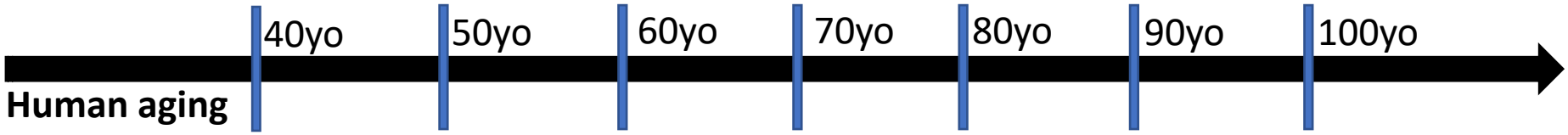
Alzheimer's & Dementia®
THE JOURNAL OF THE ALZHEIMER'S ASSOCIATION

Different cohort, disparate results: Selection bias is a key factor in autopsy cohorts

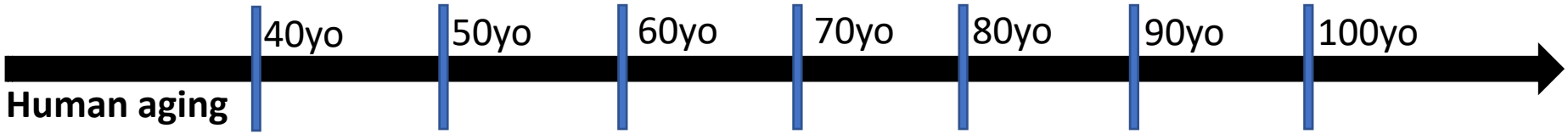
Kathryn Gauthreaux[1] | Walter A. Kukull[1] | Karin B. Nelson[2] | Charles Mock[1] | Yen-Chi Chen[1,3] | Kwun C. G. Chan[1,4] | David W. Fardo[5,6] | Yuriko Katsumata[5,6] | Erin L. Abner[5,6,7] | Peter T. Nelson[5,8]

VERY DIFFERENT STUDY DESIGNS…

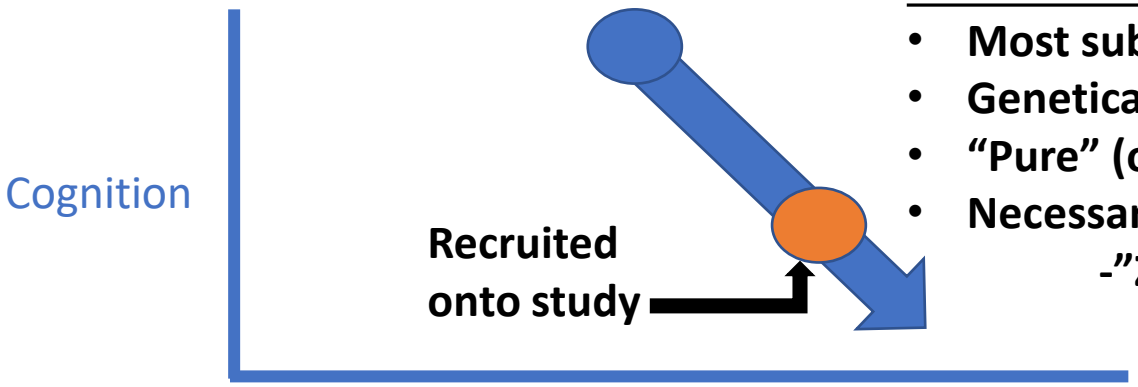
…can be conceptualized along human aging spectrum



VERY DIFFERENT STUDY DESIGNS…



Paradigm 1:



Clinic-based studies

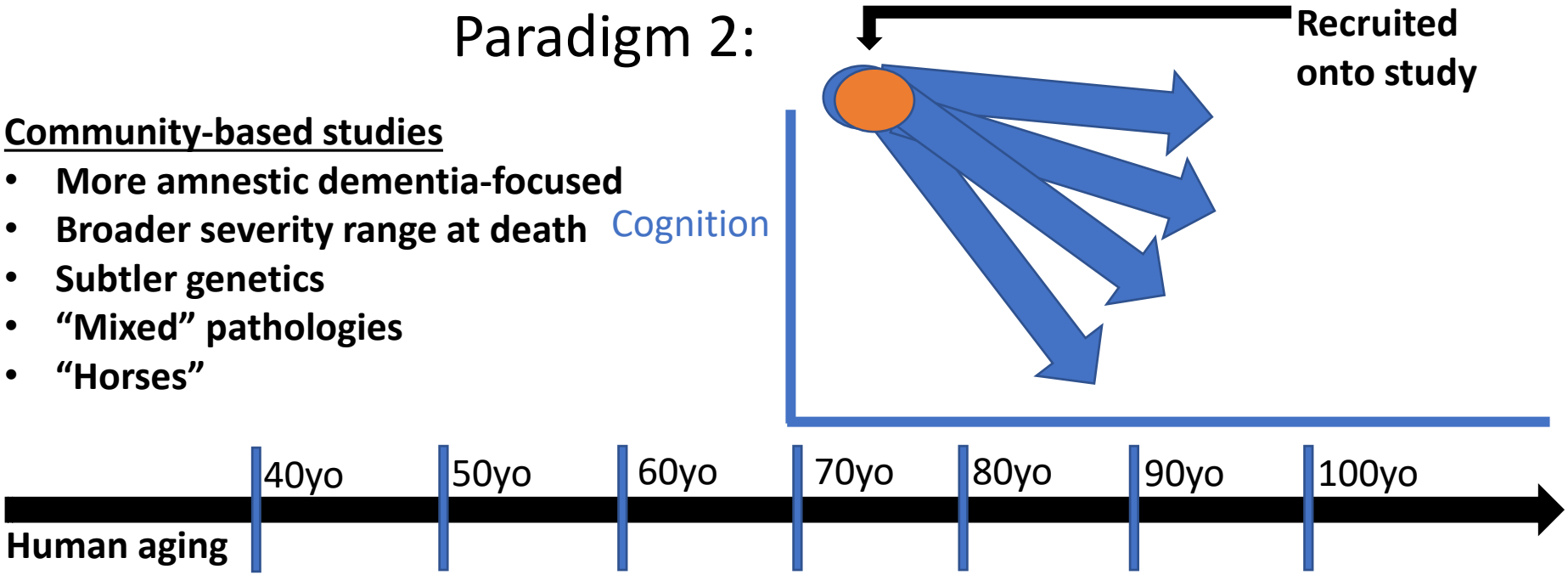
- **Most subjects die with severe dementia**
- **Genetically driven diseases**
- **"Pure" (one disease is predominant)**
- **Necessary, to study these diseases**
  - **-"Zebras"**

VERY DIFFERENT STUDY DESIGNS…

Paradigm 2:

Community-based studies

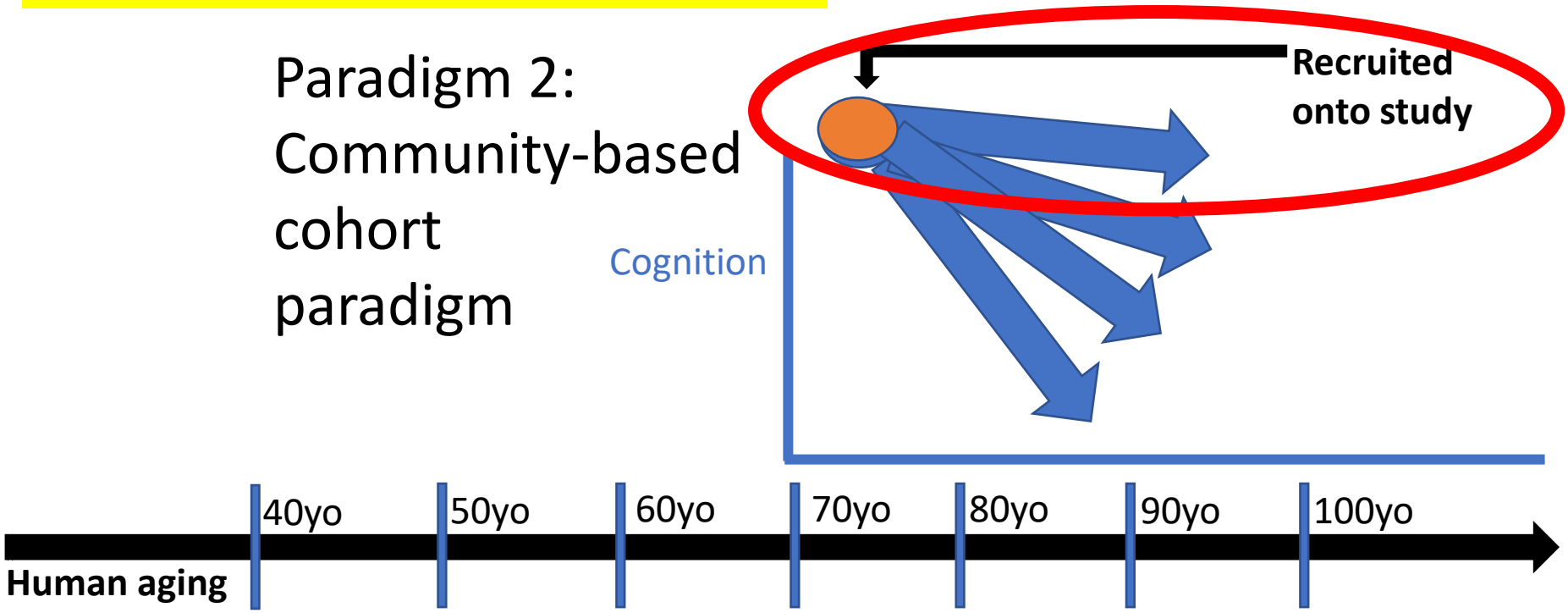
- **More amnestic dementia-focused**
- **Broader severity range at death**
- **Subtler genetics**
- **"Mixed" pathologies**
- **"Horses"**



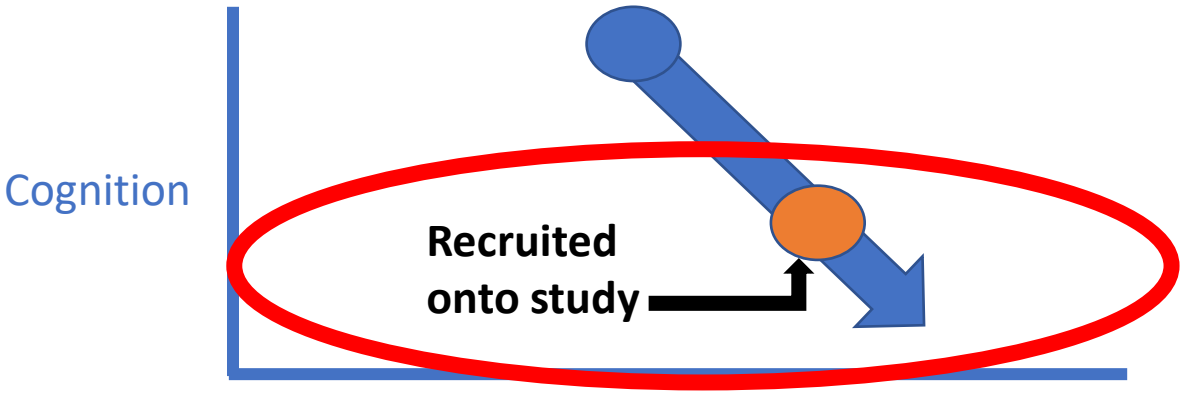
Despite these “extreme” examples, many study cohorts fall somewhere betwixt/between

HOW CAN YOU COMPARE?

Paradigm 2: Community-based cohort paradigm



Paradigm 1: Clinic-based cohort paradigm



DOI: 10.1002/alz.13422

RESEARCH ARTICLE

Alzheimer's & Dementia®
THE JOURNAL OF THE ALZHEIMER'S ASSOCIATION

Gauthreaux et al, 2023

Different cohort, disparate results: Selection bias is a key factor in autopsy cohorts

Kathryn Gauthreaux[1] | Walter A. Kukull[1] | Karin B. Nelson[2] | Charles Mock[1] | Yen-Chi Chen[1,3] | Kwun C. G. Chan[1,4] | David W. Fardo[5,6] | Yuriko Katsumata[5,6] | Erin L. Abner[5,6,7] | Peter T. Nelson[5,8]

NACC:

A great context to study different cohort types

28 ADRCs met inclusion criteria

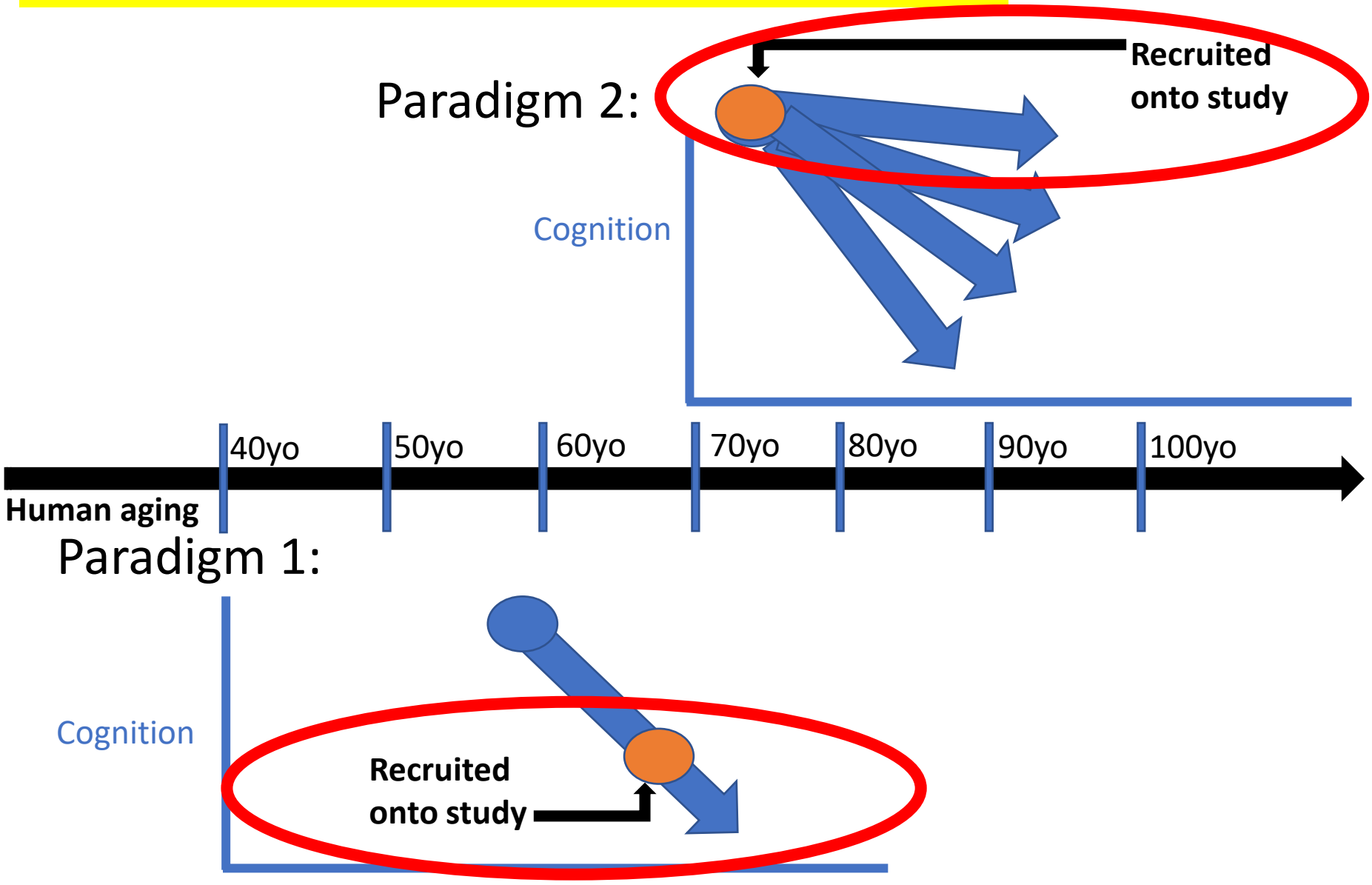
>30 study participants each

-Followed longitudinally

-Came to autopsy

-Detailed data available

HOW TO COMPARE THE PARADIGMS?



NACC:

28 ADRCs met inclusion criteria
(Avg~150 cases each)

- 8 ADRCs: More normal (>30% normal) at recruitment
- 10 ADRCs: Middle (15-30% normal) at recruitment
- 10 ADRCs: More impaired (<15% normal) at recruitment

NACC:

28 ADRCs met inclusion criteria (>30 cases each)

- 8 ADRCs: More normal (>30% normal at recruitment)
- 10 ADRCs: Middle (15-30% normal at recruitment)
- 10 ADRCs: More impaired (<15% normal at recruitment)

NACC:

28 ADRCs met inclusion criteria (>30 cases each)

- 8 ADRCs: More normal (>30% normal at recruitment)
- 10 ADRCs: Middle (15-30% normal at recruitment)
- 10 ADRCs: More impaired (<15% normal at recruitment)

28 ADRCs, and over 4000 participants, included

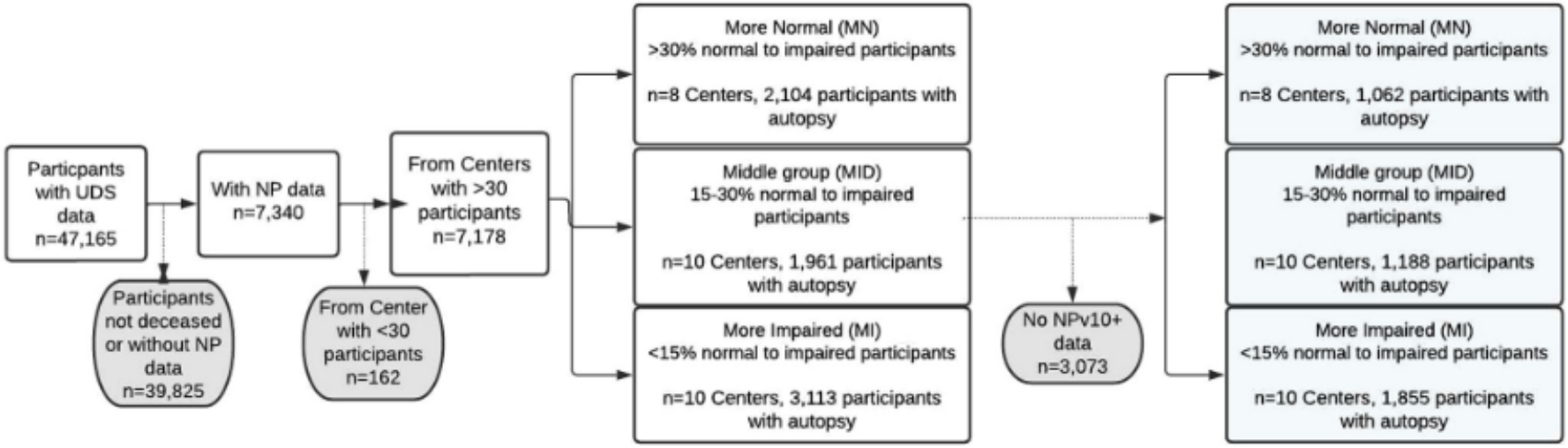


FIGURE 1 Study cohort exclusions, inclusions, sample sizes, and delineation of study groups. NP, National Alzheimer's Coordinating Center Neuropathology form; UDS, Uniform Data Set

Gauthreaux et al, 2023

2.2 | Statistical analyses

To compare the demographic characteristics, clinical profiles, and neuropathologic features between the M-Norm, M-Imp, and Mid groups, we report both descriptive statistics (mean proportion, min, max) and *P*-values from trend test using regression models. The analysis was done at the center level; thus, all statistics are reported for the ADRCs within each group. Continuous measures were compared with linear models, while counts were compared using Poisson models with a log offset for the number of participants with autopsy in each ADRC. Robust standard errors were used for the calculation of *P*-values, which allows non-constant variance across groups and overdispersion of count outcomes. Participant characteristics examined included

Thanks to Drs Yen-Chi Chen, Gary Chan, Dave Fardo, Yuri Katsumata, & Erin Abner

Gauthreaux et al, 2023

Gauthreaux et al, 2023

TABLE 1 Demographic characteristics among participants with NP data available.

Center group, *n*	More impairment (N/MCI&D < 15%) *n* = 10 centers, 3113 participants with autopsy	Middle group (N/MCI&D 15%–30%) *n* = 10 centers, 1961 participants with autopsy	More normal (N/MCI&D > 30%) *n* = 8 centers, 2104 participants with autopsy	*P*-values
Characteristic				
Number of unique 3-digit zip codes in ADRC, mean (min, max)	34.6 (0, 101)	16.8 (0, 26)	23.1 (0, 66)	0.22
Average number of UDS visits, mean (min, max)	3.6 (2.5, 4.7)	4.5 (3.4, 5.7)	4.5 (3, 5.5)	**0.01**
Average follow-up time (in years), mean (min, max)	3.0 (1.5, 4.2)	4.1 (2.9, 5)	4.1 (2.5, 5.5)	<**0.01**
% deceased participants with autopsy, mean % (min%, max%)	60.0 (33.7, 86.3)	60.5 (42.2, 76.0)	62.8 (47.3, 82.2)	0.68
Age at baseline, mean (min, max)	72.1 (66.2, 78.3)	74.9 (69.6, 78.5)	79.3 (75.1, 83.9)	<**0.01**
Age at death, mean (min, max)	77.1 (70.7, 83.8)	80.3 (75.4, 84.4)	85.1 (80.0, 89.8)	<**0.01**
Sex, mean % (min%, max%)				
Male	58.9 (51.2, 66.7)	54.7 (47.3, 66.7)	46.7 (38.8, 56.9)	<**0.01**
Female	41.1 (33.3, 48.9)	45.4 (33.3, 52.7)	53.3 (43.1, 61.2)	<**0.01**
Race, mean % (min%, max%)				
White	94.4 (89.1, 97.4)	91.6 (85.4, 96.6)	88.0 (43.9, 97.8)	0.46
Black or African American	2.4 (0.6, 4.8)	4.4 (0.5, 10.2)	8.2 (0.0, 45.9)	0.09
Other	2.1 (0.7, 3.4)	3.5 (1.2, 8.8)	3.7 (1.6, 9.2)	0.32
APOE ε4, mean % (min%, max%)				
ε4 carrier	39.7 (22.2, 53.8)	40.4 (16.7, 54.7)	35.7 (26, 46.3)	0.36
No ε4 allele	42.4 (14.3, 58)	46.3 (22.5, 56.9)	55.7 (40.8, 71.8)	**0.01**
Missing/unknown/ not assessed	17.9 (1.2, 63.5)	13.3 (0.0, 60.9)	8.7 (0.5, 21.7)	0.09

Demographics and genetics:

ADRCs which recruited higher % normal:

Longer followup, older at recruitment and death, more female

Lower *APOE* ε4 allele

Clinical findings:

ADRCs which recruited higher % impaired:

More DLB, FTD

ADRCs which recruited higher % normal:

More CVD

AD—not different!

TABLE 2 Primary clinical diagnoses at initial and last standardized Uniform Data Set visit among participants with NP data available.

Center group, *n*	More impairment (N/MCI&D < 15%) *n* = 10 centers, 3113 participants with autopsy	Middle group (N/MCI&D 15-30%) *n* = 10 centers, 1961 participants with autopsy	More normal (N/MCI&D > 30%) *n* = 8 centers, 2104 participants with autopsy	
Characteristic				*P*-values
Initial UDS visit				
Primary clinical diagnoses, mean % (min%, max%)				
AD	51 (30.7, 69.2)	54.5 (28.2, 82.8)	41.3 (22.3, 57.1)	0.16
DLB	8.8 (2.3, 20.1)	4.2 (1.2, 7.4)	1.6 (0.0, 3.8)	<0.01
CVD	0.8 (0.0, 3.3)	1.5 (0.0, 5.5)	1.3 (0.0, 3.4)	0.12
FTD[a]	17.5 (3.1, 42.2)	9.5 (0.0, 38.9)	3.4 (0.0, 8.2)	<0.01
Clinical diagnoses,[b] mean % (min%, max%)				
AD	54.8 (37.9, 72.8)	59.0 (44.7, 82.8)	43.1 (23.5, 60.2)	0.08
DLB	11.5 (3.1, 24.3)	6.0 (1.2, 10.1)	4.5 (2.1, 6.0)	<0.01
CVD	3.3 (0.0, 8.0)	4.6 (0.9, 10.7)	3.7 (1.3, 6.6)	0.90
FTD				
Any FTD[a]	21.4 (3.6, 53.7)	12.7 (1.2, 53.2)	5.0 (1.9, 10.2)	<0.01
bvFTD	10.7 (2.6, 20.7)	6.6 (0.0, 23.3)	3.5 (1.3, 9.2)	<0.01
PPA	6.8 (1.0, 15.4)	5.2 (0.0, 30.2)	1.4 (0.0, 2.7)	<0.05
Last UDS visit				
Primary clinical diagnoses, mean % (min%, max%)				
AD	57.9 (34.3, 76.2)	64.1 (37.1, 85.1)	58.7 (43.2, 66.8)	0.56
DLB	10.2 (3, 22.5)	6.0 (2.2, 11.6)	2.9 (0.0, 5.8)	<0.01
CVD	1.2 (0.0, 3.1)	2.3 (0.0, 10.2)	4.3 (1.0, 7.4)	<0.01
FTD[a]	16.9 (1.5, 40.7)	10.1 (0.0, 37.6)	4.0 (0.9, 8.2)	<0.01
Clinical diagnoses,[b] mean % (min%, max%)				
AD	62.3 (41.2, 80)	69.2 (48.8, 88.5)	61.9 (45.3, 69.8)	0.65
DLB	15.0 (4.1, 30.8)	10.2 (6.8, 15.4)	6.6 (3.2, 9.8)	<0.01
CVD	5.0 (0.0, 11.9)	8.9 (4.0, 19.1)	10.0 (4.1, 16.7)	<0.01
FTD				
Any FTD[a]	21.6 (3.1, 53.9)	13.2 (1.2, 53.5)	5.0 (1.9, 9.6)	<0.01
bvFTD	10.7 (1.5, 22.0)	6.4 (0.0, 20.1)	3.3 (0.4, 8.2)	<0.01
PPA	7.3 (1.0, 16)	4.8 (0.0, 29.7)	1.1 (0.0, 2.9)	0.03

Gauthreaux et al, 2023

TABLE 3 Neuropathological features.

Center group, *n*				
Characteristic, mean % (min%, max%)	More impairment (N/MCI&D < 15%) *n* = 10 centers, 3113 participants	Middle group (N/MCI&D 15-30%) *n* = 10 centers, 1961 participants	More normal (N/MCI&D > 30%) *n* = 8 centers, 2104 participants	*P*-values
Alzheimer's disease pathology, mean % (min%, max%)				
Pathological NIA-AA criteria[b]				
No ADNC	14.2 (6.6, 25.7)	11.1 (4.3, 17.4)	13.0 (6.9, 19.5)	0.47
Low ADNC	14.5 (0.0, 30.0)	14.9 (7.5, 26.3)	18.5 (10.0, 28.5)	0.48
Intermediate ADNC	16.2 (10, 30.7)	23.1 (10.1, 52.5)	28.3 (9.8, 40)	**<0.01**
High ADNC	50.0 (34.7, 68.2)	48.9 (23.2, 67.5)	39.7 (28.8, 54.1)	**<0.05**
Braak stage				
0	9.5 (1.3, 27.1)	3.7 (0.0, 12.6)	4.2 (0.5, 10.4)	**<0.01**
I	7.7 (3.1, 17.7)	6.6 (3.5, 11.4)	6.6 (1.9, 13.2)	0.49
II	10.1 (3.3, 18.9)	10.6 (6.5, 14.8)	11.6 (2.2, 19.2)	0.27
III	9.5 (5.5, 20.6)	9.4 (4.9, 16.4)	13.4 (7.6, 31.6)	**0.01**
IV	10.6 (7.3, 19.1)	10.3 (7.2, 13.5)	18.5 (4.7, 35.3)	**0.01**
V	15.6 (5.8, 25.4)	21.4 (12.7, 37)	20.4 (5.7, 36.2)	0.06
VI	34.5 (12.7, 51.8)	36.7 (17.1, 51.8)	24.6 (12.2, 56.6)	**<0.01**
CERAD score				
None	23.9 (9.5, 37.3)	17.6 (12.2, 26.2)	23.9 (12.1, 33.9)	0.89
Sparse	10.3 (2.6, 19.5)	8.8 (2.8, 18.0)	17.4 (4.2, 45.4)	**0.01**
Moderate	18.3 (9.5, 34.9)	19.1 (3.8, 35.1)	19.0 (5.8, 42.4)	0.71
Frequent	46.7 (23.8, 73.4)	54.2 (34.3, 69.8)	39.7 (6.6, 59.0)	0.16

Gauthreaux et al, 2023

Pathology findings (AD neuropathologic changes/ADNC):

ADRCs which recruited higher % impaired:

More Braak NFT stages 0, VI

ADRCs which recruited higher % normal:

More Braak NFT stages III, IV

TABLE 3 Neuropathological features.

Center group, *n*				
Characteristic, mean % (min%, max%)	More impairment (N/MCI&D < 15%) *n* = 10 centers, 3113 participants	Middle group (N/MCI&D 15-30%) *n* = 10 centers, 1961 participants	More normal (N/MCI&D > 30%) *n* = 8 centers, 2104 participants	*P*-values
FTLD-TDP-43[b]	7.9 (0.0, 20.8)	4.3 (0.0, 14.6)	2.5 (0.0, 5.0)	**0.01**
FTLD-tau[b]	14.4 (0.0, 39.0)	10.2 (1.0, 29.5)	15.9 (0.0, 54.6)	0.97
Argyrophilic grains[b]	2.4 (0.0, 7.89)	2.4 (0.0, 11.74)	6.1 (0.0, 24.68)	0.72
Cerebrovascular pathology, mean % (min%, max%)				
Infarcts or lacunes	15.2 (8.5, 30.1)	15.6 (9.0, 31.6)	23.2 (10.4, 41.8)	**0.03**
Microinfarcts	17.2 (3.1, 32.1)	19.6 (8.4, 37.8)	21.4 (5.3, 31.7)	0.13
Mod/severe arteriolosclerosis	46.5 (15.9, 76.7)	41.0 (8.7, 72.4)	22.0 (0.0, 45.7)	0.09

Gauthreaux et al, 2023

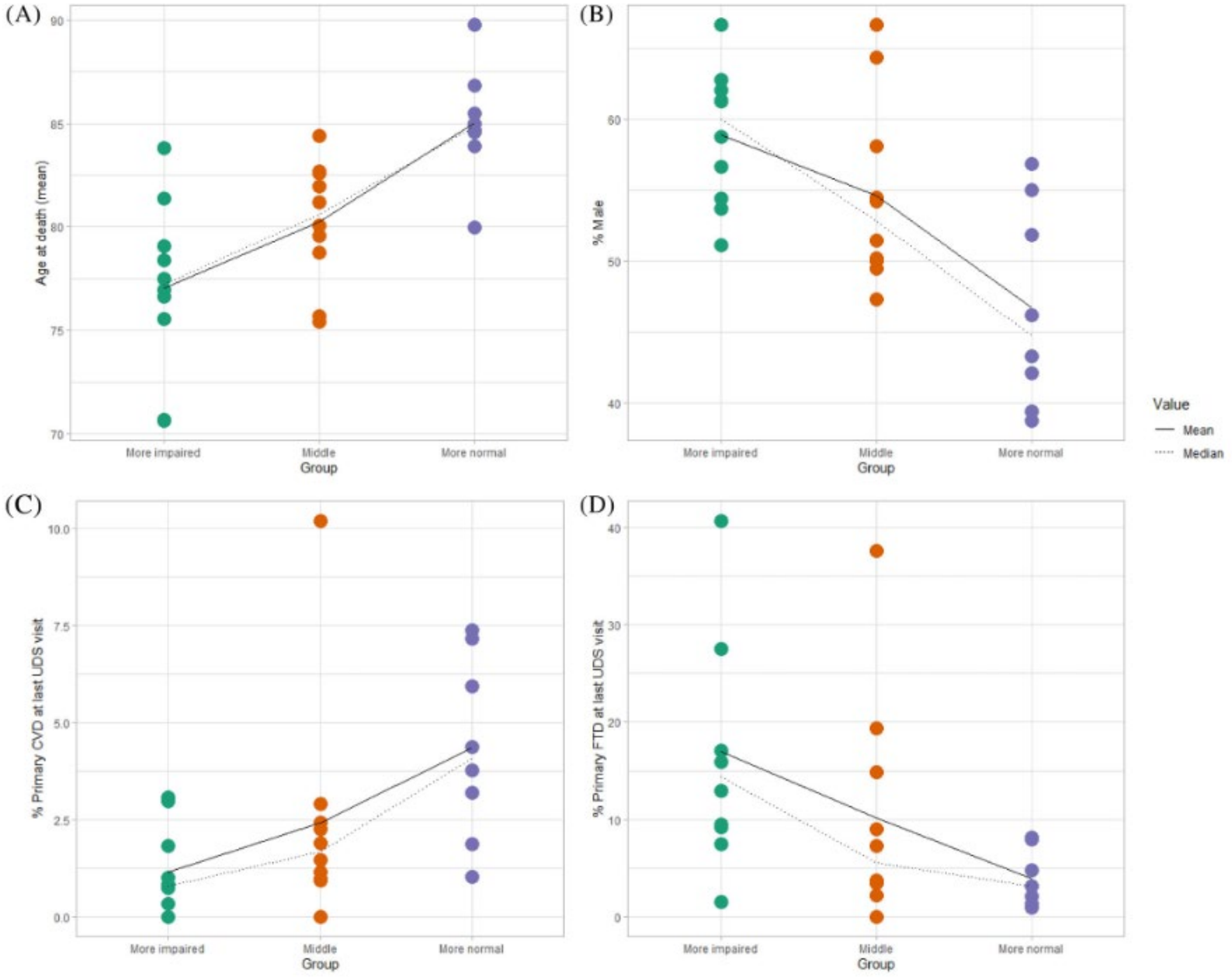
Pathology findings (non-ADNC):

ADRCs which recruited higher % impaired:

More FTLD-TDP

ADRCs which recruited higher % normal:

More infarcts or lacunes



Gauthreaux et al, 2023

TABLE 4 Summary of findings.

Cohort recruitment criteria	Findings
More participants recruited while cognitively normal	Older participants (at intake and death)
	Longer study follow-up
	More females
	More CVD clinically
	More infarcts at autopsy
	More intermediate ADNC
More participants recruited while cognitively impaired	Younger participants
	More males
	More *APOE* ε4 (though not statistically significant at the p=0.05 level)
	More severe ADNC
	More clinical DLB at last UDS visit
	More clinical FTD at last UDS visit and FTLD-TDP at autopsy

Gauthreaux et al, 2023

Conclusion:

For an autopsy cohort used in dementia research, the % of subjects cognitively normal at recruitment into the study is a measure (related to ascertainment bias) associated with a set of clinical and pathological observations.

This parameter may also provide a proxy for where a cohort falls along the clinic/community-based spectrum of study designs.

THANK YOU!

Kathryn Gauthreaux

DOI: 10.1002/alz.13422

RESEARCH ARTICLE

Alzheimer's & Dementia®
THE JOURNAL OF THE ALZHEIMER'S ASSOCIATION

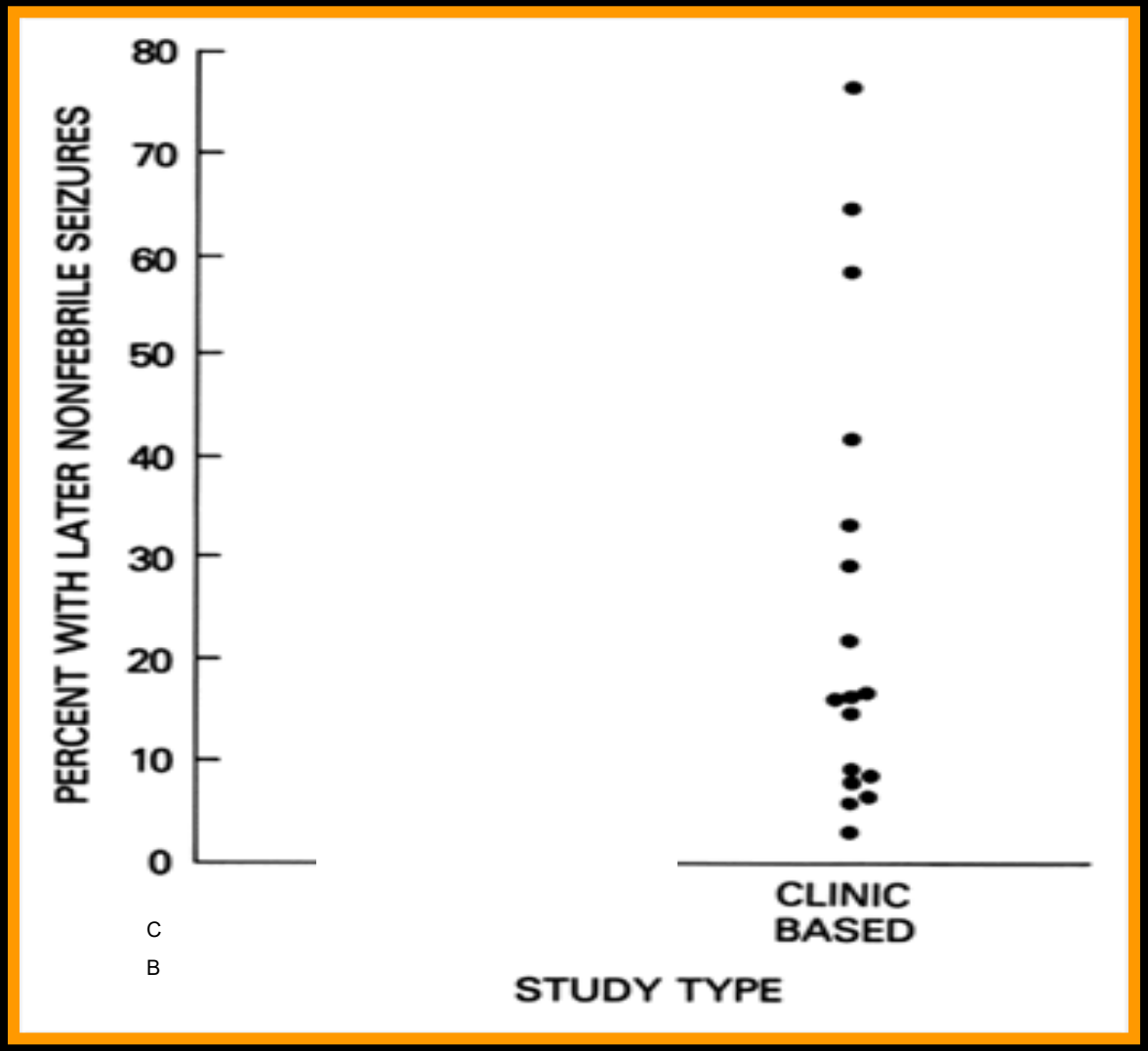
Different cohort, disparate results: Selection bias is a key factor in autopsy cohorts

Kathryn Gauthreaux[1] | Walter A. Kukull[1] | Karin B. Nelson[2] | Charles Mock[1] | Yen-Chi Chen[1,3] | Kwun C. G. Chan[1,4] | David W. Fardo[5,6] | Yuriko Katsumata[5,6] | Erin L. Abner[5,6,7] | Peter T. Nelson[5,8]

Population based vs. clinical samples

% Febrile→
Nonfebrile
seizures:

Give
phenobarbital?



Ellenberg & Nelson
JAMA 1980

Average Stanford–Binet IQ Scores at the Two-Year Visit, According to Expected IQ Level and Treatment Group.

Table 2. Average Stanford–Binet IQ Scores at the Two-Year Visit, According to Expected IQ Level and Treatment Group.

Expected IQ Level*	Phenobarbital Group			Placebo Group		
	No. at Base Line	No. at 2 Yr	Average IQ	No. at Base Line	No. at 2 Yr	Average IQ
1	24	11	85.64	19	11	97.73
2	17	11	92.00	26	21	98.00
3	22	17	97.47	21	17	105.35
4	21	18	97.33	22	20	107.65
5	24	20	114.30	21	20	115.70
Total	108	77	—	109	89	—
Horvitz–Thompson average†	—	—	95.54	—	—	103.95

*The higher the number, the higher the expected Stanford–Binet IQ score of the children at that level.

†The Horvitz–Thompson average is the weighted average of the averages for each level. The weight is the inverse of the proportion of the number of observations. For example, the weight for the third level of the placebo group is 21 divided by 17. The Horvitz–Thompson average of the placebo group is greater by 8.41, and its t statistic is 3.48 (P adjusted for three interim analyses, 0.0057).

The NEW ENGLAND JOURNAL of MEDICINE