The Value of Representative Populations for Accurate Inferences

Seattle Symposium

Eric B. Larson, MD, MPH
Vice President for Research, Group Health
Executive Director, Group Health Research Institute
October, 2016
Outline

• Personal stories: consequences of research in non-representative populations

• Examples of population based research

• Solutions and challenges of big data related to convenience samples vs population based samples
Personal Perspective from Seeing Markers for Alzheimer’s Disease

Early studies focused on small samples form highly filtered specialty samples. Examples:

- Platelet membrane Fluidity (1992)
- Amyloid deposits in skin biopsies

Then rare mutations (App717, App 693 and PRIP gene mutations) 1991

Conclusions – there are clearly hazards using small samples from specialized populations. Can these be avoided with overwhelmingly large samples?
**Differences in Community Populations**

Populations from samples recruited from AMCs and specialty clinics are younger - generally have more severe disease with stronger “genetic’ finger print in comparison to more community-based samples:

- All community recruited populations are not equal - some are more representative than others.
- Populations and samples recruited from AMCs and Specialty clinics: Younger, have more severe disease, higher frequency of apoE e4 allele.
- "Population based" subjects (from ADPR/ACT) are older, have shorter duration of symptoms when identified, milder disease.

Strength of Associations and Diagnostic Performance Changes Based on Population: The Home Visit Issue (Crane, et. al., 2016)

• Importance of home visit capacity in dementia studies (Alzheimer’s and Dementia 2016;12:419-426)

• Unique opportunity presented when ACT study began to enroll subjects in “ACT+” – UW’s ADRC which included NACC requirements.

• ACT (an epi study) supplements in clinic with home visits vs ADRC requires in clinic visit

• RESULTS: In “full data” Risk of AD ApoE 4 (1.66) vs. clinic only (2.28) p=.008

• Conclusion: “studies that only include research clinic data may lead to biased conclusion” Using data missing not at random (MNAR) can provide the wrong answer and you can’t know direction of bias.
Population Based Studies in Life Course Epidemiology Provide Unique Insights – a Dementia Example

Diagnostic accuracy will be less in everyday populations:

- THERE HAVE TO BE FALSE POSITIVES

- WHY? PERSONS SURVIVE INTO THEIR 90s with plaques and tangles and no dementia. (Sonnen et al., Arch Neurology 2011)

- Illustrates the complexity and overlap of brain aging and neurodegenerative diseases.
Figure 1. Brain autopsy results from 336 cognitively normal individuals expressed as summary neuropathology scores (range, 0-9) ranked from lowest to highest. Each stacked bar shows an individual’s burden of Alzheimer disease (AD) (blue), Lewy body disease (LBD) (green), and microvascular brain injury (µVBI) (red). A, One hundred sixteen Adult Changes in Thought study (ACT) participants. B, One hundred six Nun Study (NS) participants. C, Fifty-nine Honolulu-Asia Aging Study (HAAS) participants. D, Fifty-five Oregon Brain Aging Study (OBAS) participants.
Diagnostic Accuracy Likely Overrated

Diagnostic accuracy will be less in everyday populations:

- THERE HAVE TO BE FALSE POSITIVES
- WHY? PERSONS SURVIVE INTO THEIR 90s with plaques and tangles and no dementia. (Sonnen et al., Arch Neurology 2011)
- Illustrates the complexity and overlap of brain aging and neurodegenerative diseases.

WILL THERE ALSO BE FALSE NEGATIVES?

For a condition with a prevalence of 50% and more in the fastest growing segment of the population this is a difficult question but the answer must be YES.
Recent Lessons Learned


- Using inverse probability weighting to adjust for selection bias and bootstrap techniques to assess uncertainty one can assess generalizability of autopsy based inferences to general population from which sample was drawn and adjust associations based on differences between autopsy sample and reference population. (Haneuse S, et al., Adjustment for selection bias in observational studies with application to the analysis of autopsy data. Neuroepidemiol 2009;32(3):229-39)
Examples of GHRI Work

Group Health /UW Alzheimer’s disease patient registry now Adult Changes in Thought

• A 30 year journey: 1986 – today – 2021

• A source for many companion projects and shared data
Living Laboratory

New News from the Adult Changes in Thought (ACT) study:
A long standing living laboratory of aging funded for five more years


Background
Increasing number of older people with multiple chronic conditions
Research on multiple chronic conditions benefits from natural circumstances among elderly populations
Adult Changes in Thought (ACT) study recently awarded $5 million of funding
One of the largest, continuously funded studies on aging

Methods
RECRUITMENT
- Cohort of randomly selected people over age 65 without dementia
- Established in 1994
- Now 2035, minimum resident cohort of approximately 2000 living persons
- Follow participants every 2 years
- Replace participants who die, become demented, or are lost to follow-up

DATA
- Biobanks with active genome-wide single nucleotide polymorphism (SNP), exome sequence, and gene expression values
- Neuropathology biobank with brain aging data
- Cognitive and functional measures from Cognitive Abilities Screening Instrument (CANS) at all ages
- Baseline chart abstraction data on clinical care
- Laboratory and pharmacy records from GH automated data and medical record review
- actiWALK (high-tech) and a triaxial wrist worn accelerometer to capture all measures of activity, sleeping, standing, and sitting behaviors, and physical activity

PARTICIPANT CHARACTERISTICS
- Current enrollment: 5,194
- 1,809 cases with incident dementia
- 506 cases with Alzheimer's disease (AD) at entry
- 3.1% of participants died in 8 years
- 46.8% of participants followed up
- 535 autopsy cases with extensive brain tissues

STUDY OUTCOMES
- Dementia and AD (based on cognitive testing and consensus criteria)
- Cognitive functioning
- Neuropathological measures (neuropathology; tangles; neurofibrillary tangles; neuritic plaques, cerebral microinfarcts, systemic infections, amyloid angiopathy)
- Resilience (avoids cognitive decline and frailty in later life)

LIVING LABORATORY
ACT COHORT 5,194
Autopsy cohort (538)

Study Aims
1. MULTICORRELATE - Examining the effects of cardiovascular risk factors and their treatments on the aging brain
2. RESILIENCE - Determining the impact of physical and mental health on cognitive capacities and physical performance, identifying factors associated with robust aging and identifying whether neureptroprotective strategies associated with resilience
3. LIVING LABORATORY - Improving ACT infrastructure and resources in the high demand scientific community for research

Recruitment and data collection timeline

Feasibility of measuring physical and sedentary activity related to cognitive health in older adults
NIH NIA grant #5R01AG058005-3060064 - 72% response

Conclusions
The ACT study is a platform for a population-based living laboratory on aging.
New funding cycle focuses on the science of aging and multimorbidity, resilience and robust aging, and data sharing.
Effective partnerships, including widespread data sharing and specimen sharing, are foundational and critical for optimal success.

CONTACT FOR DATA REPOSITORY:
ACTproposals@uw.edu

Living Lab
Genetics
Neuropathology
Neuroimaging
Pharmacology
Epigenetics
Imaging of brain injury
Neuroinflammation
Neurobiology
Treatment Trials
Cognitively
Conditioned
Resilience

Participant Summary

Feasibility of measuring physical and sedentary activity related to cognitive health in older adults
NIH NIA grant #5R01AG058005-3060064 - 72% response

Conclusions
The ACT study is a platform for a population-based living laboratory on aging.
New funding cycle focuses on the science of aging and multimorbidity, resilience and robust aging, and data sharing.
Effective partnerships, including widespread data sharing and specimen sharing, are foundational and critical for optimal success.

CONTACT FOR DATA REPOSITORY:
ACTproposals@uw.edu
ACT Living Laboratory of Aging and Brain Aging

LIVING LABORATORY

ACT COHORT 5,194
Autopsy cohort (638)

- MEDICAL RECORDS Chart review
- MEDICAL RECORDS Automated data
- INTERVIEWS Biennial in-person visits
- BIOLOGIC DATA SNPs, blood, brains
- PHYSICAL MEASURES Active and sedentary time

GENETICS
NEUROIMAGING
PHARMACO EPIDEMIOLOGY
TRAUMATIC BRAIN INJURY
NEUROPATHOLOGY
TREATMENT TRIALS
COMORBID CONDITIONS
RESILIENCE
Examples of GHRI Work, cont.

- eMERGE – Electronic Medical Records and Genomics – based on existing medical records and biobanks

- Sources: ACT and NW institute for Genomics Medicine (NWIGM)

- NWIGM is a population based sample of persons randomly selected over 50 – 65 to complement the ACT subjects who were recruited age 65+

- Our eMERGE is very different from others – offers more complete EMR data capture and possibility of life course epidemiology; Unique opportunities to understand phenotyping from EMR including using NLP

- Long history of using registries in other areas
Some Issues

• The most ideal approach: population based samples and a life course epidemiology approach

• To what extent can challenges to external validity of findings be overcome simply by having giant populations, albeit convenience and in the case of PMI volunteers who may have a health use bias?

• How can we come up with statistical techniques that could help detect data missing not at random and more robust methods to address this issue.
Conclusions

• Ideally research is set in populations recruited from a known population base and followed over time with complete availability of relevant outcomes

• We need to know generalizability of very large convenience samples

• Methods work in this area will be valuable
QUESTIONS?