Uncertainties in *APOE*-Related Risk of Alzheimer's Disease: *Implications for Interpreting Direct to Consumer Genetic Testing*

Deborah Blacker, MD, ScD

Massachusetts ADRC
Professor of Psychiatry
Harvard Medical School
Professor and Deputy Chair, Epidemiology
Harvard T.H. Chan School of Public Health

OVERVIEW

- Background: *APOE*-4 and risk
- Context
- Methods
- Findings
- Discussion

PLOS | MEDICINE

OPEN ACCESS PEER-REVIEWED

RESEARCH ARTICLE

APOE-related risk of mild cognitive impairment and dementia for prevention trials: An analysis of four cohorts

Jing Qian, Frank J. Wolters, Alexa Beiser, Mary Haan, M. Arfan Ikram, Jason Karlawish, Jessica B. Langbaum, John M. Neuhaus, Eric M. Reiman, J. Scott Roberts, Sudha Seshadri, Pierre N. Tariot, Beth McCarty Woods, Rebecca A. Betensky, Deborah Blacker

Published: March 21, 2017 • http://dx.doi.org/10.1371/journal.pmed.1002254

CAVEAT AUDITOR

Upon encountering an epidemiologist, beware not to get submerged in their methodological marshes, but keep in mind that they are the only ones who can guide you safely across

--Frank Wolters, MD, PhD,
Erasmus University PhD Thesis

BACKGROUND: *APOE* AND AD RISK

A *susceptibility* gene with three alleles with complex effect on risk for AD and vascular disease

4 allele increases risk of AD and vascular disease, reduces age of onset, and decreases longevity

Relative risk much greater for one copy than two: compared to no *E-4*: RR *E-4x* ~3-4 fold; RR *E-44* ~8-20 fold

Strongest at younger ages (so clinical populations yield higher risks), stronger in women

Absolute risk less clear—previously available estimates were modeled based on relative risks in family or case controls studies plus absolute risks in population or family samples

Some reasons for concern that sampling and statistical issues bias these estimates upward

ORIGINAL CONTEXT: THE ALZHEIMER PREVENTION INITIATIVE

Asked by Jessica Langbaum, Eric Reiman, and Pierre Tariot to develop better estimates to make truly informed consent for the API Generations prevention trial in *APOE-44* individuals

Population frequency of E-44 1-2%, so must screen 1000s to meet sample size of ~700

For informed consent in active vaccine trial, need to disclose genotype-associated risk of onset in relevant time frame(s)

To ensure appropriate disclosure setting, need also to bring in subset of subjects with ineligible genotypes

Absolute risk estimates also relevant to Direct to Consumer (DTC) Testing

HEALTHY LIVING 04/07/2017 12:34 pm ET

THE HUFFINGTON POST

FDA Allows 23andMe To Sell Genetic Tests For Disease Risk

The tests will be able to assess a person's predisposition to develop Alzheimer's and 9 other diseases.

REUTERS | Toni Clarke

PBS NEWSHOUR

SUBSCRIBE

THE RUNDOWN

A BLOG OF NEWS AND INSIGHT

HEALTH SUPREME COURT VOTE 2016

HEALTH

Before you send your spit to 23andMe, what you need to know

BY SHARON BEGLEY, STAT *April 9, 2017 at 12:20 PM EDT*

SCIENTIFIC AMERICAN.

English Cart 0 Sign In | Register

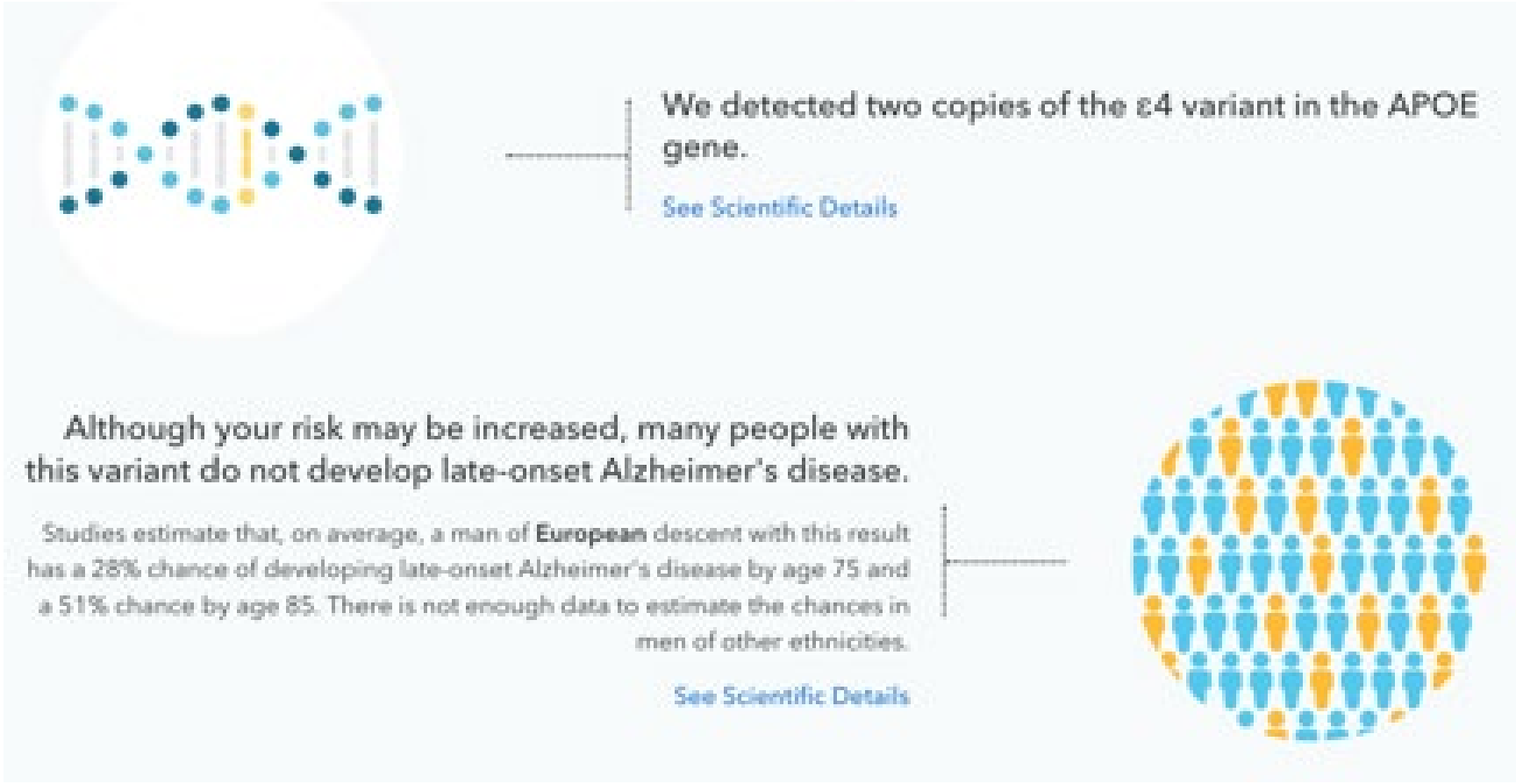
SCIENCES MIND HEALTH TECH SUSTAINABILITY EDUCATION VIDEO PODCASTS BLOGS STORE

BIOTECH

Too Much Information? FDA Clears 23AndMe to Sell Home Genetic Tests for Alzheimer's and Parkinson's

The controversial step will significantly expand direct-to-consumer testing—but what if the news is bad?

Sample 23andMe *APOE* report



Sample reports courtesy of Robert Green and Debby Tsuang

METHODS

Copied study entry criteria (normal cognition, age 60-75)

ADC samples at NACC and three population-based cohorts with baseline ages ≤60: Framingham, SALSA (Hispanic), and Rotterdam

Assessments and surveillance differ by cohort: included those identified as “normal” by the original study

Stratified cumulative incidence curves accounting for competing risk of death, by age band and gene dose , plus estimates of 5-year and (in the long-term cohorts) “lifetime” (to age 80-85) cumulative incidence

Modeled estimates from survival analyses (subdistribution hazard regression, Fine & Gray, 1999): demographics and *APOE* dose, and these plus family history, vascular risk factors, cognitive test performance, and cognitive concerns

RESULTS Table 1: Sample characteristics

Study	NACC	FHS	Rotterdam	SALSA
Description	US volunteer, 80% white	US pop, 1°ly white	European pop, white	US pop, Chicano
N	5073	4078	6399	1294
μ (sd) age (yrs)	68.7 (4.30)	62.0 (1.71)	65.4 (4.18)	67.8 (4.44)
% male	33.6%	43.2%	45.2%	41.6%
μ (sd) educ (yrs)	15.79 (2.99)	13.20 (*)	12.94 (*)	7.72 (5.42)
E-4 allele freq	0.178	0.117	0.150	0.075
% w/ fam hx	58.3%	N/A	21.7%	N/A
% w/ memory concerns	24.9%	N/A	43.1%	N/A
μ (sd) MMSE	29.0 (1.3)	28.5 (1.0)	28.8 (1.4)	86.5[#] (11.3)

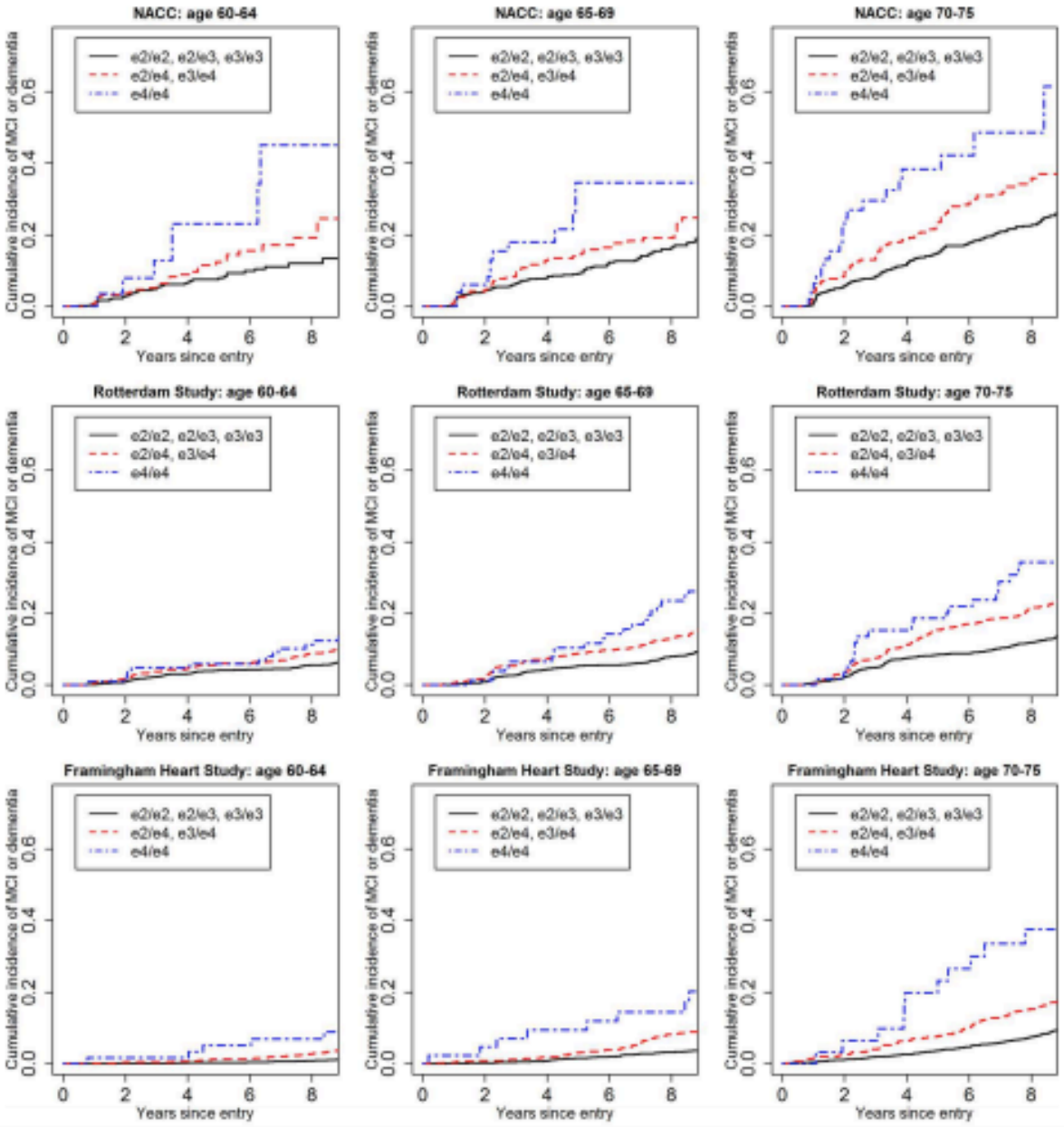
*Means estimated from distribution of ordinal education categories; [#]3MS

Cumulative Incidence of MCI/dementia by *APOE* dose

NACC

FHS

RS



Five year Cumulative Incidence of MCI or dementia in *APOE-4* homozygotes by sample and baseline age

	Age 60-64		*Age 65-69*		*Age 70-75*	
	n	*5 yr Cum Inc (%)* (95%CI)	*n*	*5 yr Cum Inc (%)* (95%CI)	*n*	*5-yr Cum Inc (%)* (95%CI)
NACC	36	23.1 (15.2, 30.9)	65	34.6 (27.1, 42.2)	57	38.3 (31.8, 44.9)
FHS	62	5.07 (2.66, 7.47)	44	9.42 (5.63, 13.2)	32	23.2 (16.6, 29.7)
RS	102	5.88 (2.68, 12.7)	77	10.4 (5.31, 19.8)	59	18.6 (10.7, 31.3)
SALSA	3	0.00 (0.00, 0.00)	5	20.0 (3.30, 36.7)	3	33.3 (2.21, 64.5)

Lifetime (80-85) cumulative incidence of MCI or dementia in *APOE-4* homozygotes by sample and baseline age

	Age 60-64		*Age 65-69*		*Age 70-75*	
	n	*20 yr Cum Inc (%)* (95%CI)	*n*	*15 yr Cum Inc (%)* (95%CI)	*n*	*10 yr Cum Inc (%)* (95%CI)
FHS	62	45.2 (31.3, 61.7)	44	46.7 (31.6, 64.7)	32	37.6 (22.4, 58.2)
RS	102	37.5 (25.1, 53.3)	77	38.1 (27.3, 51.5)	59	38.0 (26.7, 52.0)

Subdistribution hazard regression: *APOE*, family hx, and cognition*

Variable	NACC	RS	FHS	SALSA
APOE-4x	**1.49** (1.25, 1.79)	**1.63** (1.44, 1.84)	**1.75** (1.45, 2.10)	**2.15** (1.39, 3.33)
APOE-44	**2.37** (1.59, 3.53)	**2.78** (2.10, 3.69)	**4.01** (2.31, 6.96)	**1.65** (0.27, 9.93)
Standardized cognitive screen, per SD	**0.63** (0.58, 0.69)	**1.08** (1.02, 1.15)	**0.87** (0.82, 0.93)	**0.59** (0.52, 0.67)
Subjective memory concerns	**2.23** (1.87, 2.66)	**1.56** (1.39, 1.74)	**N/A**	**N/A**
Family history of dementia	**1.27** (1.06, 1.52)	**1.16** (1.01, 1.32)	**N/A**	**N/A**

These models also include age (HR ~1.1/year), gender (NS), education (HR vs. high school [HS] ~1.3 for <HS, 0.8 for >HS)

SUMMARY OF FINDINGS

APOE-44 rare even in very large samples

Incidence quite low in younger age groups over short time intervals, but higher if E4 positive, especially homozygotes

Notable variation in incidence between population based cohorts and NACC

Regression findings help to understand individual risk and variation across cohorts

- Age, *APOE-4* (*E-44>E-4x*) increase risk
- Family history increases risk beyond *E-4*
- More education generally protective
- Worse baseline memory/cognition score or memory concerns increases risk

Sources of variation across samples

Sampling frame: explicit and implicit inclusion/exclusion criteria, ascertainment methods, assessment methods and definition of normal at baseline

Overall differences in race/ethnicity, education, family history, health status, memory concerns at baseline

Volunteer participants at NACC more highly educated, more women, greater family history, and ?more memory concerns; also likely to be generally healthier

Differential impact by genotype of above factors on survival before and after initial selection

Relative risk estimates more consistent because comparisons within sample effectively control for this variation

Concluding Thoughts

Detailed estimates are relevant to DTC testing—and study design, power, and informed consent in prevention trials

The devil really is in the details: sampling, assessment, and analysis

Absolute risk varies more widely than relative risk, and short term risk more widely than longer term: different questions require different approaches

Estimates should come from a sample as similar as possible to population to which they are applied (prevention trails may draw a NACC-like sample, but DTC and screening programs better served by population based cohort data with similar demographics etc.

What do we tell participants and patients?

For both genetics and biomarkers, we need “heuristic confidence intervals” that incorporate not only statistical uncertainty but uncertainty in sampling, methods, etc.

Need to tread carefully between oversimplification and confusion

If you're interested in these and related questions, please join the Disclosure committee(s)

Thanks to:

Risk analysis team: Jing Qian, PhD, U Mass Amherst; Rebecca Betensky, PhD, HSPH/MGH; plus Frank Wolters, MD at Erasmus

Data providers: Mary Haan, PhD, SALSA; Sudha Seshadri, MD, Framingham; M. Arfan Ikram, MD, PhD, Rotterdam

API team: Jessica Langbaum, PhD, Banner; Pierre Tariot, MD, Banner; Eric Reiman, MD, Banner; Jason Karlawish, MD, Penn

Funding: National Institute on Aging

EXTRAS

Prior estimates of absolute risk

Based on estimation procedures from case control studies

REVEAL (Cupples *Genet Medicine 2004)*: Risk curves for incidence derived from relatives and spouses in family sample (Lautenschlager *Neurology* 1996); RRs by gender, age, and genotype applied from a large meta-analysis done primarily in clinically ascertained, younger onset families (Farrer *JAMA* 1997)

23andMe (Genin *Molec Psychiatry* 2011): RRs from cases and controls European GWAS (Lambert *Nat Genet* 2009) modeled with incidence estimates from cases and controls in Rochester (Rocca *Am J Epidemiol* 1998) and PAQUID (Letenneur *J Neurol Neurosurg* 1999) cohorts

Cases/probands often from clinical samples with younger ages, controls often younger and in better health

In addition, statistical methods may bias estimates

Modeled vs. Observed Estimates

Modeled estimates typically apply RR from one study to observed risk in another

Cupples (REVEAL) used RR from largely clinic-based samples in early E4 meta-analyses and applied to observed data in MIRAGE sample of relatives and married ins; did not account for competing risks

Genin (23andMe) used RR from case control GWAS and absolute risk from two cohort studies; assumed controls representative of general population

Five year Cumulative Incidence of dementia in APOE-4 homozygotes by sample and baseline age

	Age 60-64		*Age 65-69*		*Age 70-75*	
	n	*5 yr Cum Inc (%)* (95%CI)	*n*	*5 yr Cum Inc (%)* (95%CI)	*n*	*5-yr Cum Inc (%)* (95%CI)
NACC	36	0.00 (0.00, 0.00)	65	4.36 (1.09, 16.6)	57	12.4 (5.25, 27.9)
FHS	62	0.00 (0.00, 0.00)	44	4.76 (1.19, 18.0)	32	6.67 (1.67,24.6)
RS	102	2.94 (0.95, 8.89)	77	5.19 (1.97, 13.3)	59	11.9 (5.80, 23.4)
SALSA	3	0.00 (0.00-0.00)	5	0.00 (0.00, 0.00)	3	0.00 (0.00, 0.00)

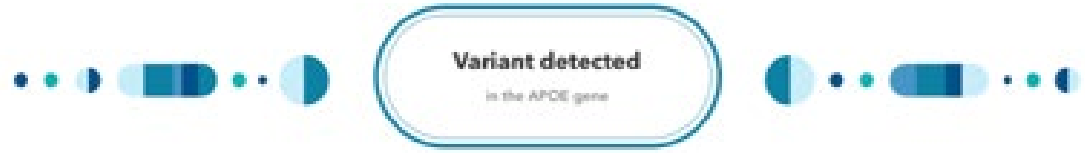
Lifetime (80-85) cumulative incidence of dementia in APOE-4 homozygotes by sample and baseline age

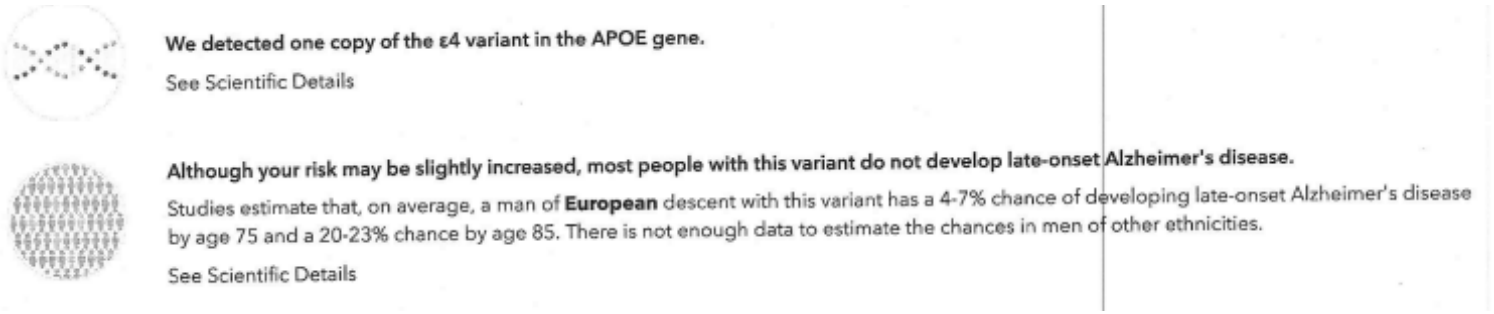
	Age 60-64		*Age 65-69*		*Age 70-75*	
	n	*20 yr Cum Inc (%)* (95%CI)	*n*	*15 yr Cum Inc (%)* (95%CI)	*n*	*10 yr Cum Inc (%)* (95%CI)
FHS	62	38.5 (25.5, 55.2)	44	40.3 (25.8, 59.0)	32	35.2 (20.3, 56.3)
RS	102	34.7 (22.8, 50.5)	77	30.8 (20.7, 44.1)	59	33.3 (22.5, 47.4)

Subdistribution hazard regression predicting MCI/dementia, full model: demographic factors

Variable	NACC	RS	FHS	SALSA
Age (per year)	**1.08** (1.05, 1.10)	**1.08** (1.07, 1.09)	**1.15** (1.12, 1.17)	**1.07** (1.03, 1.12)
Male	**1.14** (0.96,1.36)	**0.92** (0.81, 1.03)	**0.93** (0.79, 1.10)	**0.84** (0.56,1.25)
Educ <HS	**1.41** (0.91, 2.19)	**1.24** (1.06, 1.46)	**1.33** (1.06, 1.65)	**0.80** (0.43, 1.49)
Educ HS	**REF**	**REF**	**REF**	**REF**
Educ some coll	**0.90** (0.66, 1.22)	**0.83** (0.72, 0.95)	**1.10** (0.89, 1.36)	**1.01** (0.42, 2.43)
Educ college	**0.92** (0.73, 1.16)	**0.62** (0.50, 0.77)	**0.87** (0.69, 1.11)	**1.61** (0.72, 3.62)

Sample 23andMe APOE report: 1 *E-4*





We detected one copy of the ε4 variant in the APOE gene.

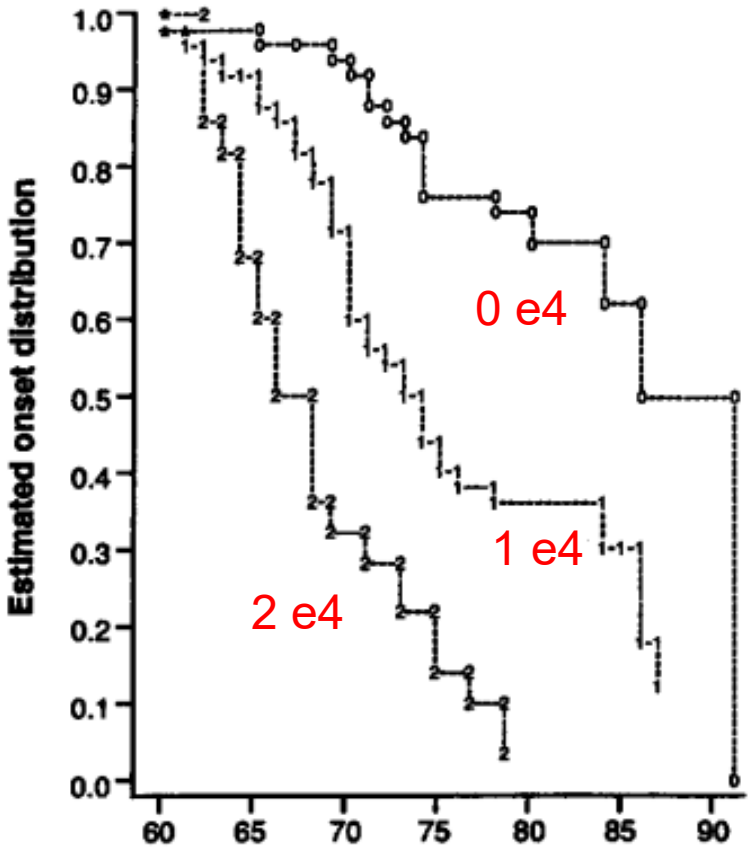
See Scientific Details

Although your risk may be slightly increased, most people with this variant do not develop late-onset Alzheimer's disease.

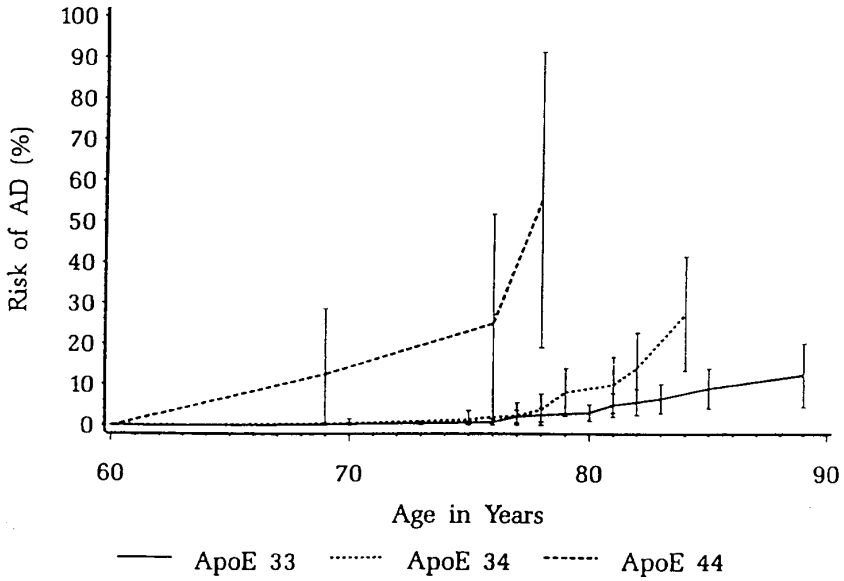
Studies estimate that, on average, a man of **European** descent with this variant has a 4-7% chance of developing late-onset Alzheimer's disease by age 75 and a 20-23% chance by age 85. There is not enough data to estimate the chances in men of other ethnicities.

See Scientific Details

Sampling matters: Family vs. population e4/e44 impact



Corder et al, Science, 1993



Myers et al, Neurology, 1996

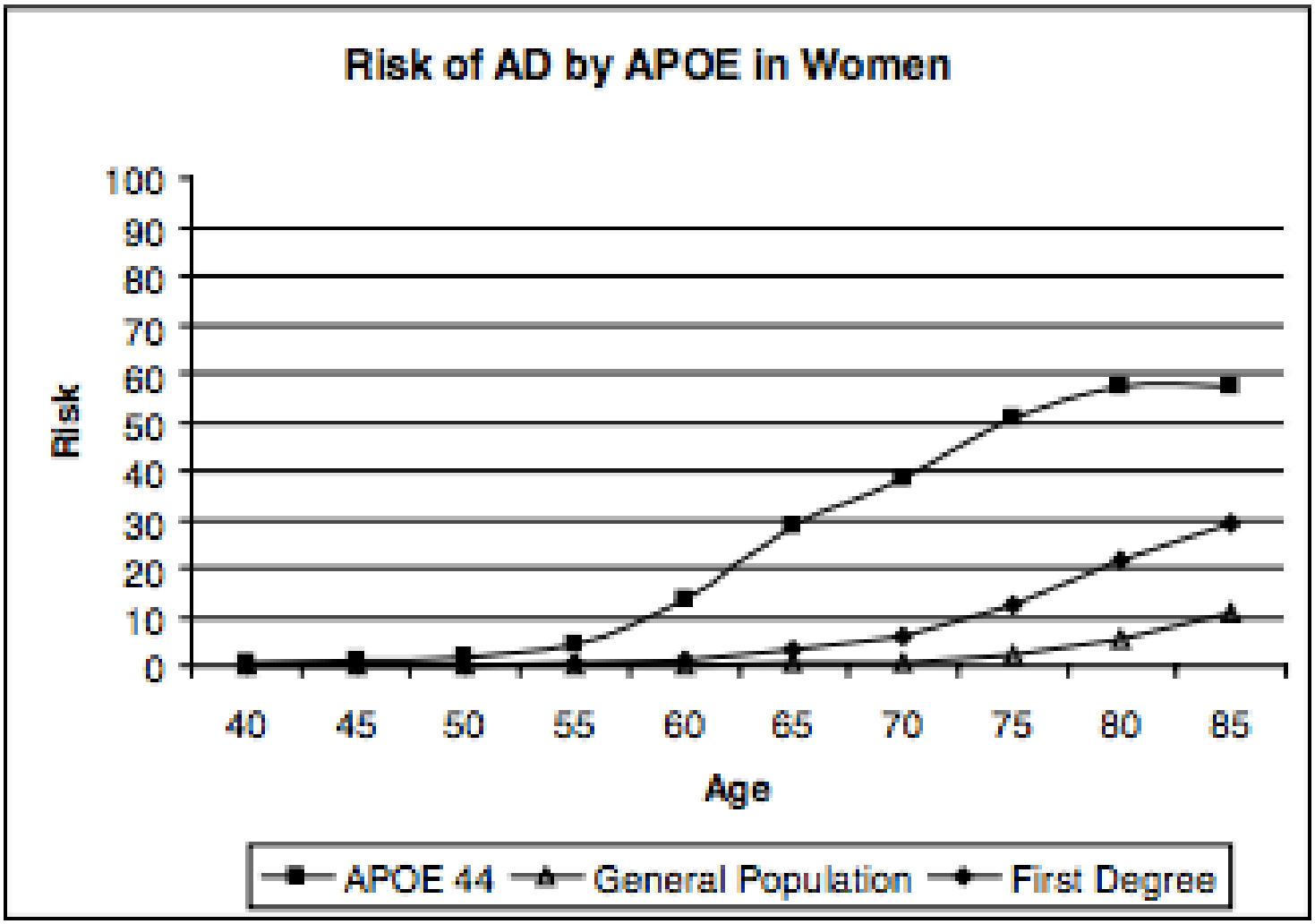
BACKGROUND: *APOE* AND AD RISK

A *susceptibility* gene with three alleles with complex effect on risk for AD and vascular disease

4 allele increases risk, reduces age of onset, and decreases longevity; much greater effect for *44*

#	Freq	Effect on AD risk	Effect on vascular risk
2	0.08	Decreased	Decreased
3	0.78	--	--
4	0.16	Increased	Increased

REVEAL risk curves



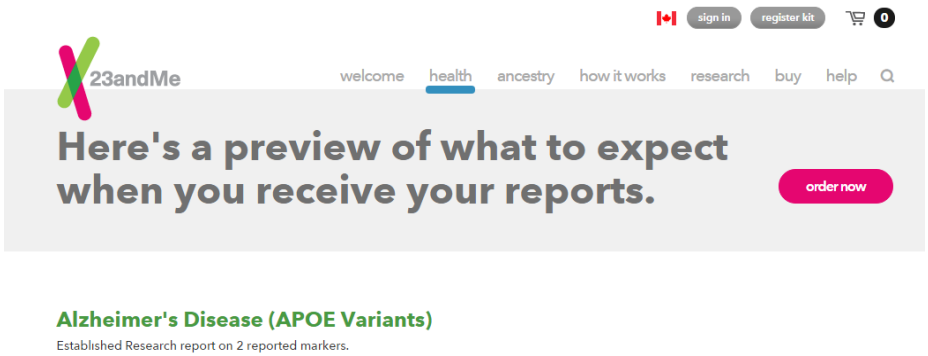
Roberts et al 2005

From 23andMe Canada

Table 3: Risk of developing Alzheimer's disease by 85 years of age

	Men	Women
All genotypes	10-11%	14-17%
ε2/ε2 or ε2/ε3	4-5%	6-8%
ε3/ε3	7-8%	10-12%
ε2/ε4	18-20%	27-31%
ε3/ε4	22-23%	30-35%
ε4/ε4	51-52%	60-68%

Lifetime risk estimates come from two large studies of Alzheimer's disease risk in individuals with European ancestry. The Rochester Study considered 17,483 Caucasian Americans, whereas the PAQUID project considered 2,881 French individuals.



https://www.23andme.com/en-ca/health/i_alzheimers/

AIMS: PROSPECTIVE RISK CALCULATIONS BY *APOE* GENOTYPE

Goal 1: to use prospective data to improve estimates of *APOE*-44-associated MCI/dementia risk for API participants

Goal 2: to inform risk communication re risk with other genotypes for ineligible subjects

Simulating API study entry criteria (normal cognition, age 60-75) within longitudinal cohort studies, and measuring relevant attributes at appropriate baseline

Use (generally) population cohorts with data collection starting at or before age 60 to address needs of trial

Estimate five-year and 'lifetime' cumulative incidence

Meta-analyses ideal given small *44* group, but substantial differences in sampling and assessment methods across cohorts

Methods: Samples

First analyses in ADC samples at NACC; many sampling issues, but may be representative of trial patients

Recruited population-based cohorts with baseline ages ≤60 with help from Sudha Seshadri and the CHARGE Consortium: Framingham, SALSA (Hispanic), and Rotterdam

Included those identified as “normal” by the original study

Assessments and surveillance differ by cohort; Rotterdam ongoing dementia surveillance with cognitive battery q 4 yrs

Sampled visits starting age 60; for FHS and Rotterdam, which have longer follow-up, subjects contributed to multiple baseline age groups in cumulative incidence analyses

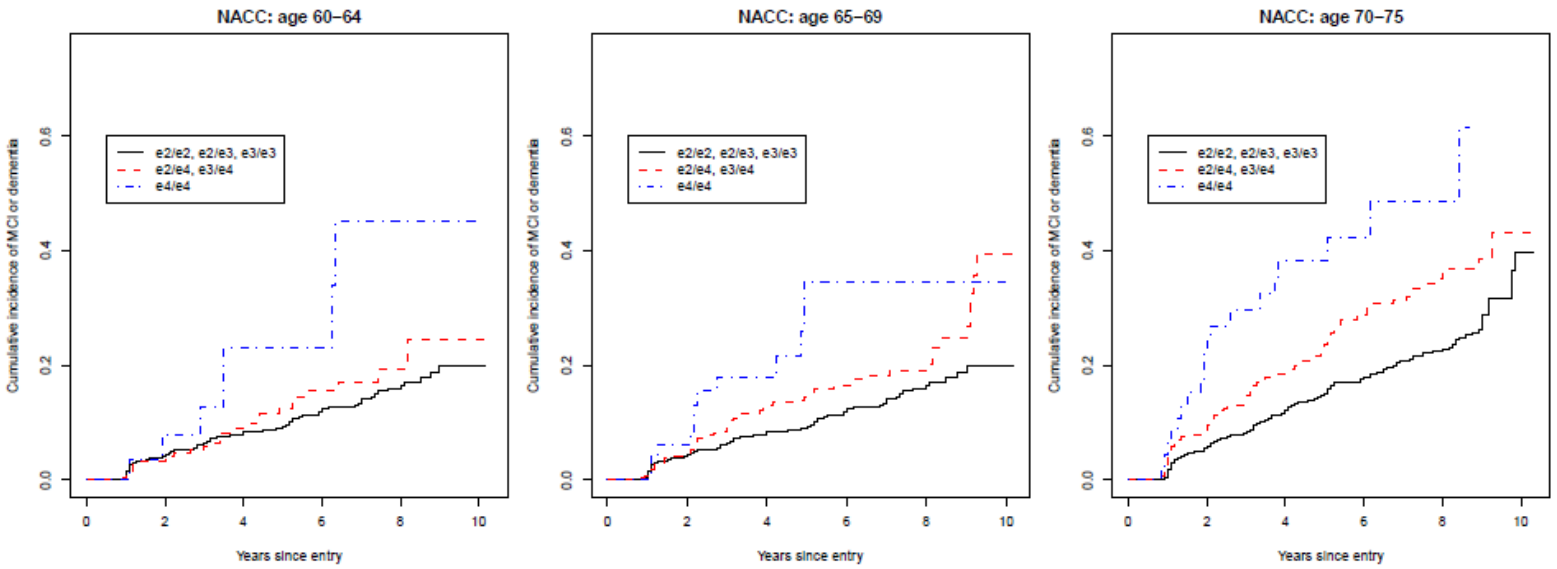
Methods: Analysis

Stratified cumulative incidence curves accounting for competing risk of death, done by age band and gene dose for each study, with estimates of 5-year and (in the long-term cohorts) “lifetime” (to age 80-85) cumulative incidence

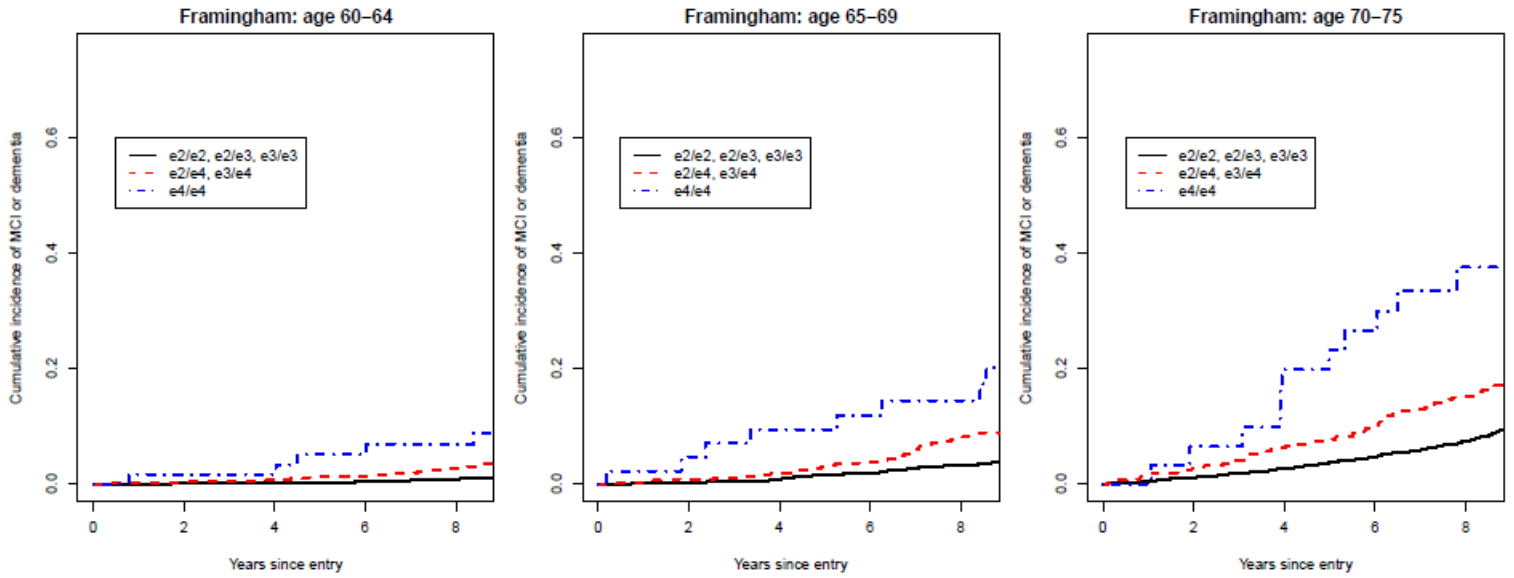
Modeled estimates from survival analyses: univariate, demographics and *APOE* dose, and these plus family history, vascular risk factors, and cognitive test/sx

Chose to use subdistribution hazard regression (Fine & Gray, 1999), in which coefficients can be directly linked to the cumulative incidence function (Haller et al, 2013); results from cause-specific competing risk regression (Prentice et al, 1978) very similar

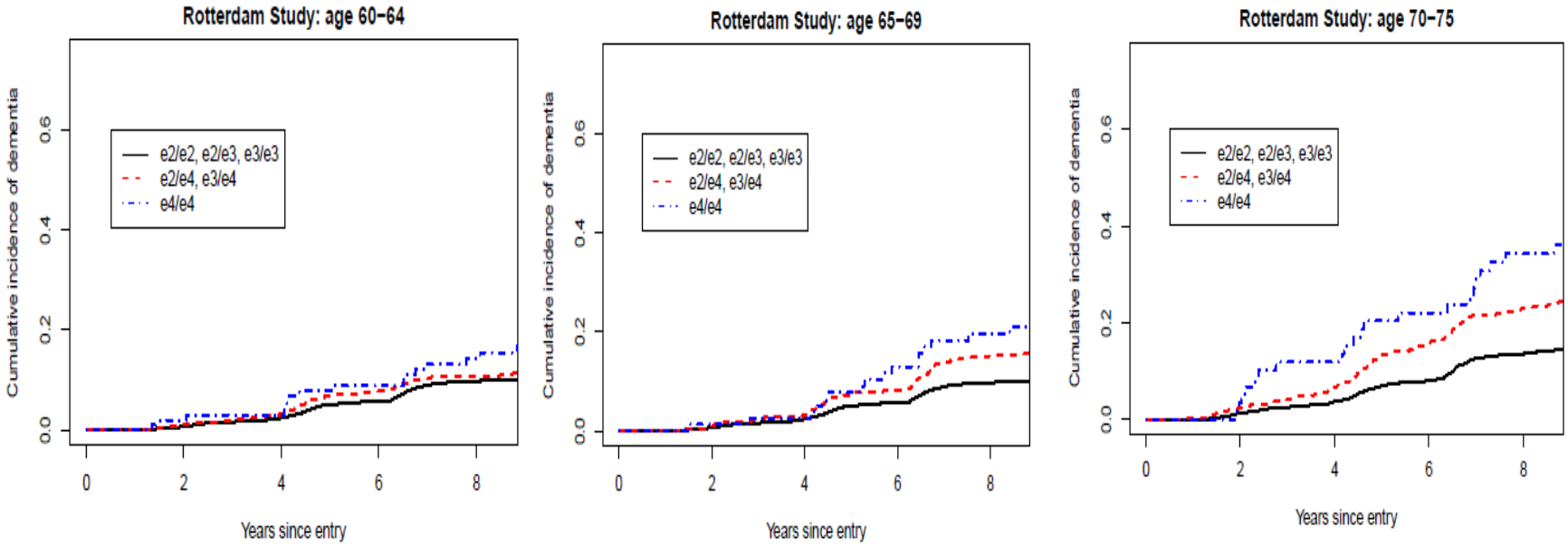
NACC: Cumulative Incidence of MCI or dementia by *APOE* dose



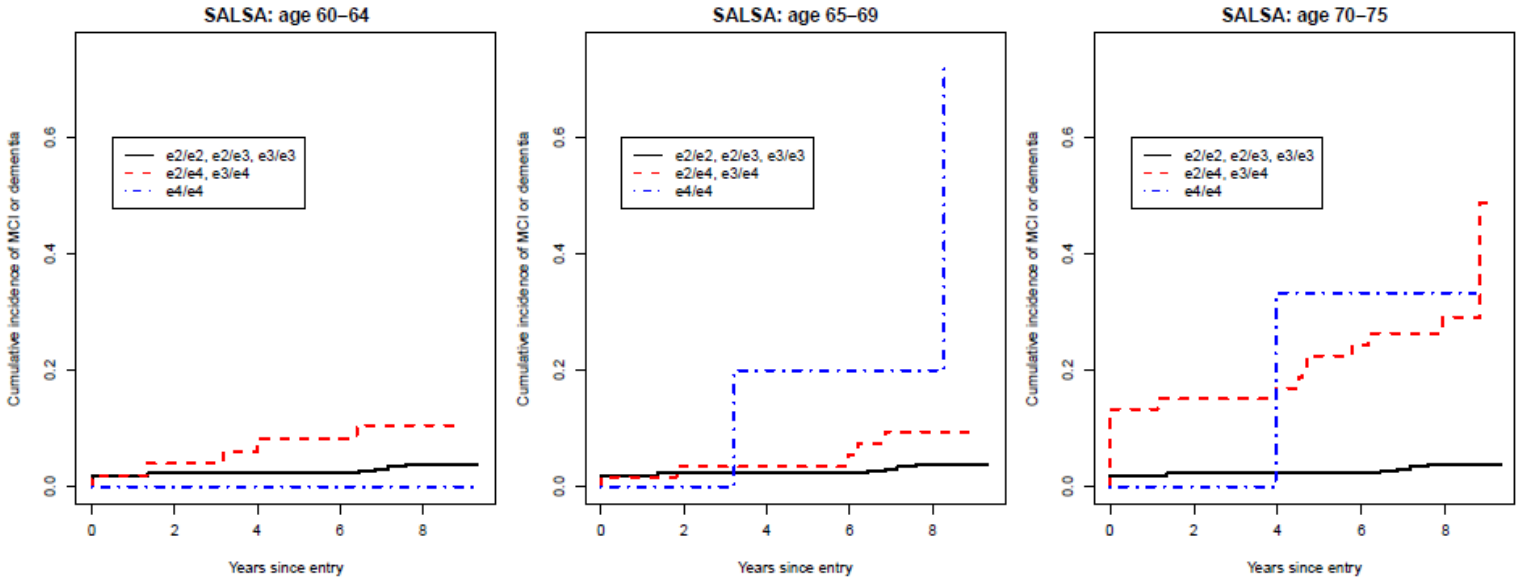
FHS: Cumulative Incidence of MCI or dementia by *APOE* dose



RS: Cumulative Incidence of MCI or dementia by *APOE* dose



SALSA: Cumulative Incidence of MCI or dementia by *APOE* dose



GROUP'S DISCLOSURE PLANS

Extensive discussion of how best to integrate varying estimates across cohorts and detailed modeling

Chose lifetime risks, which they felt were more stable: 0 *E4*: 10-15%; 1 *E4*: 20-25%; *2 E4: 30-55%*

Stressing chances of NOT getting the disease and highlighting uncertainty

Adding that family history and less education may increase risk (and vice versa); not including memory scores or concerns for logistical reasons

Offering relative risks for context, but compared to the general population rather than to just non-carriers (0 *E4*: 0.80-0.85; 1 *E4*: 1.4-1.9; *2 E4:* 2.5-3)