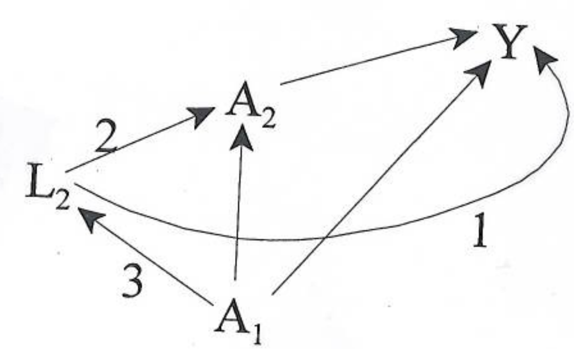
Structural Nested Models and g-estimation

Dylan Small
Department of Statistics, University of Pennsylvania
Acknowledgement: Slides draw from materials from Marshall Joffe

Time-varying treatments : Example

- Observational study for treatment of high blood sugar, measured by hemoglobin A1C (measures long term burden of blood sugar)
- Outcome of interest is
  $Y = \text{A1C at end of second time period}$
- Treatments:
- $A_1 = \text{Treatment in first time period},$
  $A_2 = \text{Treatment in first time period}$
- Post-baseline covariate:
  $L_2 = \text{Ill at beginning of second time period}$
- Illness at beginning of second time period associated with higher blood sugar
- True effect: Treatment in each time period decreases blood sugar in A1C by 10 units



Structural Nested Models

$Y_i^{a_1,a_2}$ =outcome person *i* would experience if she were to receive treatment level a_1 in first period, a_2 in second period.

Structural Nested Model:

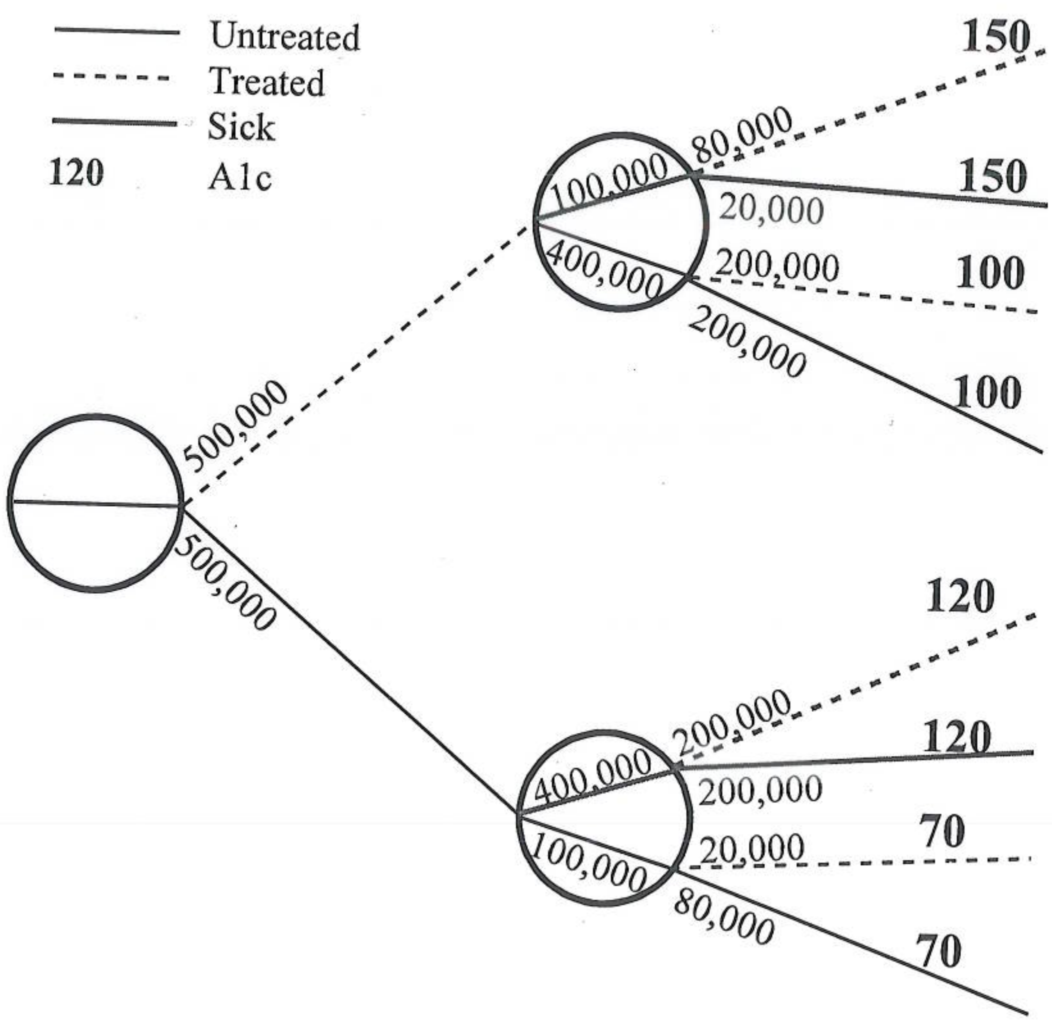
$Y^{a_1,a_2} = Y^{0,0} + \Psi(a_1 + a_2)$

Ψ = effect of one unit of treatment on final outcome

(Could consider alternative models that allow for effect of treatment to differ by time period, e.g.,

$Y^{a_1,a_2} = Y^{0,0} + \Psi_1 a_1 + \Psi_2 a_2$)

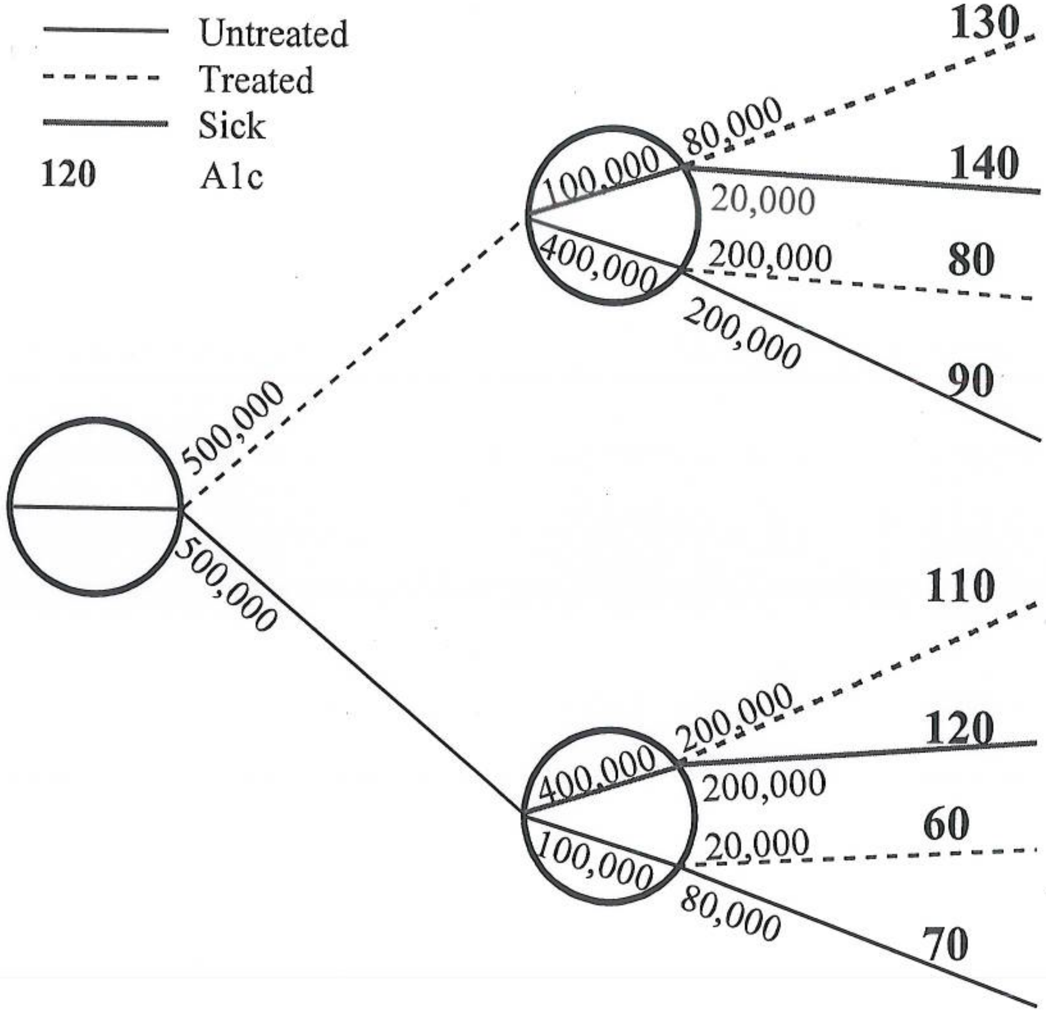
Hypothetical data: potential outcomes $y^{0,\bar{0}}$



right circumference points: comparable groups

effect of each increment of treatment: lower A1c by 10 (see next panel)

Observed data Y^{A_1,A_2}



g-estimation of Structural Nested Models

Structural Nested Model:

$$Y^{a_1,a_2} = Y^{0,0} + \Psi(a_1 + a_2)$$

Ψ = effect of one unit of treatment on final outcome

<u>g-estimation</u>

Compute putative potential outcome if the true effect is Ψ.

$$Y^{0,0}(\Psi) = Y - (A_1 + A_2)\Psi$$

where Y, A_1, A_2 are observed outcome and treatments.

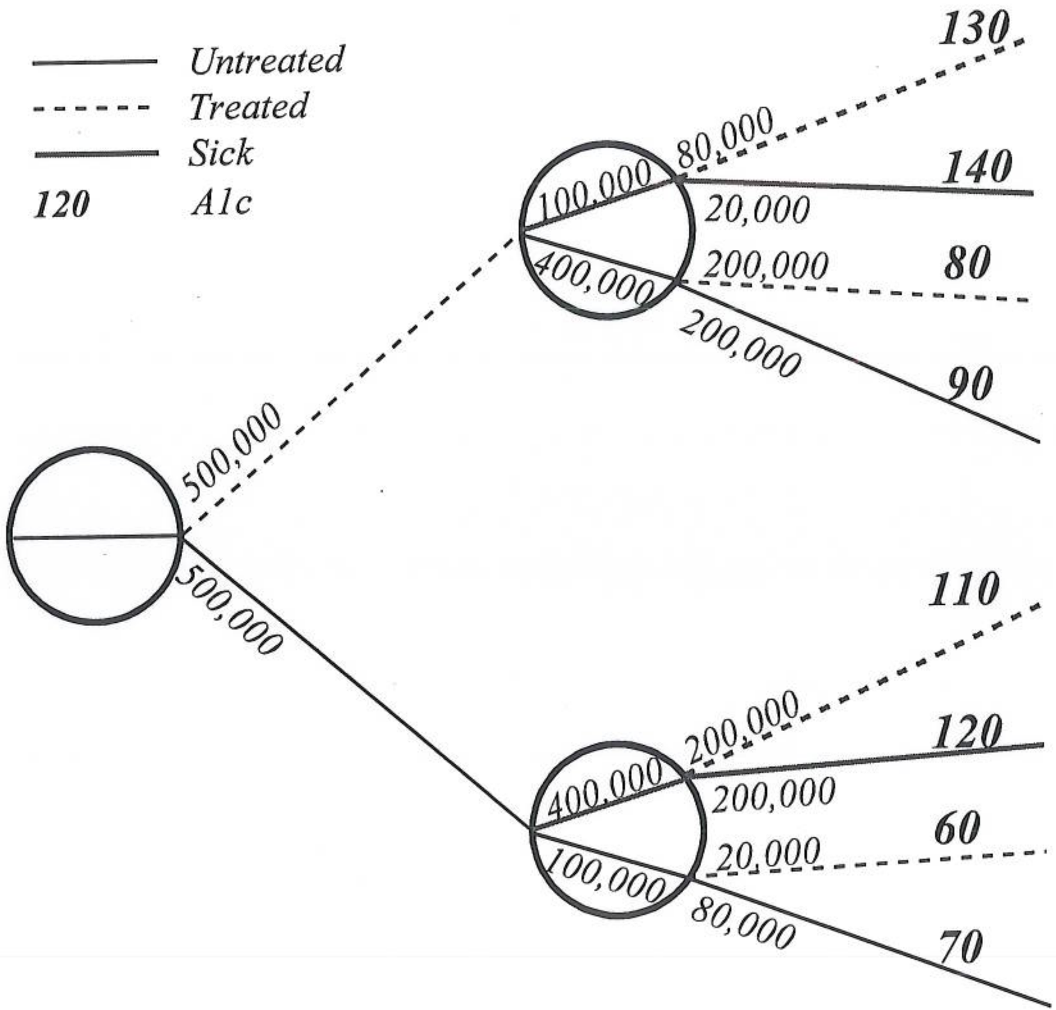
Key assumption:

Sequential ignorability (sequential effective randomization): Treatment at each time point effectively randomized given previous treatment history and covariates up to that point.

Under sequential ignorability, $Y^{0,0}(\Psi)$ is independent of treatment at each time point given previous treatment history and covariates up to that point.

g-estimation: Test this independence using logistic regression or estimating equations.

Reference:
Robins, J.M. et al. (1992), G-estimation of the effect of prophylaxis therapy for pneumocystic carinii pneunmonia on the survival of AIDS patients, *Epidemiology*, 3, 319-36.



for $\Psi = 0$, test independence of A_2, $Y^{00}(\Psi)$ (Compare 130 vs, 140, etc.)

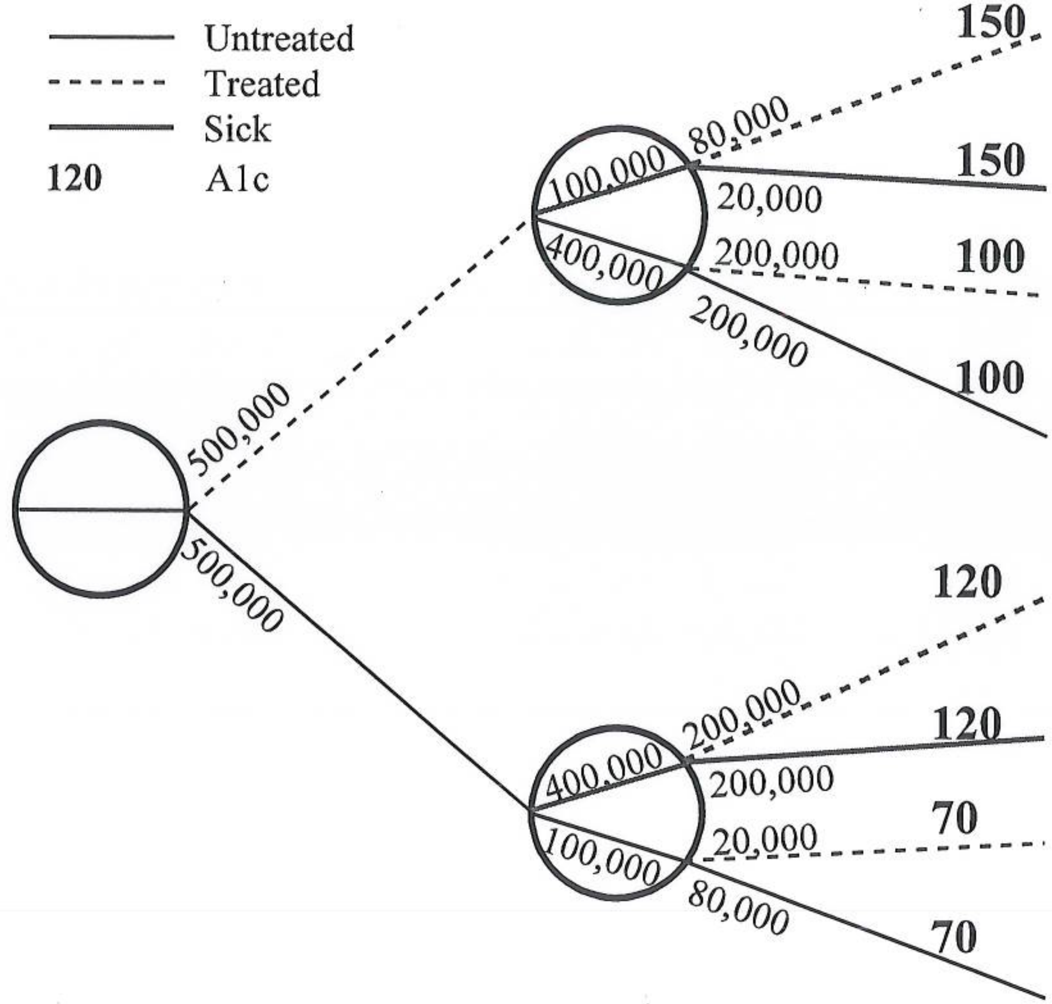
also test independence of A_1, $Y^{00}(\Psi)$:

(130*8+140*2+80*20+90*20)/50 = 94.4

vs.

(110*20+120*20+60*2+70*8) /50= 105.6

(Reject)



for $\Psi = -10$, test independence of A_1, $Y^{00}(\Psi)$

(150*8+150*2+100*20 +100*20)/50 =110

vs.

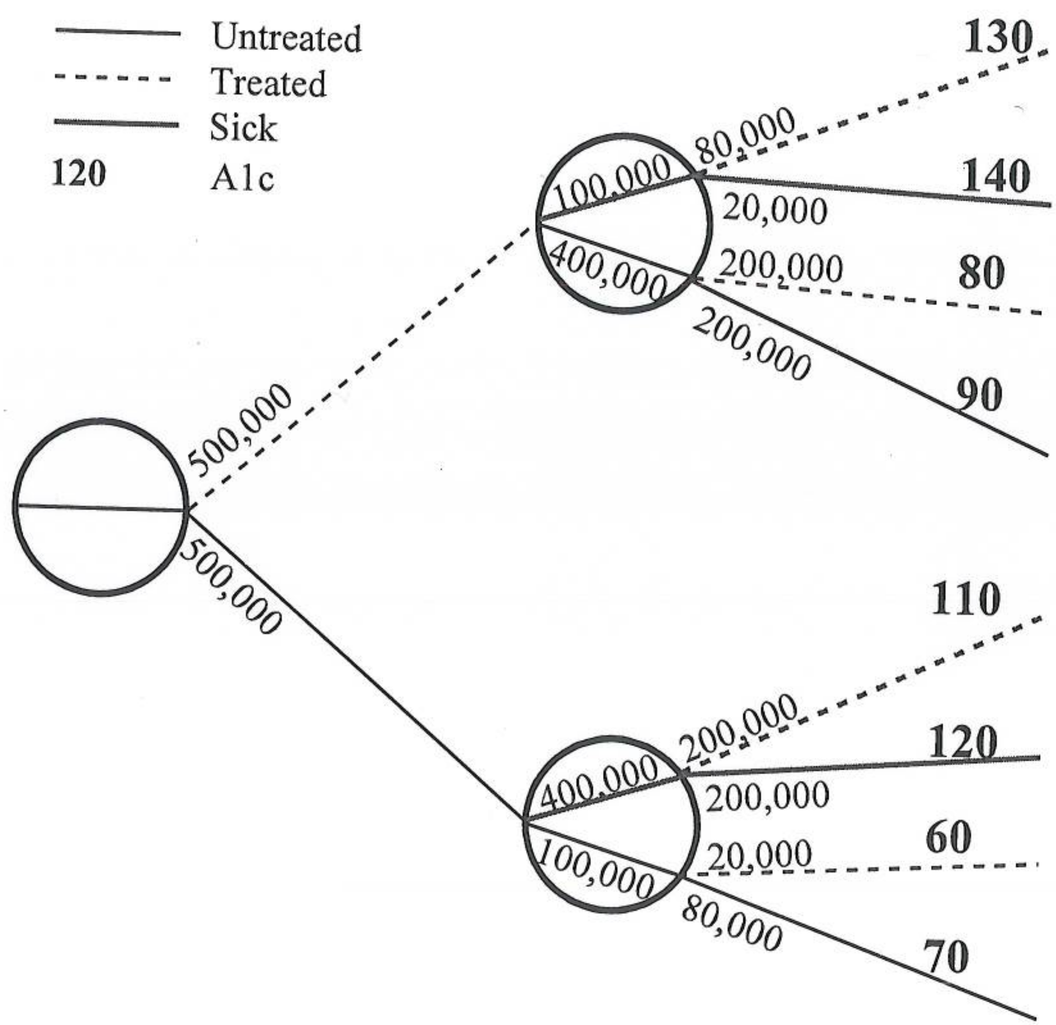
(120*20+120*20+70*2 +70*8) /50= 110

(Fail to reject)

Comparison of structural nested model/g-estimation with standard regression approaches

- Standard regression approaches will be biased for estimating joint effects of treatments when
  - (A1) there exists a time-dependent risk factor for or predictor of the event of interest that also predicts subsequent treatment
  - (A2) past treatment history predicts subsequent risk factor level.
- g-estimation provides unbiased estimates under (A1) and (A2) under sequential ignorability. Other methods which can provide unbiased estimates are (i) g formula with associated method of g computation; (ii) marginal structural models with associated method of inverse probability weighting

standard approaches



$E(Y|A_1) = 105.6 - 11.2A_1$

attenuated as measure of overall effect

$$E(Y|A_2)=100.8-1.6A_2$$
$$E(Y|A_1,A_2)=105.71429-11.16883A_1-0.25974A_2$$
$$E(Y|A_1+A_2)=105.71429-5.71429(A_1+A_2)$$
$$E(Y|A_1,L_2)=68+17A_1+47L_2$$
$$E(Y|A_2,L_2)=84.28571-6.10390A_2+37.53247L_2$$
$$E(Y|A_1,A_2,L_2)=70+20A_1-10A_2+50L_2$$
$$E(Y|A_1+A_2,L_2)=78.59873+2.42038(A_1+A_2)+37.96178L_2$$

The only correct estimate is effect of A_2 when controlling for A_1,L_2

expected under sequentially ignorable treatment assignment
can estimate component effect of last bit of treatment

cannot estimate joint effects of treatments received at different times

Comparison of g-estimation with marginal structural models and g-computation

- Comparison of g-estimation to marginal structural models:
  - When there is a lack of overlap (for some covariate combinations, almost everybody takes treatment or almost everybody takes control), g-estimation estimates average treatment effect in region of overlap with reasonable variance, while MSM tries to estimate overall average treatment with very high variance.
  - Structural nested models with g-estimation can be used to accommodate failure of sequential ignorability through use of instrumental variables and other methods.
- Comparison of g-estimation to g-computation
  - Structural nested model with g-estimation directly models the effect of treatment on the outcome in a parsimonious way.
- Disadvantage of g-estimation: Lack of off the shelf software. SAS and STATA macros available at http://www.hsph.harvard.edu/causal/software/.

Reference: Vansteelandt, S. and Joffe, M. (2014) Structural nested models and g-estimation: the partially realized promise. *Statistical Science*, 29, 707-731.

Structural Nested Failure Time Model (SNFTM)

- SNFTM: Structural nested model for a failure time outcome.
- SNFTM can correct for immortal time bias and healthy worker survivor effect

Oscar Winners Live Longer?

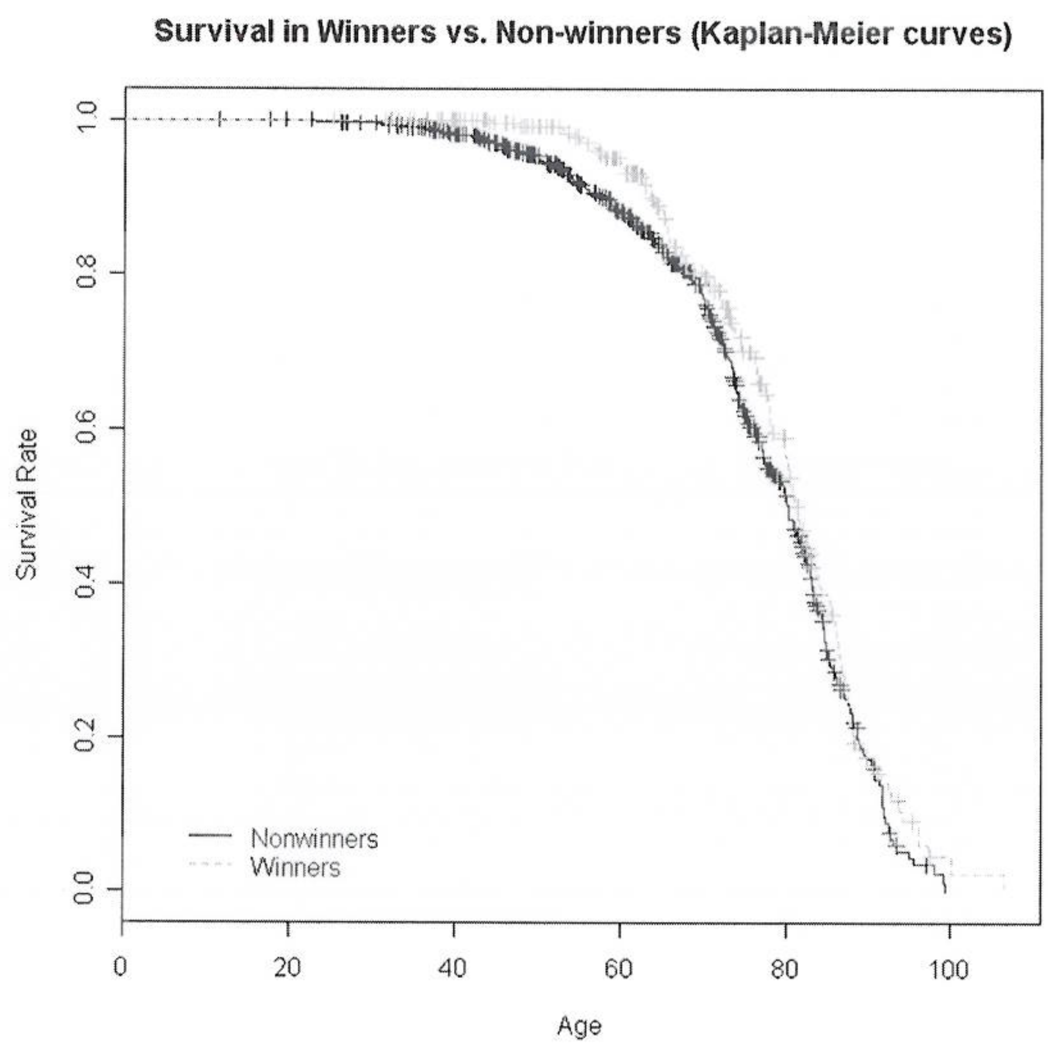
From *Science*, May 25, 2001

- Katharine Hepburn celebrated her 94th birthday this month, and her four Oscars may have something to do with it:

According to a study published in the 15 May issue of the *Annals of Internal Medicine*, Academy Award winners live longer than other successful actors.

Analysis in the Annals of Internal Medicine Paper

- Main analysis:
  - Considered all actors and actresses who have been nominated for an Oscar (Academy Award).
  - Classified an actor or actress who won an award at some point in his/her lifetime as a winner and all other actors/actresses as nonwinners (controls).
  - Fit Cox proportional hazards model to find effect of winning on mortality rate.



Cox proportional hazards model: The reduction in mortality rate associated with winning was estimated to be 25% with a 95% CI of 5%-41%.

Immortal Time Bias

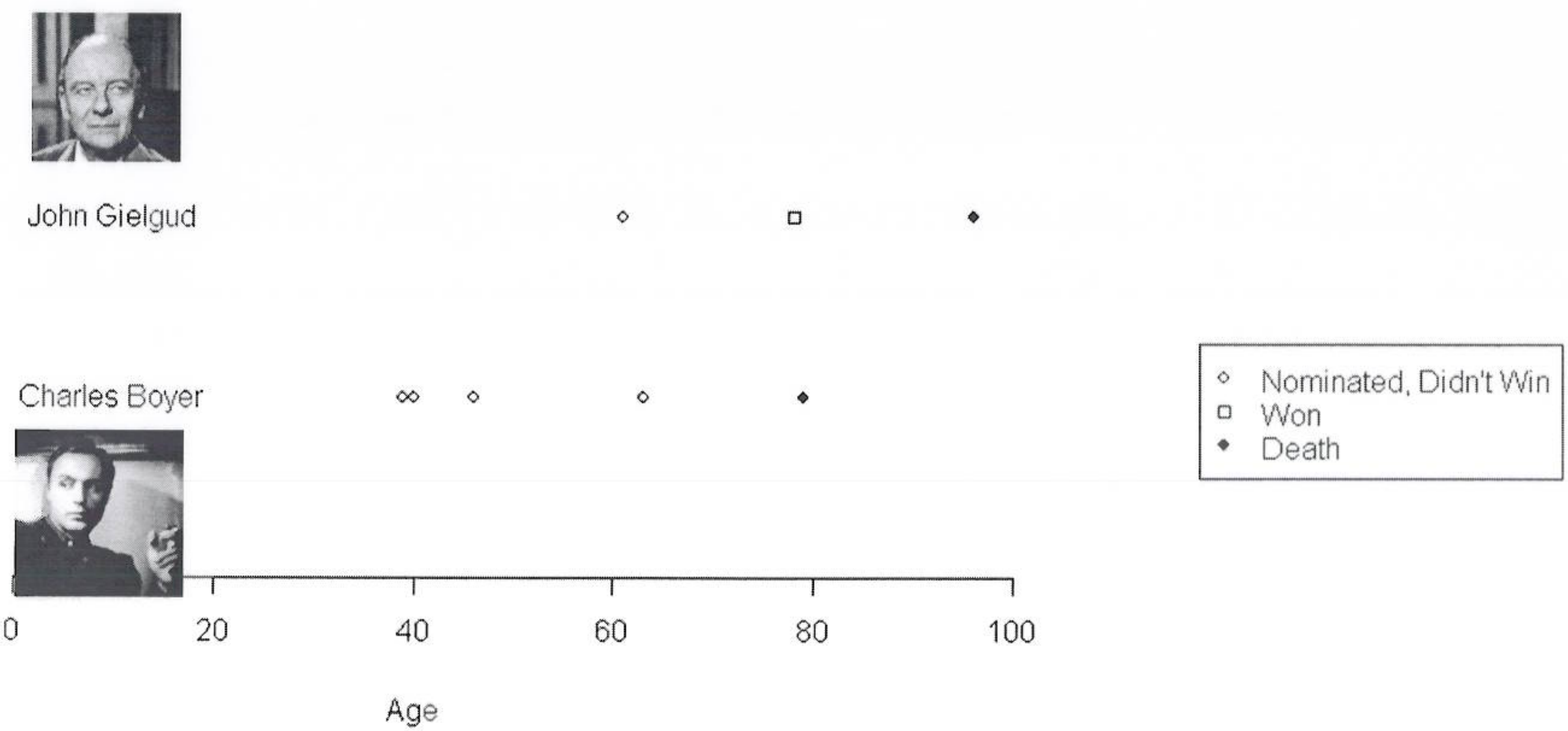
- The nonwinning nominees can die anytime after their first nomination.
- The winners cannot die until they win. The time between the winners' first nomination and first win is "immortal time."
- Example:

  - Henry Fonda was first nominated in 1941 for his role in "The Grapes of Wrath." Fonda won his first Oscar for his role in "On Golden Pond" in 1981.
  - James Stephenson was first nominated in 1940 for his role in "The Letter" and did not win. Stephenson died in 1941.

Healthy Worker Survivor Bias

- Winning is an indication of health:

  After the first nomination, winning an award at a later time is an indication that the worker was health at that later time.



Structural Nested Failure Time Model

- Treatments are win or not win in each year.
- Among nominees in a given year, test (using logistic regression) whether winning is independent of potential survival time without winning conditional on actor/actress' age and previous nomination history.

Results

- There is suggestive but inconclusive evidence ($p=0.07$) that winning has an effect on lifetime
- We obtain a 95% confidence interval that winning multiplies lifetime by a factor of (0.99,1.27).
- To translate this into lifetime, we estimated the median lifetime for winners compared to their median putative lifetime if they did not win under a given treatment effect:

95% confidence interval for effect of winning on median lifetime: (-0.4 years, 8.4 years).

Reference: Han, X., Small, D., Foster, D. and Patel, V. (2011). The effect of winning an Oscar award on survival: correcting for healthy performer survivor bias with a rank preserving structural accelerated failure time model. *Annals of Applied Statistics*, 5, 746-772.

Summary

- Structural nested models and the associated method of g-estimation are an approach to modeling the joint effects of a sequence of treatments or exposures.
- Standard regression approaches are biased when (A1) there exists a time-dependent risk factor for or predictor of the event of interest that also predicts subsequent treatment and (A2) past treatment history predicts subsequent risk factor level.
- Structural nested models with g-estimation provide unbiased estimates under (A1) and (A2) when enough time varying covariates have been collected so that treatment is sequentially ignorable.