Average Causal Effects (ACE) on Binary Outcomes: Measures, Collapsibility, Estimation by Propensity Scoring

Yi Huang
Ph.D Candidate

Karen Bandeen-Roche
Professor

Constantine Frangakis
Associate Professor

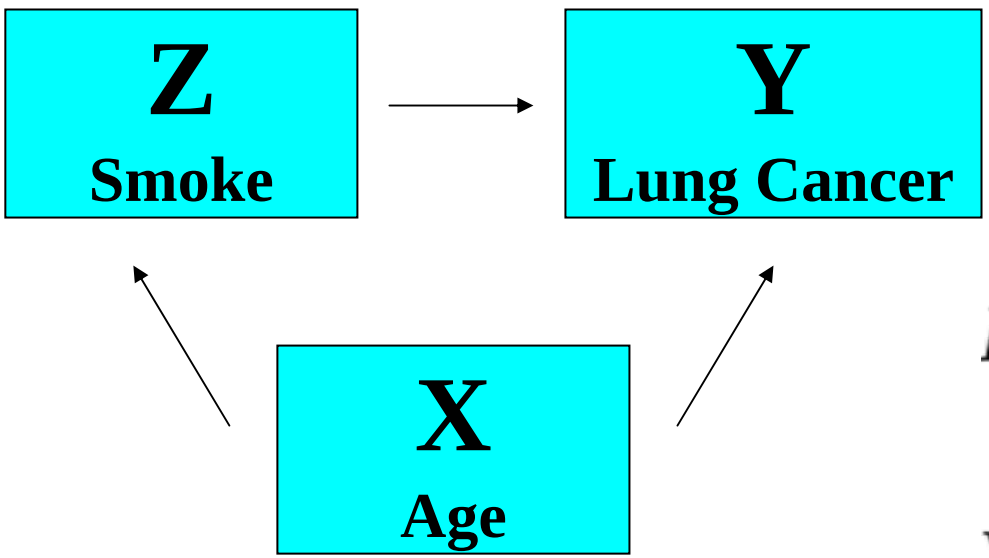
Department of Biostatistics,
Johns Hopkins Bloomberg School of Public Health

Abstract

Propensity score approach (Rosenbaum, Rubin, 1983,1984) become increasingly popular in biomedical applications to find ATE, especially when our prior knowledge on the relationship between outcome and covariates is not strong. For binary outcomes, different bin-specific parameters + different ways of combining them have been proposed to estimate ATE. However, some of them are poor, which could potentially give very misleading inferences. In order to tell the better ones, collapsibility property related with different ATE measures using propensity scoring have been discussed. Propensity scoring estimands for average risk difference and marginal relative risk appear to have better interpretation than marginal odds ratio. And, ATE estimators should be better constructed from estimated population average risk, not bin-specific log(OR). The procedure of combining bin-specific risk to estimate marginal relative risk is also attached. Furthermore, we show that it is not always correct to say "average treament effect is a weighted average of bin-specific treatment effect", even under rare disease assumption. To visualize it, some simulation results are shown. At last, we find the general mathematical requirement for the form of any collapsible ATE measure, under constant individual treatment effect assumption. Enjoy the poster ! ☺

Questions ??

- **How to determine the average treatment/ risk factor effect (ATE) on reducing / increasing disease risk, in a large observational study with a lot of observed confounders ?**



- $Y_i^{(0,1)}$: Potential outcome
- Population average risk:

$$P^{(Z)} = \frac{1}{N}\sum_{i=1}^{N} Pr(Y_i^{(Z)} = 1)$$

where, Z = 0, 1

Propensity Scoring (P.S.)

- **Define: $e_T(X) = P_r(Z=1|X)$**

  (e.g. the risk of being a smoker at giving age.)

- **$e_T(x)$ is a balancing score → $X \perp\!\!\!\perp Z \mid e_T(x)$.**

  (e.g. for people with same $e_T(X)$, the distribution of age is same across smoking groups.)

- Typical tool for causal inference – **average** effect.

- **Three conditions** for valid causal inference:

  1). [Z | X] is not degenerate.

  2). Treatment assignment is strongly ignorable

  3). Close to correctly specified: Z relationship to X.

Rosenbaum, Rubin, 1983, 1984; Rubin, 1997.

P.S. Procedure

1. **Estimate e(x).**
2. **Take overlapped $\widehat{e(x)}$, then subclassify into bins.**
3. **Checking: $f_{X|Z=1,\ jth\ bin} \approx f_{X|Z=0,\ jth\ bin}$**
4. **Target parameter:** $P_j^{(Z)} = Pr(Y^{(Z)} = 1 | j^{th}\ \text{bin})$

$$\widehat{P_j^{(0)}} = average(Y|Z=0, j^{th}\ \text{bin})$$

$$\widehat{P_j^{(1)}} = average(Y|Z=1, j^{th}\ \text{bin})$$

$$\widehat{P^{(1)}} = \sum_{j=1}^{J} \omega_j \widehat{P_j^{(1)}}$$

$$\widehat{P^{(0)}} = \sum_{j=1}^{J} \omega_j \widehat{P_j^{(0)}}$$

5. **Choose** ATE measure, and estimate it.

Collapsibility

- Hypothetical example: No confounding

	$P(Y^{(1)}=1)$	$P(Y^{(0)}=1)$
Age<65	0.4	0.2
Age≥65	0.8	0.6
Overall	0.6	0.4

OR – not collapsible.

→ $OR_1 = 2.67$

→ $OR_2 = 2.67$

→ $OR_{all} = 2.25 \neq \frac{1}{2}2.67 + \frac{1}{2}2.67$

- **Collapsible:**

$$\sum_{j=1}^{J} w_j f(P_k^{(1)}, P_k^{(0)}) = f(\sum_{j=1}^{J} w_j P_k^{(1)}, \sum_{j=1}^{J} w_j P_k^{(0)})$$

- **Average Risk Difference = $P^{(1)}$-$P^{(0)}$, fully collapsible**

Example from: Greenland, 1999, also see: 2002, 1987, 1985, ; Whittaker, 1990

$P^{(1)}/P^{(0)}$ – Marginal Relative Risk

- **Not Collapsible, in general.**

$$\frac{P^{(1)}}{P^{(0)}} = \frac{\sum_{j=1}^{J} \omega_j P_j^{(1)}}{\sum_{j=1}^{J} \omega_j P_j^{(0)}} \neq \sum_{j=1}^{J} \omega_j \frac{P_j^{(1)}}{P_j^{(0)}}$$

- **Collapsible, w/ constant treatment effect assumption.**

$$\text{If } \frac{Pr(Y_i^{(1)} = 1)}{Pr(Y_i^{(0)} = 1)} = r \quad for \quad i = 1, 2, \ldots N \text{ , then}$$

$$\frac{P^{(1)}}{P^{(0)}} = \sum_{j=1}^{J} \omega_j \frac{P_j^{(1)}}{P_j^{(0)}} = r$$

Population Subgroup Individual

P.S.: $\frac{P^{(1)}/(1-P^{(1)})}{P^{(0)}/(1-P^{(0)})}$ – Marginal Odds Ratio (OR)

- **Not collapsible, even under:**

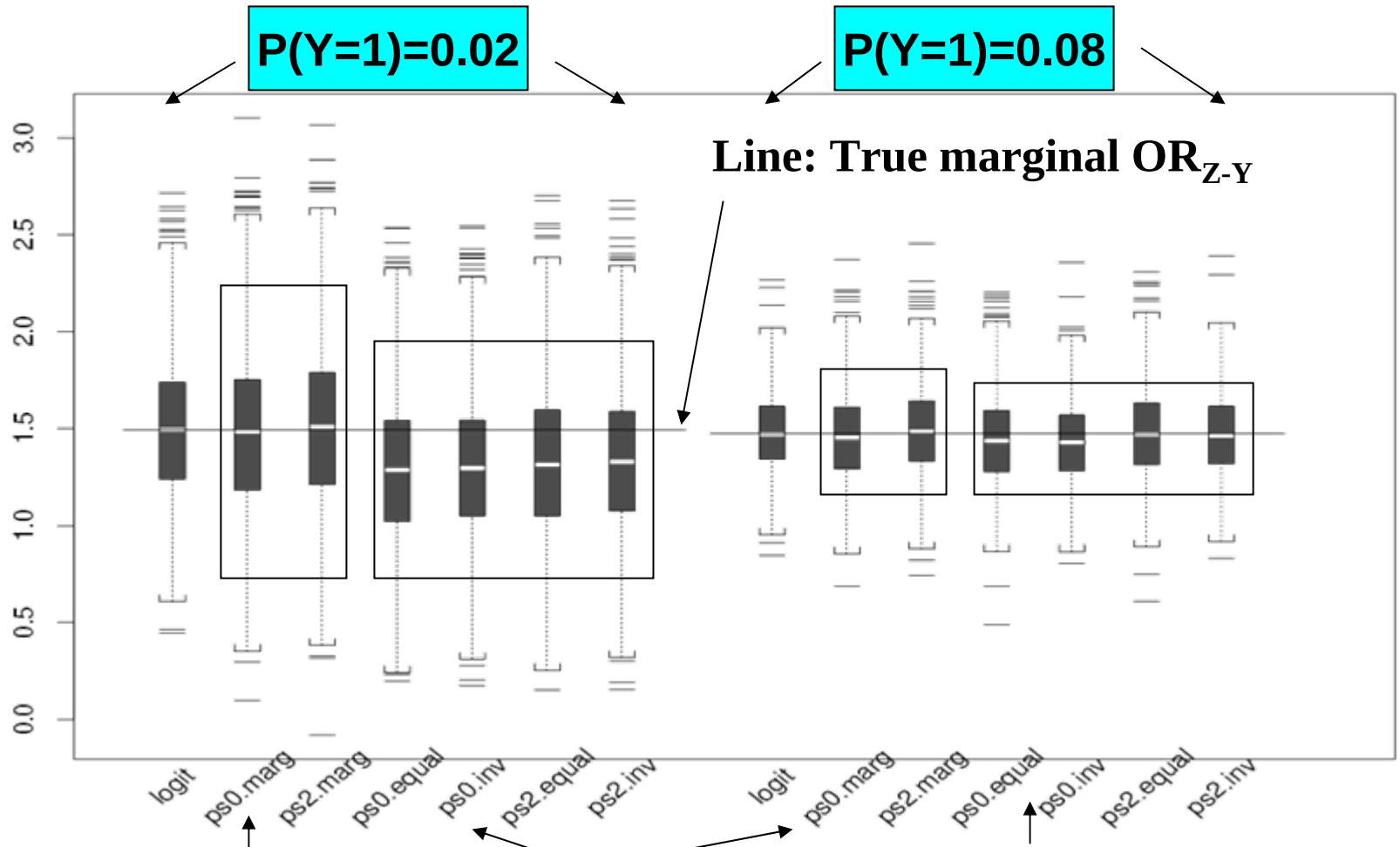
  constant treatment effect assumption

$$\frac{P^{(1)}/(1-P^{(1)})}{P^{(0)}/(1-P^{(0)})} = \frac{\sum_{j=1}^{J}\omega_j P_j^{(1)}/(1-\sum_{j=1}^{J}\omega_j P_j^{(1)})}{\sum_{j=1}^{J}\omega_j P_j^{(0)}/(1-\sum_{j=1}^{J}\omega_j P_j^{(0)})}$$

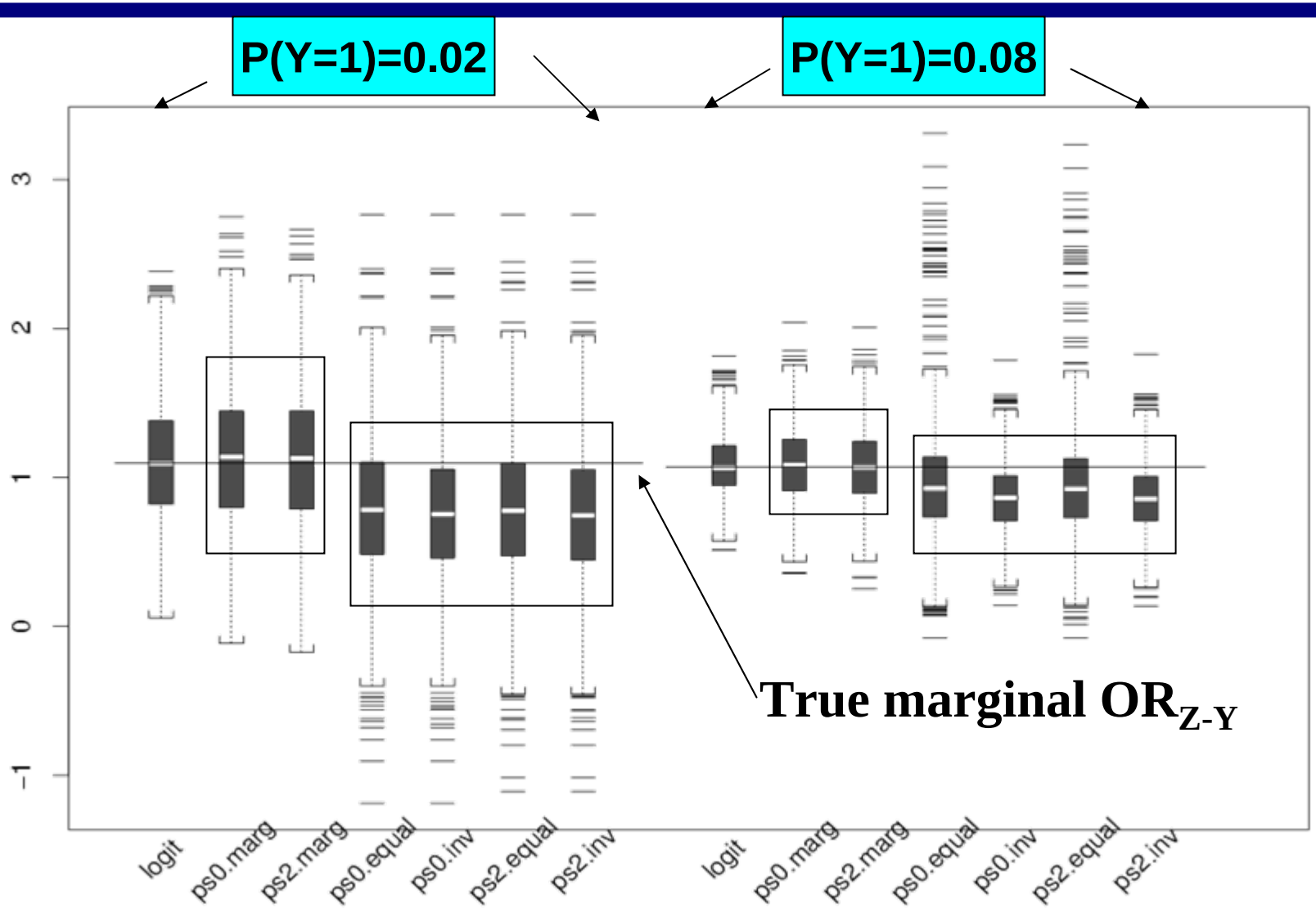
$$\neq \sum_{j=1}^{J}\omega_j \frac{P_j^{(1)}/(1-P_j^{(1)})}{P_j^{(0)}/(1-P_j^{(0)})} \quad \textbf{(1)}$$

Constant OR_i → $$= \frac{\sum_{j=1}^{J}\omega_j P_j^{(1)}/(1-P_j^{(1)})}{\sum_{j=1}^{J}\omega_j P_j^{(0)}/(1-P_j^{(0)})} \quad \textbf{(2)}$$

Log(Marginal $OR_{Z\text{-}Y}$), + constant treatment effect

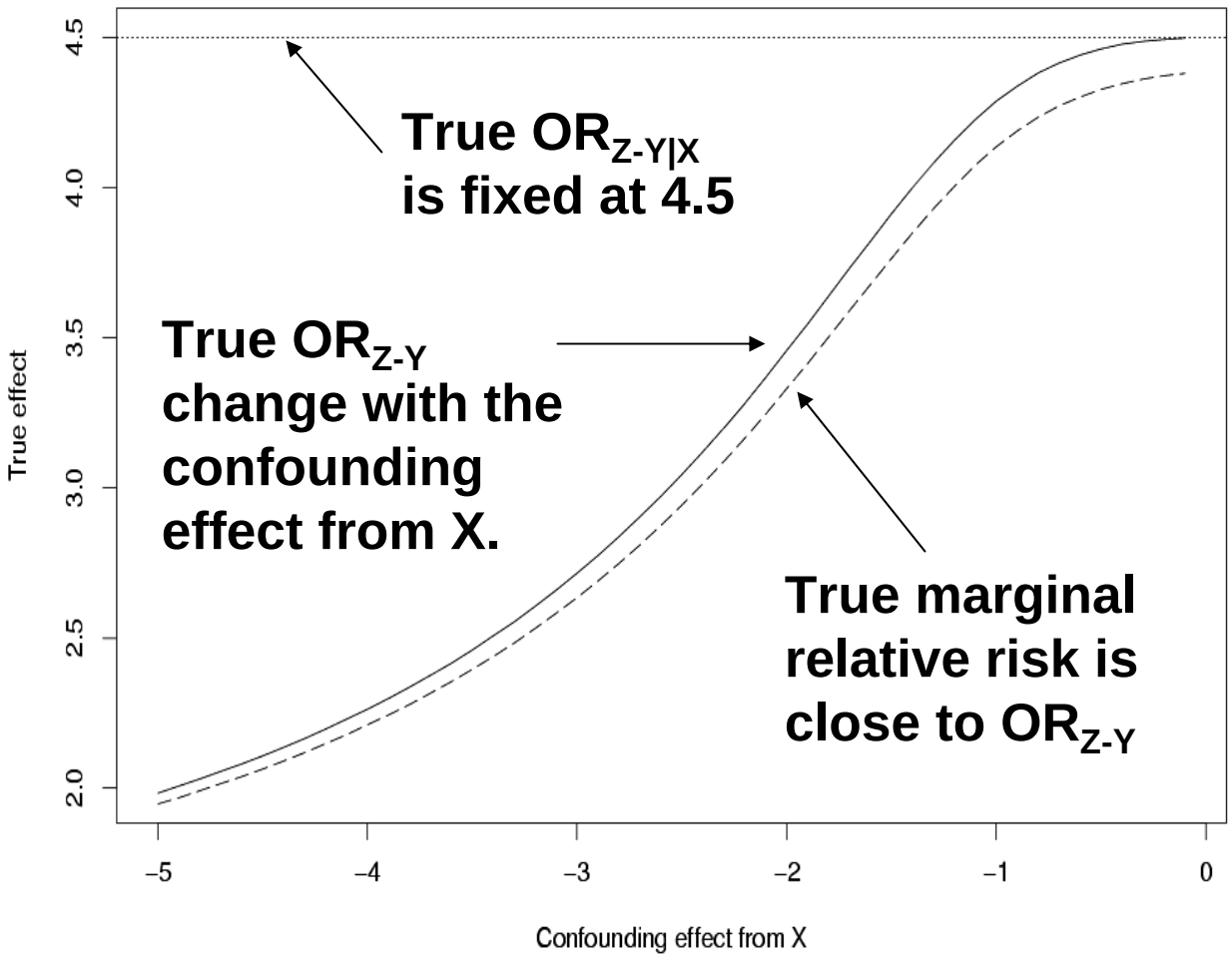


Log(Marginal OR_{Z-Y}), No constant treatment effect



$OR_{Z\_Y|X} \neq OR_{Z\_Y}$, even when disease is rare

Comparing $OR_{Z\text{-}Y|X}$ & $OR_{Z\text{-}Y}$



Setting:

$$X \sim N(0,1)$$

$$logit[E(Y|,X)] = \beta_0 + log(4.5)Z + \beta_x X$$

$Pr(Y = 1) = 0.02$, by adjusting β_0

N = 8000

Requirement: Collapsible ATE Measure

- under constant individual treatment effect

- Individual effect = $f(a_i, b_i)$
- ATE = $f(\bar{a}, \bar{b})$
- **If** for all i,

$$f(\bar{a}, \bar{b}) = f(a_i, b_i) = \lambda$$

then, **f** is a function, s.t. :

$$b_i = r(\lambda)a_i + \delta(\lambda)$$

$r(\lambda)\ \delta(\lambda)$ continuous fxn

w/ $r(\lambda_0) = 1$ $\delta(\lambda_0) = 0$

where, λ_0 = no treatment effect.

Special Cases:

$$a_i = P(Y_i^0 = 1)$$
$$b_i = P(Y_i^1 = 1)$$

1, Set $r(\lambda) \equiv 1$, we show

If, $f(a_i, b_i) = f(b_i - a_i)$

then, it is fully collapsible.

2, Set $\delta(\lambda) \equiv 0$, we show

If, $f(a_i, b_i) = f\left(\frac{b_i}{a_i}\right)$

then, it is fully collapsible.

Summary + Discussion

- If marginal RR/OR are target estimands, estimators should be **better constructed** from the **estimated population average risk**, not bin-specific log(OR).
- P.S: **Not always correct** to say – "average treament effect is a weighted average of bin-specific treatment effect", even under rare disease assumption. **It really depends on your choice of treatment effect measure.**
- $P^{(1)}$-$P^{(0)}$ & $P^{(1)}/P^{(0)}$ have **better** **collapsibility + interprebility** property.
- With **constant treatment effect + the increasing of disease prevalence,** the performance on weighted average type estimators becomes better. **Without** constant treatment effect, their performance is poor.
- Finding → **general mathematical requirement** for any collapsible effect measure under constant individual treatment effect assumption.