Adjustments for Unmeasured Confounders with External Information using Propensity Score Calibration

Jin-Liern Hong

Department of Epidemiology
UNC at Chapel Hill

October 24th, 2014

THE UNIVERSITY
of NORTH CAROLINA
at CHAPEL HILL

Some of the slides are kindly provided by Dr. Til Stürmer
(Department of Epidemiology, UNC at Chapel Hill)

THE UNIVERSITY
of NORTH CAROLINA
at CHAPEL HILL

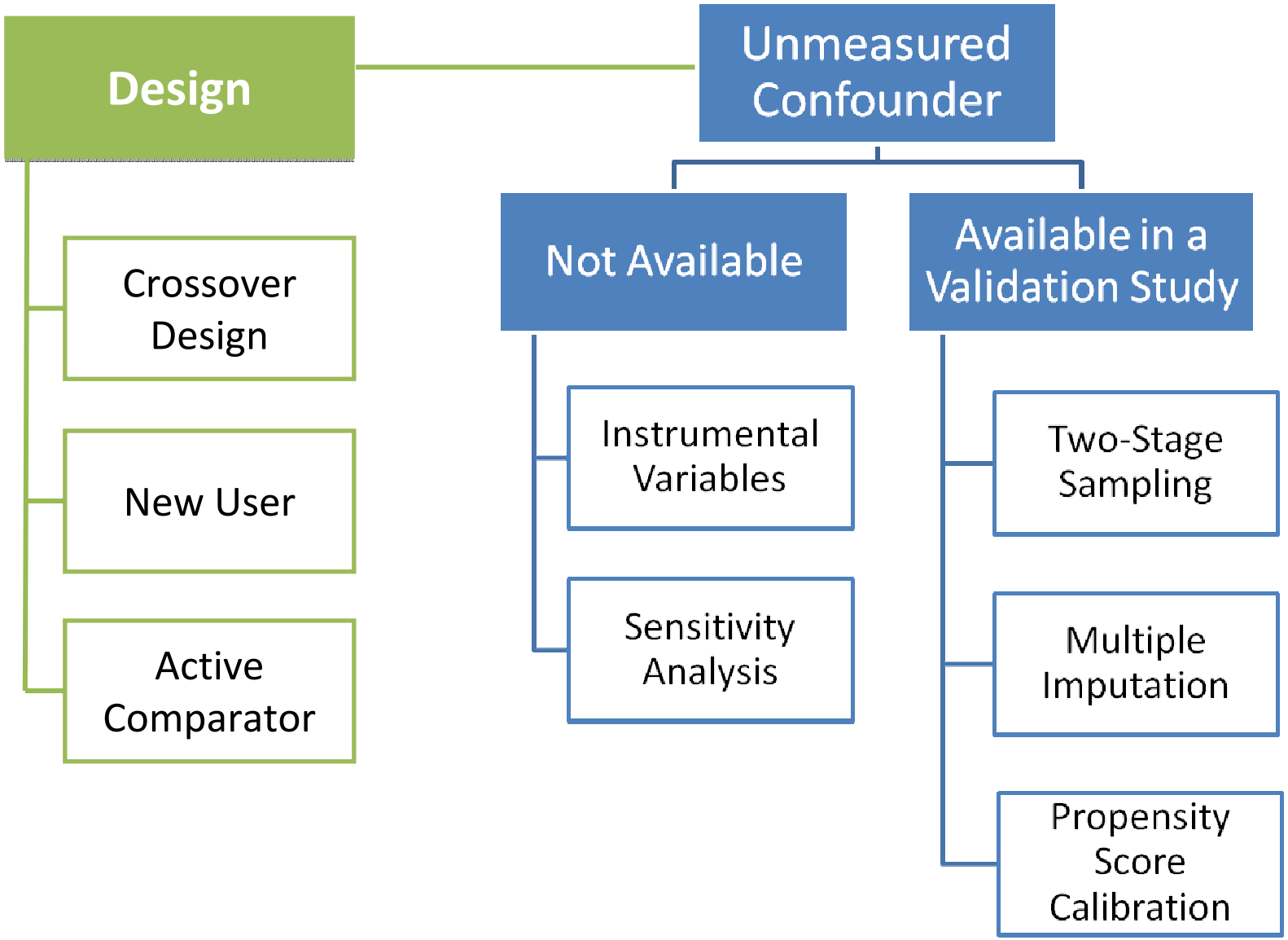
Overview

- Unmeasured confounding in PE claims data
- Strategies to control for unmeasured confounding
- Propensity Score Calibration
  - Propensity Score
  - Propensity Score Calibration
- Example 1: Metformin and Breast Cancer Risk
- Example 2: NSAIDs and 1-year Mortality

Unmeasured Confounding in PE Claims Data

- Unmeasured Confounding is often of concern in observational study using claims data
  - Lack information on BMI, smoking status, alcohol use, over-the-counter medication, or laboratory data
  - These variables are likely to be risk factors for disease (outcome), and may lead to selective prescribing of drugs.
- Examples:
  - NSAIDs and 1-year Mortality: Frailty
  - Benefits of Lipid-lowering agents on all-cause mortality: Frailty
  - Selective COX2 inhibitor and Myocardial Infarction: BMI, Smoking Status, etc

Strategies to Control for Unmeasured Confounders in PE



Internal vs External Validation Study

TABLE 1. Classification of Validation Studies for External Adjustment of Confounding Unmeasured in the Main Study and General Notions About Their Availability, Possible Use of Different Analytic Strategies, and Possible Use for Multiple Associations

	Validation Study		
	Internal	External	
Information on disease outcome	Yes	Yes	No
Availability	Rare*	Rare	Frequent
Analytic methods			
Multiple imputation	Yes†	Yes†	No
2-stage sampling	Yes	No	No
Propensity score calibration	Yes	Yes	Yes
Validation study can be used to adjust multiple associations	Yes (if random sample)	Yes	Yes
Transportability of parameters‡	Yes	—	—

Ref. Stürmer, T., etc(2007). Adjustments for unmeasured confounders in pharmacoepidemiologic database studies using external information. *Medical care*, *45*(10 SUPL), S158.

Data Availability

- Main Study: Claims
  - e.g., Medicare Beneficiaries, Taiwanese National Health Insurance
- External Validation Study:
  - Survey data (e.g., the Medicare Current Beneficiary Survey) or Electronic Medical Records (EMR)
  - Disadvantage: no exactly the same variables; representative of the main study, usually lack of outcome information

Variables	Main Study (Claims)	Validation Study (Claims+ Survey/EMR)
Exposure (Prescription)	X	X
Outcome (Disease)	X	?
Demographics	X	X
Clinical Variables	X	X
Unmeasured Variables: e.g., BMI, Smoking, alcohol, OTC use, etc.	--	X

Propensity Score Calibration (PSC)

- In external validation study:
  - Cross-sectional Study
  - Get information on joint confounding by unmeasured covariates (≥ 1)
- **Propensity Score Calibration =**

Propensity Score

+

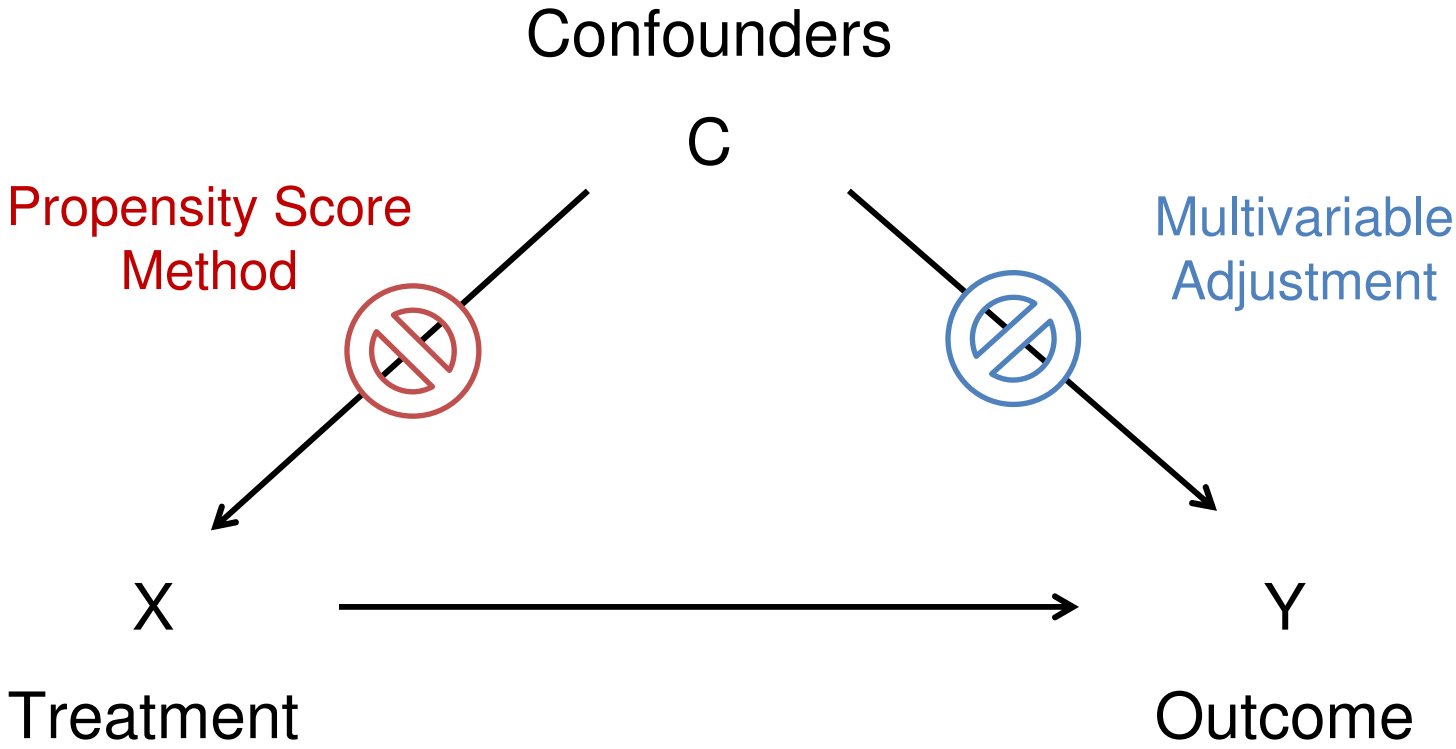
Regression Calibration

(Single Imputation)

Propensity Score vs Multivariable Adjustment

- Propensity score (PS) is the probability of receiving treatment given C.

$$PS = \Pr(X = 1|C)$$



Propensity Score

- Estimating PS:

$$\Pr(X = 1|C) = logit(\beta_0 + \beta_1 C_1 + \beta_2 C_2 + \cdots + \beta_n C_n)$$

- Variable Selection:
  - Confounders
  - Risk factors for outcome
  - Calendar year/Healthcare use
- Using PS: Matching
  - Treated and untreated patients paired on the same PS
  - Among patients with the same PS, treatment is effectively randomized

3 Steps of Propensity Score Calibration (PSC)

STEP 1: In Main Study

- **Error-Prone PS ($PS_{EP, Main}$):** Based on variables available from claims

$$PS_{EP,Main} = \Pr(X = 1|C_1, C_2,\ldots C_n)$$

STEP 2: In Validation Study:

- **Error-Prone PS ($PS_{EP, MCBS}$):** Based on variables available from claims

$$PS_{EP,Validation} = \Pr(X = 1|C_1, C_2,\ldots C_n)$$

- **Gold-Standard PS ($PS_{GS, MCBS}$):** Based on variables available from claims and **additional variables (e.g., BMI)** from survey/EMR

$$PS_{GS,Validation} = \Pr(X = 1|C_1, C_2,\ldots C_n, C_{BMI})$$

- Linear Regression Model:

$$E[PS_{GS,Validation}|X, PS_{EP,Validation}] = \beta_0 + \beta_1 X + \beta_2 PS_{EP,Validation}$$

3 Steps of Propensity Score Calibration (PSC)

From Step 2:

$$E\left[PS_{GS,Validation}\middle|X, PS_{EP,Validation}\right] = \boldsymbol{\beta_0} + \boldsymbol{\beta_1}(X) + \boldsymbol{\beta_2}(PS_{EP,Validation})$$

STEP 3: In Medicare Sample (Main study):

- Impute $PS_{GS,\ Medicare}$

$$PS_{GS,\ Main} = \boldsymbol{\beta_0} + +\boldsymbol{\beta_2}(X) + \boldsymbol{\beta_1}(PS_{EP,Main})$$

- Control confounding by imputed $PS_{GS,\ Main}$

Surrogacy Assumption

- PS_{EP} independent of Y given PS_{GS} (and Treatment)

- If validation study contains information on disease-outcome, we can test for surrogacy. But then we can also use i.e., multiple imputation

Simulation Study on Surrogacy Assumption

Parameter† varied and value	Direction of confounding		Median OR_{AY}‡		Surrogacy		Propensity score calibration		
	X_1, X_2	C	Crude	Adjusted for X_1, X_2	p value from likelihood ratio test§	% of variance explained¶	OR_{AY} Median	Bias reduction# (%)	Coverage** (%)
OR_{CA}									
0.5	↓	↓	*0.49*	*0.65*	*0.5*	*90.9*	*1.00*	*86*	*96.5*
1.0	↓	→	0.74	1.00	0.05	43.6	1.00	—††	96.1
2.0	↓	↑	1.13	1.54	0.01	8.6	2.09	−75	NA
OR_{CY}									
0.5	↓	↑	1.14	1.54	0.01	8.2	2.12	−76	NA
1.0	↓	→	0.75	1.00	0.05	39.7	1.50	—	NA
2.0	↓	↓	*0.49*	*0.65*	*0.5*	*90.9*	*1.00*	*86*	*96.5*

Stürmer, T., Schneeweiss, S., Rothman, K. J., Avorn, J., & Glynn, R. J. (2007). Performance of propensity score calibration—a simulation study. *American journal of epidemiology*, *165*(10), 1110-1118.

Summary

- PSC results largely depend on surrogacy
- If surrogacy holds, major bias reduction (nominal coverage dependent on rare disease)
- If surrogacy violated, PSC may increase rather than decrease bias
- Direction of confounding by unobserved variable(s) differs from direction of observed confounding
- In validation studies with sufficient information on disease violations of surrogacy can be detected

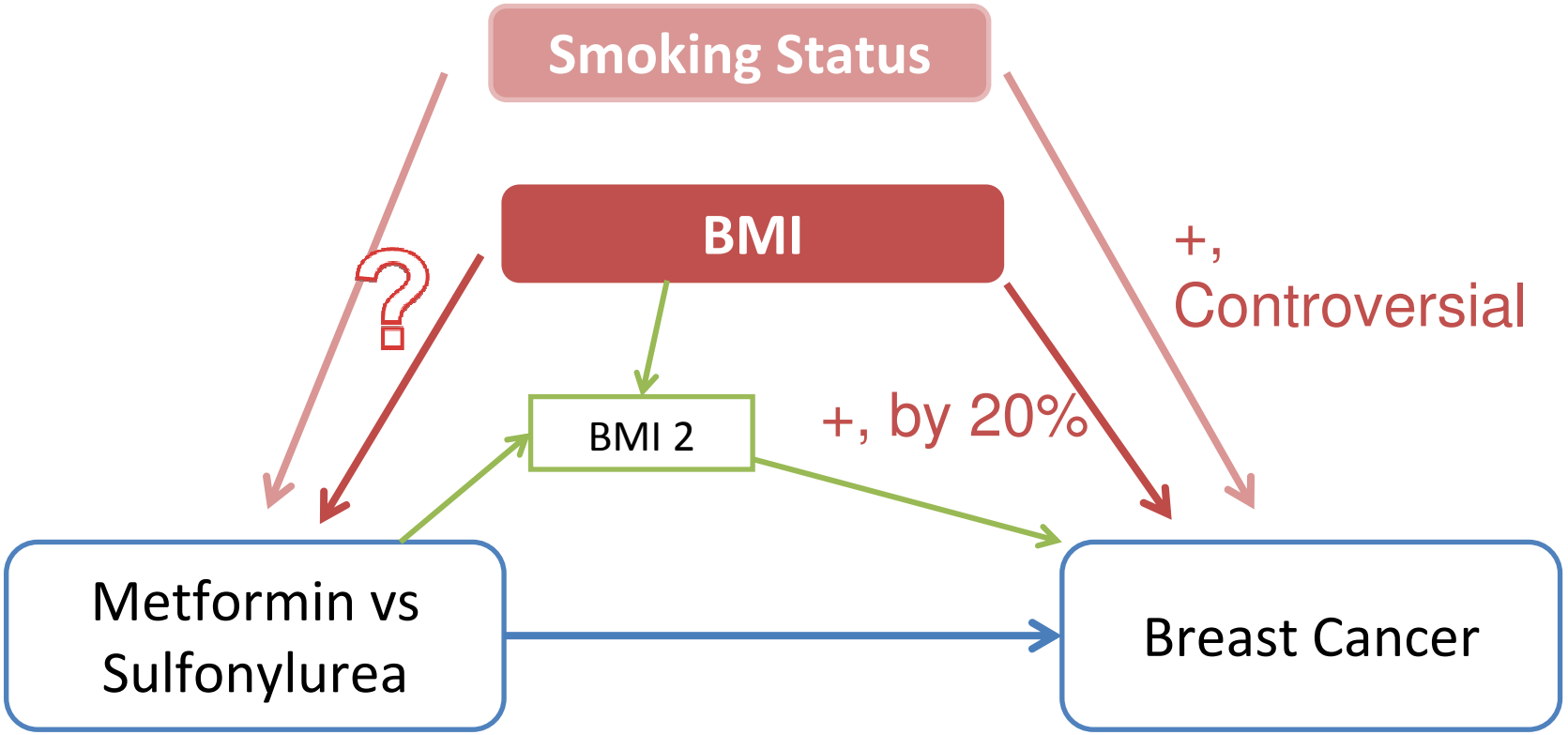
Example 1:

Unmeasured Confounding by BMI and Smoking Status on the Association Between Metformin and Breast Cancer Risk

THE UNIVERSITY of NORTH CAROLINA at CHAPEL HILL

Metformin and Breast Cancer Risk

- Unmeasured confounding by body mass index (BMI) and smoking status on association between metformin and breast cancer risk



External Validation for BMI & Smoking

- Medicare Current Beneficiary Survey (MCBS) Data:
  - Questionnaire: BMI and smoking status
  - Claims: Medicare Part ABD data available for the years of survey

Variables	Medicare (Claims)	MCBS (Claims+ Survey)
Exposure (Prescription)	X	X
Outcome (Disease)	X	--
Demographics	X	X
Disease	X	X
Medication	X	X
HealthCare Use	X	X
BMI	--	X
Smoking Status	--	X

- External control for unmeasured confounding by BMI and Smoking status with Propensity Score Calibration

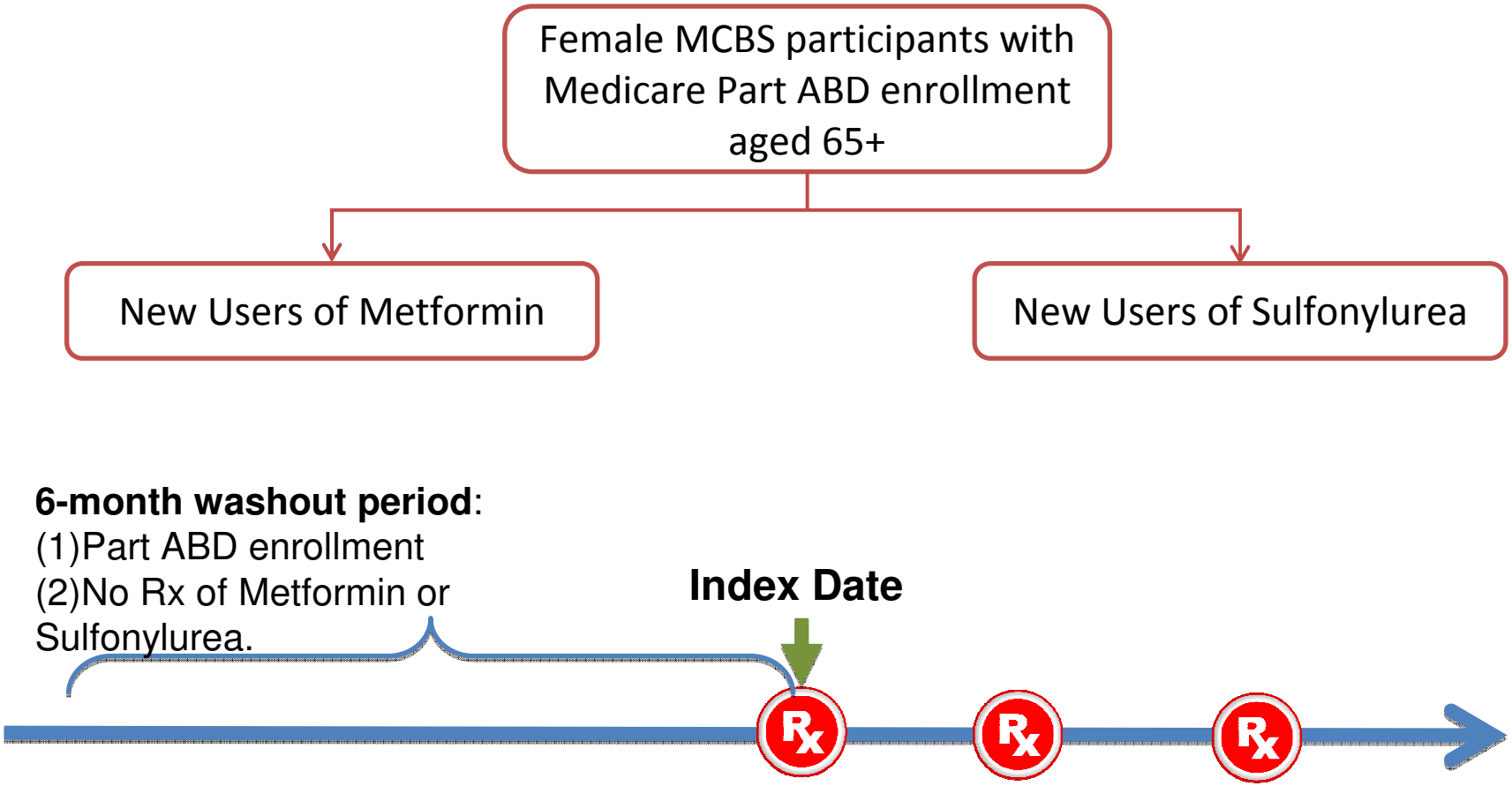
Study Design: Main Study

- To examine the risk of breast cancer after metformin initiation compared with sulfonylurea initiation in older women
  - Cohort study: female new user of metformin versus sulfonylureas who were aged 65+ years
  - Data from fee-for-service Medicare Beneficiaries 2007-2011
    - 20% random sample from Medicare Beneficiaries
  - Outcome: incident breast cancer, including *in situ* and invasive breast cancer
  - Using propensity score method to control for confounding

Study Design: External Validation

- To examine the impact of body mass index (BMI) and smoking status on physicians' choice of metformin versus sulfonylurea
  - A Cross-Sectional Study
  - Data from MCBS 2006-2009
  - Exposure: BMI and smoking status
  - Outcome: Initiation of metformin vs sulfonylurea
- If BMI and smoking are not balanced, propensity score calibration (PSC) is used to adjust unmeasured confounding by BMI and smoking externally in the Medicare data

Study Population



Statistical Analysis

- **Body Mass Index (BMI):**

$$weight\ (kg) \Big/ [height\ (m)]^2$$

  - Continuous Variable: 1 unit increase
  - Categorical Variable: (1) <25 as normal; (2) ≥25 and <30 as overweight; (3) ≥30 and <35 as obese; and (4) ≥35 as morbidly obese.

- **Self-reported Smoking Status**
  - Questionnaire: Ever smoked (Yes/No); Current smoker (Yes/No)
  - Binary Variable: Never, and Ever Smoker

- The prevalence of BMI and smoking status calculated for initiators of metformin vs sulfonylurea, respectively
- A logistic regression model used to estimate odds ratio (OR) of

External Control for Unmeasured Confounders

- To quantify the extent of potential confounding by BMI and smoking status, **independent of other covariates**
  - A logistic regression used to estimate OR
  - Fitting a propensity score model equivalent to the one in the main study (Project 1; as similar as possible).
- Propensity Score Calibration (PSC) used to further correct the effect estimates in the main study

Results: Baseline Characteristics

	MET	SUL	Crude OR
Total	118 (100.0)	79 (100.0)	
Median Age (IQR)	74.0 (70.0-80.0)	78.0 (75.0-84.0)	0.92 (0.88- 0.96)
Race			
White	89 (75.4)	59 (74.7)	1.04 (0.54-2.01)
Other	29 (24.6)	20 (25.3)	1.00
Median of BMI (IQR)	29.9 (25.6-34.0)	28.6 (25.1-33.1)	--
Mean of BMI (Stdev)	30.5 (6.5)	29.9 (6.9)	1.01 (0.97-1.06)
BMI Category			
<25	24 (20.3)	18 (22.8)	1.00
25 to <30	35 (29.7)	30 (38.0)	0.87 (0.40-1.91)
30+	58 (49.2)	29 (36.7)	1.50 (0.70-3.20)
Smoking Status			
Never	61 (51.7)	48 (60.8)	1.00
Ever Smoking	57 (48.3)	28 (35.4)	1.60 (0.89-2.89)

Results: Propensity Score Calibration

	Main Study	Validation Study	
	Error-Prone PS	Error-Prone PS	Gold Standard PS
	OR (95% CI)	OR (95% CI)	OR (95% CI)
Age (1 year)	0.95 (0.95, 0.95)	0.93 (0.88, 0.97)	0.94 (0.89, 0.99)
Race			
White vs others	1.19 (1.13, 1.25)	0.90 (0.43, 1.87)	0.85 (0.40, 1.82)
Congestive Heart Failure	0.79 (0.73, 0.85)	0.41 (0.13, 1.28)	0.43 (0.14, 1.33)
Ischemic Heart Disease	0.78 (0.74, 0.83)	1.43 (0.54, 3.80)	1.25 (0.46, 3.38)
Hypertension	1.12 (1.06, 1.18)	0.92 (0.44, 1.94)	0.93 (0.43, 1.99)
Loop Diuretics	0.71 (0.67, 0.75)	0.66 (0.30, 1.41)	0.66 (0.29, 1.47)
Beta Blockers	0.99 (0.95, 1.04)	1.31 (0.63, 2.69)	1.30 (0.62, 2.75)
Estrogen	1.25 (1.13, 1.39)	0.87 (0.20, 3.79)	0.90 (0.21, 3.93)
statins	1.40 (1.34, 1.46)	1.46 (0.72, 2.98)	1.57 (0.75, 3.30)
Hospitalization (Yes/No)	0.73 (0.68, 0.78)	0.40 (0.14, 1.15)	0.43 (0.14, 1.31)
Physician Visit (4+ vs <4 per year)	1.34 (1.28, 1.41)	0.82 (0.39, 1.70)	0.83 (0.38, 1.78)
Mammogram	1.29 (1.21, 1.38)	1.96 (0.65, 5.98)	1.95 (0.64, 5.94)
BMI Category			
25-30 vs <25			0.84 (0.34, 2.06)
30+ vs <25			1.27 (0.52, 3.10)
Smoking Status			
Ever vs Never			1.41 (0.72, 2.74)

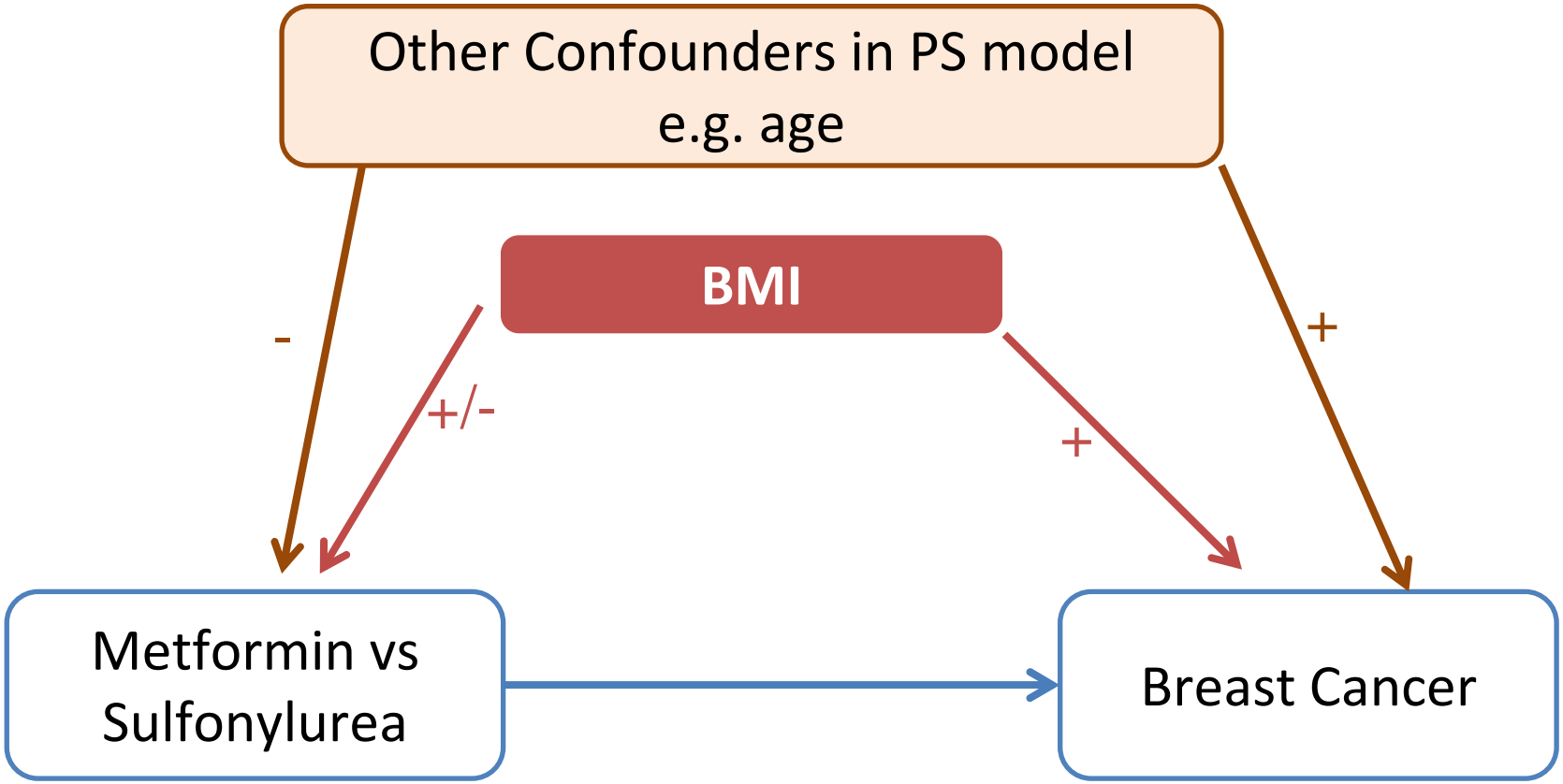
$$PS_{GS} = 0.031363 + 0.950909 \times (PS_{EP}) + 0.01263 \times (Treatment)$$

Results: Propensity Score Calibration

$$PS_{GS} = 0.031363 + 0.950909 \times (PS_{EP}) + 0.01263 \times (Treatment)$$

Cohort (Medicare)	HR (95% CI)		
	Crude	Error-Prone (EP) PS SMR weighted	Gold Standard (GS) PS SMR weighted
AT analysis			
MET	1.04 (0.81-1.34)	1.04 (0.78-1.38)	1.04 (0.79-1.36)
SUL	1.00	1.00	1.00
ITT analysis			
MET	1.04 (0.85-1.27)	1.04 (0.83-1.29)	1.03 (0.83-1.28)
SUL	1.00	1.00	1.00

Violation on Surrogacy Assumption?



Summary

- Our results showed that BMI and ever exposure to smoking was associated with a higher probability of receiving metformin than sulfonylurea, but a little association of BMI/Smoking-Metformin after controlling for other variables.
  - Thus, a little residual confounding by BMI and Smoking in the main study
- After implementing PSC, we still observed no association between metformin and breast cancer risk.

Example 2:

Unmeasured Confounding by Frailty on the Association between NSAIDs and Mortality

Stürmer, T., etc(2007). Adjustments for unmeasured confounders in pharmacoepidemiologic database studies using external information. *Medical care*, *45*(10 SUPL), S158.

THE UNIVERSITY of NORTH CAROLINA at CHAPEL HILL

NSAIDs and Mortality

- NSAIDs were associated with a strong reduction in risks for short-term mortality (RR=0.74), after controlling for variety of variables.
- No known biological mechanism on NSAIDs-Mortality relationship
- Plausible explanation:
  - Due to selection bias leading to a strong unmeasured confounding by frailty
  - Physicians are less likely to prescribe NSAIDs in frail older adults
    - Measured: age, cardiovascular disease, etc
    - Unmeasured : BMI, smoking, daily activity, etc

Study Design (Main Study)

- Study Cohort (N=103,133):
  - New Jersey residents aged 65+,
  - Filled prescription with Medicaid or the pharmaceutical Assistance to the Aged and Disabled program
  - hospitalized in 1995 to 1997
- Exposure: NSAIDs use in prior 4 month before the date of hospital admission
- Outcome: All-Cause Mortality in 12 months from the date of hospital admission

External Validation Study

- Using data from the Medicare Current Beneficiary Survey (MCBS)
  - Randomly selected 5,108 MCBS participants who were aged 65+, matched by the age and sex distribution in the main cohort
  - Variables were extracted from survey and the linked Medicare data
    - Medication data was assess based on survey (Prescription program started from 2006)

	Main Study		Validation Study			
	"Error-Prone"		"Error-Prone"		"Gold-Standard"	
	OR*	95% CI*	OR*	95% CI*	OR*	95% CI*
Age (1 yr)	0.98	0.98–0.99	0.98	0.97–1.00	0.98	0.97–1.00
Female	1.2	1.2–1.3	1.2	1.0–1.5	1.1	0.9–1.4
Race						
Black	1.6	1.5–1.7	1.4	1.1–1.9	1.2	0.9–1.7
Other	2.0	1.9–2.2	1.5	1.0–2.2	1.6	1.1–2.5
Diagnoses based on claims data						
Myocardial infarction	0.9	0.8–0.9	1.1	0.8–1.5	1.1	0.7–1.5
Congestive heart failure	0.9	0.9–1.0	0.9	0.6–1.3	0.8	0.6–1.3
Diabetes	1.0	1.0–1.0	0.9	0.6–1.3	0.7	0.5–1.0
Cancer	0.8	0.8–0.8	0.6	0.4–0.9	0.6	0.4–1.0
Arthritis (RA or OA)	2.1	2.0–2.2	2.4	1.7–3.4	1.8	1.3–2.5
Health care system use						
No. physician visits†	1.3	1.3–1.3	1.1	1.0–1.4	1.1	0.9–1.3
No. hospitalizations†	0.9	0.9–0.9	1.1	0.9–1.3	1.0	0.9–1.2
Medications						
Thiazides	1.3	1.2–1.3	1.6	0.9–2.5	1.5	0.9–2.5
Steroids	1.0	0.9–1.0	1.5	1.2–2.0	1.3	1.0–1.8
Anticoagulants	0.5	0.5–0.6	0.5	0.3–0.8	0.5	0.3–0.7
Body Mass Index (1 kg/m^2)	—	—	—	—	1.05	1.03–1.06
Education‡	—	—	—	—	1.0	0.8–1.1
Income§	—	—	—	—	1.1	1.0–1.2
Smoking						
Current	—	—	—	—	1.0	0.7–1.4
Past	—	—	—	—	1.1	0.9–1.3
Activities of daily living						
Difficulties with¶	—	—	—	—	1.2	1.1–1.3
Unable to perform¶	—	—	—	—	1.2	1.0–1.3

* Table has been modified, not showing all the variables

Results

TABLE 3. Association Between Nonsteroidal Antiinflammatory Drug Use and 1-yr Mortality in a Population-Based Cohort of 103,133 Elderly—Propensity Score Calibration Adjustment Based on Data From 5108 Participants of the Medicare Current Beneficiary Survey as External Cross-Sectional Validation Study

	Hazard Ratio*	95% CI*
Unadjusted model	0.68	0.66–0.71
Conventional multivariate outcome model		
Age and gender adjusted	0.74	0.71–0.77
Fully adjusted†	0.80	0.77–0.83
Propensity score (main study) adjusted†	0.81	0.78–0.84
Propensity score calibration adjusted	1.06	1.00–1.12

PSC Conclusions

- Observational study can include some information on assessing unmeasured confounding, to evaluate the robustness of the findings
- External control for unmeasured confounding using Propensity score calibration depends on
  - Availability of data:
    - EMR and survey provide opportunities
  - Transportability of models (Surrogacy Assumption)

Reference

- Glynn, R. J., Schneeweiss, S., Wang, P. S., Levin, R., & Avorn, J. (2006). Selective prescribing led to overestimation of the benefits of lipid-lowering drugs. Journal of clinical epidemiology, 59(8), 819-828.
- Stürmer, T., Glynn, R. J., Rothman, K. J., Avorn, J., & Schneeweiss, S. (2007). Adjustments for unmeasured confounders in pharmacoepidemiologic database studies using external information. Medical care, 45(10 SUPL), S158.
- Schneeweiss, S. (2006). Sensitivity analysis and external adjustment for unmeasured confounders in epidemiologic database studies of therapeutics. *Pharmacoepidemiology and drug safety*, *15*(5), 291-303.
- Stürmer, T., Schneeweiss, S., Rothman, K. J., Avorn, J., & Glynn, R. J. (2007). Performance of propensity score calibration—a simulation study. *American journal of epidemiology*, *165*(10), 1110-1118.
- Lunt, M., Glynn, R. J., Rothman, K. J., Avorn, J., & Stürmer, T. (2012). Propensity score calibration in the absence of surrogacy. *American journal of epidemiology*, *175*(12), 1294-1302.

THE UNIVERSITY
of NORTH CAROLINA
at CHAPEL HILL