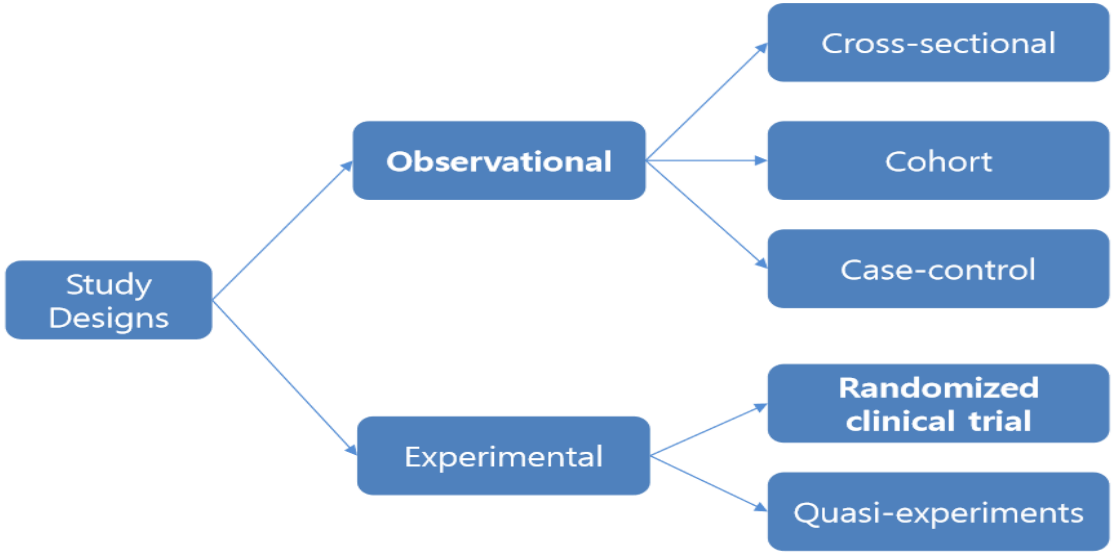
PSM

(propensity score matching)

Type of study designs



Observational study

장점

1.대규모 자료 활용 가능 (예 : 건강보험공단 데이터)

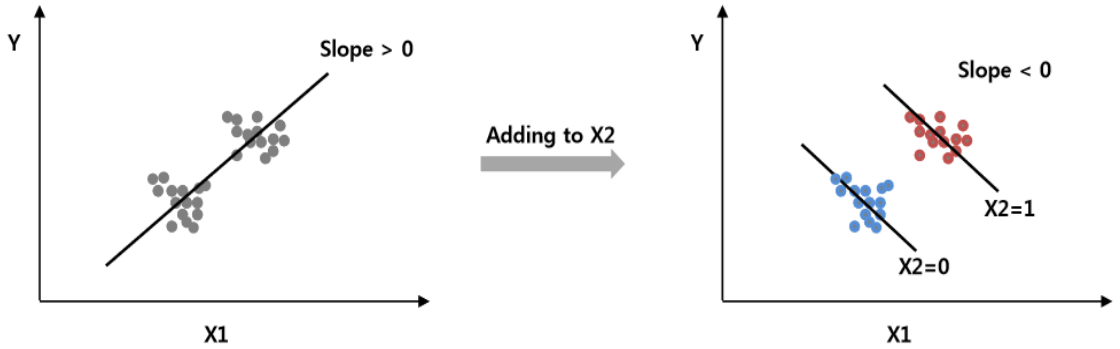
2.Rare event분석할 때 용이

단점

1.그룹 간 비교시 관찰된 다른 공변량들이 outcome에 영향을 끼침

2.특정 그룹이 다수로 관측된 경우 selection bias와 같은 문제 발생

Covariates



→ **Adjusting confounders by modeling**

Adjusting confounders via the multivariate model

공변량을 보정하기 위한 방법

-Treatment/control 그룹 간 비교 뿐 아니라 모형 수립이 연구 목적일 때

단점

1.Sample size 및 outcome 변수의 event 발생율에 따른 모형 내 변수 개수 제한

2.모형 외 demographic table을 통한 covariates보정 후 비교분석 어려움

3.Unmeasured confounders로 인하여 selection bias를 완전히 해결 어려움

Randomized clinical trial (RCT)

장점

1.Outcome에 영향을 미치는 다양한 공변량의 효과를 사전에 제어함

2.Randomization을 통해 selection bias 제어

단점

1.비용적 측면과 윤리적인 비현실적 제약

2.Rare event에 대해서는 어려움이 따름

PS(Propensity score)

Observational study -> RCT 처럼 하기 위한 방법

1.Adjustment in a regression model (PS를 하나의 변수로 보정)

2.Stratification(PS를 grouping 하여 stratified approach 적용)

3.Matching

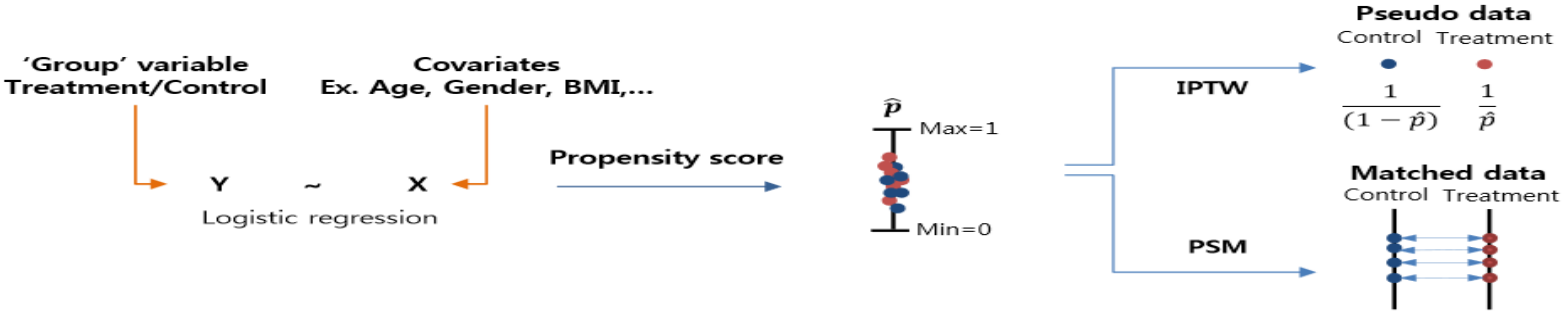
4.Weighting approaches

5.Combinations of the above-mentioned methods

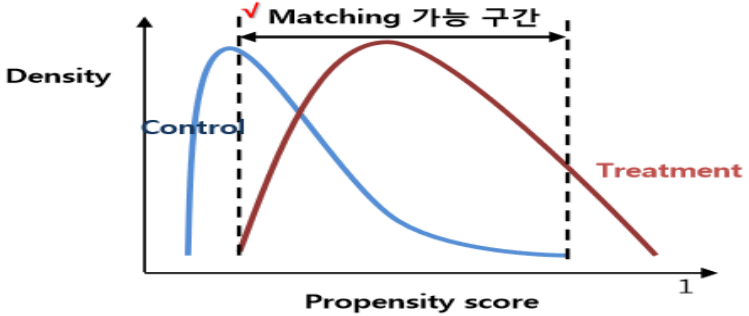
PS(propensity score)

Primary study analysis method	Pros	Cons
Traditional covariate adjustment	• Performed well • Provides prognostic model for outcome of interest	• May not be suitable with many covariates in smaller studies
Propensity score (PS) stratification	• Retains data from all study participants • Opportunity to explore interactions between treatment and PS on outcome risk • Provides effect estimates for every stratum	• Performs less well in datasets with few outcomes, particularly when the number of strata is large • May not account for strong confounding
PS matching	• Reliable; provides excellent covariate balance in most circumstances • Simple to analyze, present and interpret	• Some patients are unmatched leading to information excluded from the analysis • Less precise
PS inverse probability weighting	• Retains data from all study participants • Easy to implement • Creates a pseudo population with perfect covariate balance	• Can be unstable when extreme weights occur
PS covariate adjustment (use of PS as a covariate)	• Performed well	• Adding the PS as an additional covariate produced results very similar (and not necessarily superior) to traditional covariate adjustment

PSM(propensity score matching)



√ 일반적으로는 # of treatment < # of control



√ **Matching type**

1 : 1
1 : 2
⋮
1 : n

√ **Matching method**

stratified
nearest neighbor
radius
caliper
Mahalanobis metric
⋮

⇒ **Select the optimal 'n'**
⇒ **PS 분포가 그룹 간 얼마나 overlap 되는지가 중요!**
⇒ **Treatment의 N이 충분하고**
이와 PS가 중첩되는 control의 수가 충분할 때 Best!

What variables are ‘Matching variables’?

매칭 변수를 선정하는데 통계적인 measure은 없지만 matching을 저해하는 변수들은 제외하는 것이 좋다

- Rubin, D.B. (2001) & Brookhart, M.A. et al. (2006)

  > Outcome과 group 변수에 영향을 미치는 모든 변수들은 matching variables에 포함해야 하지만 결과변수와의 관계가 강한 변수들을 넣은 것이 우선이다.

  > Treatment assignment에 의해 영향을 받는 변수와 해당 변수를 완전히 예측할 수 있는 변수들은 제외하는 것이 좋다.

  > 특히, small study의 경우 group 변수에만 영향이 큰 변수를 포함하게 되면 모형의 bias가 커지기 때문에 주의해야 한다.

PSM (cont'd)

1)Treatment와 control 간의 분포 이질성

Matching 변수 = 2개

	Un-matched				Matched			
	All patients N=557	Treatment N=223	Control N=334	p-value	All patients N=424	Treatment N=212	Control N=212	p-value
Operation				0.014				√ 0.329
Operation1	464(83.3)	174(78)	290(86.8)		332(78.3)	164(77.4)	168(79.2)	
Operation2	34(6.1)	20(9)	14(4.2)		34(8)	20(9.4)	14(6.6)	
Operation3	5(0.9)	4(1.8)	1(0.3)		5(1.2)	4(1.9)	1(0.5)	
Operation4	54(9.7)	25(11.2)	29(8.7)		53(12.5)	24(11.3)	29(13.7)	
Matching variable 1				0.003				√ 0.05
Grade1	364(66.9)	133(62.7)	231(69.6)		261(61.6)	133(62.7)	128(60.4)	
Grade2	149(27.4)	58(27.4)	91(27.4)		132(31.1)	58(27.4)	74(34.9)	
Grade3	31(5.7)	21(9.9)	10(3)		31(7.3)	21(9.9)	10(4.7)	

Matching 변수 = 3개

	Un-matched				Matched			
	All patients N=557	Treatment N=223	Control N=334	p-value	All patients N=346	√ Treatment N=173	Control N=173	p-value
Operation				0.014				0.05
Operation1	464(83.3)	174(78)	290(86.8)		303(87.6)	149(86.1)	154(89)	
Operation2	34(6.1)	20(9)	14(4.2)		34(9.8)	20(11.6)	14(8.1)	
Operation3	5(0.9)	4(1.8)	1(0.3)		0(0)	0(0)	0(0)	
Operation4	54(9.7)	25(11.2)	29(8.7)		9(2.6)	4(2.3)	5(2.9)	
Matching variable 2				<.001				√ <.001
Level1	246(50.6)	55(29.4)	191(63.9)		117(33.8)	50(28.9)	67(38.7)	
Level2	240(49.4)	132(70.6)	108(36.1)		229(66.2)	123(71.1)	106(61.3)	
Matching variable 1				0.002				√ 0.031
Grade1	355(66.5)	125(61.6)	230(69.5)		185(53.5)	97(56.1)	88(50.9)	
Grade2	148(27.7)	57(28.1)	91(27.5)		131(37.9)	56(32.4)	75(43.4)	
Grade3	31(5.8)	21(10.3)	10(3)		30(8.7)	20(11.6)	10(5.8)	

PSM (cont'd)

2)Treatment의 환자 수가 적을 때 -> statistical power 감소

	All patients	Un-matched		p-value	All patients	Matched		p-value
		Treatment	Control		√	Treatment	Control	
	N=391	N=322	N=69		N=138	N=69	N=69	
Matching variable1				0.215				√ >.999
Level 1_1	297(77.1)	238(75.3)	59(85.5)		118(85.5)	59(85.5)	59(85.5)	
Level 1_2	18(4.7)	16(5.1)	2(2.9)		4(2.9)	2(2.9)	2(2.9)	
Level 1_3	70(18.2)	62(19.6)	8(11.6)		16(11.6)	8(11.6)	8(11.6)	
Matching variable2				0.263				√ >.999
Level 2_1	247(65.2)	196(63.2)	51(73.9)		102(73.9)	51(73.9)	51(73.9)	
Level 2_2	107(28.2)	92(29.7)	15(21.7)		30(21.7)	15(21.7)	15(21.7)	
Level 2_3	25(6.6)	22(7.1)	3(4.3)		6(4.3)	3(4.3)	3(4.3)	

PSM (cont'd)

3)Matching으로 인해 N이 줄어들어 특정 변수의 특정 level이 없어지는 현상

	Un-matched				Matched			
	All patients	Treatment	Control	p-value	All patients	Treatment	Control	p-value
	N=391	N=322	N=69		N=138	N=69	N=69	
Matching variable1				0.215				√ >.999
Level 1_1	297(77.1)	238(75.3)	59(85.5)		118(85.5)	59(85.5)	59(85.5)	
Level 1_2	18(4.7)	16(5.1)	2(2.9)		4(2.9)	2(2.9)	2(2.9)	
Level 1_3	70(18.2)	62(19.6)	8(11.6)		16(11.6)	8(11.6)	8(11.6)	
Matching variable2				0.263				√ >.999
Level 2_1	247(65.2)	196(63.2)	51(73.9)		102(73.9)	51(73.9)	51(73.9)	
Level 2_2	107(28.2)	92(29.7)	15(21.7)		30(21.7)	15(21.7)	15(21.7)	
Level 2_3	25(6.6)	22(7.1)	3(4.3)		6(4.3)	3(4.3)	3(4.3)	

4)Matching으로 인해 Outcome의 비율이 극단적으로 편향되어지는 경우

Balance checking for matched data

1)연속형 (d < .1)

$$d = \frac{(\bar{x}_t - \bar{x}_c)}{\sqrt{\frac{{s_t}^2 + {s_c}^2}{2}}}$$

$\bar{x}$ and s^2 denote the sample mean and the sample variance of the covariate

2)이분산형

$$d = \frac{(\hat{p}_t - \hat{p}_c)}{\sqrt{\frac{\hat{p}_t(1-\hat{p}_t) + \hat{p}_c(1-\hat{p}_c)}{2}}}$$

$\hat{p}$ denotes the prevalence of the dichotomous variable.

Balance checking for matched data

3)다범주형

가변수 생성을 통해 이분산형과 동일

범주형 변수 X1 : 1 2 3 -> X1′ : 0 1 / X1″ : 0 1 / X1‴ : 0 1

$$d = \frac{(\hat{p}_t - \hat{p}_c)}{\sqrt{\frac{\hat{p}_t(1-\hat{p}_t) + \hat{p}_c(1-\hat{p}_c)}{2}}}$$

Analysis for matched data

- Comparing groups for statistical differences
  - > Paired t-test (normality condition O) or Wilcoxon signed-rank test (normality condition X)
  - > Stratified chi-square test (Cochran-Mantel-Haenszel method, McNemar test)
  - ** Small-sample로 인해 p-value가 커진다고 matching variables의 balance가 맞았다는 오류는 주의

- Stratified regression
  - > Linear mixed model (LMM), Generalized estimating equation (GEE)

- Survival analysis
  - > Stratified Cox proportional hazard model (stratified log-rank test)
  - > Cox model with the robust sandwich variance estimate

층화변수로 Matched index를 넣어 준다
ex) Sample1, Sample 10 Matchced -> 1
Sample2, Sample 15 Matched -> 2

Analysis for matched data

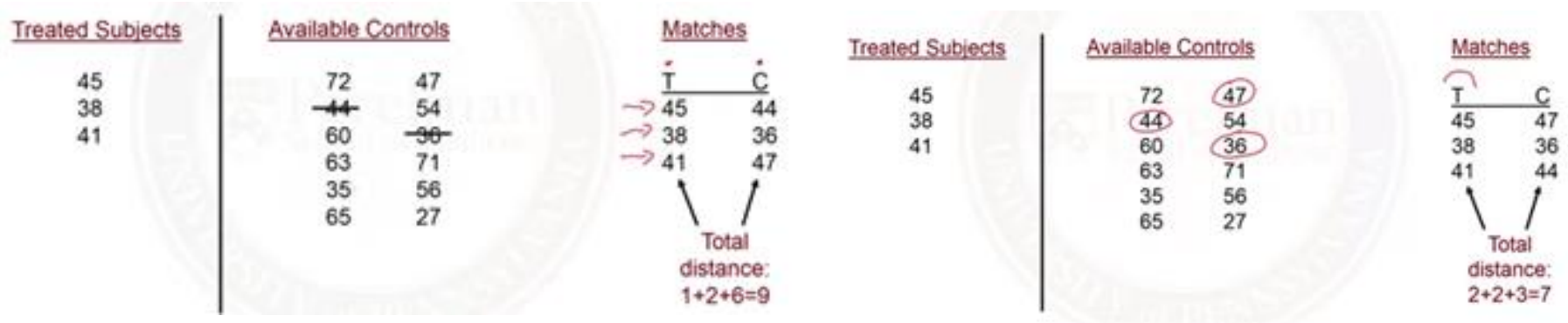
	Un-matched				Matched			
	All patients	Treatment	Control	p-value	All patients	Treatment	Control	p-value
	N=391	N=322	N=69		√ N=138	N=69	N=69	
Matching variable1				0.215				√ >.999
Level 1_1	297(77.1)	238(75.3)	59(85.5)		118(85.5)	59(85.5)	59(85.5)	
Level 1_2	18(4.7)	16(5.1)	2(2.9)		4(2.9)	2(2.9)	2(2.9)	
Level 1_3	70(18.2)	62(19.6)	8(11.6)		16(11.6)	8(11.6)	8(11.6)	
Matching variable2				0.263				√ >.999
Level 2_1	247(65.2)	196(63.2)	51(73.9)		102(73.9)	51(73.9)	51(73.9)	
Level 2_2	107(28.2)	92(29.7)	15(21.7)		30(21.7)	15(21.7)	15(21.7)	
Level 2_3	25(6.6)	22(7.1)	3(4.3)		6(4.3)	3(4.3)	3(4.3)	

Matching techniques

Optimal matching method

-유사한 PS를 가진 대조군과 처치군을 하나의 계층으로 분류

-기준은 아래와 같이 global distance를 줄이는 방향



Matching techniques

Optimal matching method

1:N matching 중 적절한 N을 골라주는 알고리즘 존재

-> 대부분 N = 1

But N > 1 일때

->Treat 와 Matching된 Control군은 반복측정된 값이라는 문제점

->정해진 답이 아직 없음. (현재 문제점)

Matching techniques

[Ann. Surg. Oncol.] Shim, J.H. et. al., 2012.

"Is serum alpha-fetoprotein useful for predicting recurrence..."

Endpoint	Crude		Adjusted		IPTW		Propensity matched (152 pairs)	
	HR[a] (95 % CI)	*P* value	HR (95 % CI)	*P* value	HR (95 % CI)	*P* value	HR (95 % CI)	*P* value
Time to recurrence	0.88 (0.70–1.11)	0.29	0.83 (0.65–1.05)	0.12[b]	0.86 (0.66–1.12)	0.28	0.91 (0.65–1.25)	0.55

HR hazard ratio, *CI* confidence interval, *AFP* alpha-fetoprotein, *IPTW* inverse probability of treatment weighting

[a] Hazard ratios are for the AFP-positive group, compared with the AFP-negative group

[b] Adjusted for variables such as sex, etiology, liver cirrhosis, tumor size, and microvascular invasion that were significant in the univariable analysis

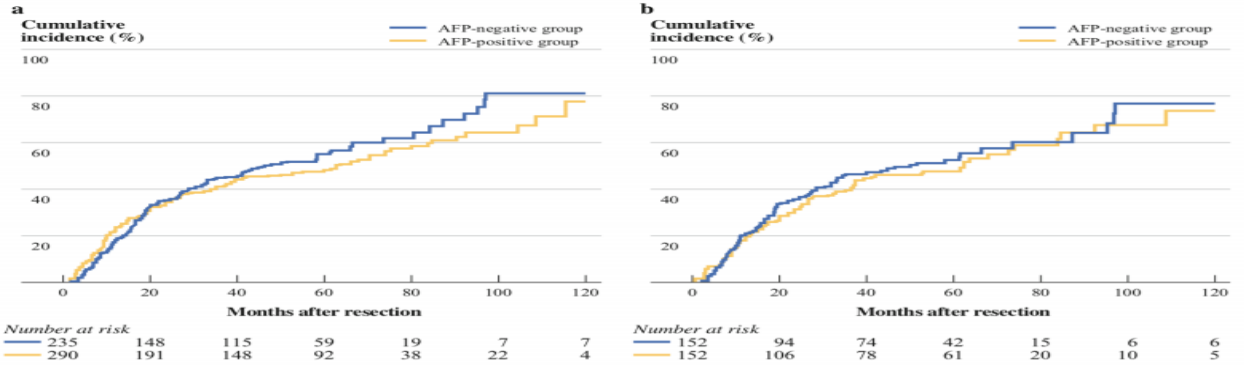
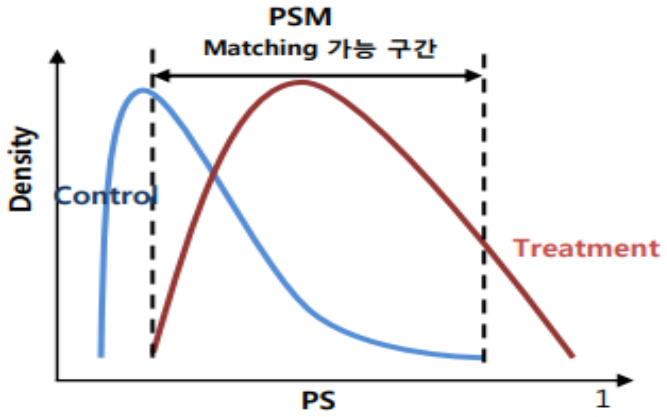
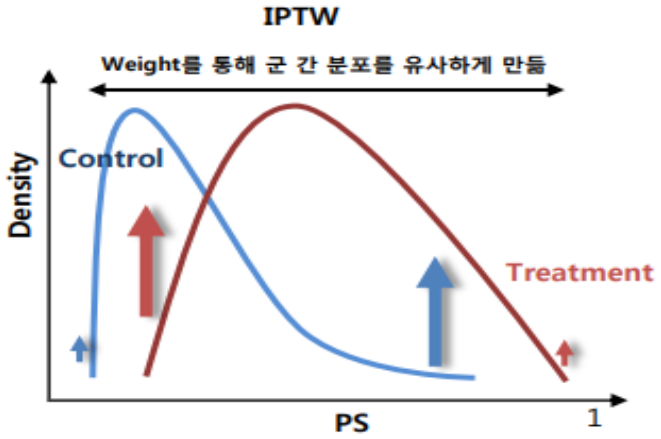


FIG. 1 Kaplan-Meier curves for time-to-recurrence in the AFP-negative and AFP-positive groups involving **a** all patients and **b** propensity-matched patients. Both nonmatched and matched models revealed no significant difference in time-to-recurrence between the 2 groups ($P = 0.29$ and 0.55, respectively)

Inverse Probability of Treatment Weighting(IPTW)



Weighting



$$\arg\min_{\beta} \sum_{i=1}^{m} w_i \left| y_i - \sum_{j=1}^{n} X_{ij}\beta_j \right|^2 = \arg\min_{\beta} \left\| W^{1/2}(\mathbf{y} - X\boldsymbol{\beta}) \right\|^2.$$

⇒ 그룹 간 낮은 density 부분을 높은 weight를 주어 유사한 density로 변경

Inverse Probability of Treatment Weighting(IPTW)

- 매칭 후 가중치를 고려한 최소자승법(weighted least squares; WLS)을 사용한 회귀 모형으로 추정함

- 각 집단에서 극단 값의 비율이 많다면 IPTW는 왜곡된 정보를 도출할 수 있기에 IPTW 세부 방법 중에 극단 값을 보정하는 algorithm을 추가적으로 적용

Matching techniques

[JACC] Elze, M,C. et. al., 2017.

"Comparison of propensity score methods and covariate adjustment"

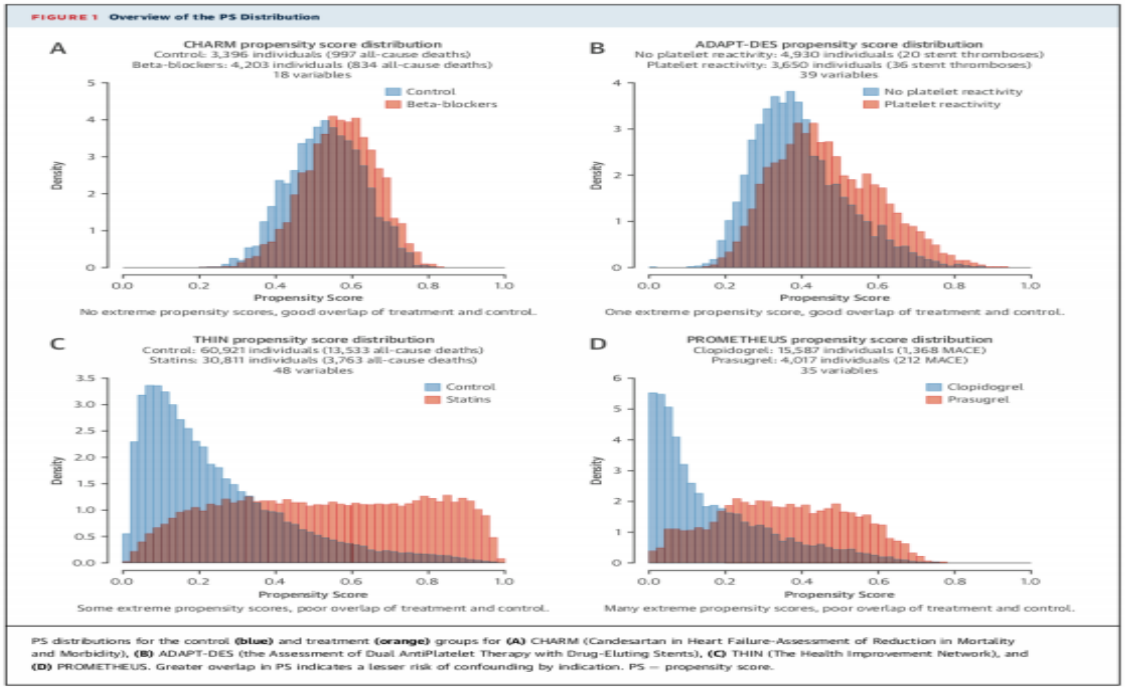


FIGURE 2 Comparison of Hazard Ratios From Different PS Methods and Covariate Adjustment

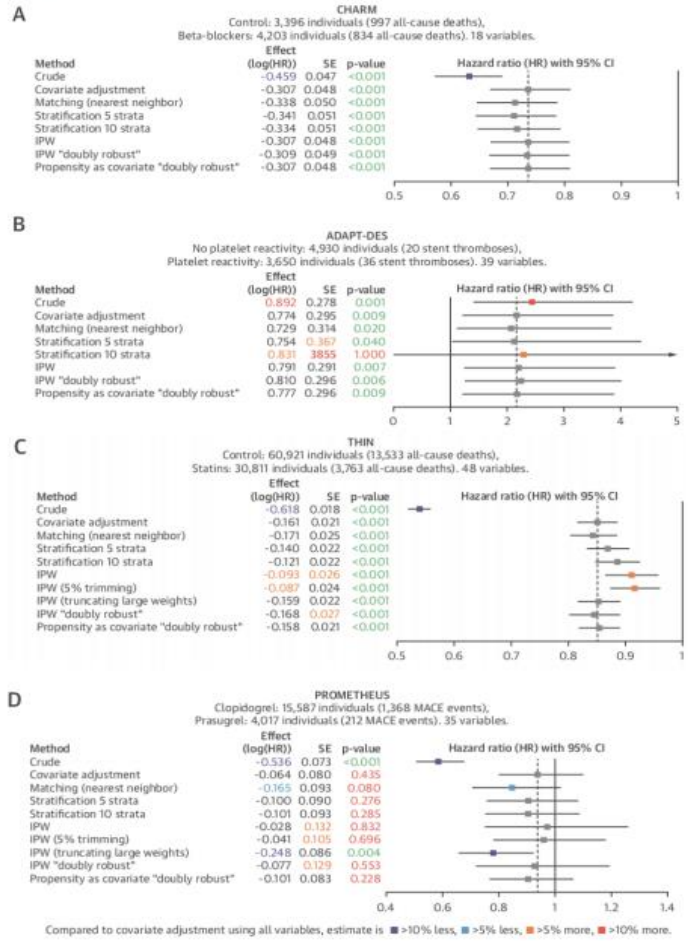
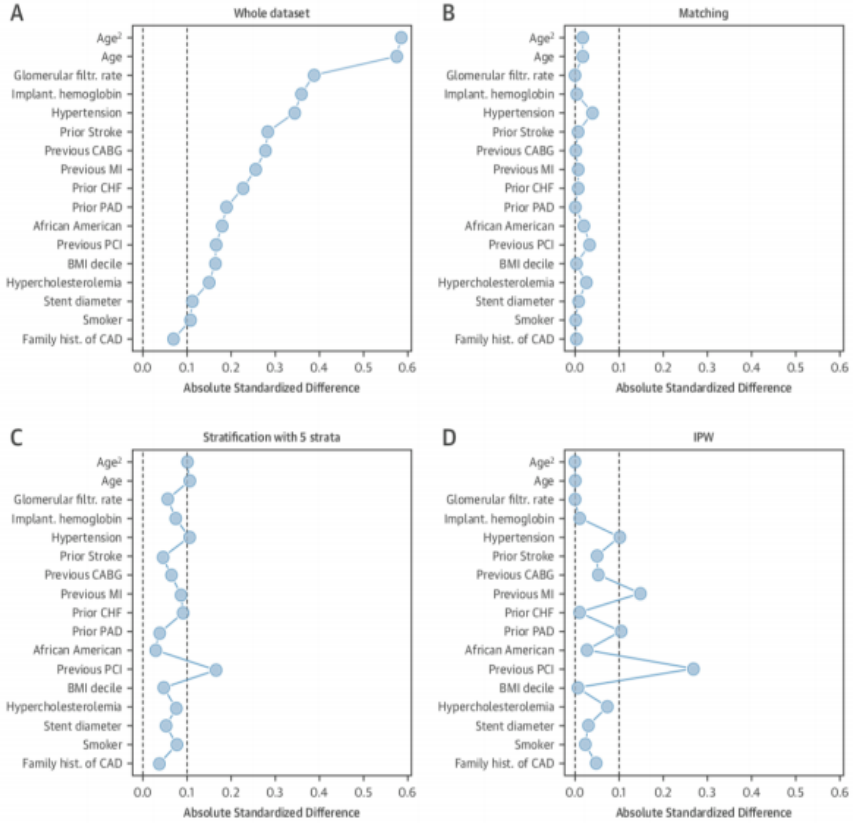


FIGURE 4 Comparison of the Extent of Covariate Imbalance in PROMETHEUS



Comparisons using **(A)** crude comparisons, **(B)** propensity matching, **(C)** propensity stratification, and **(D)** IPW. Graphs show the absolute standardized difference between treatment and control; values <0.1 are conventionally considered acceptable. BMI = body mass index; CABG = coronary artery bypass graft; CAD = coronary artery disease; CHF = congestive heart failure; MI = myocardial infarction; PAD = peripheral arterial disease; PCI = percutaneous coronary intervention; other abbreviations as in Figures 1 to 3.

How to utilizing PSM

Treatment / Control 변수 X1 + Gene expression 변수 X2, X3.....X30000

X1변수만의 Y에 대한 영향력 -> Gene expression 변수 보정

Traditional adjusting의 경우 sample n < p 로 인해 분석이 어렵기 때문에

PSM을 사용

-> 연구중이라고 합니다 .

	Balance checking	Statistical test
PS stratification		CMH / stratified t test
PS matching		1:1 Chi-square / paired t test 1:N ????
IPTW		X