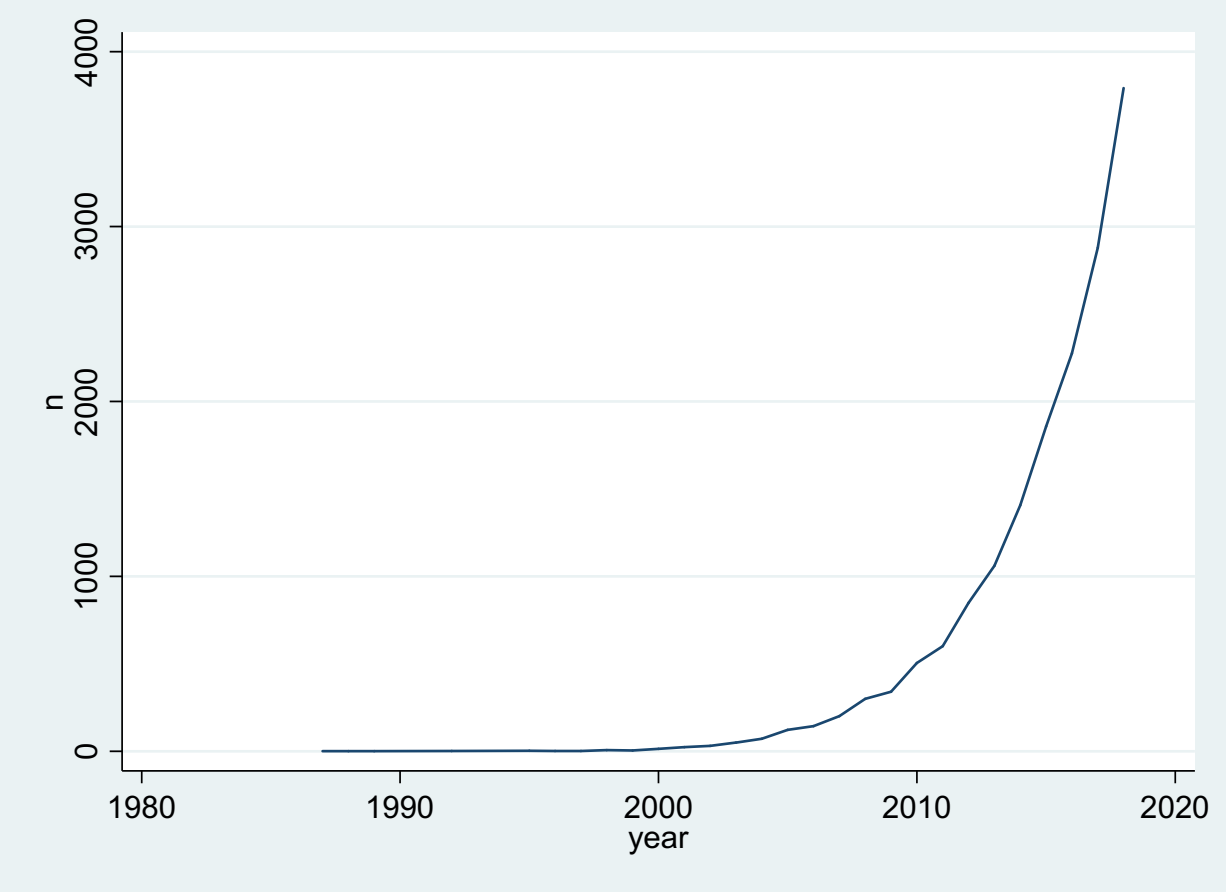
Propensity score methods in pharmacoepidemiology

Jesper Hallas, professor, DrMedSc
Clinical Pharmacology
Odense Universitetshospital/SDU

Disposition

- Re-capturing confounding
- Definition of PS, description
- Properties
- Use of PS
  - Matching, adjustment, stratification, weighting
    - Matching tecniques
  - Diagnostics of covariate balance
  - Scenarios where PS are particularly useful
  - Choice of variables for PS
  - Fringe benefits of PS
  - Trimming
- Disease risc scores
- Exercises

Propensity scores (PS) are popular



Confounding, definition

Bias generated by the lack of comparability between exposed and non-exposed

Formal definition

1. Independent determinant of the outcome
2. Associated with the exposure (positive or negative association)
3. Not caused by the exposure (Not in the causal pathway)

Exercise: Guess the confounder

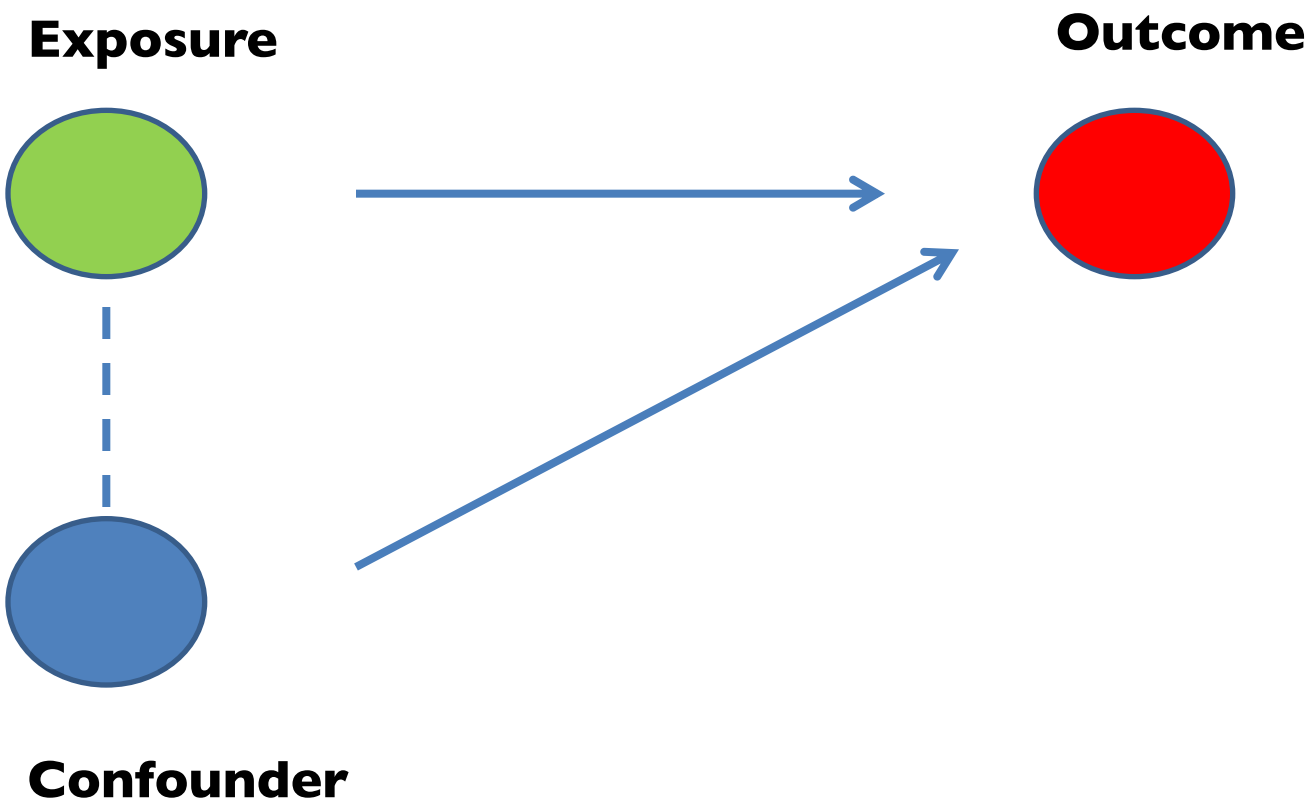
Persons using a bra have breast cancer more frequently than others

Persons with a high alcohol consumption have an increased risk of lung cancer

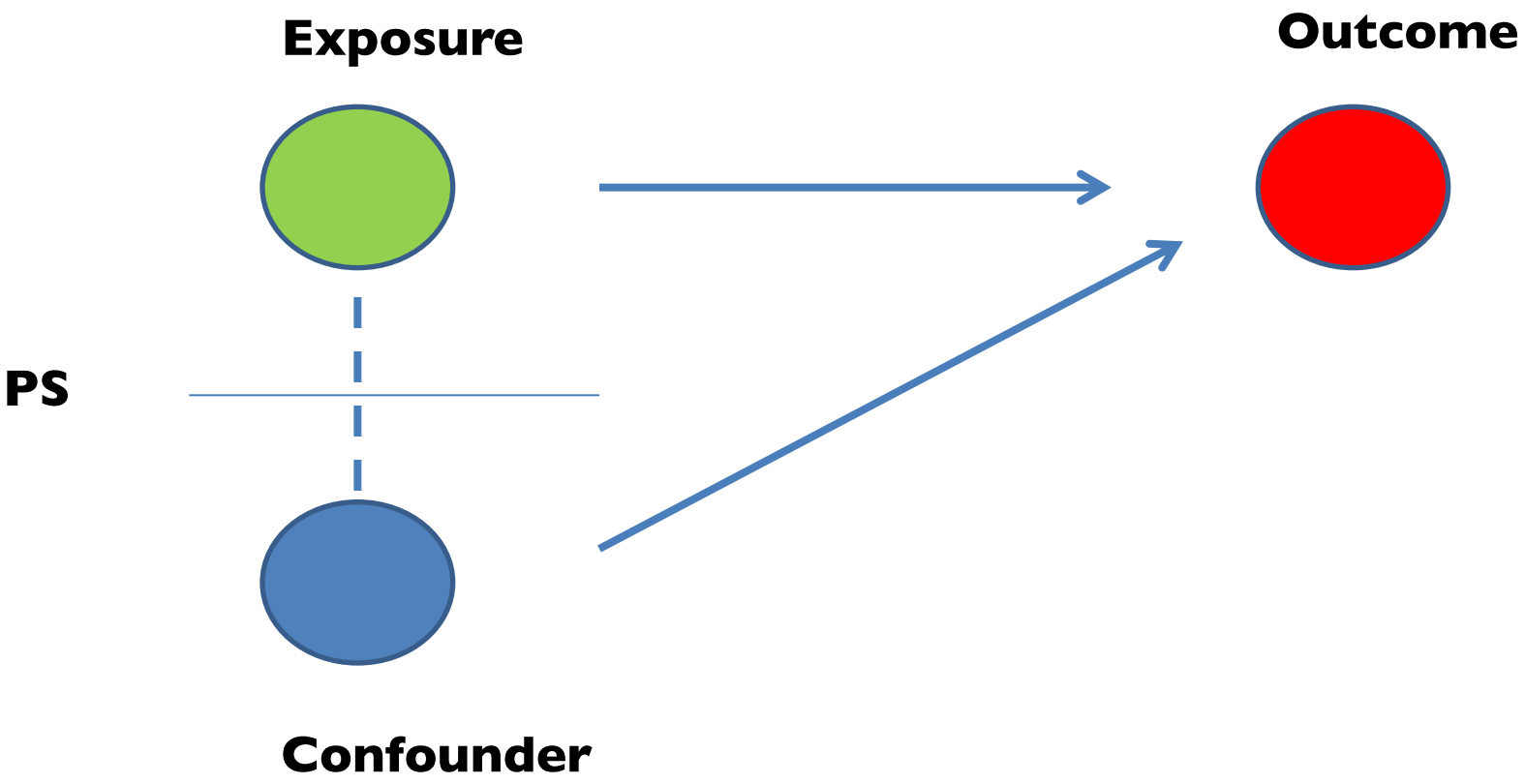
Users of diet products have fewer hip fractures than non-users of the same age.

Users of low-dose ASA have AMI more frequently than others.

Confounding



Confounding



PS, description

- Multivariat modeling of the probability (propensity) of being treated, e.g. the probability of being a statin user.
- Entails a comparison of persons with the same PS, where some are treated and some are not.
- Adjusts for the the variables included in the PS model and not for variables that are not included.

Probability of being a statin users depends on:

- Age
- Sex
- Diabetes dx
- Antidiabetics rx
- Renal disease
- MI
- Stroke
- PAD
- Smoking
- Hypertension dx
- Antihypertensives rx
- …
- ….

PS matching

- Generates ”pairs” with the same PS, where one is treated and the other is not.
- Achieves (aggregate) balance on all covariates included in the PS model
- Usually a loss of subjects who cannot be matched
- Usually followed by a crude analysis, not taking matching or other covariates into account

PS matching, PPI-survival

Table 1. Baseline and Propensity Score–Matched Baseline Characteristics at Inclusion*

Characteristic	Patients Not Receiving Clopidogrel			Propensity Score–Matched Patients Not Receiving Clopidogrel		
	Patients Not Receiving a PPI (n = 22 815)	Patients Receiving a PPI (n = 8889)	P Value	Patients Not Receiving a PPI (n = 8437)	Patients Receiving a PPI (n = 8437)	P Value
Mean age (SD), y	70.3 (13.4)	73.3 (12.4)	<0.001	73.2 (12.8)	73.1 (12.5)	0.72
Men	13 811 (60.5)	4739 (53.3)	<0.001	4535 (53.7)	4516 (53.5)	0.77
Income group			<0.001			0.25
0	3630 (15.9)	1482 (16.7)		1527 (18.1)	1405 (16.7)	
1	4914 (21.5)	2316 (26.1)		2145 (25.4)	2178 (25.8)	
2	4717 (20.7)	2108 (23.7)		1966 (23.2)	1986 (23.5)	
3	4741 (20.8)	1725 (19.4)		1569 (18.6)	1650 (19.6)	
4	4813 (21.1)	1258 (14.2)		1230 (14.6)	1218 (14.4)	
Shock	251 (1.1)	196 (2.2)	<0.001	146 (1.7)	163 (1.9)	0.33
Diabetes with complications	1226 (5.4)	606 (6.8)	<0.001	527 (6.3)	565 (6.7)	0.23
Peptic ulcer	160 (0.7)	546 (6.1)	<0.001	160 (1.9)	172 (2.0)	0.52
PCI	2169 (9.5)	727 (8.2)	<0.001	665 (7.9)	713 (8.5)	0.177
Pulmonary edema	303 (1.3)	145 (1.6)	0.04	114 (1.4)	134 (1.6)	0.20
Cerebral vascular disease	1309 (5.7)	709 (8.0)	<0.001	583 (6.9)	646 (7.7)	0.062
Cancer	69 (0.3)	57 (0.6)	<0.001	49 (0.6)	50 (0.6)	0.92
Cardiac dysrhythmias	3094 (13.6)	1295 (14.6)	0.092	1160 (13.8)	1212 (14.4)	0.29
Acute renal failure	200 (0.9)	209 (2.4)	<0.001	147 (1.7)	171 (2.0)	0.174
Chronic renal failure	269 (1.2)	314 (3.5)	<0.001	240 (2.8)	249 (3.0)	0.68
Loop diuretic	10 135 (44.4)	5143 (57.8)	<0.001	4838 (57.3)	4804 (56.9)	0.60
Spironolactone	2665 (11.7)	1324 (14.9)	<0.001	1214 (14.4)	1255 (14.9)	0.37
Aspirin	11 218 (49.2)	3953 (44.5)	<0.001	3704 (43.9)	3826 (45.4)	0.058
Statin	11 492 (50.4)	3970 (44.6)	<0.001	3754 (44.4)	3841 (45.4)	0.178
β-Blocker	16 056 (70.4)	5915 (66.5)	<0.001	5645 (66.9)	5655 (67.0)	0.29
ACE inhibitor	9761 (42.8)	3754 (42.2)	0.79	3535 (41.9)	3597 (42.6)	0.33
Diabetes medication	2868 (12.6)	1301 (14.6)	<0.001	1152 (13.7)	1235 (14.6)	0.067

Charlot, Ann Intern Med 2010

Matching techniques

- Nearest neighbor
- Sequential, balanced nearest neighbor
- greedy

Matched PS analysis

- stcox main_exposure

What to do about covariate imbalance

- ”Eyeball” table 1
- Standardised difference

$$d = \frac{(\overline{X_1} - \overline{X_2})}{\sqrt{\frac{s_1^2 + s_2^2}{2}}}$$

- Where
- d = standardised difference
- $\overline{X}_1$ and $\overline{X}_2$ are means of sample X_1 and X_2
- s_1 and s_2 are standard deviations of sample X_1 and X_2

What to do about covariate imbalance

- If there is imbalance on an essential variable:
  - Include it in parallel with the propensity score
    - ”Doubly robust”
  - Up-weigh it in the regression model that builds the PS
    - logistic exposure 2*variable_of_interest other_variables …..

Critique of PS matching

- For persons with high PS (e.g. >0.50) not enough untreated for matching
  - -> loss of treated sujects
  - -> loss of representativeness

- Solution: matching with replacement

Using PS, as covariate

- Only covariate, apart from main exposure
  - stcox main_exposure PS
- Used in high-dimensional PS
- No loss of subjects

Using PS, as stratification criterion

- For example as deciles of PS spectrum
  - stcox main_exposure, by(PSdecile)
- No loss of subjects

Using PS,as a weighting criterion

- Inverse probability of treatment weighting (IPTW)
- Subjects are weighted by the reciprocal value of the probality of receiving the treatment that they actually receive
  - 1/PS for treated
  - 1/(1-PS) for untreated
- Creates two equally large psedo-cohorts, one treated and one untreated
- Increasingly popular

Properties of IPTW

- Contrafactual:
  - What if we treated everyone in the whole population vs treating none of them? ATE
  - (PS matching : What if hadn't treated those who actually got the treatment?) ATT
- Subjects who are treated contrary to prediction are given very large weights
- Statistically unstable
  - "Stabilizing" e.g., by having an upper limit on the allowed weight

Alternative approaches for confounding adjustment in observational studies using weighting based on the propensity score: a primer for practitioners

Rishi J Desai,[1] Jessica M Franklin[1]

[1]Division of Pharmacoepidemiology and Pharmacoeconomics, Brigham and Women's Hospital and Harvard Medical School, 1620 Tremont Street, Boston, MA 02120, USA

Correspondence to: R J Desai
rdesai@bwh.harvard.edu
(ORCID 0000-0003-0799-7773)

Additional material is published online only. To view please visit the journal online.

Cite this as: *BMJ* 2019;367:l5657
http://dx.doi.org/10.1136/bmj.l5657

Accepted: 5 August 2019

This report aims to provide methodological guidance to help practitioners select the most appropriate weighting method based on propensity scores for their analysis out of many available options (eg, inverse probability treatment weights, standardised mortality ratio weights, fine stratification weights, overlap weights, and matching weights), and outlines recommendations for transparent reporting of studies using weighting based on the propensity scores.

Propensity scores[1] have become a cornerstone of confounding adjustment in observational studies evaluating outcomes of treatment use in routine care. Propensity score based methods target causal inference in observational studies in a manner similar to randomised experiments by facilitating the measurement of differences in outcomes between the treated population and a reference population.[2] Despite the conceptual equivalence between randomised experiments and observational studies using propensity scores, randomised experiments can successfully achieve exchangeability between treated and reference populations with respect to both using observed data based on a statistical model such as a logistic regression model. After estimation, confounding adjustment through conditioning on the propensity scores can be done in many ways, including matching, stratification, adjustment as a regressor, and weighting.[3] Previous research has suggested that the traditional outcome regression model provides generally equivalent confounding adjustment to various propensity score based approaches in cohort studies with a large sample size and sufficient number of outcome events to support multivariable model fit.[4] However, some key advantages of propensity scores, including the ability to clearly define the target population of inference and the ability to identify and exclude patients in atypical circumstances with near zero probability of receiving a certain treatment,[5] have made use of these scores a method of choice for analysing observational data for many researchers.

Matching each treated observation to a fixed number of reference observations if their propensity scores are within a prespecified range (the caliper) has often been the preferred approach of using propensity scores for confounding adjustment.[6] However, this method has an important limitation of discarding unmatched observations falling within the caliper after a pre-specified number of observations are found for each treated observation. More recently, a paradoxical phenomenon of increasing rather than decreasing covariate imbalance after propensity score matching has been described by King and Nielsen.[7] Notably, other methods of using propensity scores in analysis (including stratification, adjustment as a regressor, and weighting) are not affected by this problem

Table 1 | Alternative approaches for weighting based on propensity scores

Method	Weight calculation: Treated patients	Weight calculation: Reference patients	Target of inference (estimand)	Features	Interpretation
Inverse probability of treatment weights	1/PS	1/(1 – PS)	ATE in the whole population	Clear target of inference, which mimics the target of inference from randomised controlled trials, is a strength. However, because the PS is directly used to create weights, extreme weights are commonly observed. Weight trimming is routinely necessary to address extreme weights and prevent variance inflation	ATE estimates can be interpreted as effect of the treatment when the whole study population is treated with the treatment under investigation versus the reference treatment
Fine stratification weights (ATE)	$(N_{\text{total in PS stratum}}/N_{\text{total}}) / (N_{\text{exposed in PS stratum}}/N_{\text{total exposed}})$	$(N_{\text{total in PS stratum}}/N_{\text{total}}) / (N_{\text{reference in PS stratum}}/N_{\text{total reference}})$	ATE in the whole population	Does not use the PS directly to calculate weights; instead, the scores are used to create fine stratums and weights are subsequently calculated to account for stratum membership. As a result, extreme weights due to PSs that are very close to 0 or 1 are unlikely: an important strength in circumstances where exposure prevalence is low. Clear target of inference is another strength	
Standardised mortality ratio weighting	1	PS/(1 – PS)	ATT	Weighting is conducted by the odds in the reference group, can naturally extend to circumstances with >2 treatment arms. Weight trimming might be necessary to address extreme weights and prevent variance inflation. Clear target of inference is a strength	ATT estimates can be interpreted as effect of the treatment when patients receiving treatment in the study population (that is, the exposed group) were treated with the treatment under investigation versus the reference treatment
Fine stratification weights (ATT)	1	$(N_{\text{exposed in PS stratum}}/N_{\text{total exposed}}) / (N_{\text{reference in PS stratum}}/N_{\text{total reference}})$	ATT	Does not use the PS directly to calculate weights; instead, the scores are used to create fine stratums and weights are subsequently calculated to account for stratum membership. As a result, extreme weights due to PSs that are very close to 0 or 1 are unlikely: an important strength in circumstances where exposure prevalence is low. Clear target of inference is another strength	
Matching weights	(Minimum (PS, 1 – PS)) / PS	(Minimum (PS, 1 – PS)) / (1 – PS)	ATE in a subset	Extreme weights are impossible because weights are bound between 0 and 1 by design, eliminating the need for weight trimming. Can naturally extend to circumstances with more than two treatment arms	Target of inference is close to ATE in the whole population when groups are equally sized and PS distributions have good overlap, and is close to the ATT in the smaller group when groups are unequally sized but PS distribution have good overlap. In circumstances of limited overlap in PS distribution, could lead to treatment effect estimation in a subpopulation that does not reflect patients receiving the treatment of interest in routine care or the whole study population
Overlap weights	(1 – PS)	PS	ATE in the overlap population	Extreme weights are impossible because weights are bound between 0 and 1 by design, eliminating the need for weight trimming. Yields exact covariate balance between treated and reference groups by construction	Estimates can be interpreted as ATE when patients with a realistic probability of receiving either treatment were treated with the treatment under investigation versus the reference treatment. The target population in this approach can be described as the overlap population or population with reasonable clinical equipoise for treatment decision. However, this approach could lead to treatment effect estimation in a subpopulation that does not reflect patients receiving the treatment of interest in routine care or the whole study population, especially when PS overlap is limited

ATE=average treatment effect; ATT=average treatment effect among the treated population; PS=propensity score.

Desai & Frankling, BMJ 2019

Few outcomes

- E.g. a large cohort study with a rare outcome (n)
- Number of variables that can be included in a Cox regression model = n/6
  - -> small sample bias
- Not a limitation in PS matching!

Fringe benefits of PS

- Insight into what drives treatment
- Matching excludes persons with absolute indications or contraindications
- Statistically robust if there are few outcomes
- By so-called asymmetrical trimming, there is a certain level of adjustment non-measured confounders

Persons in the tails of PS distribution

TABLE 2. Proportion of deaths among 6,269 ischemic stroke patients registered in a German stroke registry between 2000 and 2001 who were treated or not treated with tissue plasminogen activator, according to percentiles of the propensity score for the entire study population

Percentile	Treated ($n = 212$)				Not treated ($n = 6,057$)				Empirical OR*
	Score†	No.	Deaths		Score†	No.	Deaths		
			No.	%			No.	%	
99 to 100	0.5809	36	3	8.3	0.5474	26	7	26.9	0.25
95 to <99	0.3143	73	13	17.8	0.2912	178	27	15.2	1.21
90 to <95	0.1393	55	8	14.6	0.1363	258	19	7.4	2.14
75 to <90	0.0585	31	3	9.7	0.0459	910	82	9.0	1.08
50 to <75	0.0115	10	4	40.0	0.0084	1,558	87	5.6	11.27
25 to <50	0.0017	5	2	40.0	0.0014	1,561	54	3.5	18.60
10 to <25	0.0004	2	1	50.0	0.000267	940	36	3.8	25.11
5 to <10		0	0	0	0.000066	313	6	1.9	
1 to <5		0	0	0	0.000027	251	8	3.2	
0 to <1		0	0	0	0.000007	62	1	1.6	
Overall	0.2521	212	34	16.0	0.0262	6057	327	5.4	3.35

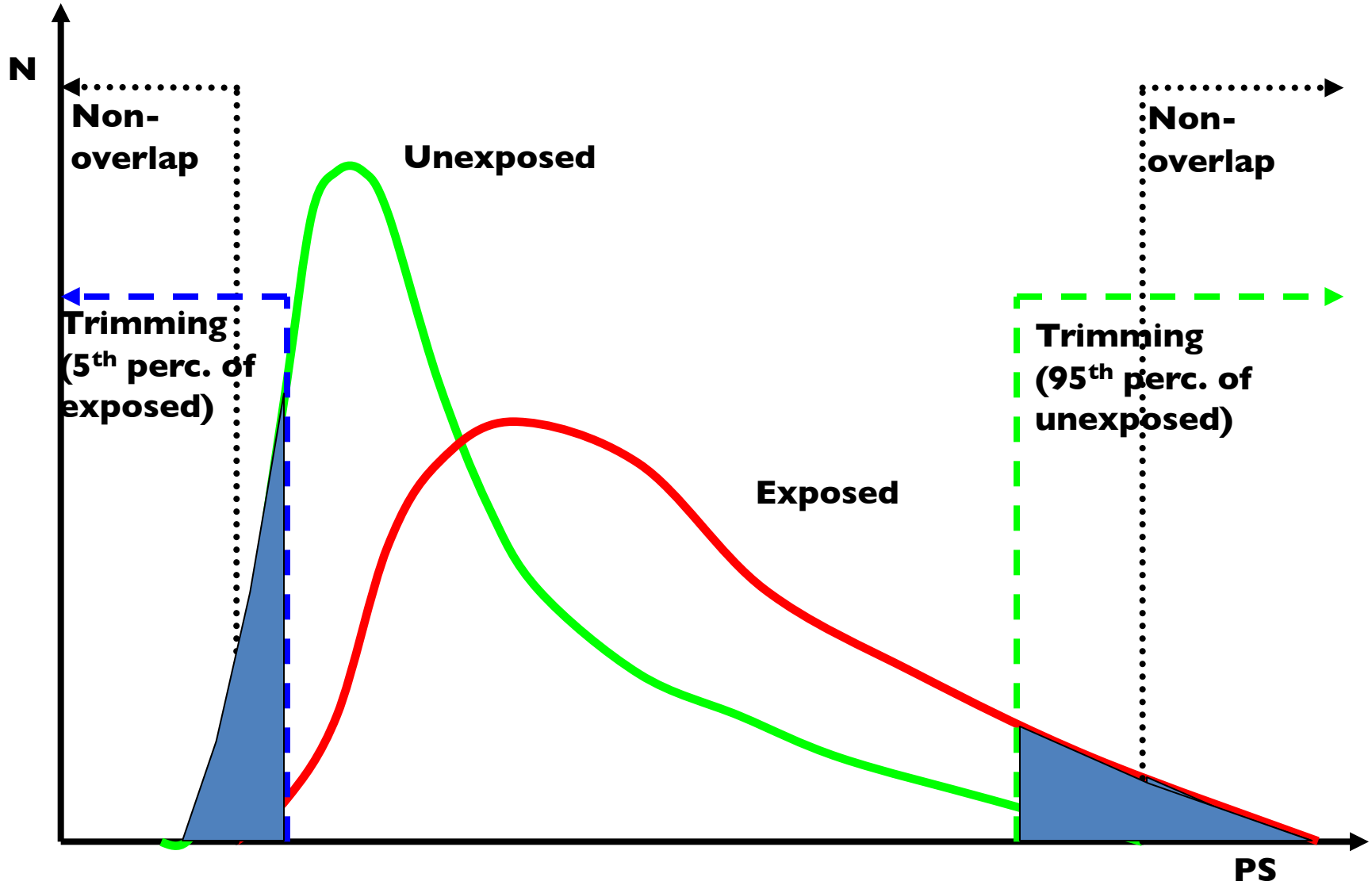
* Propensity-stratum-specific-treatment–mortality odds ratio.
† Mean propensity score in percentile.

Kurth T et al. Am J Epidemiol 2006

Persons in the tails of PS distribution

- Persons with absolute indications or contraindication
- If treated contrary to prediction often last resort treatment (= very bad prognosis)

PS Range Restriction, „Trimming“



Stürmer T

”Asymmetrical” trimming

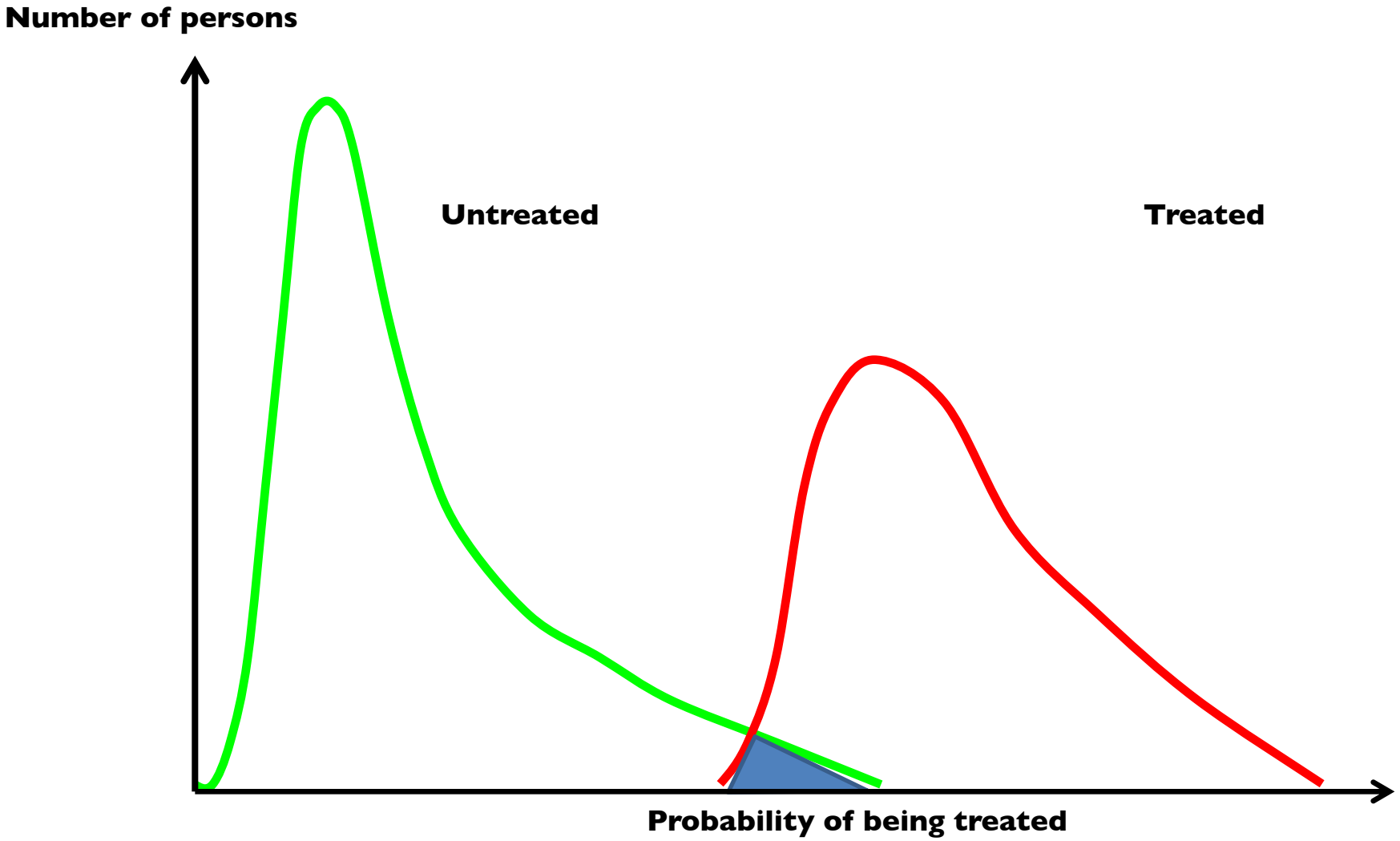
- Removes subjects treated contrary to prediction
- Surprisingly inexpensive in terms of power
- Provides some protection against unmeasured confounders

What variables shold be put into a PS model?

- Risc factors for the outcome, associated with exposure (confounders) Yes -> confounder adjustment
- Risc factors for the outcome, not associated with exposure (risc factors) Yes-> improve precision
- Not risc factors for the outcome, associated with exposure (instruments) No -> lose precision
- Not risc factors for the outcome, not associated with exposure (noise) No -> lose precision

- I.e., all risc factors for outcome

PS model is too"good"



Epidemiology. 2009 July ; 20(4): 512–522. doi:10.1097/EDE.0b013e3181a663cc.

High-dimensional propensity score adjustment in studies of treatment effects using health care claims data

Sebastian Schneeweiss, **Jeremy A. Rassen**, **Robert J. Glynn**, **Jerry Avorn**, **Helen Mogun**, and **M. Alan Brookhart**
Division of Pharmacoepidemiology and Pharmacoeconomics, Department of Medicine Brigham and Women's Hospital/Harvard Medical School

Abstract

Background—Adjusting for large numbers of covariates ascertained from patients' health care claims data may improve control of confounding, as these variables may collectively be proxies for unobserved factors. Here we develop and test an algorithm that empirically identifies candidate covariates, prioritizes covariates, and integrates them into a propensity-score-based confounder adjustment model.

Methods—We developed a multi-step algorithm to implement high-dimensional proxy adjustment in claims data. Steps include (1) identifying data dimensions, e.g. diagnoses, procedures, and medications, (2) empirically identifying candidate covariates, (3) assess recurrence of codes, (4) prioritizing covariates, (5) selecting covariates for adjustment, (6) estimating the exposure propensity score, and (7) estimating an outcome model. This algorithm was tested in Medicare claims data, including a study on the effect of Cox-2 inhibitors on reduced gastric toxicity compared to nonselective nonsteroidal anti-inflammatory drugs (NSAIDs).

Results—In a population of 49,653 new users of Cox-2 inhibitors or nonselective NSAIDs, a crude relative risk (RR) for upper GI toxicity (RR = 1.09 [95% confidence interval = 0.91–1.30]) was initially observed. Adjusting for 15 predefined covariates resulted in a possible gastroprotective effect (0.94[0.78–1.12]). A gastroprotective effect became stronger when adjusting for an additional 500 algorithm-derived covariates (0.88[0.73–1.06]). Results of a study on the effect of statin on reduced mortality were similar. Using the algorithm adjustment confirmed a null finding between influenza vaccination and hip fracture (1.02[0.85–1.21]).

hdPS

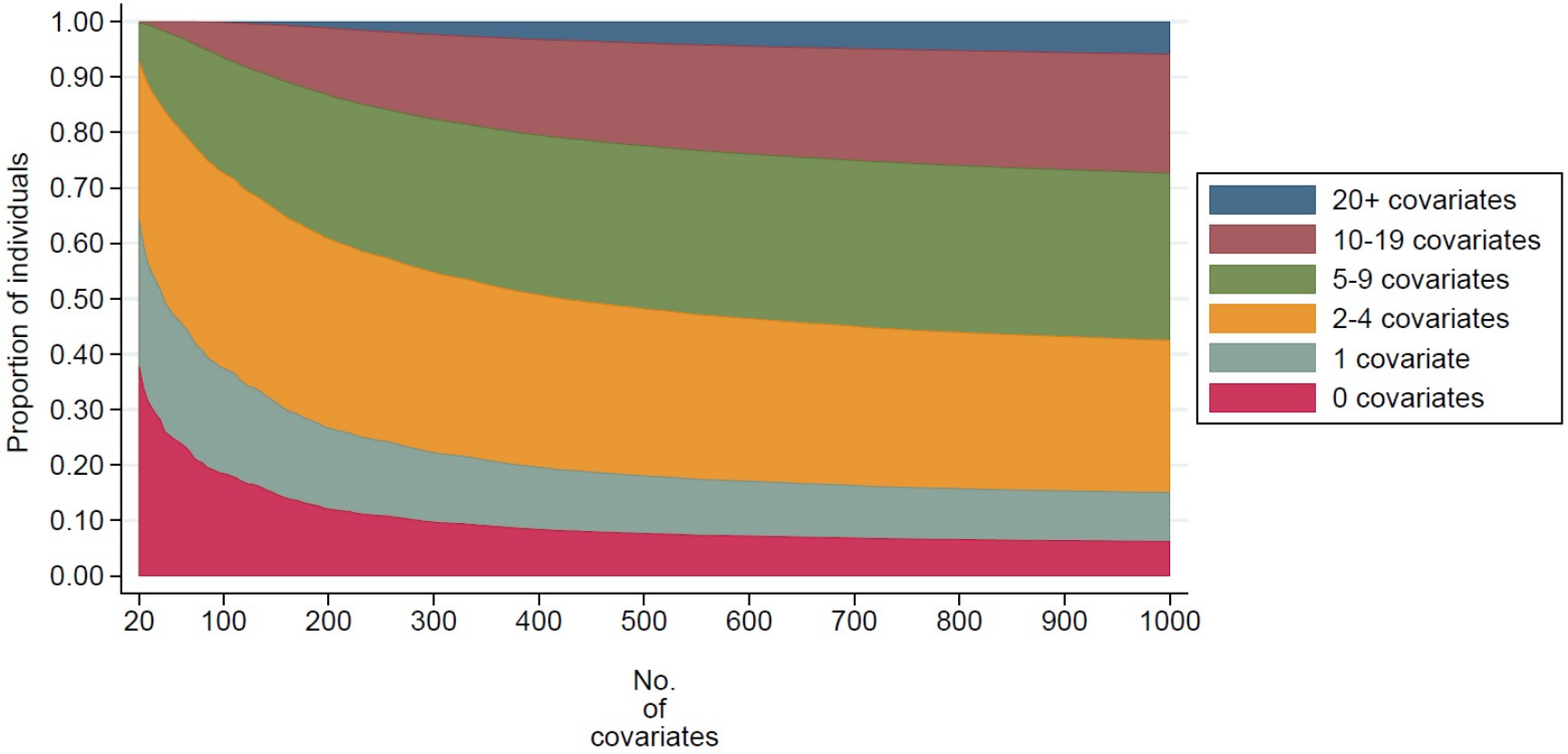
- High-dimension PS
- A lot of variables (n>200) included in the PS model
- Typically the 200 most frequently occurring diagnoses and treatments
- Can be automated
- Exploits the inherent capability of confounder adjustment for the data set

HdPS in a Nordic setting, GI bleeding, Coxibs vs traditional NSAIDs in DK

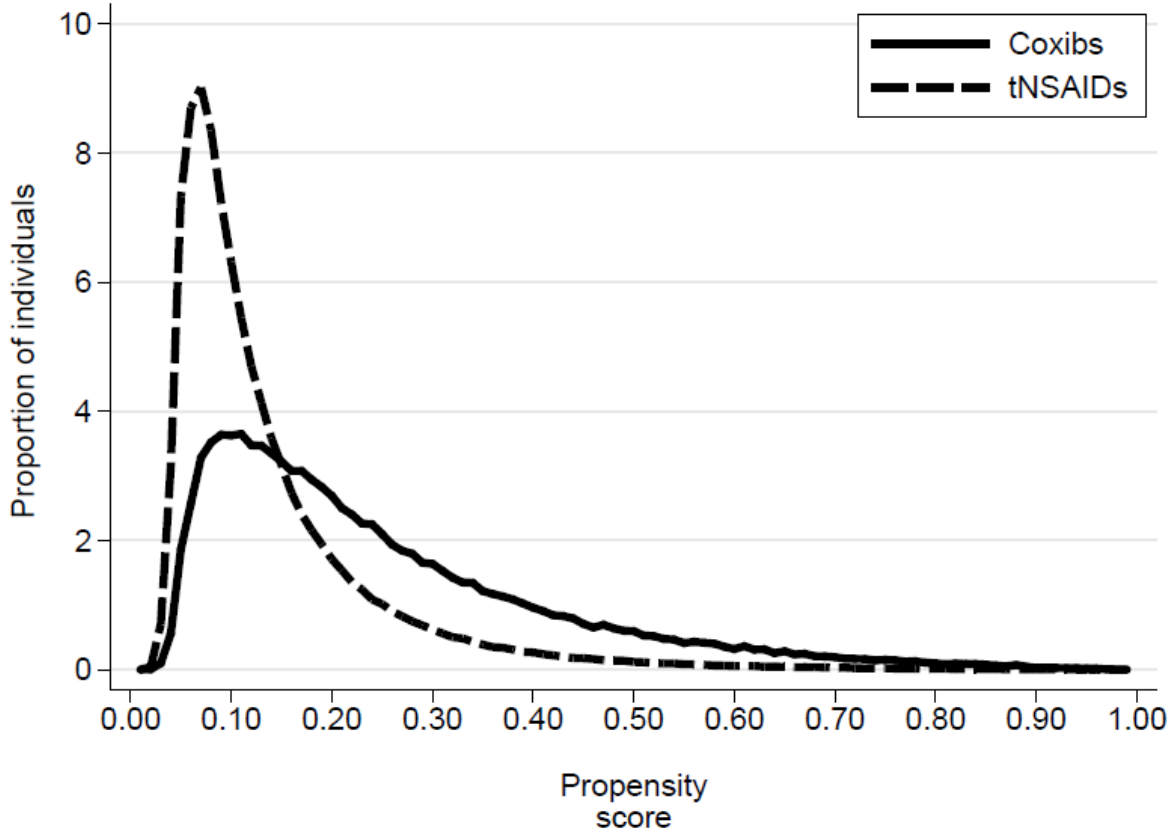
Analytic approach	Diagnosis covariates included	Prescription covariates included	Hazard ratio (95% CI)
Crude analysis	NA	NA	1.76 (1.57 - 1.97)
Adjusted by age and sex	NA	NA	1.12 (1.00 - 1.26)
Adjusted by age, sex and clinically selected covariates	3	8	0.99 (0.88 - 1.12)
Adjusted by PS model	10	10	0.89 (0.77 - 1.02)
	20	20	0.86 (0.75 - 0.99)
	50	50	0.86 (0.76 - 0.99)
	100	100	0.89 (0.77 - 1.01)
	200	200	0.88 (0.77 - 1.01)
	500	500	0.90 (0.79 - 1.03)

Hallas, Basic Clin Pharmacol Toxicol 2017

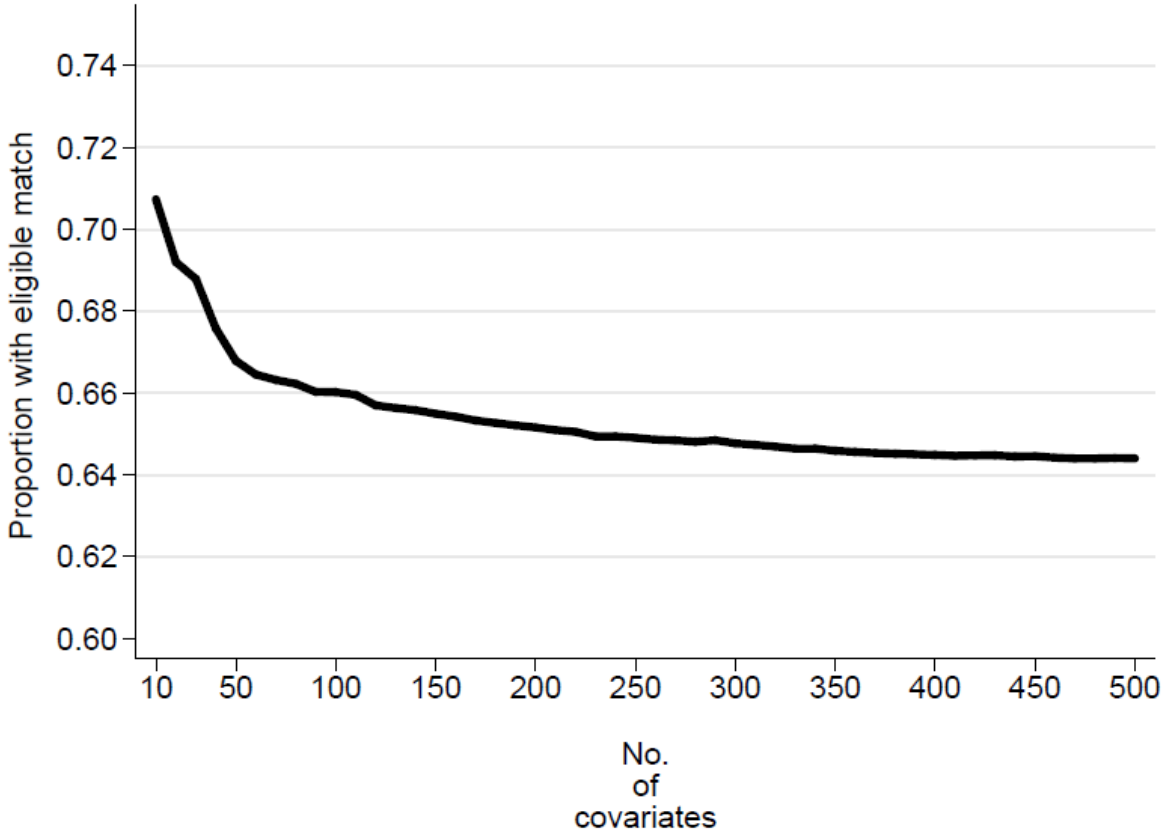
Positive covariates, dx and rx



PS distributions, 500 dx and 500 rx



PS distribution overlap



HdPS in DK, conclusions

- HdPS does work in Danish data
- It works better than a GI specialist
- Early saturation re number of variables

Exercise

- What will happen if you employ PS matching in a case-control study ?

THYROID
Volume 30, Number 11, 2020
Mary Ann Liebert, Inc.
DOI: 10.1089/thy.2019.0589

Association Between Thionamides and Acute Pancreatitis: A Case–Control Study

Jia-Yin Guo,[1,2] Chia-Ling Chang,[3,4] and Ching-Chu Chen[2,5]

Background: Thionamides have been extensively used to treat patients with hyperthyroidism worldwide. Recent pharmacovigilance studies have revealed a safety signal between carbimazole or methimazole and pancreatitis. The associated risk remains unclear.
Methods: We identified patients with newly diagnosed acute pancreatitis from 2000 to 2013 as the case group from the Taiwan Longitudinal Health Insurance Database 2000, which contains data from 1996 to 2013. Each patient with acute pancreatitis was matched for age, sex, comorbidities, and cancer with four controls through propensity score matching. A total of 52 patients without matched controls were excluded. Sensitivity analyses including the 52 excluded patients were performed using a matching ratio of 1:2. Odds ratios (ORs) along with 95% confidence intervals (CIs) for the association were estimated using multivariate logistic regression.
Results: We included 9204 and 36,816 patients in the case and control groups, respectively. The proportions of patients who had used thionamides, carbimazole, methimazole, and propylthiouracil were similar in these two groups. In addition, the adjusted OR (CI) for the association of acute pancreatitis with thionamides was 1.03 (0.86–1.24), with carbimazole it was 0.90 (0.63–1.30), with methimazole it was 1.05 (0.84–1.31), and with propylthiouracil it was 1.00 (0.74–1.34). The sensitivity analysis results were unchanged.
Conclusions: We were unable to demonstrate an association between acute pancreatitis and usage of thionamides.

Keywords: thionamides, carbimazole, methimazole, propylthiouracil, pancreatitis, hyperthyroidism

are used to define disease diagnoses from outpatient and inpatient data in the LHID2000. The ICD-9-CM codes used in this study are presented in Supplementary Table S1.

We identified patients with new diagnoses of acute pancreatitis—defined as ≥2 outpatient visit diagnoses or ≥1 diagnosis during hospitalization or an emergency visit—as the case group. The exclusion criteria included the following: index date not within the study period (2000–2013), age younger than 20 years, age older than 100 years, and missing sex or age data. The diagnosis date was defined as the index date. We used the same exclusion criteria to identify patients without a history of acute pancreatitis as the control group; these controls were matched for age, sex, comorbidities (including alcoholic liver disease, gallbladder stone, hyperlipidemia, and type 2 diabetes mellitus), and cancer with the case group through propensity score matching at a ratio of 1:4. A total of 52 patients without matched controls were excluded. Sensitivity analyses including the 52 excluded patients were performed at a matching ratio of 1:2 to clarify the effect of excluding the 52 unmatched patients on the association. The status of antithyroid drug use was categorized as never use and ever use.

Statistical analyses

Chi-square tests were used to compare categorical variables. Odds ratios (ORs) along with 95% confidence intervals (CIs) for the association were estimated using multivariate logistic regression. The ORs were adjusted for age, sex, alcoholic liver disease, gallbladder stone, and cancer. All statistical analyses were performed using STATA/SE version 14.0 (STATA Corp., College Station, TX). Results with a two-sided p-value of <0.05 were considered significant.

Results

As shown in Figure 1, from the total of 1 million patients in the LHID2000, 10,963 patients with newly diagnosed acute pancreatitis were included in the case group. We excluded 1536 patients for whom the date of diagnosis was not within the study period; 171 patients without sex or age data, aged <20 years, or aged >100 years; and 52 patients without matched controls (including 3 patients who had been prescribed thionamides). Finally, a total of 9204 patients were included in the case group. For the control group, we identified 989,037 patients without a history of acute pancreatitis. From these patients, 234,744 were excluded using the same

When not to use PS

- In case-control studies
  - Tricks to circumvent (Månsson et al, AJE 2007)

Disease risc scores

- Summary score, like the PS
- Models the risk of having the outcome, based on baseline covariates
  - E.g., Charlson comorbidity index
- Usually based on the untreated
  - Inclusion of treated will bias the DRS if there is effect modification

Disease risc scores, properties

- Summary score, just like the PS
  - Covariate, matching, stratification and weighting
- DRSs are ”biological” and PSs are ”behavioural”
  - DRSs are useful for newly marketed drugs
- Does not generate covariate balance
  - No table 1
  - ”Dry-run analyses”, Wyss et al, AJE 2017
- Can be used in case-control studies

PS for dummies

- PS models the likelihood of receiving the treatment, based on covariates
- Adjusts for variables included in the model, not for others.
- Can be applied with matching, as a covariate, by stratification and by weighting
- Particularly useful if there are many covariates and few outcomes
- Fringe benefits include insight into prescribing patterns and exclusion of subjects who have absolute indications or contraindications
- HdPS expoits the inherent capability of the data to adjust for confounders
- Choose variables that are determinants of the outcome and not others

Litterature

- Stürmer T at al. Treatment effects in the presence of unmeasured confounding: dealing with observations in the tails of the propensity score distribution--a simulation study. Am J Epidemiol 2010; 172: 843-54.
- Kurth T et al. Results of multivariable logistic regression, propensity matching, propensity adjustment, and propensity-based weighting under conditions of nonuniform effect. Am J Epidemiol 2006; 163: 262-70.
- Schneeweiss S et al. High-dimensional propensity score adjustment in studies of treatment effects using health care claims data. Epidemiology 2009; 20: 512-22.
- Hallas J et al. Performance of the High-dimensional Propensity Score in a Nordic Healthcare Model. Basic Clin Pharmacol Toxicol. 2017 Mar;120(3):312-317.
- Brookhart MA et al. Variable selection for propensity score models. Am J Epidemiol 2006; 163: 1149-56.
- Glynn RJ et al. Indications for propensity scores and review of their use in pharmacoepidemiology. Basic Clin Pharmacol Toxicol; 98: 253-9.
- Desai RJ, Franklin JM. Alternative approaches for confounding adjustment in observational studies using weighting based on the propensity score: a primer for practitioners. BMJ. 2019 Oct 23;367:l5657. doi: 10.1136/bmj.l5657. PMID: 31645336.

?