CENTRE OF REGISTERS
VÄSTRA GÖTALAND

Matched Cohort designs.

Stefan Franzén PhD

Lund 2016-10-13

VÄSTRA GÖTALANDSREGIONEN

Registercentrum Västra Götaland

Purpose: improved health care

25+	**30+**	**70+**
Registries	Employed	Papers

- Statistics
- IT
- Project management
- Registry development
- Communication

CENTRE OF REGISTERS
VÄSTRA GÖTALAND

Agenda

- Why match?
- The estimand
- Selecting controls
- Matching options
- Checking balance
- Post matching analysis

CENTRE OF REGISTERS
VÄSTRA GÖTALAND

Why match

1. Reduce confounding
2. Set an index date for the ”controls” in longitudinal data
3. Reduce variability and model dependence

CENTRE OF REGISTERS
VÄSTRA GÖTALAND

The estimand

ATE

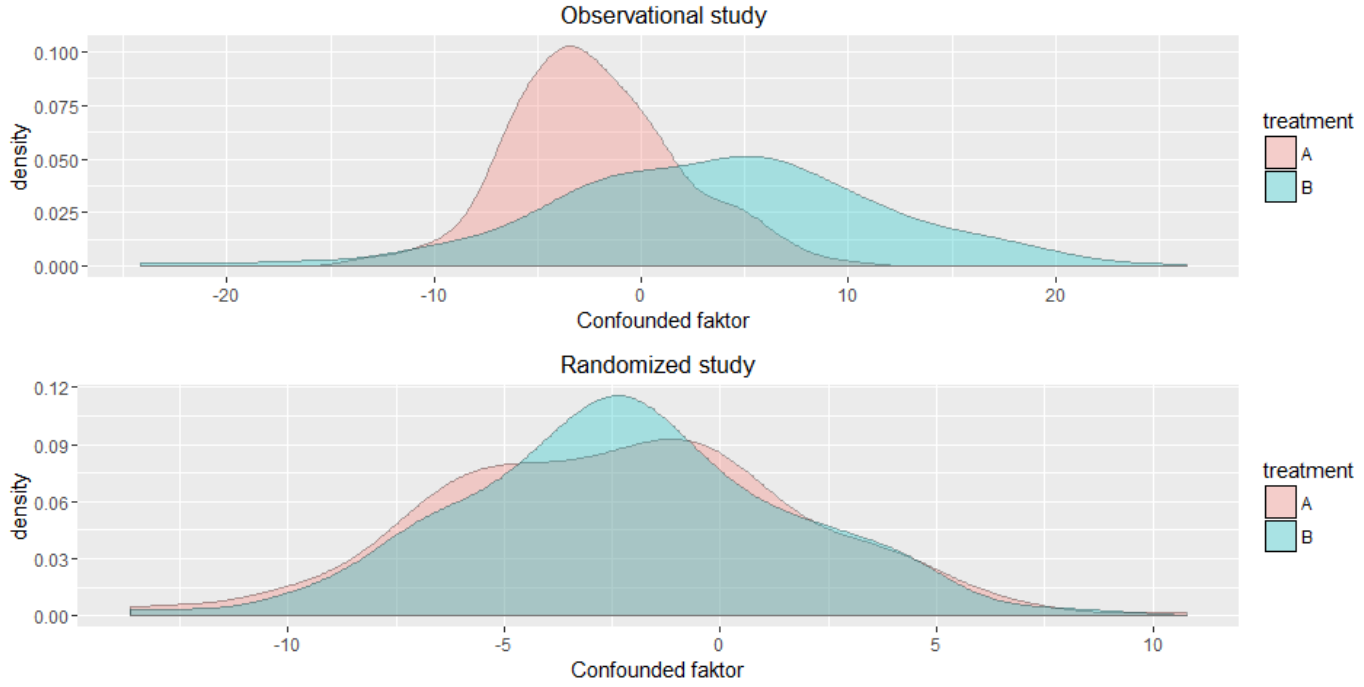
The average treatment effect for everyone

ATT

The average treatment effect for the treated

CENTRE OF REGISTERS
VÄSTRA GÖTALAND

The estimand



ATT is weighted by the “treated population” (red)
ATE is weighted by the total population (red and green combined)
An RCT estimated ATE. –Why?

CENTRE OF REGISTERS
VÄSTRA GÖTALAND

The estimand, example

Compare mortality after PCI and CABG post MI

ATE: What is the effect of treating everyone with PCI vs everyone with CABG?

ATT_{PCI} What is the effect of treating the PCI-patients with CABG?

ATT_{CABG} What is the effect of treating the CABG-patients with PCI?

CENTRE OF REGISTERS
VÄSTRA GÖTALAND

Selecting controls

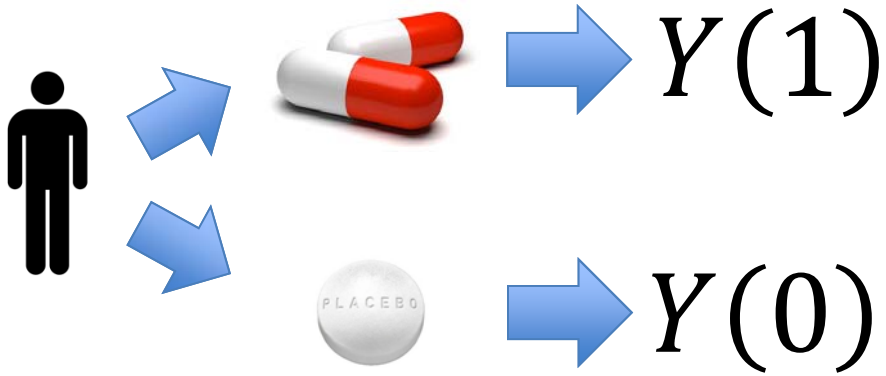
What is the exposure?

What is the alternative to exposure?

- One specific alternative exposure e.g. PCI or CABG
- ”Unexposed” – standard care e.g. not yet exposed to a particular drug

What would have happend if the patient was not exposed?

Potential outcomes:



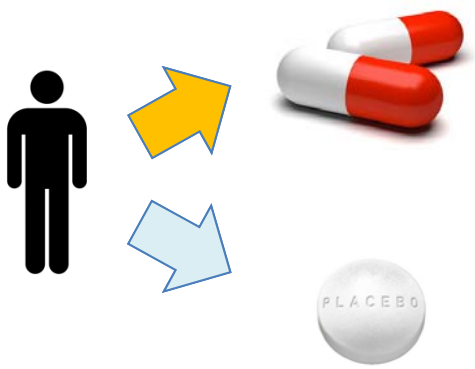
$Y(1)$

$Y(0)$

CENTRE OF REGISTERS
VÄSTRA GÖTALAND

Selecting controls counterfactuals

What would have happend if the patient was not exposed?



$Y(1)$:observed

$Y(0)$:missing

We can only observe one of the potential outcomes

The potential outcome under control remains missing for a treated patient

Matching is away to find that missing outcome

CENTRE OF REGISTERS
VÄSTRA GÖTALAND

Selecting controls, example 1

PCI or CABG

1: Determine the estimand: ATT_{CABG}

2: For each CABG-patient, find a matching PCI-patient

CABG is only an option for 3 or 4 vessel disease and these patients may have been treaed with PCI before for 1 or 2 vessel disease

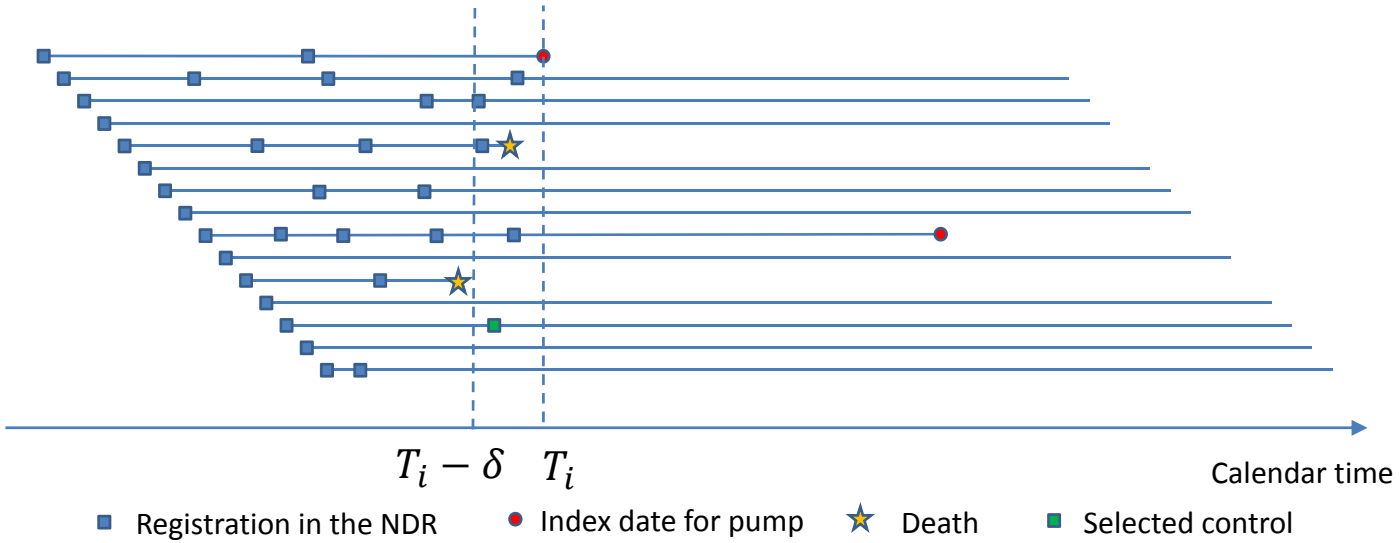
3: A potential control is a patien treated with PCI for 3 or 4 vessel disease

CENTRE OF REGISTERS
VÄSTRA GÖTALAND

Selecting controls, example 2

Insulin pump or injections

For each exposed, select without replacement the closest unexposed patient from the risk set, who are alive at T_i and who have the latest registration no earlier than δ days



The selected control is subsequently followed from T_i with respect to outcomes and may be censored if exposed to pump at a later date

CENTRE OF REGISTERS
VÄSTRA GÖTALAND

Matching options

- Attribute matching
- Mahalanobis distance
- Propensity score
- Disease risk score
- Combinations

CENTRE OF REGISTERS
VÄSTRA GÖTALAND

Attribute matching

Create matched (pairs) such that $\boldsymbol{X}_i = \boldsymbol{X}_j$ i.e. both individuals have the same values on ”all” matched variables.

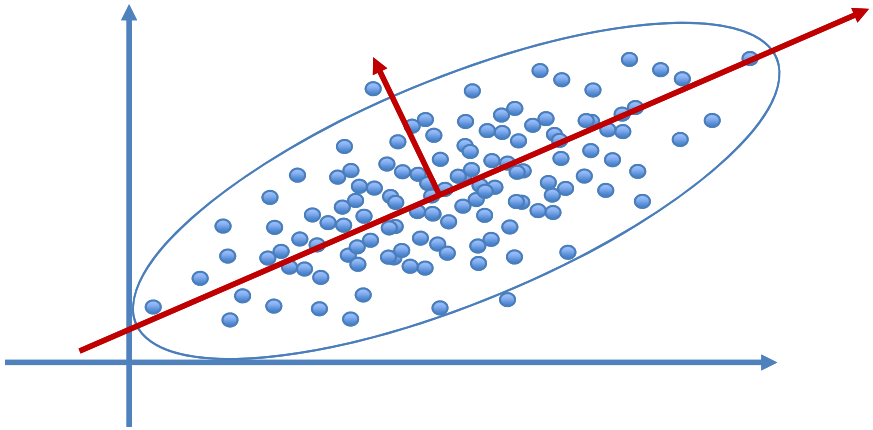
- Creates exact balance
- Curse of dimentionality
- Continuous variables has to be discretized

CENTRE OF REGISTERS
VÄSTRA GÖTALAND

Mahalanobis distance matching

Idea: Create a one dimensional summary measure

$$\rho = \sqrt{(X - \bar{X})^T \Sigma^{-1} (X - \bar{X})}$$



Find patients such that $\rho(X_i) \approx \rho(X_j)$ implying $X_i \approx X_j$

Works best for small number of continuous variables Leon & Carriére 2005

Considers all interaction terms equally important

Claimed to be less sensitive to model dependence
(in the analysis) than propensity score.
(King & Nielsen 2016)

CENTRE OF REGISTERS
VÄSTRA GÖTALAND

Propensity score matching

Definition:

$e_i(\boldsymbol{X}_i) = P(T = 1|\boldsymbol{X}_i)$ where $T \in \{0,1\}$ indicates treatment

Point:

$$\big(Y(0), Y(1)\big) \perp T \mid e(\boldsymbol{X}) \xrightarrow[assumptions]{} \big(Y(0), Y(1)\big) \perp T \mid \boldsymbol{X}$$

- If the treatment isn't confounded by $e(\boldsymbol{X})$ then it isn't confounded by $\boldsymbol{X}$.
- But there may be variables in $\boldsymbol{X}$ that are weekly related to T and these will not be so similar within a matched pair i.e. given $e(\boldsymbol{X})$.

CENTRE OF REGISTERS
VÄSTRA GÖTALAND

Disease risk score matching

For a new treatment we may not have enough treated patients to reliably estimate a propensity score

Idea: use a disease risk score as a balancing score

If $Y_C \perp X \mid \Psi(X)$ then $\Psi(X)$ is a prognostic score

1. Estimate a model for the outcome based on the non exposed
2. Create predictions (scores) for all patients
3. Create a matched analysis data using the scores

Hansen 2007

CENTRE OF REGISTERS
VÄSTRA GÖTALAND

Combinations of matching measures

Attribute matching on a few important variables

- Matching on summary measure on the rest
- Handle remaining confounding by regression modelling

Mathing on propensity and disease risk score combined

1. Estimate propensity score
2. Estimate disease risk score
3. Match on Mahalanobis distance of both

CENTRE OF REGISTERS
VÄSTRA GÖTALAND

Ways to match (and alternatives)

Method	Effect	
Match 1-1	ATT	Reduced sample size. Potential loss of treated patients
Match 1-n	ATT	Reduced sample size. Potential loss of treated patients
Match m-n	ATT	Reduced sample size. Potential loss of treated patients
Full matching	ATE	Complicated matching process
IPT-Weighting	ATE or ATT	Some observations may carry a very large weight
Stratify	ATE (ATT)	Simple– watch out for very unbalanced strata
Regression	ATE	• May hide an extrapolation • Assumes a specific functional from • Uses the outcome in the model • May not be possible to include all potential confounders

CENTRE OF REGISTERS
VÄSTRA GÖTALAND

Greedy 1-k matching

1. Select a treated patient at random
2. Select the closest k (often k=1) untreated patient

Aim: to estimate ATT

Number of controls: 2 is better than 1… (law of diminishing return)

Calipers: Reject matches where $|e(X_i) - e(X_{i^*})| > \delta$

- Helps to improve balance
- Rubin & Thomas: $\delta = 0.25\sigma\big(e(X)\big)$ or $0.5\sigma\big(e(X)\big)$
- Prunes the data by dropping ”treated” and no longer estimates ATT

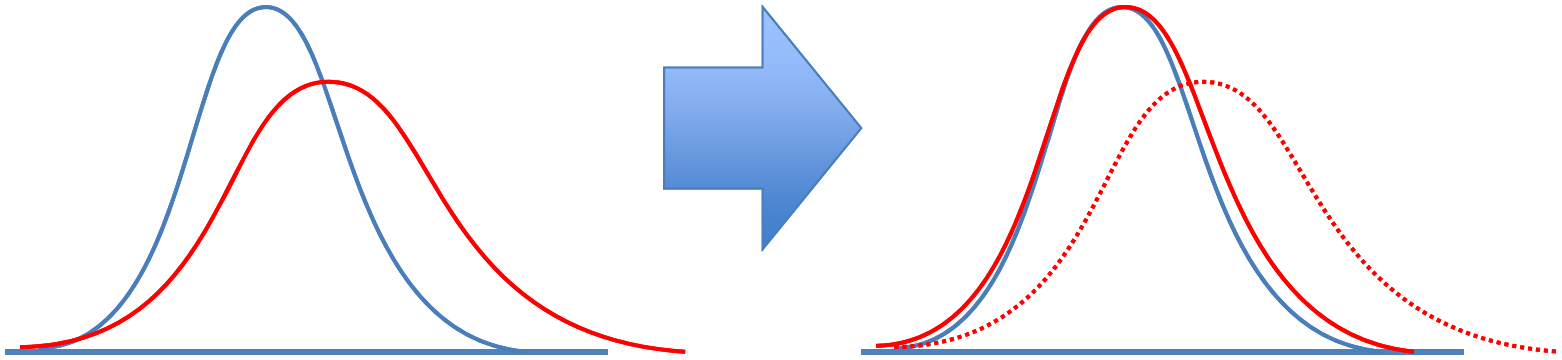
Replacements: Match with or without replacement

Replacements may improve the balance but the observations are no longer independent which mess up the analysis

CENTRE OF REGISTERS
VÄSTRA GÖTALAND

Why greedy 1-k matching estimates ATT

Starting with the treated preserves the distribution for the treated



but only if we don’t loose and ”treated”

CENTRE OF REGISTERS
VÄSTRA GÖTALAND

Alternatives to greedy 1-k matching

Greedy m-k matching

Simlilar to 1-k but takes m treated at time.

Estimates ATT

Full matching

Find small sets of exposed and controls minimizing some global distance measure

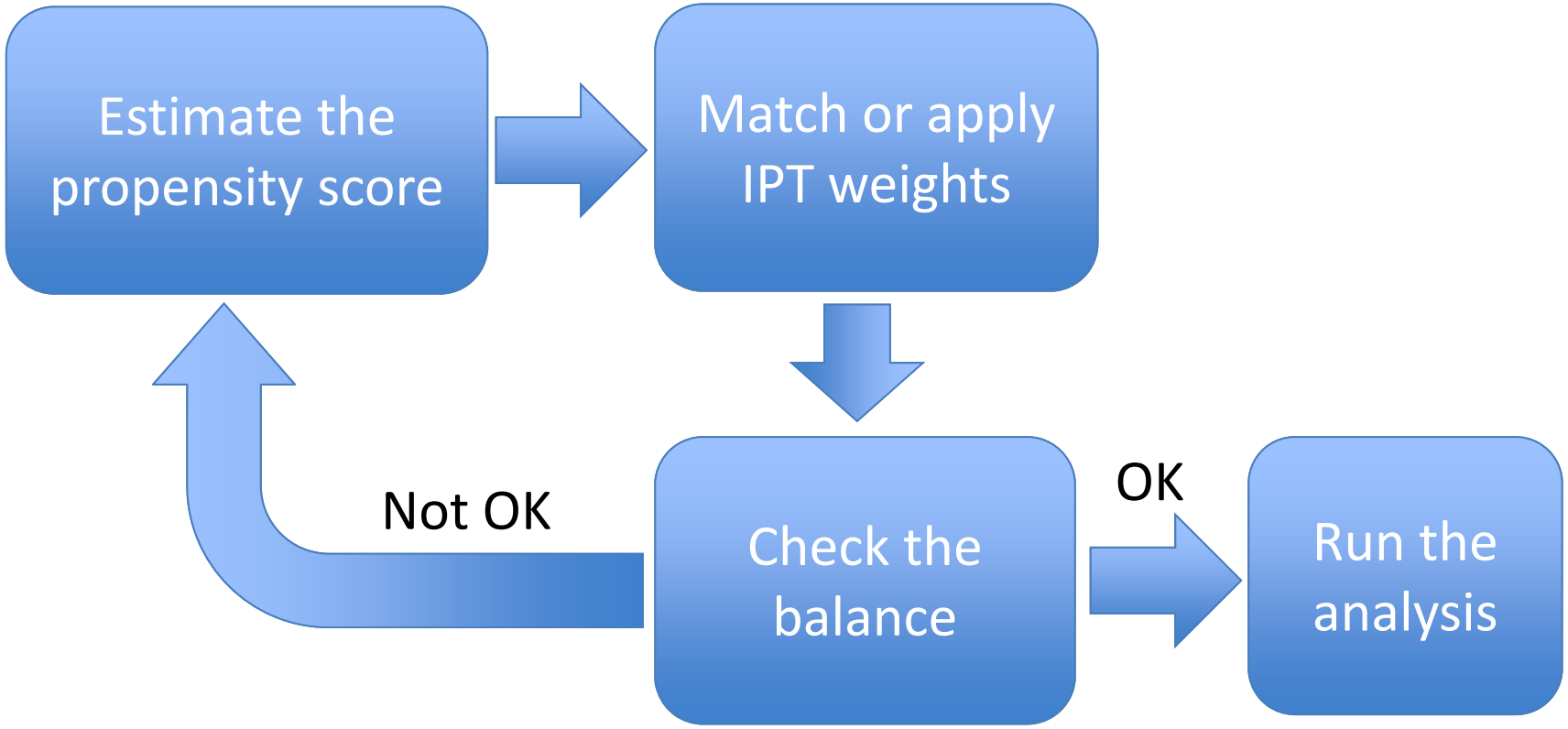
The number of exposed and controls in each set varies

Computer intensive iterative process

May mess up the analysis

Estimates ATE

Creating balance, an iterative process



Since the propensity score model doesn't contain the outcome you can fiddle around with the model until you are satisfies with the balance you get

CENTRE OF REGISTERS
VÄSTRA GÖTALAND

Evaluating the balance

Unfortunately it is (still) common practice to compare groups at index using a hypothesis test

$$p>5\%$$

CENTRE OF REGISTERS
VÄSTRA GÖTALAND

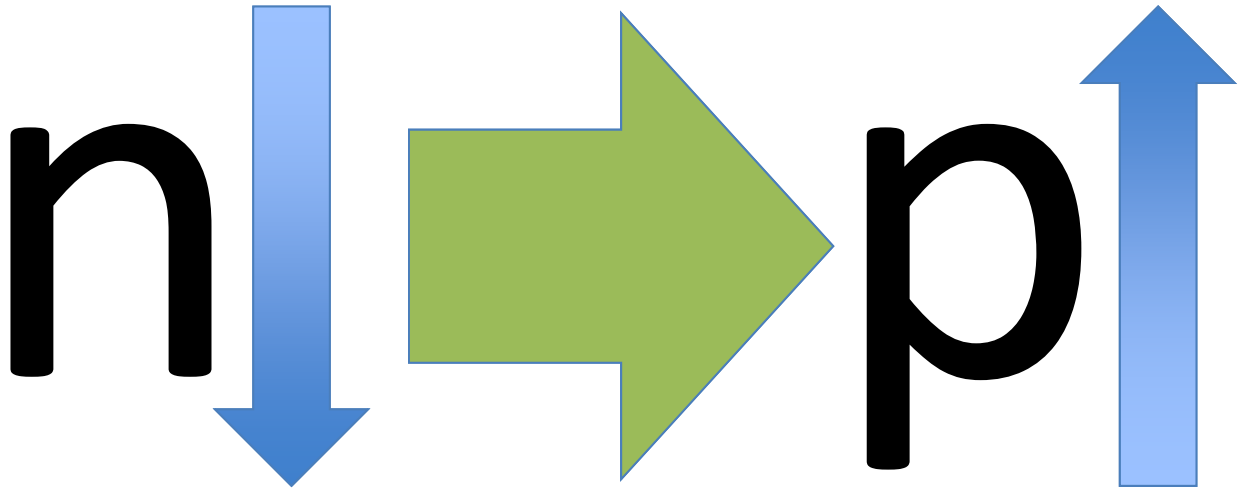
Evaluating the balance

Balance between groups is a property of the data and not of the underlying true population averages

CENTRE OF REGISTERS
VÄSTRA GÖTALAND

The p value and the sample size

The p-value is heavily affected by the number of observations



Using the p-value as balance metric we can create a balance between groups by randomly deleting observations

CENTRE OF REGISTERS
VÄSTRA GÖTALAND

The standardized difference

The standardized difference is not influenced by the number of observations and is a much better balance metric

$$\frac{|\overline{x_t} - \overline{x_c}|}{\hat{\sigma}}$$

Some authors have suggested that 10% or 20% indicates sufficient balance, but this is as arbitrary as the famous 5%

CENTRE OF REGISTERS
VÄSTRA GÖTALAND

Other balance metrix

The prognostic risk score

Hansen 2007:

If $Y_C \perp X \mid \Psi(X)$ then $\Psi(X)$ is a prognostic score

Stuart, Lee & Leacy 2013:

Model $\mathrm{E}[Y \mid X]$ based on unexposed, eg glm…

Predict outcomes for all patients

Evaluate the balance based on the predictions

CENTRE OF REGISTERS
VÄSTRA GÖTALAND

What if I can't match?

Failing to find matching controls for all treated is a property of the data, not of the method

CENTRE OF REGISTERS
VÄSTRA GÖTALAND

Post matching analysis

Convensional wisdom:

Analyze as designed

Attribute matching:

Include matching variables in the analysis

Propensity scores:

Greedy 1-k:

doesn’t influence the analysis

Full matching n-m:

”Clusters” has to be accounted for in the analysis

CENTRE OF REGISTERS
VÄSTRA GÖTALAND

Post matching analysis

Fit the analysis model just as if matching hadn't happaned

ANOVA: keep variabiliy from the error term
Glm inc Cox: Avoid "hidden" covaraite bias

Combinations of proc scores and regression is not "doubly robust"

Doubly robust estimator

$$\hat{\Delta}_{dr} = \frac{1}{n}\sum_{i\in\omega}\frac{Y_i T_i - (T_i - \hat{e}_i)m_1(X_i, \hat{\beta}_i)}{\hat{e}_i} - \frac{1}{n}\sum_{i\in\omega}\frac{Y_i(1-T_i) - (T_i - \hat{e}_i)m_0(X_i, \hat{\beta}_i)}{1-\hat{e}_i}$$

CENTRE OF REGISTERS
VÄSTRA GÖTALAND

Example 1: Gastric Bypass

Analysis strategy: match as closely as the data permits and fit a Cox regression to the matched data accounting for remaining confounding

6177 GBP patients are matched 1-1 to ”not yet treated” patients from NDR selected from 440824 patients with 4884442 records

Matched on: Year: 2008-2009, 2010-2011, 2011-2012, 2013-2014
BMI: [28-35),[35, 39), [39, 44), [44,∞)
Age: [18,42), [42, 49)[49,56), [56, ,∞)
Sex: Male, Female

128 strata

Matched controls are selected from the patients at risk in the time interval

Matched controls are removed from the risk sets in following time intervals (no replacement)

Matched controls are censored when (if) treated (n=733 control patients are censored this way)

If mutiple registrations are availabe in for a selected control, one is selected randomly as index

CENTRE OF REGISTERS
VÄSTRA GÖTALAND

Example 1: Gastric Bypass

Table 1. Descriptive statistics (mean(SD) or count(%)) of soreg and matched NDR control patients after matching

Variable	Gastric bypass (n=6132)	Matched control (n=6132)	p-value	standardized difference
Sex	2364 (38.6%)	2364 (38.6%)	1	0
Age (years)	48.5 (9.8)	50.5 (12.7)	8.212E-24	0.1835435
BMI	42.0 (5.7)	41.4 (5.7)	9.039E-10	0.1107137
HbA1c(%)	7.6 (1.6)	7.6 (1.5)	0.5767475	0.0100873
LDL	2.8 (1.0)	2.8 (0.9)	0.0706582	0.0326673
HDL	1.1 (0.5)	1.2 (0.3)	2.7347E-8	0.1019714
SBP	140.0 (17.2)	133.8 (15.6)	5.48E-93	0.3730043
DBP	83.3 (10.2)	79.8 (10.2)	3.571E-80	0.3449707
Previous MI	231 (3.8%)	261 (4.3%)	0.1674394	0.1270678
Previous CHF	172 (2.8%)	254 (4.1%)	0.0000526	0.3993627
Previous stroke	131 (2.1%)	179 (2.9%)	0.0057565	0.3179059
Smoking	540 (8.8%)	1049 (17.1%)	1.784E-11	0.7474336
Type 2 diabetes	4968 (94.8%)	5524 (91.5%)	5.193E-11	0.5208766
BP medication	3644 (59.4%)	4083 (66.6%)	2.195E-16	0.3088453
Lipid lowering medication	2142 (34.9%)	3076 (50.2%)	3.025E-65	0.6382504
Diabetes medication	4975 (81.1%)	5110 (83.3%)	0.0014266	0.1508076
Duration of diabetes	7.6 (7.1)	7.7 (7.8)	0.4546706	0.0149411
Married	2914 (47.5%)	2548 (41.6%)	2.937E-11	0.2425066
Yearly income (Ksek)	202.5 (126.7)	183.5 (124.4)	5.359E-17	0.1515408
Education (Low)	1253 (20.4%)	1776 (29.0%)	6.579E-28	0.4631218
Education (Mid)	3663 (59.7%)	3256 (53.1%)	1.247E-13	0.2711535
Education (High)	1216 (19.8%)	1100 (17.9%)	0.0074437	0.1235661

Example 1: Gastric Bypass

Cox proportional hazards regression for mortality, cardiovascular death and myocardial infarction for patients treated with gastric bypass

Endpoint	Hazard ratio	95% Confidence interval	p-value
Mortality	0.46	[0.35, 0.62]	<.0001
CV mortality	0.38	[0.18, 0.79]	0.0097
Myocardial infarction	0.49	[0.29, 0.84]	0.0094

The analysis is based on a Cox proportional hazards regression including the treatment and age, sex, BMI, hba1c, ldl, hdl, SBP, DBP, smoking, previous MI, previous CHF, previous stroke, bloodpressure med, lipid med, diabetes med, income, education and marrietal status. The analysis data was matched on age, sex, BMI and calendar time period

CENTRE OF REGISTERS
VÄSTRA GÖTALAND

Example Statins and type 1 diabetics

Propensity scores

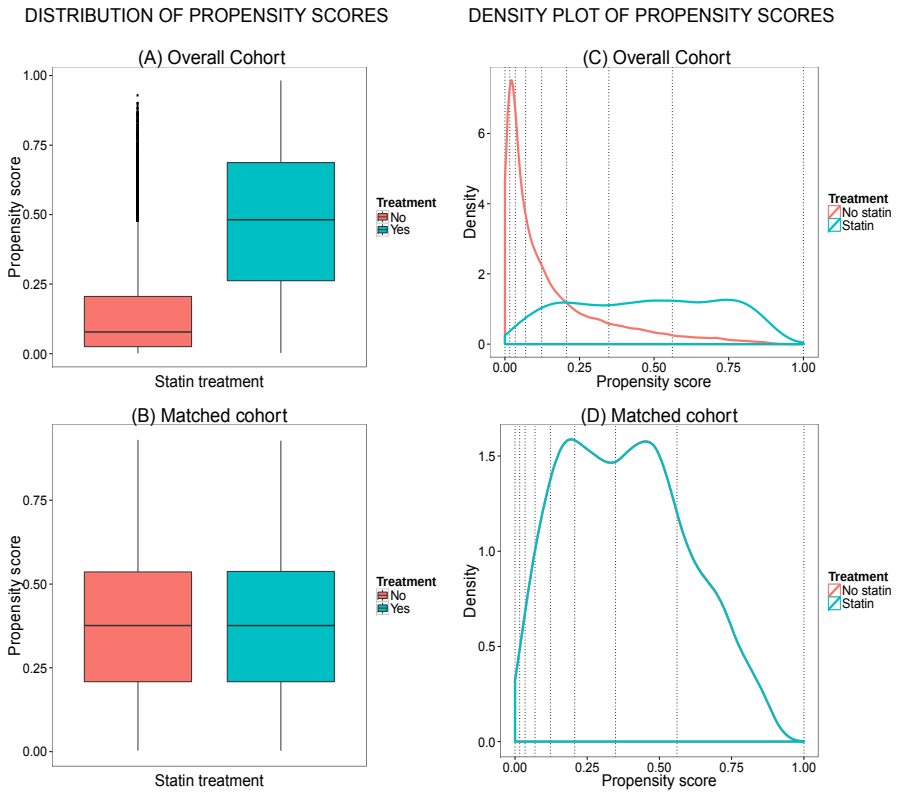
Greedy 1-1 matching

Superb balance!

4025 matched treated

1362 excluded treated

What are we estimating?

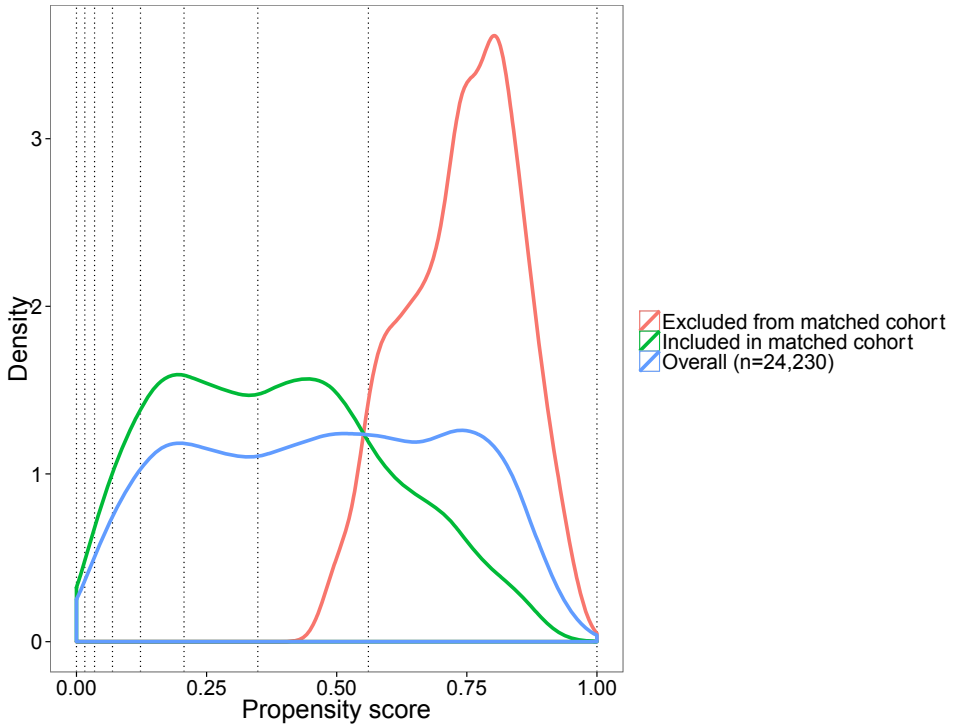


endoint	HR	95% CI	P-value
Total death	0.74	[0.62, 0.88]	0.001
CV death	0.83	[0.66, 1.03]	0.086

Hero et.al Diabetes Care 2016

CENTRE OF REGISTERS
VÄSTRA GÖTALAND

Indeed what are we estimating



Aimed to estimate ATT

Not all statin patients included

Doesn't quite estimate ATT

The use of statins has become so regulated that we can no longer properly evaluate it.

OK if regulations are based on evidence…

CENTRE OF REGISTERS
VÄSTRA GÖTALAND

Softwares

R-packages: MatchIt, twang

SAS: code and macros arround but no official proc

Stata: psmatch2, pscore, and more

Propensity score: Easy to do with standard components

Matching: greedy 1-1 easy but beyond that it gets harder

Software

http://www.biostat.jhsph.edu/~estuart/propensityscoresoftware.html

CENTRE OF REGISTERS
VÄSTRA GÖTALAND

Summary

1. Use the right estimand: ATT or ATE
2. Balance is everything
3. Try to avoid p-values
4. Usually ok to analyze as usual after matching

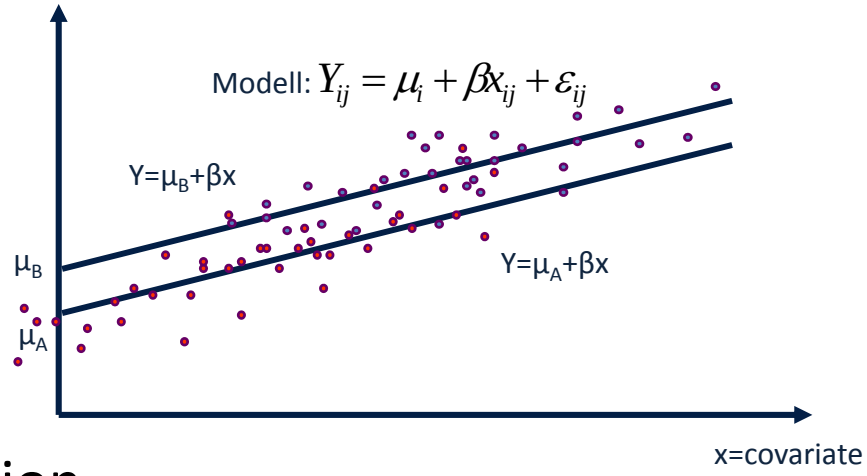
CENTRE OF REGISTERS
VÄSTRA GÖTALAND

Backups

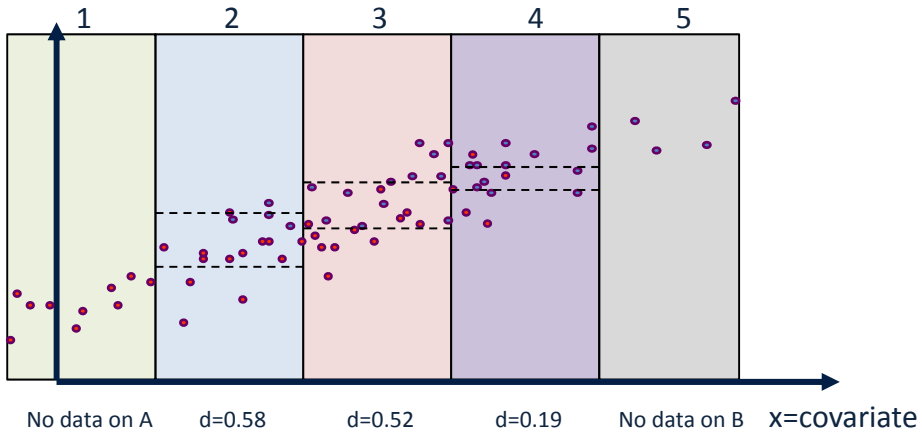
CENTRE OF REGISTERS
VÄSTRA GÖTALAND

Alternatives to matching

Regression adjustment



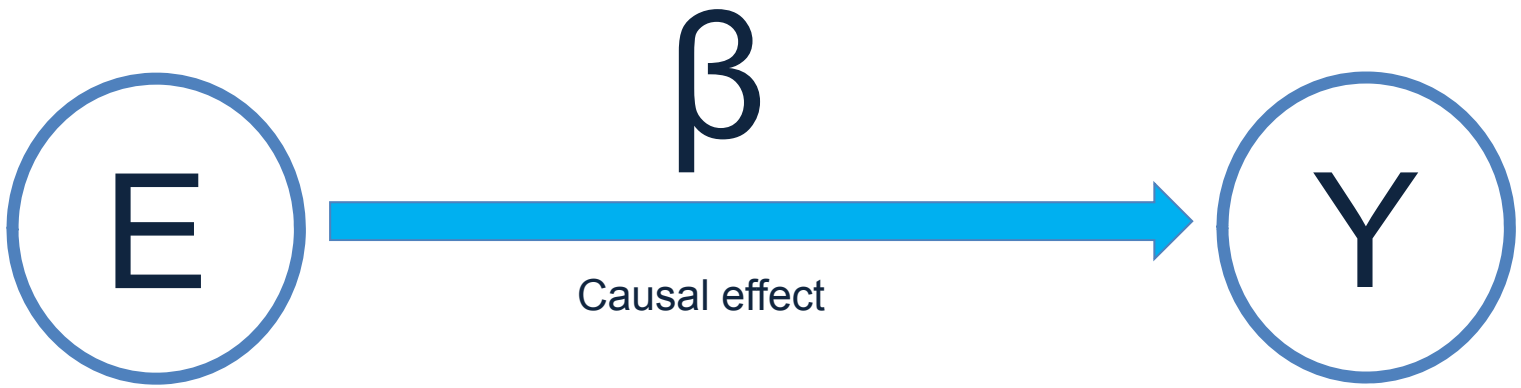
Stratification



CENTRE OF REGISTERS
VÄSTRA GÖTALAND

Causation

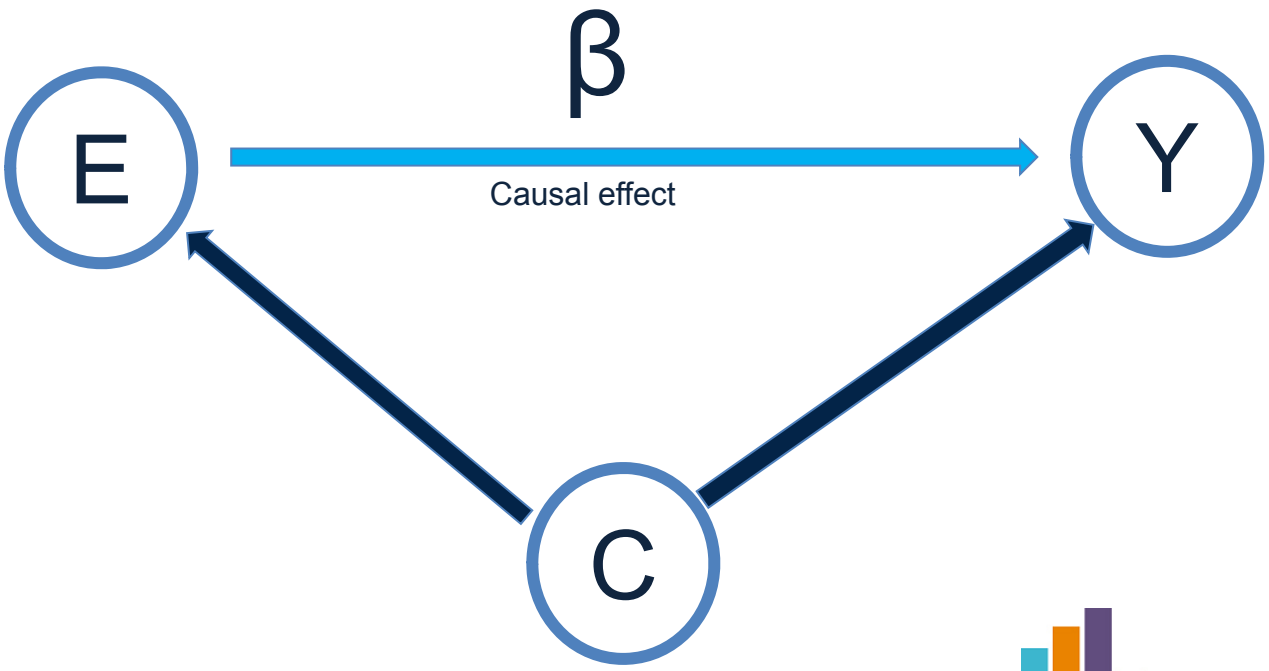
We are interested in investigating the extent β of the causal effect of exposure E on the outcome Y



CENTRE OF REGISTERS
VÄSTRA GÖTALAND

Confounding

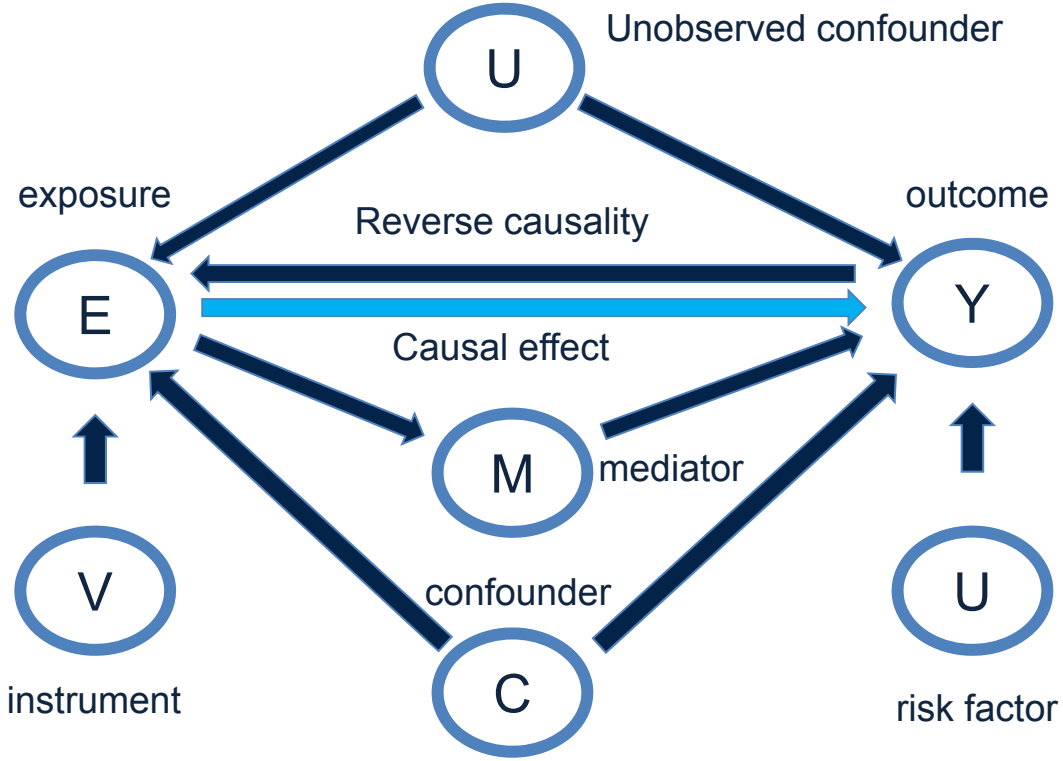
A confounder is a factor C distorts the relationships between exposure E and outcome Y. If C is not observed we are in trouble.



CENTRE OF REGISTERS
VÄSTRA GÖTALAND

Confounding

In reality the relation between the exposure and the outcome is often very complex and partially unobservable



CENTRE OF REGISTERS
VÄSTRA GÖTALAND

What to ”adjust for”

Confounders 

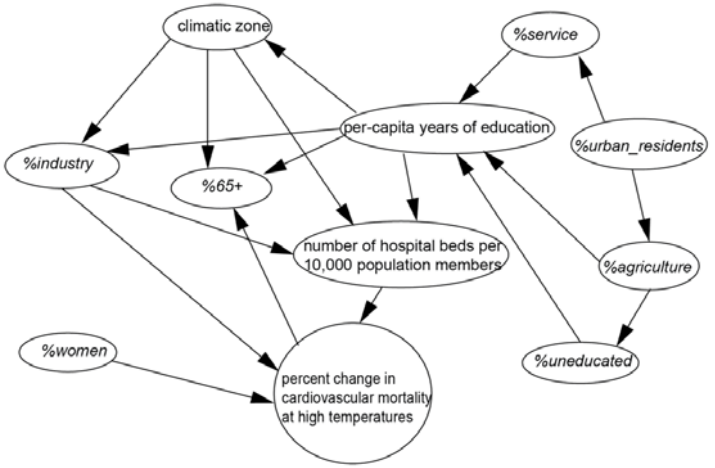
Mediators* 

Risk factors 

Instruments**

Not ok in propensity score models

Directed acyclic graph



*) Consider Causal Mediation analysis

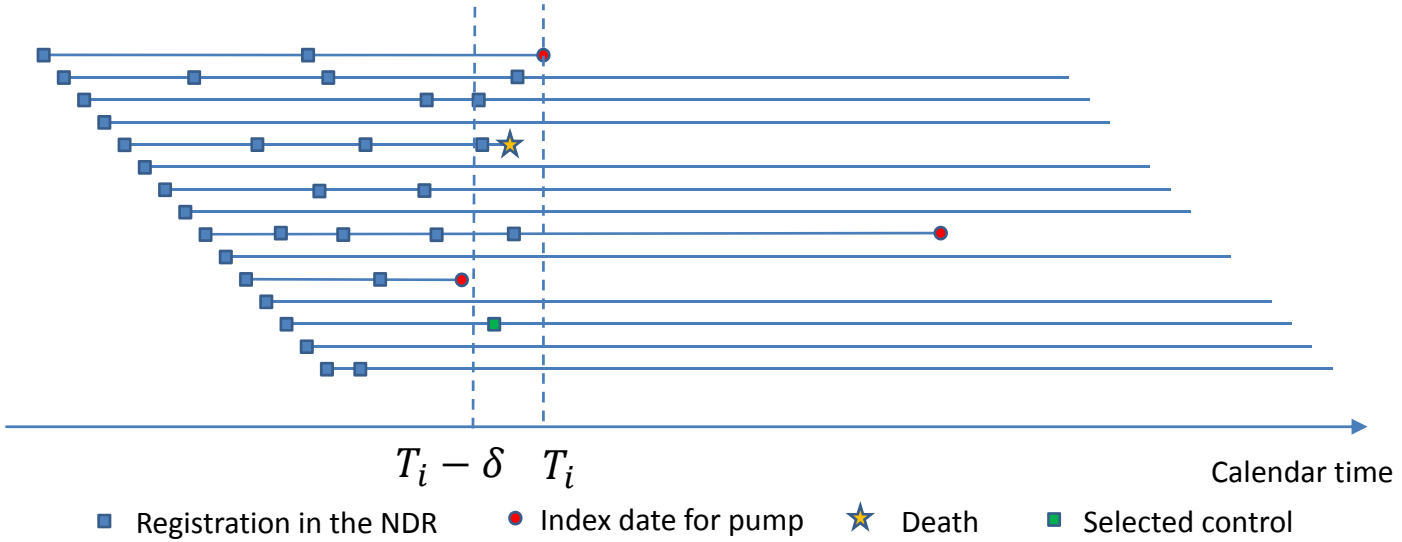
**) Leads to instrumental variable methods that can handle unobserved confounders, but good instruments are VERY hard to find

CENTRE OF REGISTERS
VÄSTRA GÖTALAND

Propensity scores with time dependent variables

Fit a Cox regression model with time updated covariates: $\lambda(t) = \lambda_0(t)\exp(\beta^T X(t))$

For each exposed, select without replacement the closest unexposed patient from the risk set, who are alive at T_i and who have the latest registration no earlier than δ days before using $\beta^T X(t)$ as a distance measure



The data from the index date for a pump user is regarded as being prior to treatment and the actual index date for the exposure is set to the date after this registration

The selected control is subsequently followed from T_i with respect to outcomes and may be censored if exposed to pump at a later date

CENTRE OF REGISTERS
VÄSTRA GÖTALAND

Propensity scores with time dependent variables

Propensity score: $P(T \mid X)$ constant over time…

often modelled using logistic regression

In the longitudinal world, consider time τ to exposure

Consider the probability of being exposed at time t given non exposed up to time t

or even $\lambda(t, X) = \lim_{\delta t \to 0} \frac{P(\tau \in [t, \delta t) \mid \tau > t, X)}{\delta t}$

Why not use Coxregression: $\lambda(t, X) = \lambda_0(t) exp(\beta X)$?

Match on βX

CENTRE OF REGISTERS
VÄSTRA GÖTALAND

Matching on attributes vs matching on PS

Exact attribute matching: $X_i = X_{i^*}$

- Virtually eliminates model dependence
- Curse of dimentionality
- Continuous variables has to be discretized

Mahanobis distance matching: $X_i \approx X_{i^*}$

- Reduces model dependence
- Works best for continuous variables

Propensity score matching: $\big(Y(0), Y(1)\big) \perp T \mid \boldsymbol{X}$

- Only gives balance on average, not between matched patients
- Model dependence

CENTRE OF REGISTERS
VÄSTRA GÖTALAND

Headline here

CENTRE OF REGISTERS
VÄSTRA GÖTALAND