CENTRE OF REGISTERS
VÄSTRA GÖTALAND

Evaluating the generalizability of an RCT using electronic health records data

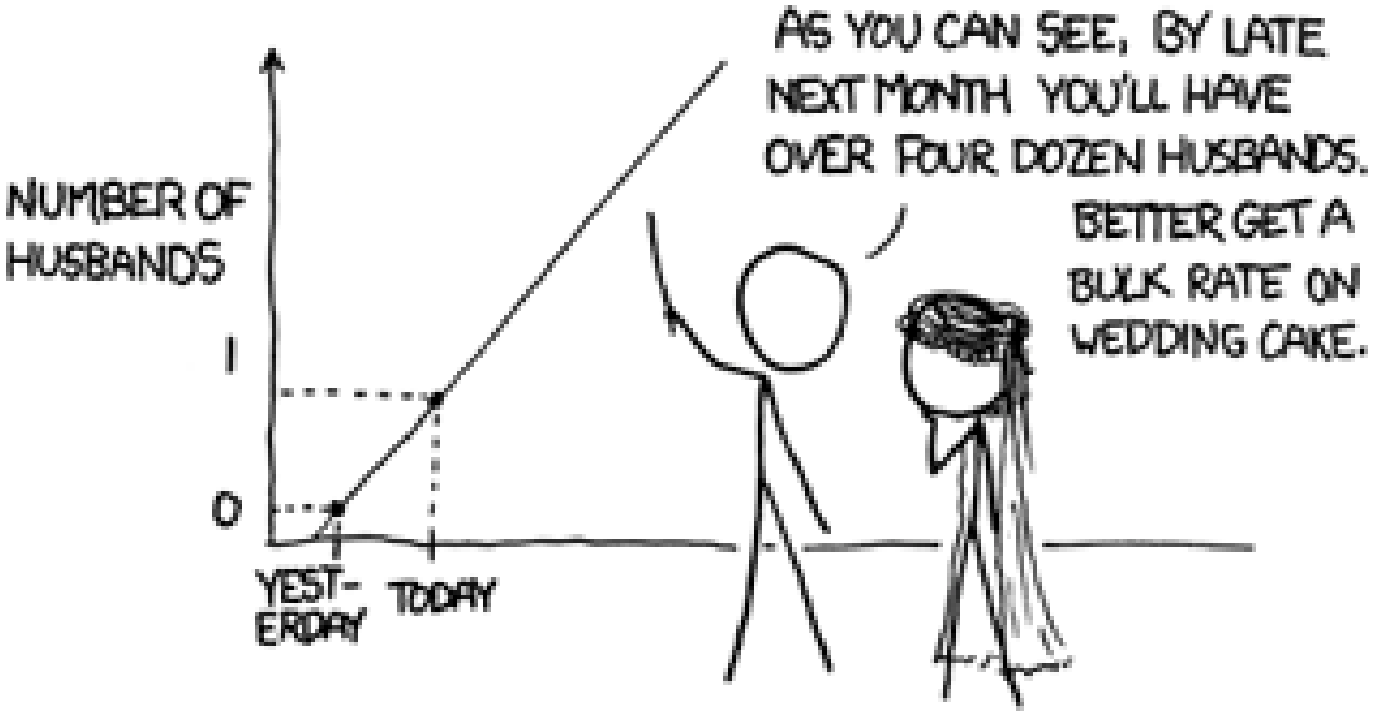
VÄSTRA GÖTALANDSREGIONEN

3 interesting questions

- Is our RCT representative?
- How can we generalize RCT results?
- Can we use EHR* data as a ”control” group?

*) Electronic Health Records

CENTRE OF REGISTERS
VÄSTRA GÖTALAND



The Danger of Overgeneralization

CENTRE OF REGISTERS
VÄSTRA GÖTALAND

Difference in data collection

The RCT data

- Patients included by criterias
- Regular visits
- Measurements according to protocol
- Extensite data cleaning
- Usually randomized

Missing could be non-informative

e.g. LDL missing due to lost sample

Placebo effect

The EHR data

- Patients included if they go to the doctor
- Visits as needed
- Measurements as needed
- Data cleaning?
- confounding

Missing is almost never non-informative

e.g. LDL not needed and thus ”missing”

CENTRE OF REGISTERS
VÄSTRA GÖTALAND

An illustrative(?) example

- One large global CV outcome study in primary prevention with liberal inclusion/exclusion criteria

- Three CPRD cohorts indicated for primary prevention of CV disease were created from a 6 year study window (2003-2008):
  - Cohort 1: patients meeting the trial's inclusion criteria and excluding those with prior CV history or CRP value indicating severe bacterial infection
  - Cohort 2: patients meeting NICE guidelines for recommended statin therapy in primary prevention of CV disease

CENTRE OF REGISTERS
VÄSTRA GÖTALAND

Sample sizes and variables

Samplesizes

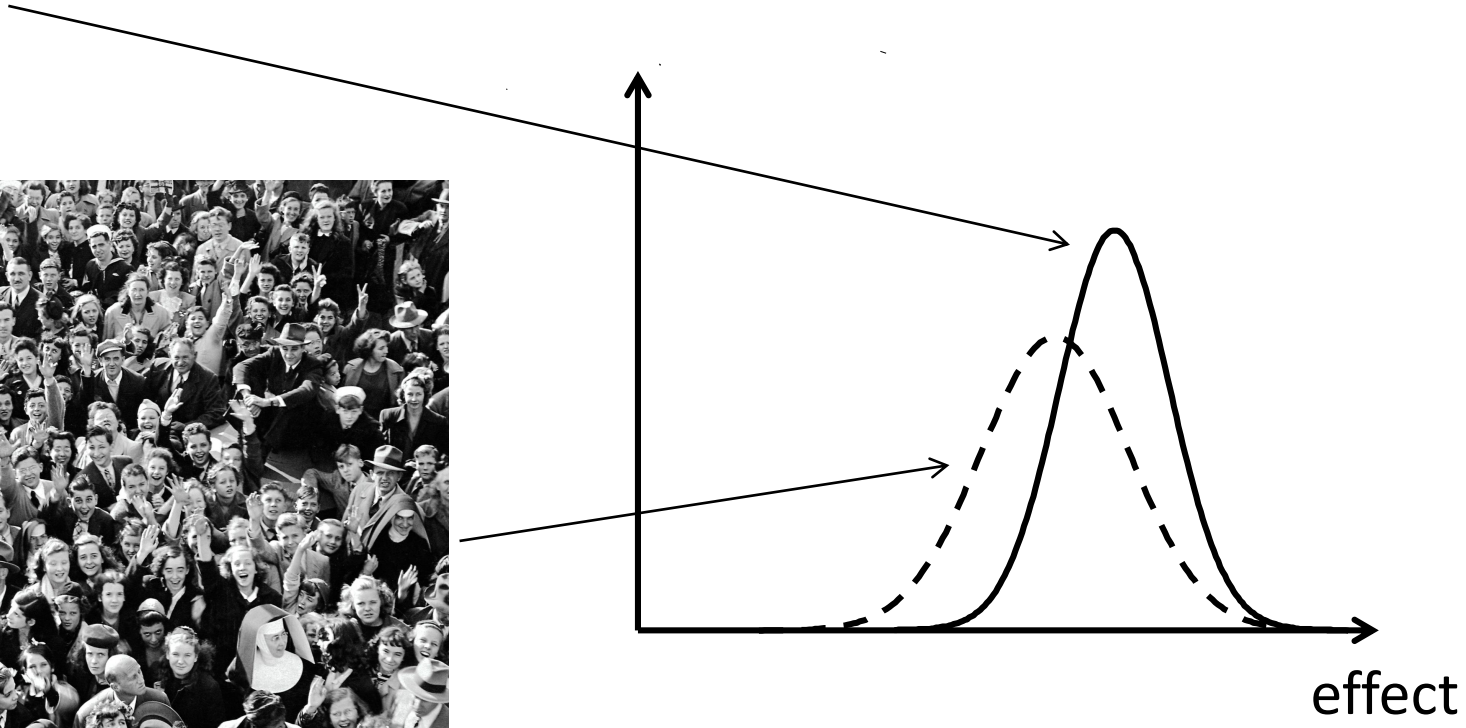
Data set	N	N complete cases	N after imputation	
RCT	17622	17622		
CPRD1	8892	894	8892	Realistic cohort
CPRD3	78008	11567	78008	Very broad cohort

The analysis was run on two different sets of variables:

- Framingham variables: AGE, BMI, TC, sex, smoking
- All* variables: AGE , ANTIHT_USE , ASA_USE , BMI , CRP , DBP , FPG , HDL , LDL , SBP , TC , TG , sex , smoking , weight

CENTRE OF REGISTERS
VÄSTRA GÖTALAND

If we have the RCT results, what can we say about the effect in ”different” population?



CENTRE OF REGISTERS
VÄSTRA GÖTALAND

Is our trial representative?

-compare patient characteristics

- One variable at a time
- All variables at the same time
  - Convex hull
  - Cross matching
  - (Linear) discriminant analysis

CENTRE OF REGISTERS
VÄSTRA GÖTALAND

Descriptive statistics per variable

	RCT mean	RCT std	EHR mean	EHR std
GENDER_MALE	0.52	0.50	0.62	0.49
AGE	69	9.6	66	7.7
WEIGHT	82	19	82	18
BMI	29	5.9	29	5.5
SBP	140	20	136	17
DBP	79	11	81	9.0
CIGS/DAY	13.6	11.6	12.9	9.5
SMOKER	0.10	0.30	0.16	0.36
FPG	106	36	95	12
GLUC	118	55	95	12
LDL	97	28	104	19
HDL	52	16	51	15
TG	158	69	138	73
TC	181	36	183	24
HBA1C	7.3	1.5	5.7	0.4
CRP	10	18	6.8	8.9
ASA USE	0.43	0.49	0.19	0.39
ANTIHT USE	0.84	0.36	0.50	0.50
MEDHIST DM	0.26	0.44	0.00067	0.023
FAMHIST CHD	0.31	0.46	0.11	0.32

Convex hull

Idea:

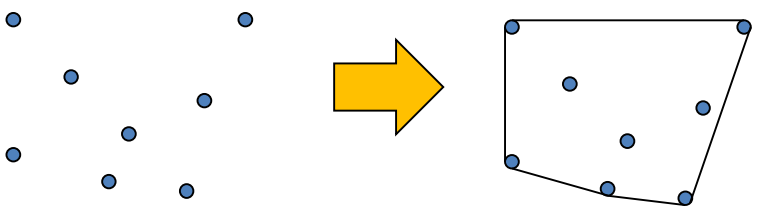
Construct the convex hull for the RCT patients and see how many RWE patients fall into that

Definition: the convex hull for a set S is the smallest convex set that contains s

1 d convex hull: The range

2 d convex hull:

K d convex hull: Wrap it in (stiff) paper…



Example:

```
matchit(formula = trial ~ AGE + BMI+ CRP + SBP + DBP +
sex + LDL + FPG + CR_CL + TC + TG  + HDL,
      data=anadata,
      method="nearest",
      discard="hull.control")
```

Sample sizes:

	EHR patients
All	3933
In the RCT convex hull	254
Discarded	3679

Almost all of the EHR patients lies outside of the convex hull for the RCT patients.

It's enough to be extreme on one variable (or in one direction)…

CENTRE OF REGISTERS
VÄSTRA GÖTALAND

Cross matching

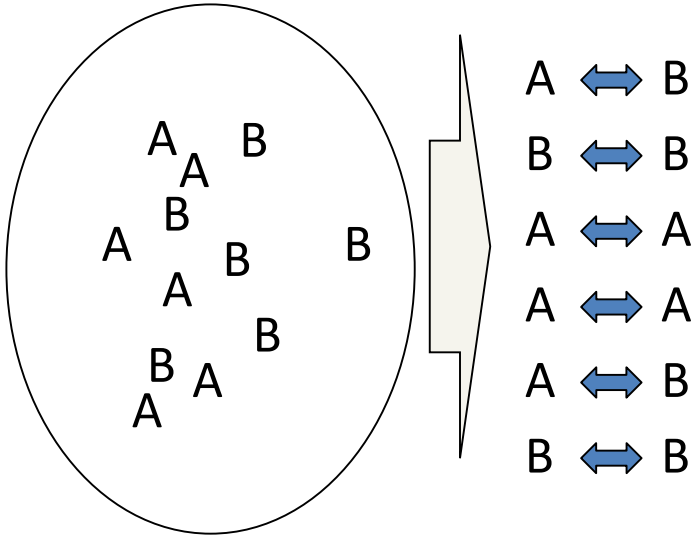
Cross matching as a test comparing multivariate distributions Rosenbaum (2005)

Idea: Merge all the data

Create match pairs using the Mahalanobis distance

Count the number of cross matches (A matched to B)

The number of cross matches has a know distribution under H_0



Observed number of cross matches: 907

Expected number of cross matches under H_0: 1183

P-value: 10**(-35)

Indicates that the RCT and RWE populations differ

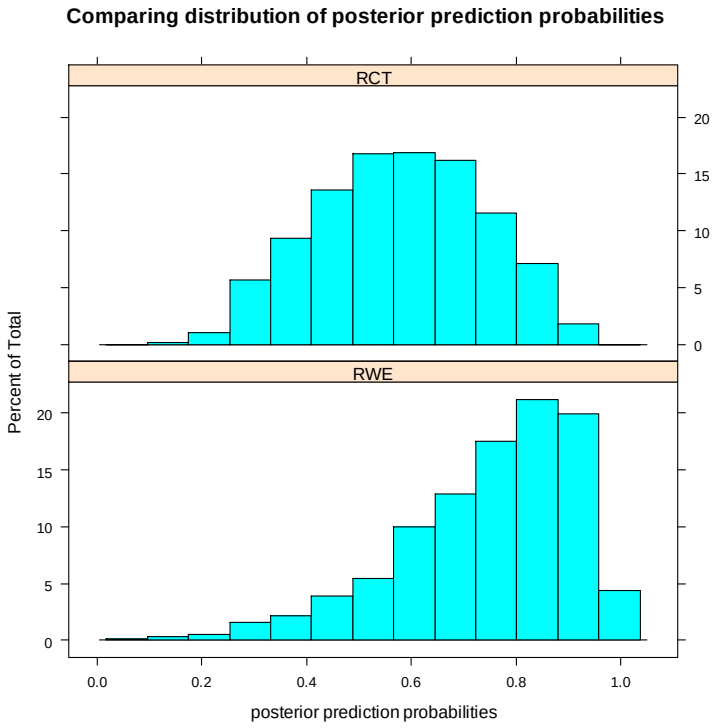
CENTRE OF REGISTERS
VÄSTRA GÖTALAND

Linear discriminant analysis

Try linear discriminant analysis to find a linear function that separates RCT and EHR patients

Coefficients of linear discriminants:

	LD1
AGE_I	-0.066845646
BMI	0.014905047
CRP	-0.011794747
SBP	-0.016440719
DBP	0.044964972
LDL	0.839770845
FPG	-0.214263226
CR_CL	-0.013896685
TC	0.005260017
TG	-0.163245948
HDL	-0.437969649



Mean=0.58

Mean=0.75

CENTRE OF REGISTERS
VÄSTRA GÖTALAND

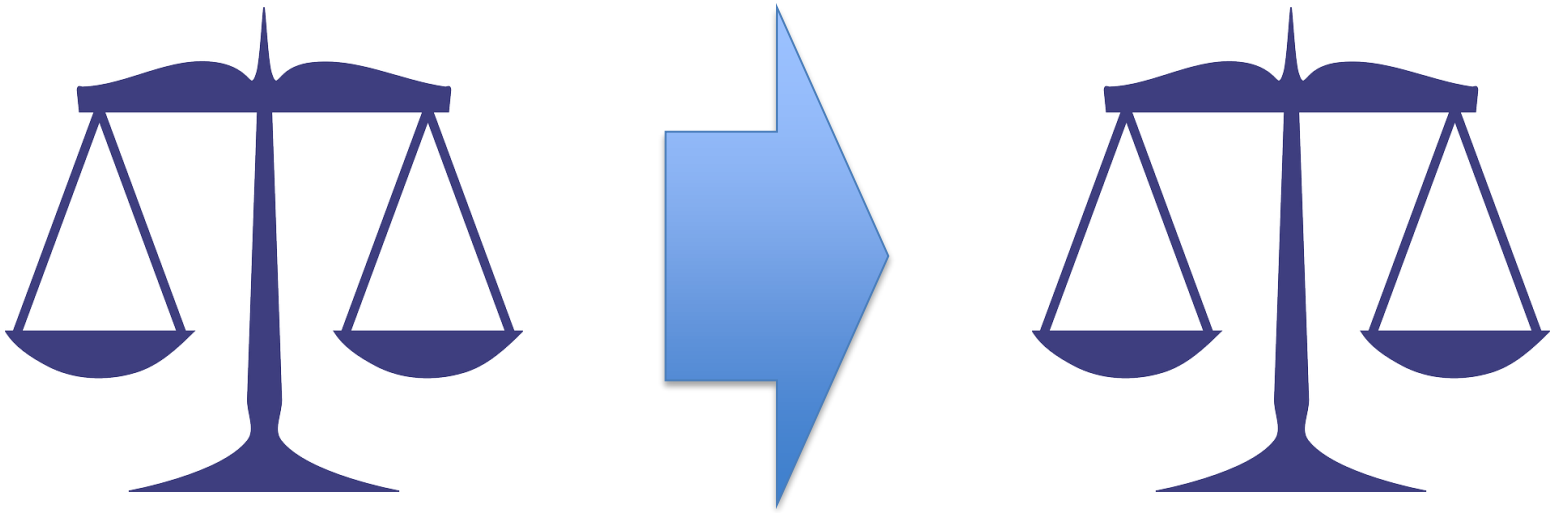
Next steps

- Propensity score (Stuart &Cole 2010, 2011)
- Cross design synthesis (Kaizar 2009)
- Hierarichical models (Prevost et al 2000)

CENTRE OF REGISTERS
VÄSTRA GÖTALAND

The nice thing about propensity score

$$e(X) = P(T_i|X_i)$$



$e(X)$ Xi

CENTRE OF REGISTERS
VÄSTRA GÖTALAND

For us…

- S_i indicates membership in the RCT sample
- T_i indicates treatment assignment $T_i = \{1,0\}$
- covariates X
- potential outcomes:
  - $Y_i(1)$ would be observed under treatment
  - $Y_i(0)$ would be observed under control

The sample average treatment effect

$$\text{SATE} = \frac{1}{n} \sum_{i \in \{s_i = 1\}} \big(Y_i(1) - Y_i(0)\big)$$

The population average treatment effect

$$\text{PATE} = \frac{1}{N} \sum_{i=1}^{N} \big(Y_i(1) - Y_i(0)\big)$$

CENTRE OF REGISTERS
VÄSTRA GÖTALAND

Key assumptions

All patients in the population have some probability of being in the trial and no patients are always in the trial

$$0 < P(S_i = 1|X_i) < 1$$

Inclusion in the trial does not depend on the potential outcomes except though the covariates X

$$S \perp \big(Y(0), Y(1)\big)|X$$

Treatment assignment does not depend on the inclusion into the trial or the potential outcomes except through X

$$T \perp \big(S, Y(0), Y(1)\big)|X$$

CENTRE OF REGISTERS
VÄSTRA GÖTALAND

Propensity score as a distance meaure

$$\Delta_p = \frac{1}{n} \sum_{i \in \{S_i = 1\}} e_i(x) - \frac{1}{N-n} \sum_{i \in \{S_i = 0\}} e_i(x)$$

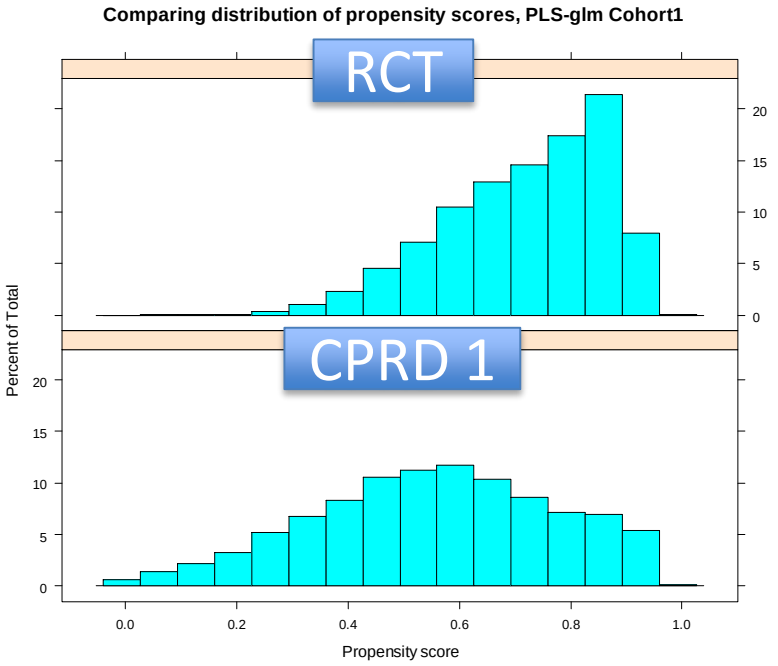
So, what consitues a ”big” differnece?

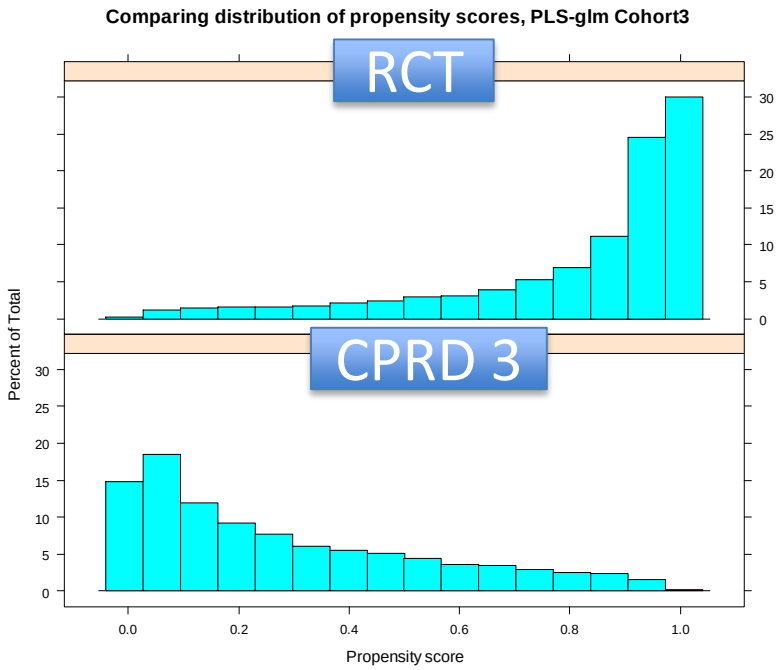
Suggestion: big if $\Delta_p > c \times \sigma(\hat{p}_i)$

C=0.25 or 0.1 has been suggested…

CENTRE OF REGISTERS
VÄSTRA GÖTALAND

Propensity score, all varibles



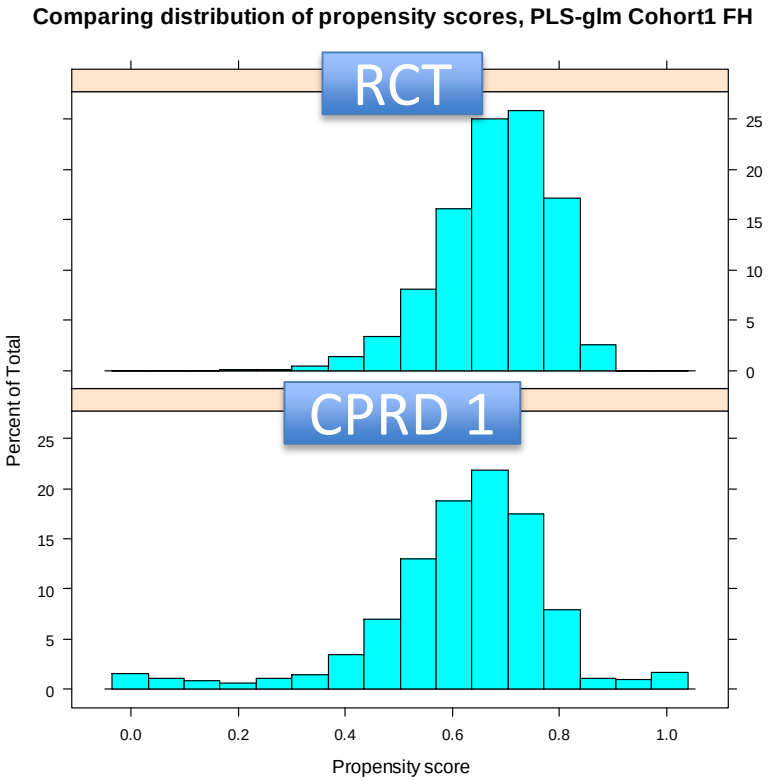


Δ=0.165

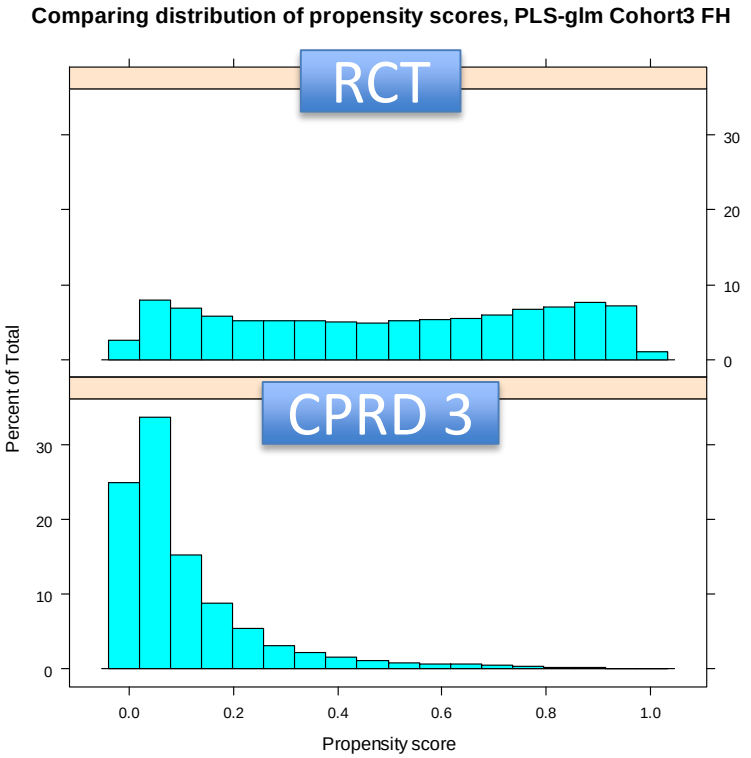
Δ=0.478

CENTRE OF REGISTERS
VÄSTRA GÖTALAND

Propensity scores based on risk factors



Δ=0.06



Δ=0.38

CENTRE OF REGISTERS
VÄSTRA GÖTALAND

Predict the results in the EHR population (PATE)

IPSW : Inverse probability of selection weight

Weight: $$w_i = \frac{P(S_i = i)}{P(S_i = 1 | X_i)}$$

In our case the endpoint is a time to event so we'll fit a weighted Cox proportional hazards model with the partial likelihood

$$L(\beta) = \prod_{i=1}^{n} \left[\frac{\exp(\beta X_i) \times w_i}{\sum_{k=1}^{n} 1_{\{i \in R_k(t_i)\}} \exp(\beta X_k) \times w_k} \right]^{Y_i}$$

Stuart & Cole 2011

CENTRE OF REGISTERS
VÄSTRA GÖTALAND

Predicted treatment effect in the example

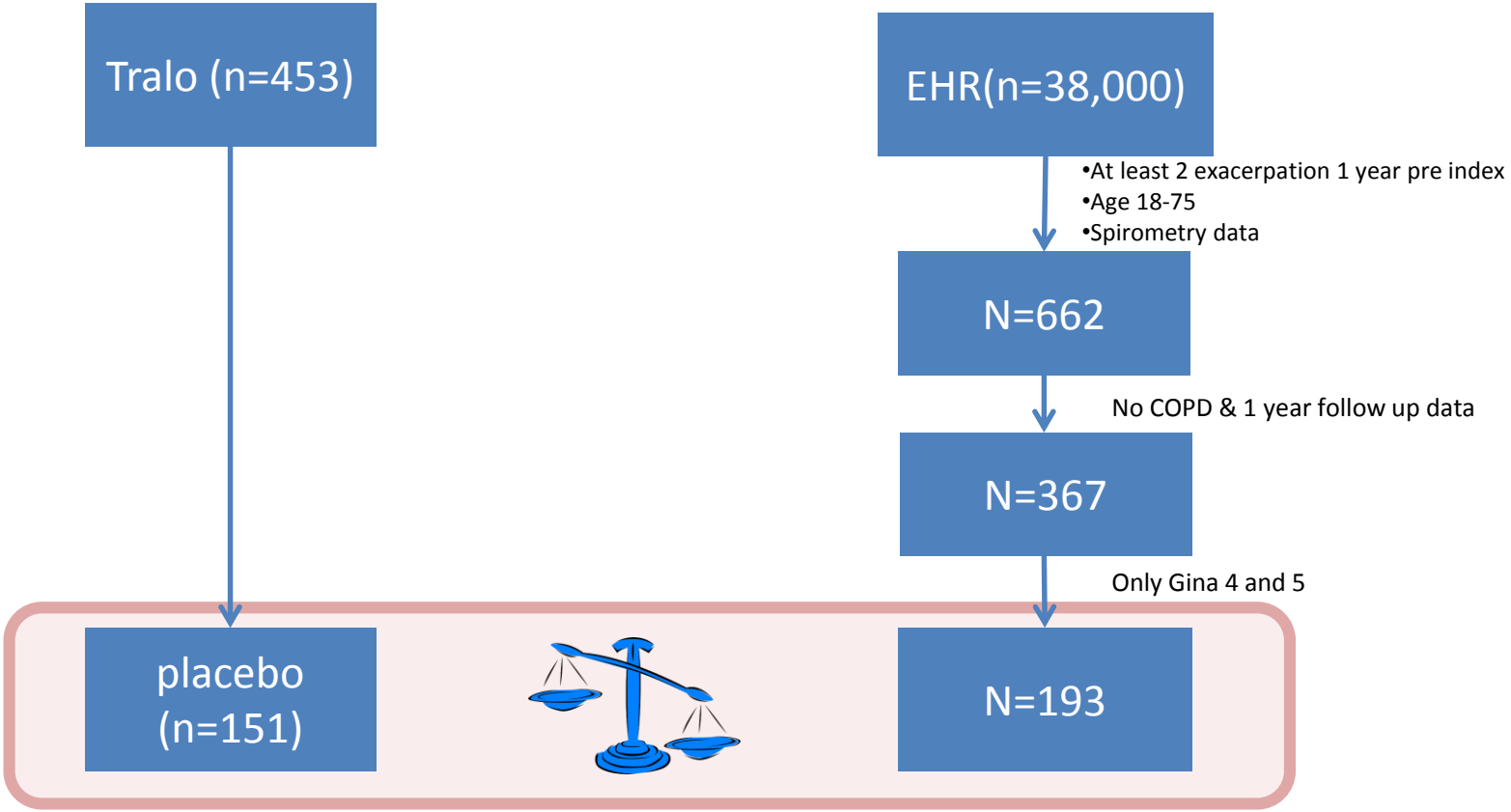
Cohort	Variables used	Hazard ratio	Lower	Upper	p-value
RCT		0.551	0.448	0.678	<0.0001
Cohort1	All	0.555	0.462	0.666	<0.0001
Cohort1	Framingham	0.557	0.460	0.675	<0.0001
Cohort3	All	0.607	0.519	0.709	<0.0001
Cohort3	Framingham	0.785	0.728	0.847	<0.0001

- The predicted treatment effect is slightly closer to 1 which indicates an under representation of ”low” risk patients in the Jupiter population
- The difference is smallest for cohort 1 and largest for cohort 3
- Risk factors don’t account for all confounding(?)

CENTRE OF REGISTERS
VÄSTRA GÖTALAND

What's next?

pacEHR



iptw

CENTRE OF REGISTERS VÄSTRA GÖTALAND

EHR data as a control group

Weight the EHR data by: $\frac{P(S_i = 1|X_i)}{P(S_i = 0|X_i)}$

Weight the HER to estimate the treatment outcome that would have been observed if the HER data had the same distribution of patients characteristics as the RCT

Sort of like estimating the ATT in an observational study….

CENTRE OF REGISTERS
VÄSTRA GÖTALAND

Doing without patient level RCT data…

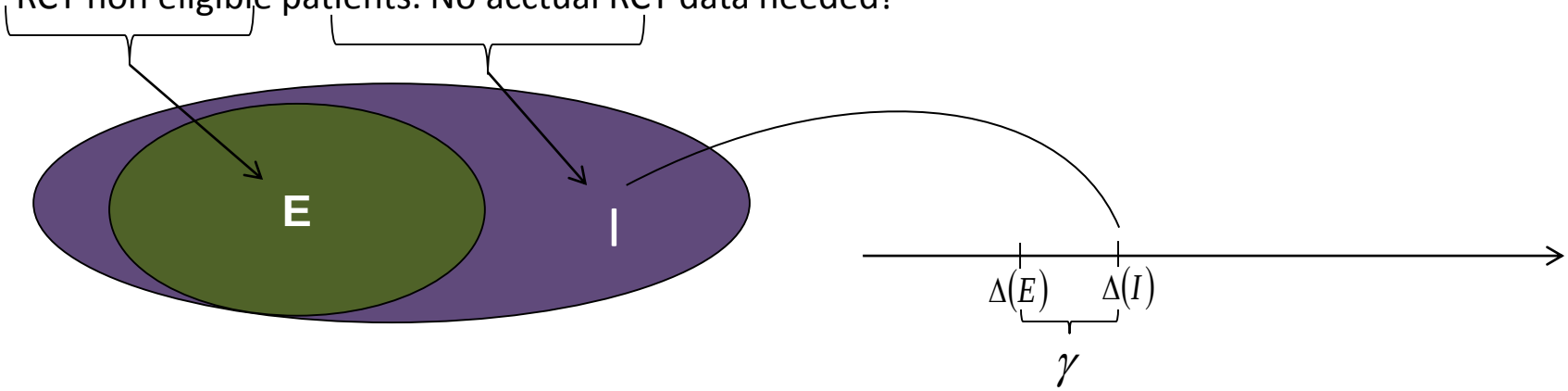
Evaluate the generalizability using Presslee's method

Use weights from the method of moments (Signorovich 2012)

CENTRE OF REGISTERS
VÄSTRA GÖTALAND

Doing without tyhe RCT patient data 1

Idea: Use the RCT inclusion/exclusion critera to split a registry cohort into RCT eligible and RCT non eligible patients. No acctual RCT data needed!



Define the sub population average treatment effect in each sub population

PATE $$\Delta = \frac{1}{N}\sum_{i=1}^{N}(Y_i(1)-Y_i(0))$$

SPATE(I) $$\Delta(I) = \frac{1}{N_I}\sum_{i=1}^{N}1_{\{i\in I\}}(Y_i(1)-Y_i(0))$$

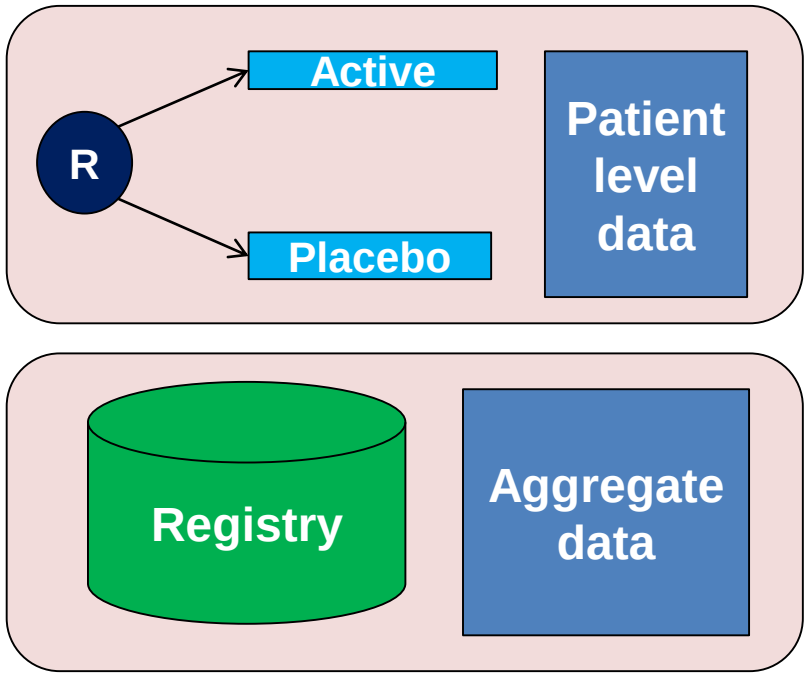
SPATE(E) $$\Delta(E) = \frac{1}{N_E}\sum_{i=1}^{N}1_{\{i\in E\}}(Y_i(1)-Y_i(0))$$

Measure the generalization error by comparing outcomes for I and the whole population

$$\gamma = \Delta(I)-\Delta$$
$$= \pi_E(\Delta(I)-\Delta(E))$$
$$\hat{\gamma} = \hat{\pi}_E\left(\hat{\Delta}(I)-\hat{\Delta}(E)\right)$$

CENTRE OF REGISTERS
VÄSTRA GÖTALAND

Doing without patient level data 2



Weight the registry on
- Gender
- Age
- BMI
- SBI
- LDL
- TG
- Smoking
- Outcome Y

x_i, y_i

$\bar{x}_C, \bar{y}_C$

Idea: Use $\hat{\theta} = \frac{\sum y_i w_i}{\sum w_i} - \bar{y}_C$ where $w_i = \frac{P(S_i = 1 | x_i)}{P(S_i = 0 | x_i)}$

Estimate the weights using logistic regression: $w_i = \exp(\alpha + x_i^T \beta)$

Method of moments, solve $\frac{\sum x_i \exp(x_i^T \hat{\beta})}{\sum \exp(x_i^T \hat{\beta})} - \bar{x}_C = 0$

CENTRE OF REGISTERS
VÄSTRA GÖTALAND

References

Imai K, King G, Stuart EA. Misunderstandings between experimentalists and observationalists about causal inference. Journal of the Royal Statistical Society A. 2008; 171: 481-502

Stuart EA, Cole SR, Bradshaw CP, Leaf PJ. The use of propensity scores to assess the generalizability of results from randomized trials. Journal of the Royal Statistical Society A. 2011; 174: 369-386

Cole SR, Stuart EA. Generalizing evidence from randomized clinical trials to target populations. American Journal of Epidemiology. 2010; 172: 107-115

Rosenbaum PR. An exact distribution-free test comparing two multivariate distributions based on adjacency. Journal of the Royal Statistical Society B. 2005; 67: 515-530

Prevost TC, Abrams KR, Jones DR. Hierarcical models in generalized synthesis of evidence: an example based on studies of breast cancer screening. Statistics in Medicine. 2000; 19: 3359-3376

Signorovich et al Matching-Adjusted Indirect Comparisons: A New Tool for Timely Comparative Effectiveness Research. Value in Health. 2012; 12: 940-947

Pressler T. R. Kaizar E. E. The use of propensity scores and observational data to estimate randomized controlled trial generalizability bias . Statistics in Medicine. 2013; 32: 3552-3568

CENTRE OF REGISTERS
VÄSTRA GÖTALAND