Propensity Score Diagnostics

Dr Emily Granger

Dr Jamie C. Sergeant

Dr Mark Lunt

41st Annual Conference of the International Society for Clinical Biostatistics, 26th August 2020

The University of Manchester

emily.granger@lshtm.ac.uk

@EGranger90

Background

- Commonly used in observational data to deal with confounding bias.

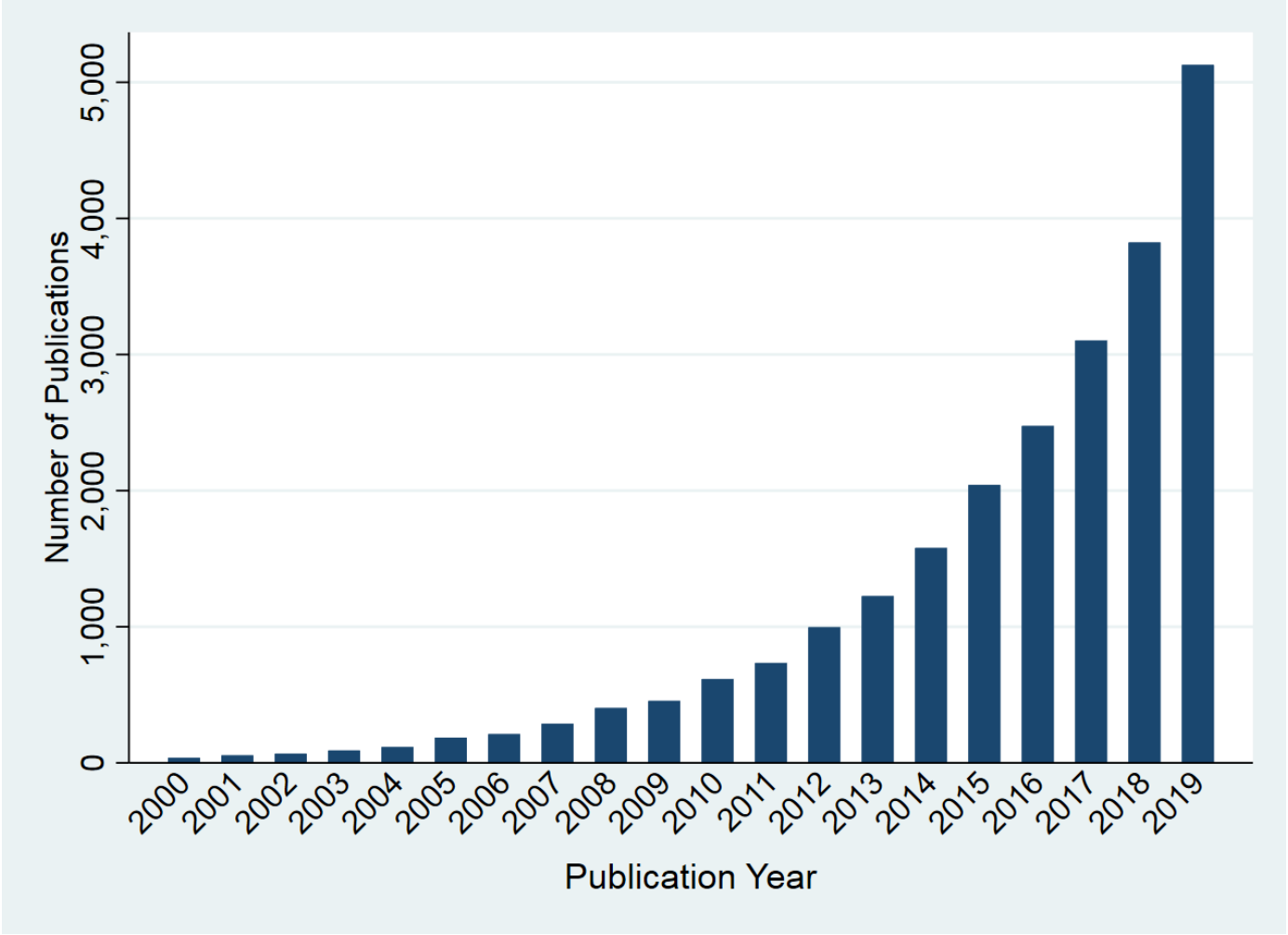


Figure 1: Number of propensity score publications in medical research by year

Background

- Commonly used in observational data to deal with confounding bias.
- Poorly estimated propensity scores may lead to biased estimates.
- Use of diagnostics to assess propensity scores is important.
- Currently unknown how best to assess propensity scores.

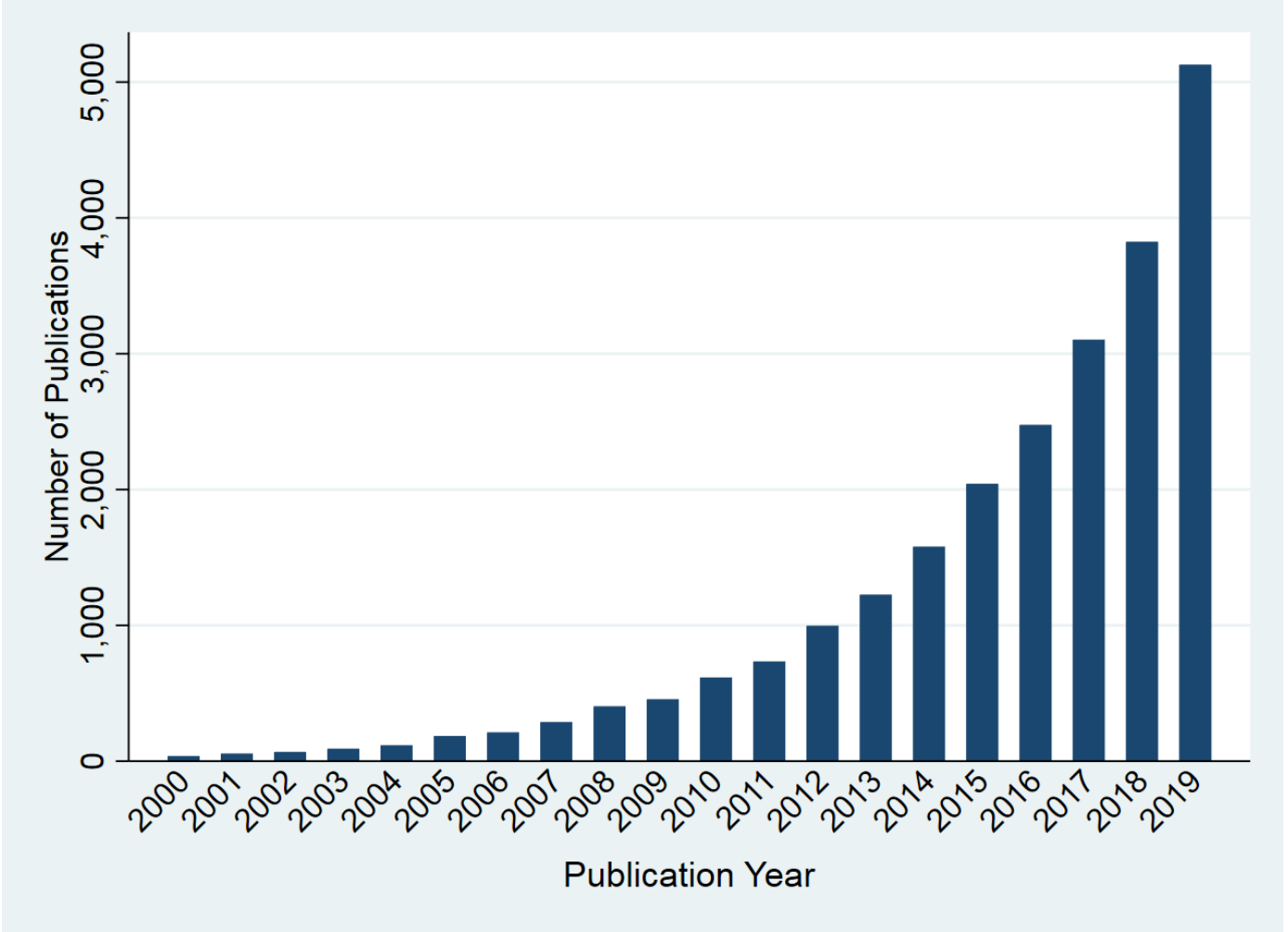


Figure 1: Number of propensity score publications in medical research by year

Aims of research

Aim 1:

Review and compare the existing propensity score diagnostics.

Aim 2:

Develop guidelines for how to build and assess propensity score models.

Aims of research

Individual Diagnostics

Overall Diagnostics

Aim 2:

Develop guidelines for how to build and assess propensity score models.

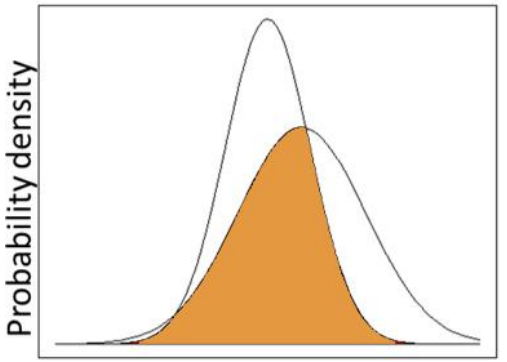
Individual diagnostics

Mean-based

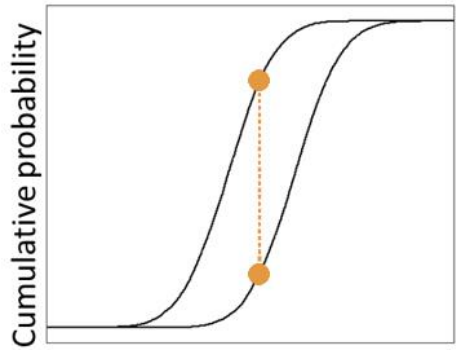
- Standardised difference (SD)
- t-test statistic (t)
- Percent reduction in mean difference (PR)

Distribution-based

- Overlapping coefficient (OVL)



- Kolmogorov-Smirnov Statistic (KS)

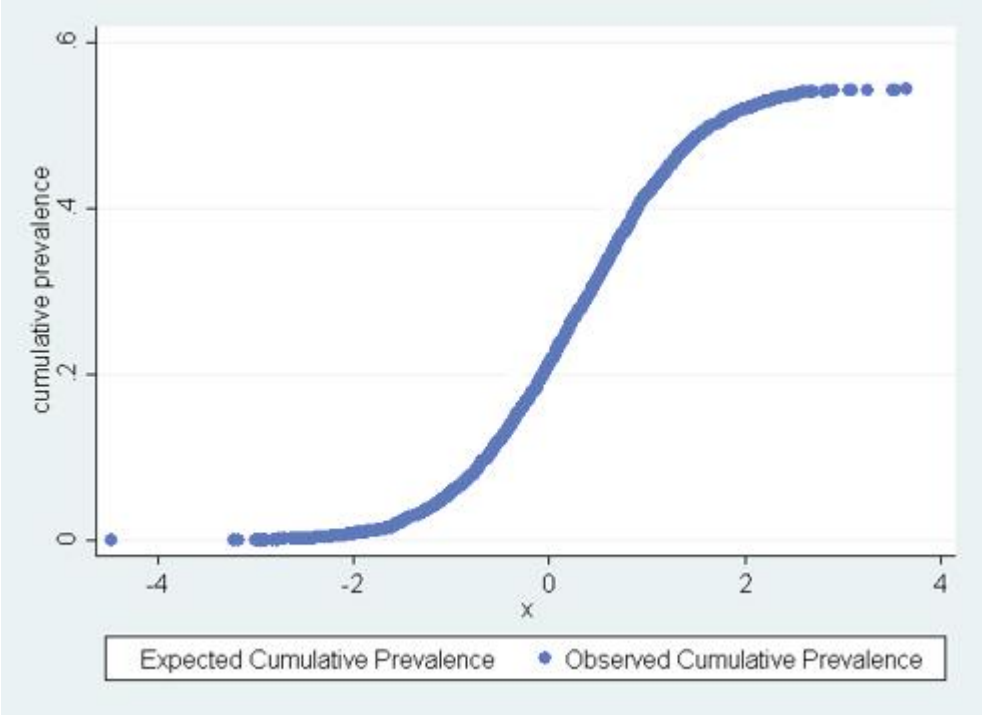


Cumulative prevalence of exposure

Notation: exposure indicator for subject i: E_i, propensity score for subject i: PS_i, sample size: n.

For continuous variable X:

- $OCP_X(X_0) = \frac{1}{n}\sum_{i:X_i \leq X_0} E_i$

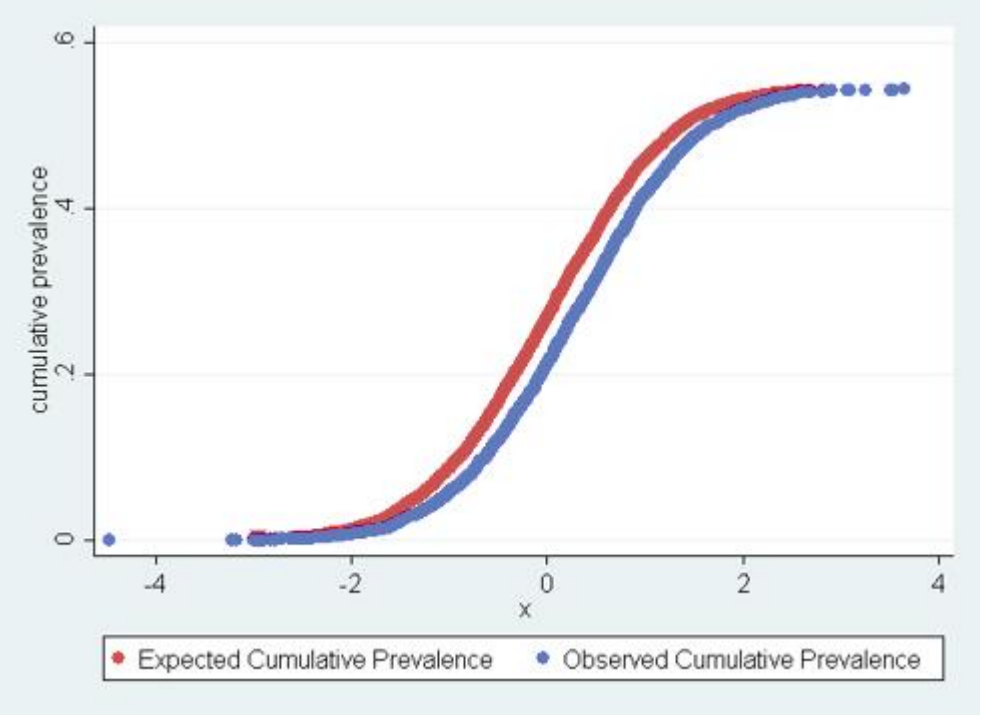


Cumulative prevalence of exposure

Notation: exposure indicator for subject i: E_i, propensity score for subject i: PS_i, sample size: n.

For continuous variable X:

- $OCP_X(X_0) = \frac{1}{n}\sum_{i:X_i \leq X_0} E_i$
- $ECP_X(X_0) = \frac{1}{n}\sum_{i:X_i \leq X_0} PS_i$

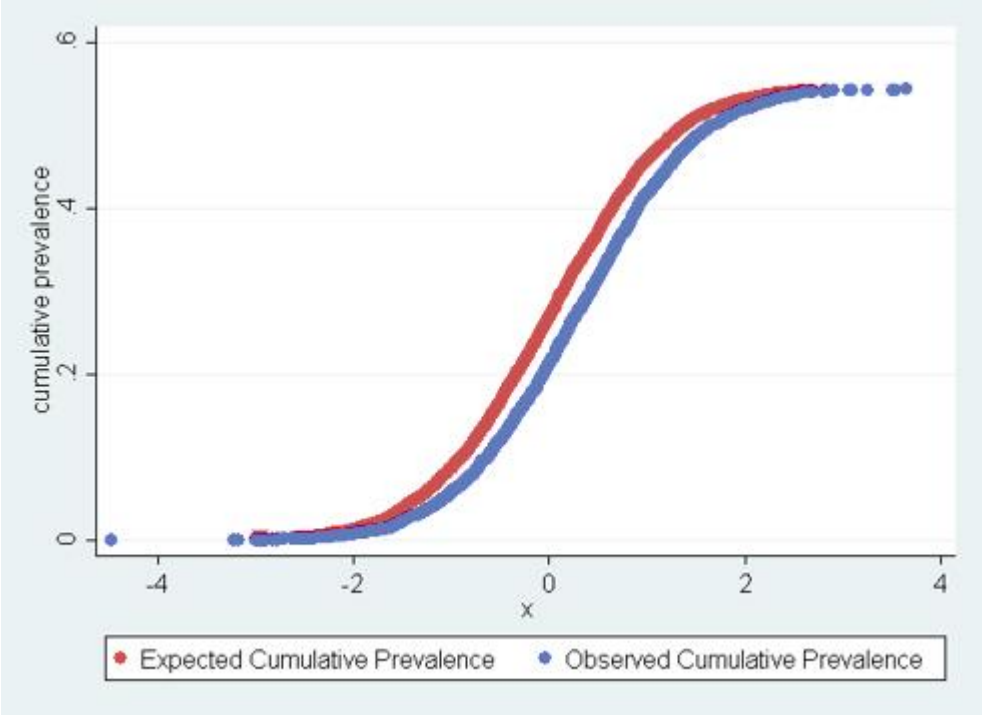


Cumulative prevalence of exposure

Notation: exposure indicator for subject i: E_i, propensity score for subject i: PS_i, sample size: n.

For continuous variable X:

- $OCP_X(X_0) = \frac{1}{n}\sum_{i:X_i \leq X_0} E_i$
- $ECP_X(X_0) = \frac{1}{n}\sum_{i:X_i \leq X_0} PS_i$
- $D_X = |\, OCP_X - ECP_X \,|$



Simulated data

Propensity score model:

- logit(PS)=$\alpha_0 + \alpha_1 X_1 + \alpha_2 X_2 + \cdots + \alpha_7 X_7 + \alpha_8 X_8$

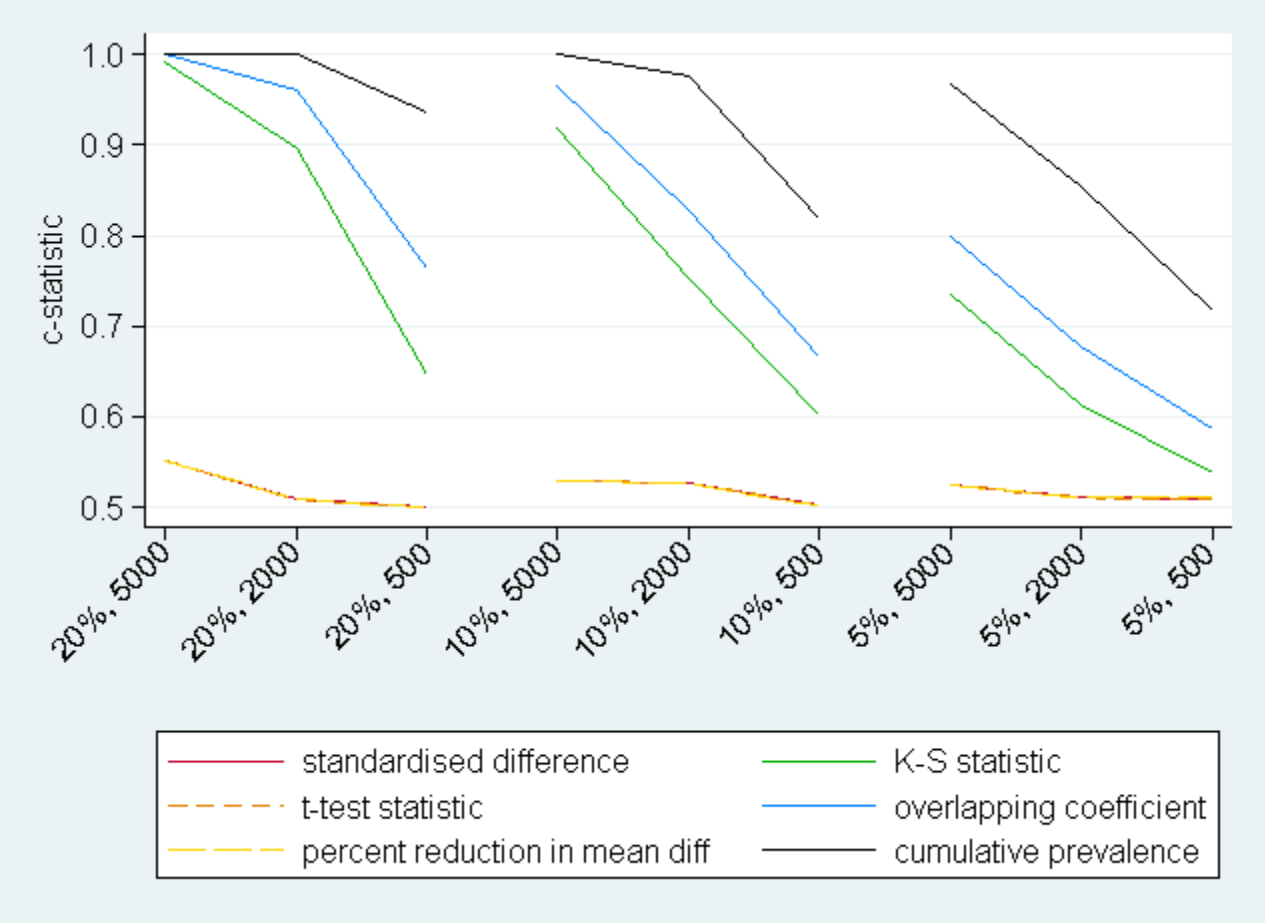
Variation between scenarios:

		Correct PS: / **Incorrect PS:**
S1: $X_8 = 0.4(3.5^{X_1} - 1)$	Nonlinearity added (monotonic)	$X_8 = 0$
S2: $X_8 = X_4 X_5$	Binary-binary interaction	$X_8 = 0$
S3: $X_8 = X_4 X_1$	Binary-continuous interaction	$X_8 = 0$
S4: $X_8 = X_1 X_2$	Continuous-continuous interaction	$X_8 = 0$

Data Analysis:

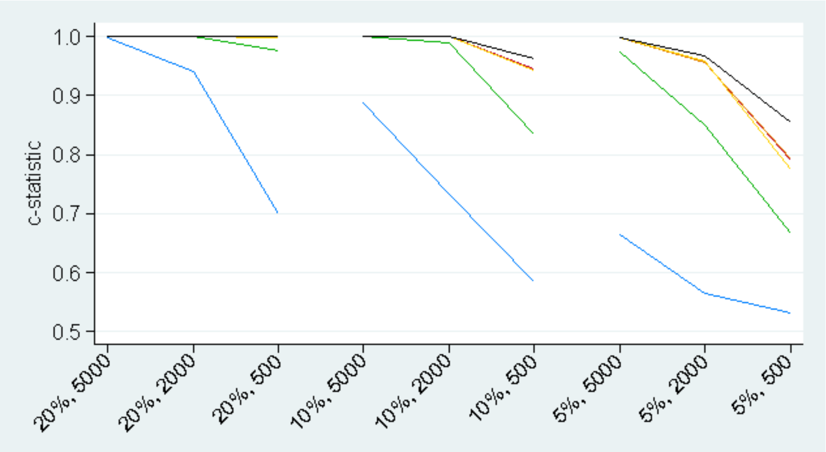
- Logistic Regression: PS indicator ~ PS diagnostic

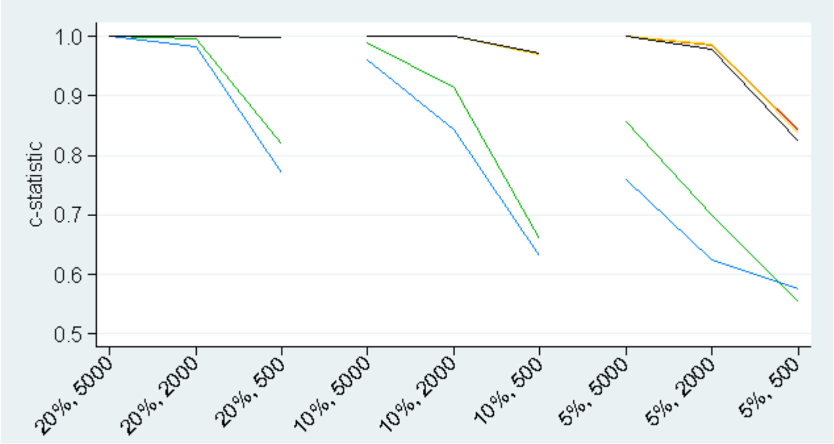
Scenario 1: Misspecification of a non-linear term

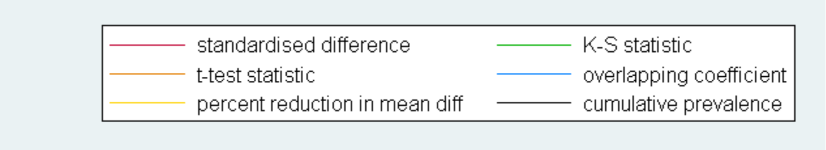


*diagnostics used to assess balance/specification the non-linear covariate X_1; balance assessed in PS-matched samples.

Scenarios 2-4: Omission of an interaction term







Figures:

- Scenario 2 (top left) binary-binary
- Scenario 3 (top right) binary-continuous
- Scenario 4 (bottom left) continuous-continuous

*diagnostics used to assess balance/specification the interaction term;
balance assessed in PS-matched samples.

Overall diagnostics

Disease risk score

- Disease risk scores (DRS) defined as predicted outcome under the control condition
- Standardised mean difference in DRS as a propensity score diagnostic [Stuart et al. 2013]

Weighted average of balance

Let w_j denote the weight for covariate X_j.
Then:

- $w_j = \gamma_j Std.Dev(X_j)$ [Caruana et al. 2015]
  - γ_j is the coefficient for X_j obtained after regressing outcome on X_j.

Balance measured using either:

- Standardised difference (SD)
- Overlapping coefficient (OVL)
- Kolmogorov-Smirnov Statistic (KS)

Simulated data

Propensity score model:

- logit(PS)=$\alpha_0 + \alpha_1 X_1 + \alpha_2 X_2 + \cdots + \alpha_7 X_7$

Outcome model:

- Y=$\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_7 X_7 + \beta_8 X_8$

Variation between scenarios:

S1: $X_8 = 0$	Independent baseline covariates
S2: $X_8 = 0$	Correlated baseline covariates
S3: $X_8 = 0.2(6.0^{X_1} - 1)$	Monotonic non-linearity

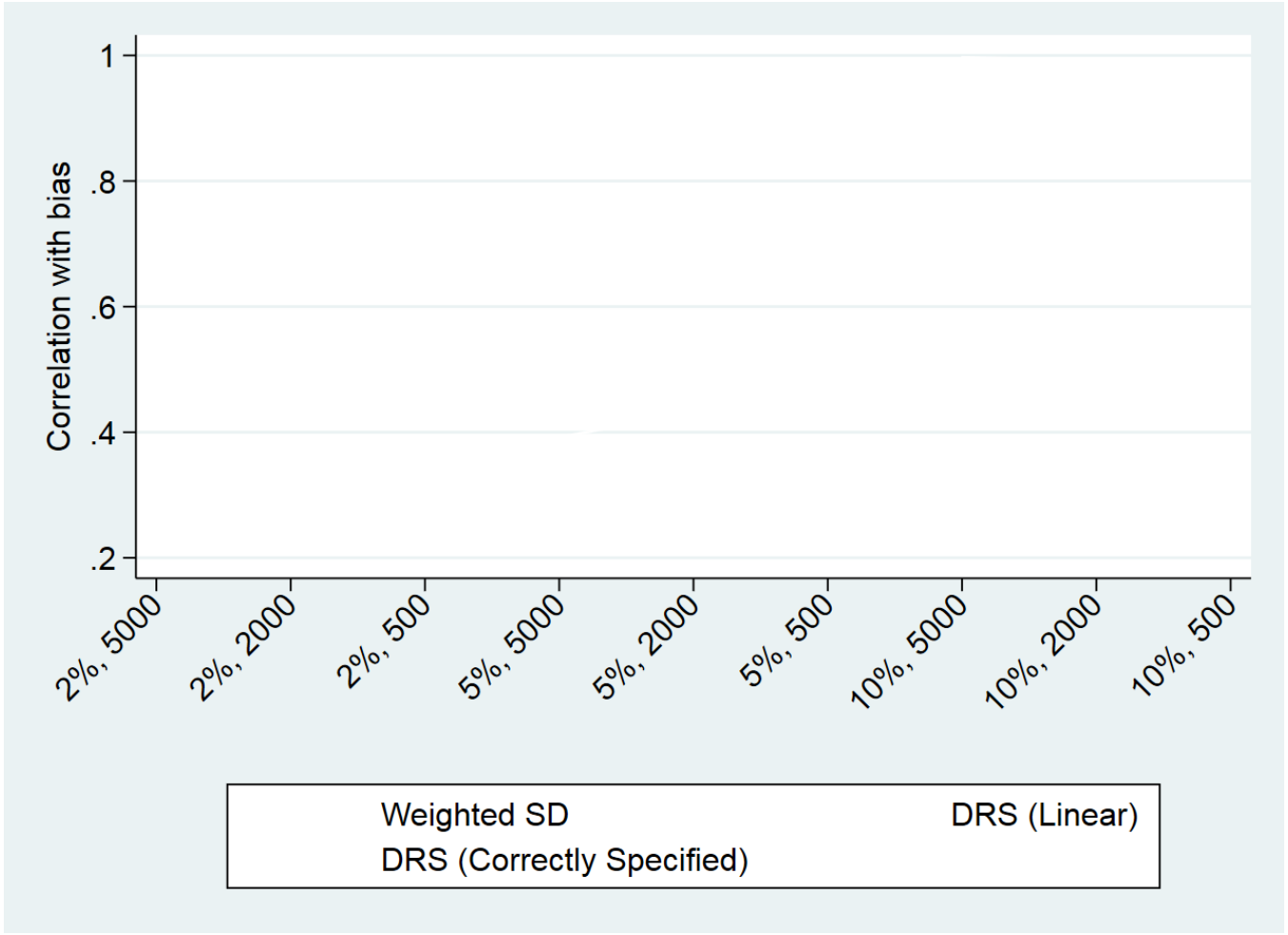
Scenarios 1 and 2: Linear outcomes

Table 1: Spearman rank correlation between overall diagnostics and bias (sample sizes 5000)

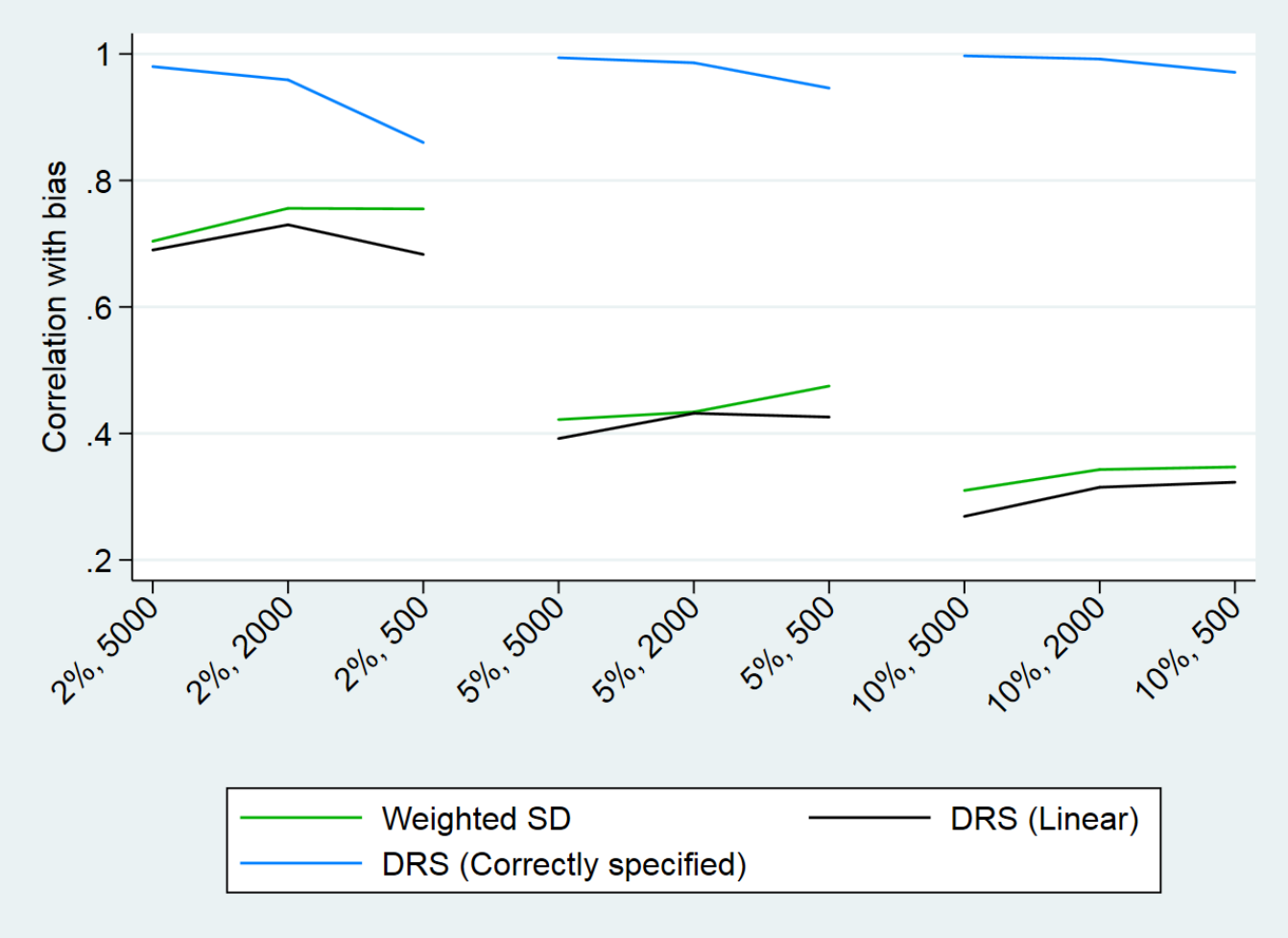
Scenario	Weighted SD	Weighted KS	Weighted OVL	SD(DRS)
Scenario 1	0.992	0.137	0.012	1.000
Scenario 2	0.129	0.102	0.031	1.000

*SD: Standardised difference; KS: Kolmogorov-Smirnov statistic; OVL: Overlapping coefficient; DRS: Disease risk score

Scenario 3: Non-linear term in outcome model



Scenario 3: Non-linear term in outcome model



Conclusions

- Cumulative prevalence (CP) diagnostic most useful for identifying all types of propensity score misspecification.
- Standardised mean difference in the disease risk score (DRS) is a promising overall diagnostic.
- Main limitation:
  - DRS not robust to misspecifications in the outcome model
  - Could use CP diagnostics to check specification
  - Future research into different estimation methods for the DRS.

Aims of research

Aim 1:

Review and compare the existing propensity score diagnostics.

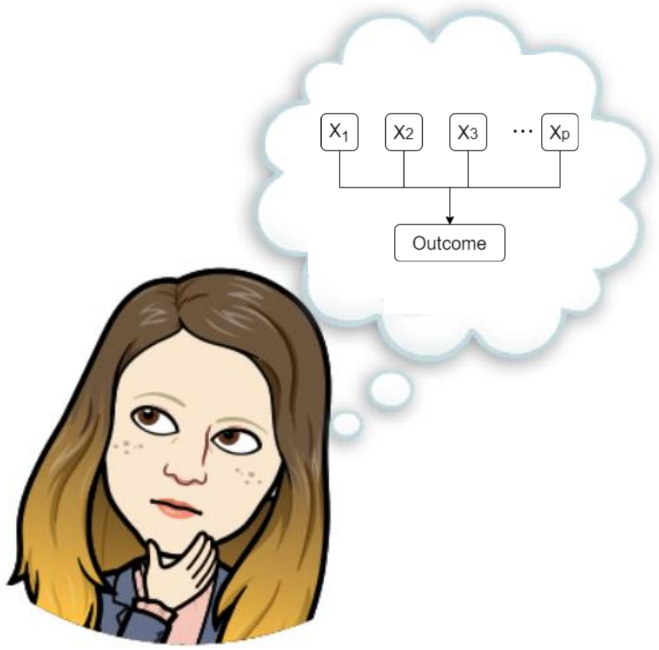
Aim 2:

Develop guidelines for how to build and assess propensity score models.

Proposed guidelines for propensity score assessment

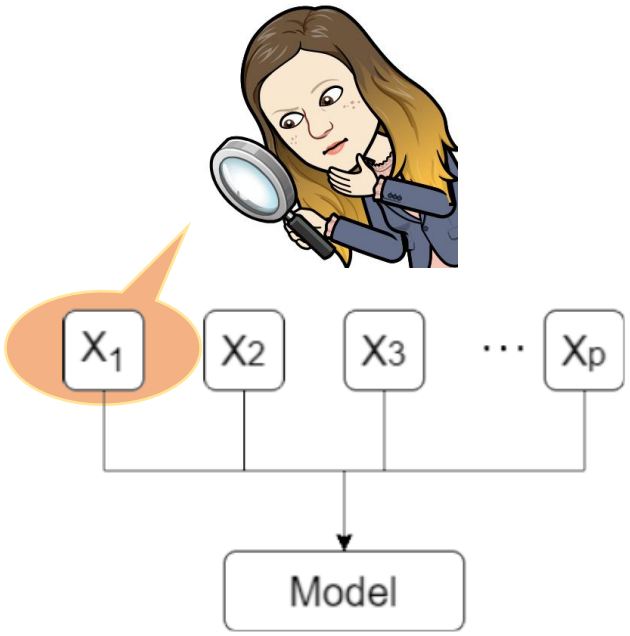
STEP 1:

Choose variables



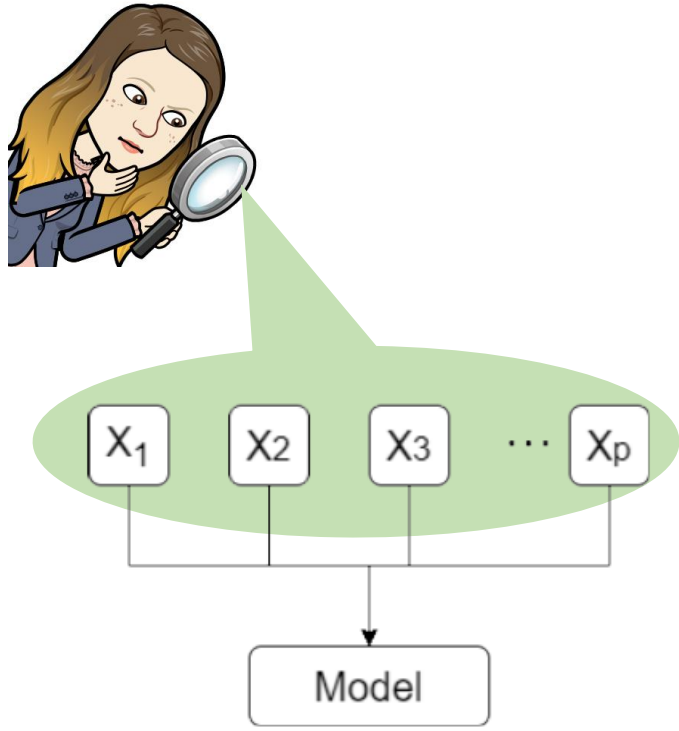
STEP 2:

Check individual covariates using CP diagnostics



STEP 3:

Check overall balance using DRS



References

[1] Stuart, EA et al. Prognostic score-based balance measures for propensity score methods in comparative effectiveness research. *Journal of Clinical Epidemiology.* 2013

[2] Caruana, E et al. A new weighted balance measure helped to select the variables to be included in a propensity score model. *Journal of Clinical Epidemiology.* 2015.

[3] Belitser, SV et al. Measuring balance and model selection in propensity score methods. *Pharmacoepidemiology and Drug Safety.* 2011.

[4] Ali, MS et al. Reporting of covariate selection and balance assessment in propensity score analysis is suboptimal: a systematic review. *Journal of Clinical Epidemiology.* 2008.

[5] Granger, E et al. A review of the use of propensity score diagnostics in papers published in high-ranking medical journals. *BMC Research Methodology.* 2020.

emily.granger@lshtm.ac.uk

@EGranger90