Quasi-experimental study designs

Editorial

The scandal of poor medical research. Doug Altman.

BMJ 1994; 308 doi: https://doi.org/10.1136/bmj.308.6924.283

- Huge sums of money are spent annually on research that is seriously flawed through the use of inappropriate designs, unrepresentative samples, small samples, incorrect methods of analysis, and faulty interpretation.
- Why are errors so common? Put simply, much poor research arises because researchers feel compelled for career reasons to carry out research that they are ill equipped to perform, and nobody stops them.
- As the system encourages poor research it is the system that should be changed. **We need less research, better research, and research done for the right reasons**. Abandoning using the number of publications as a measure of ability would be a start.

Editorial

Medical Research – still a scandal. Richard Smith (2014)

blogs.bmj.com/bmj/2014/01/31/richard-smith-medical-research-still-a-scandal/

- In his editorial entitled, “The Scandal of Poor Medical Research,” Altman wrote that much research was “*seriously flawed through the use of inappropriate designs, unrepresentative samples, small samples, incorrect methods of analysis, and faulty interpretation*.” Twenty years later *I fear that things are not better but worse*.

We've all been taught that after systematic reviews and meta-analyses, RCT's give the *highest level of evidence*

If you *enrol enough people,* & assign to groups with a *random allocation sequence*:

- *known and unknown* confounding factors will be *balanced across* groups
- allows strong inferences about *cause and effect (unbiased effect estimate)*

But…….

- Not all interventions can be assessed with an RCT
- Labour-intensive, expensive to run (monitor, verify data, comply with all ICH-GCP)
- Sample sizes often too small (unrealistic effect size assumed)
- External validity (generalisability to real world settings) is poor – multimorbidity
- Pragmatic trials?

Experimental studies: Historical development

- James Lind (1753): citrus fruit can prevent/treat scurvy
  - intervention–control allocation based on ensuring participants in both groups were as similar as possible in observable characteristics
- Johannes Fibiger (1898): – efficacy of diphtheria antitoxin
  - Alternate allocation of intervention to control selection bias
- UK MRC (1948): Streptomycin vs bed rest for treatment of TB
  - Random allocation to prevent researchers consciously or subconsciously selecting different types of patients for treatment arms
- Increases in the use of RCTs for cardiovascular disease, oncology, HIV

Experimental vs Quasi-experimental designs

- Experimental study
  - researcher intervenes in the “natural” process, to establish the causal effects of a treatment
- Quasi-experimental study
  - causal effects of a treatment are established without a researcher's intervention
  - offer opportunities for causal analyses in health research because they can generate results of high external and internal validity, and can be used when experiments are not feasible (financial, ethical, or political constraints)
  - Have been widely used in economics, social sciences

Approaches

- (Natural experiments)
- Interrupted (segmented) time series regression
- Regression discontinuity
- Stepped wedge designs
- Difference in differences analyses
- Matching and reweighting
- (Instrumental variables and extended regression models (ERMs)

- Many of these techniques can be used with administrative data or routinely collected clinical data

Caveats Assumptions

- All these model have assumptions
- As the complexity of the model increases, so do the number of assumptions

- Most important thing: the researchers need to have a good knowledge of the topic so they can understand how to make the model and model the potential confounders

Interrupted time series

Interrupted time series

- Used to evaluate the effect of an intervention
- Considered one of the strongest quasi-experimental designs
- Data are collected at equally spaced points pre- and post-intervention (weekly or monthly)
  - Need an adequate number of pre-intervention points so the trend is clear
  - Need to know when the intervention starts
  - Need an adequate number of post-intervention points
  - Calculate changes based on the pre-intervention trend
- Things to consider and adjust for: autocorrelation, seasonality, other interventions/factors which could influence the outcome

Interrupted time series

Anaesthesia 2019 doi:10.1111/anae.14570

Original Article

A sustainable method to reduce postoperative oxycodone discharge prescribing in a metropolitan tertiary referral hospital*

J. Stevens,[1] **A. Trimboli,**[2] **P. Samios,**[3] **N. Steele,**[4] **S. Welch,**[5] **P. Thompson,**[6] **C. Halvorsen**[7] and **S. Kerr**[8,9]

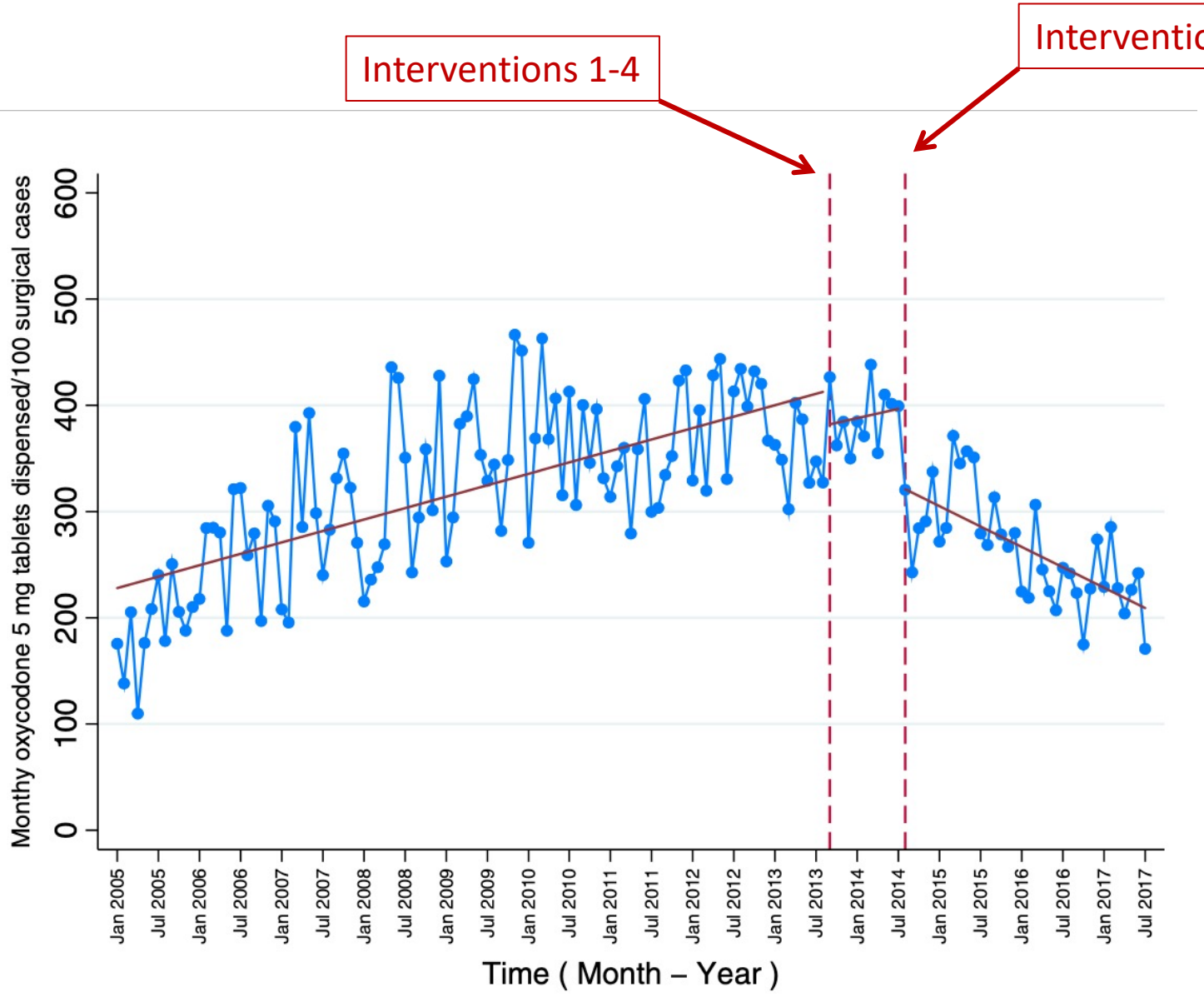
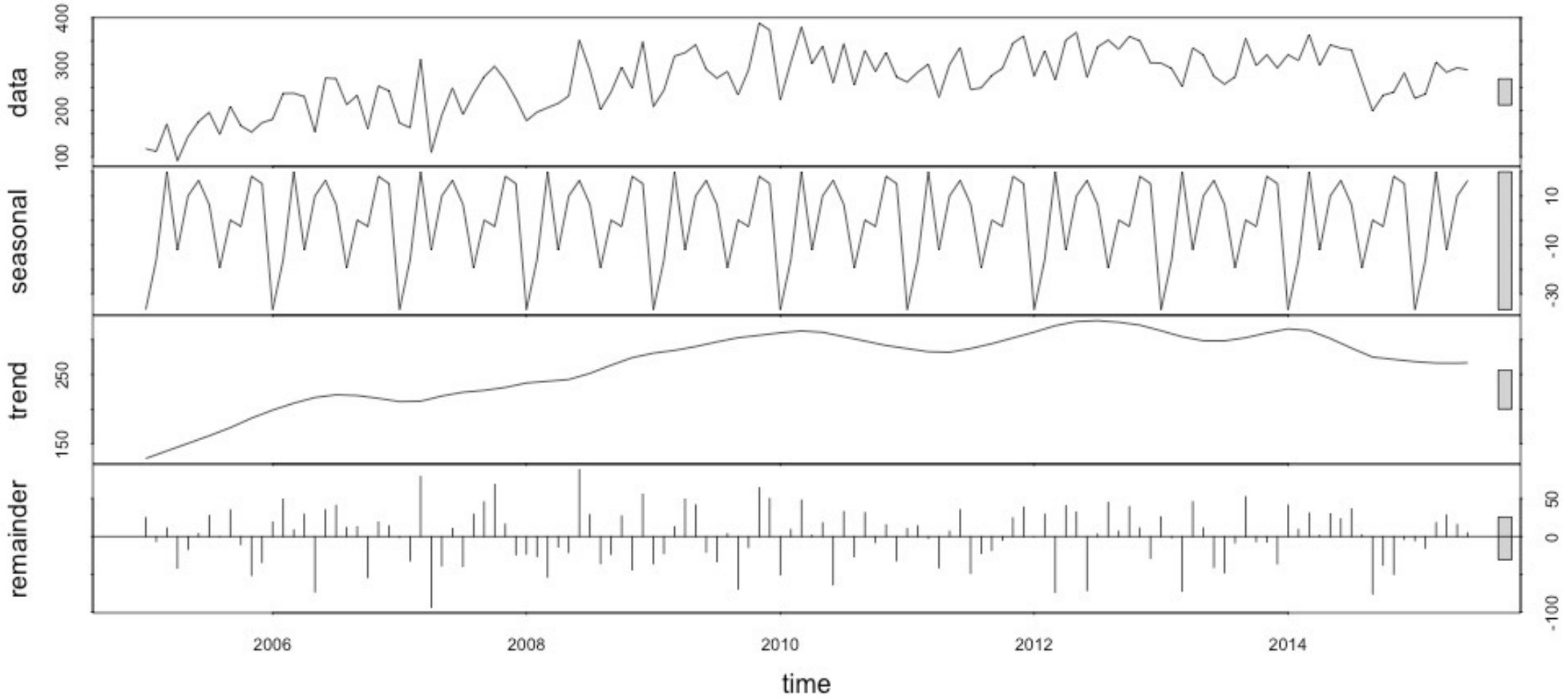


Table 1 Mean monthly oxycodone parameter estimates per 100 surgical admissions from segmented regression model. Values are coefficients (95%CI).

	Coefficient (95%CI)	p value
Intercept (baseline level)	227 (199–256)	0.001
Baseline trend	1.8 (1.3–2.3)	0.001
Level change after intervention 1	−32 (−76 to 11)	0.14
Trend change after intervention 1	−0.3 (−4.2 to 3.5)	0.87
Post-intervention 1 linear trend	**1.5 (−2.5 to 5.4)**	0.46
Level change after intervention 5	−77 (−115 to 39)	0.001
Trend change after intervention 5	−4.7 (−8.8 to −0.5)	0.03
Post-intervention 5 linear trend	**−3.2 (−4.6 to 1.8)**	0.001

Post-intervention 1 linear trend is the sum of the baseline trend and the trend change from intervention 1- 4 inclusive (in bold). Post-intervention 5 linear trend is the sum of baseline and all post-intervention trends (in bold).

STL decomposition: (Seasonal decomposition of time series by Loess)



Regression discontinuity

Regression discontinuity

- Useful when an exposure or treatment is determined by a threshold rule, causal effects can be estimated by regression discontinuity
  - Patients assigned to a regimen if they are ‘high risk’ based on a continuous biomarker (eg cholesterol)
- These continuous biomarkers are subject to random variability, so we expect patients scoring just above and just below the threshold will be similar in observable and unobservable characteristics
- If the threshold affects the outcomes, we expect to see a discontinuity in the regression around the threshold; if there’s no ‘treatment effect’ the regression line should be continuous above and below the threshold

Efficacy of Prostate-Specific Antigen Screening

Use of Regression Discontinuity in the PLCO Cancer Screening Trial

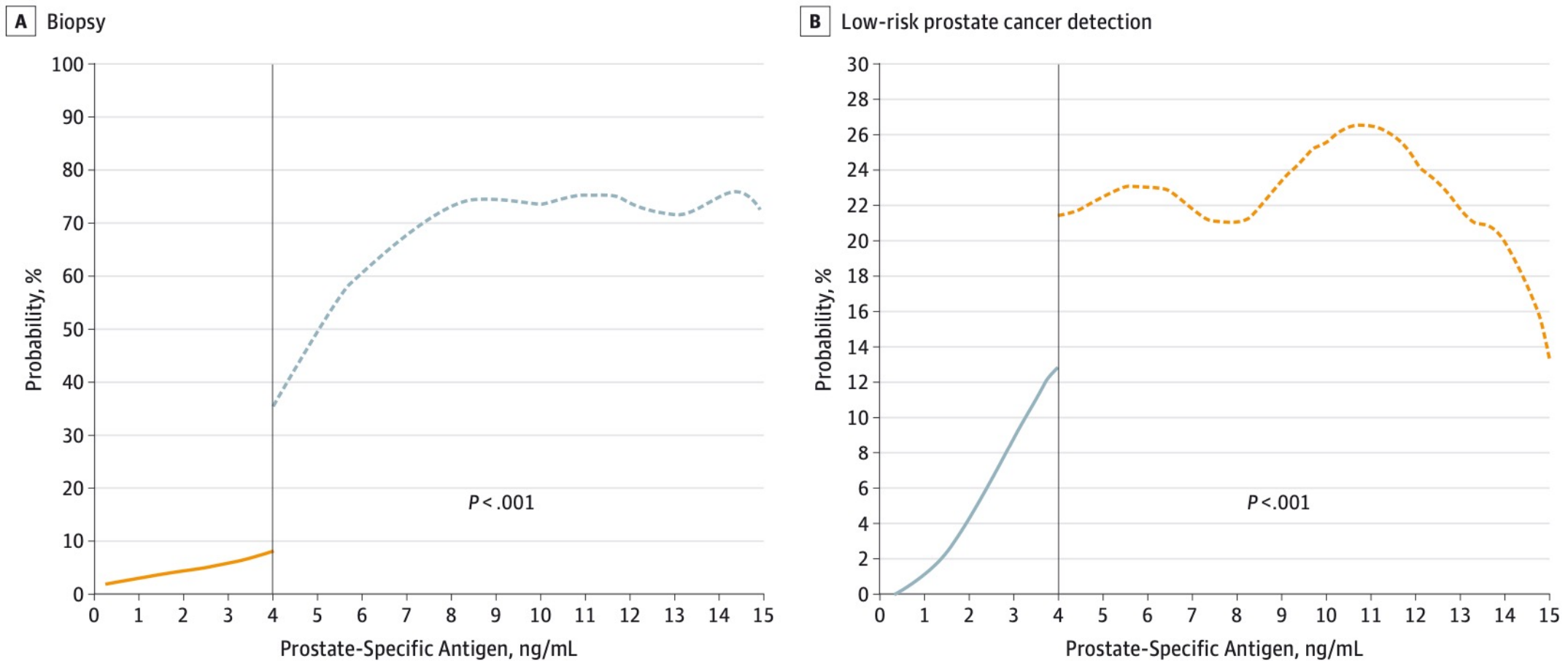
Jonathan Shoag, MD[1]; Joshua Halpern, MD[1]; Brian Eisner, MD[2]; et al

» Author Affiliations | Article Information

JAMA Oncol. 2015;1(7):984-986. doi:10.1001/jamaoncol.2015.2993

- The Prostate Lung Colorectal and Ovarian (PLCO) cancer screening trial randomized 76 693 men from 1993 to 2001 to usual care or annual prostate-specific antigen (PSA) screening for 6 years and annual digital rectal examination for 4 years.
- A PSA level of 4.0 ng/mL was used as the threshold for further workup

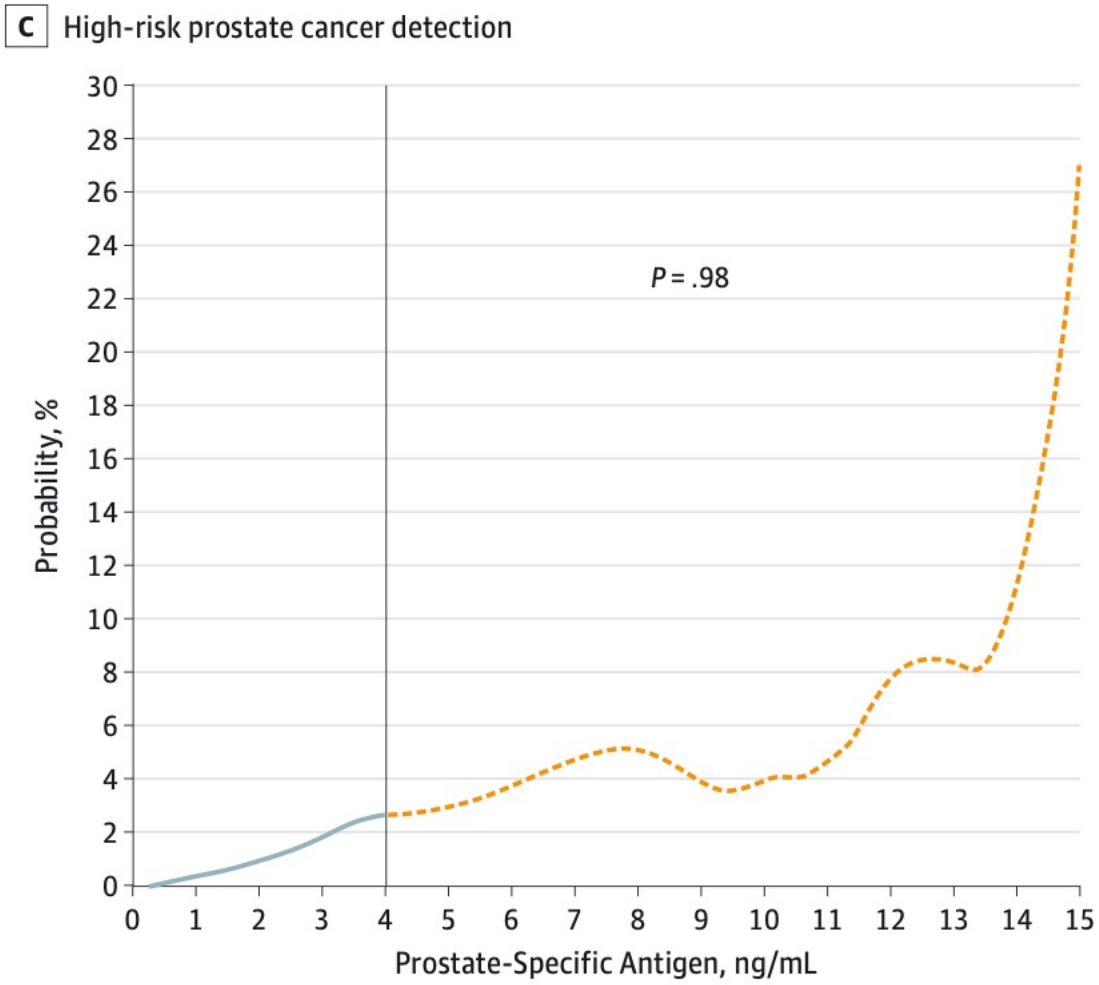
Figure. Regression Discontinuity Analyses

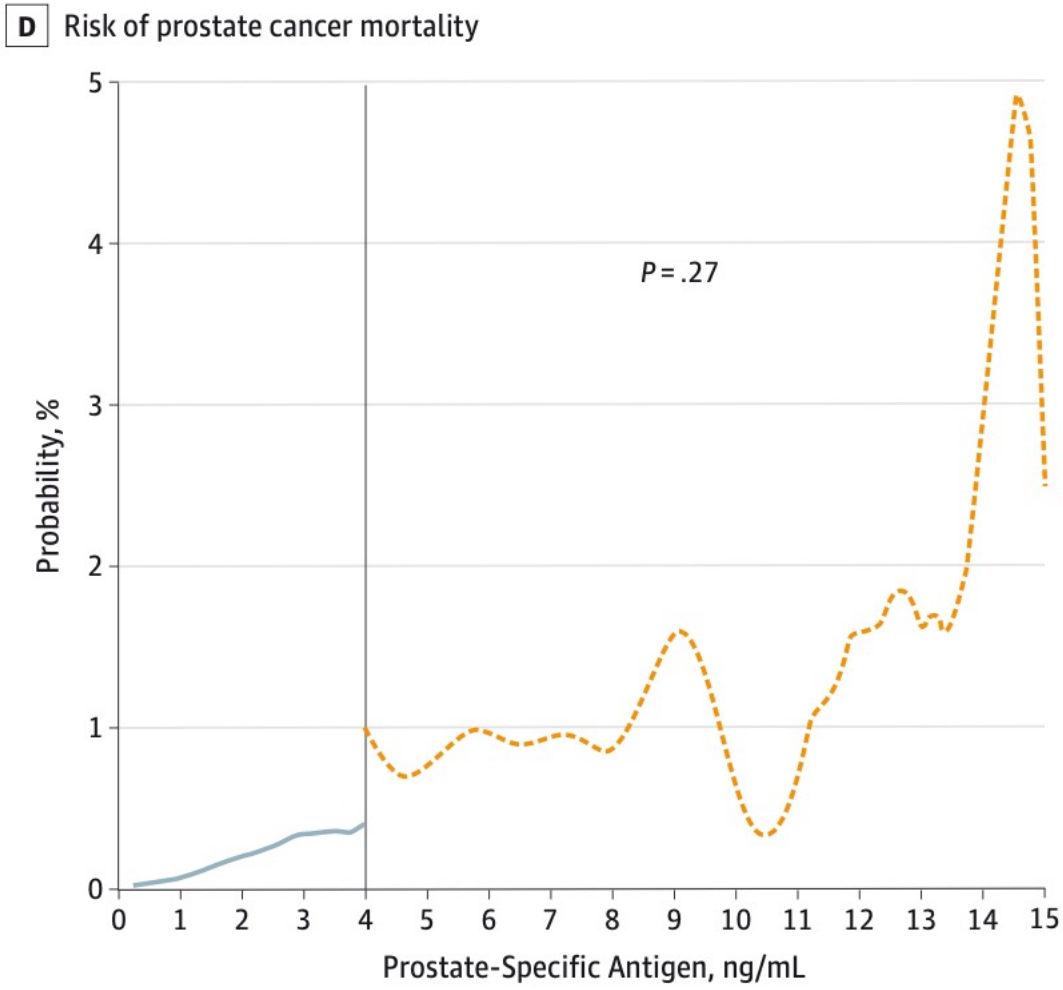


Undergo biopsy: 27.3 (95%CI 23.3 – 31.3)

LR cancer detection: 7.2 (95%CI 3.6 – 10.8)

Effect sizes are absolute changes in probability





HR cancer detection: 0.0 (95%CI -1.8 – 1.9)

Mortality: Effect size = 0.6 (95%CI -0.5 – 1.7)

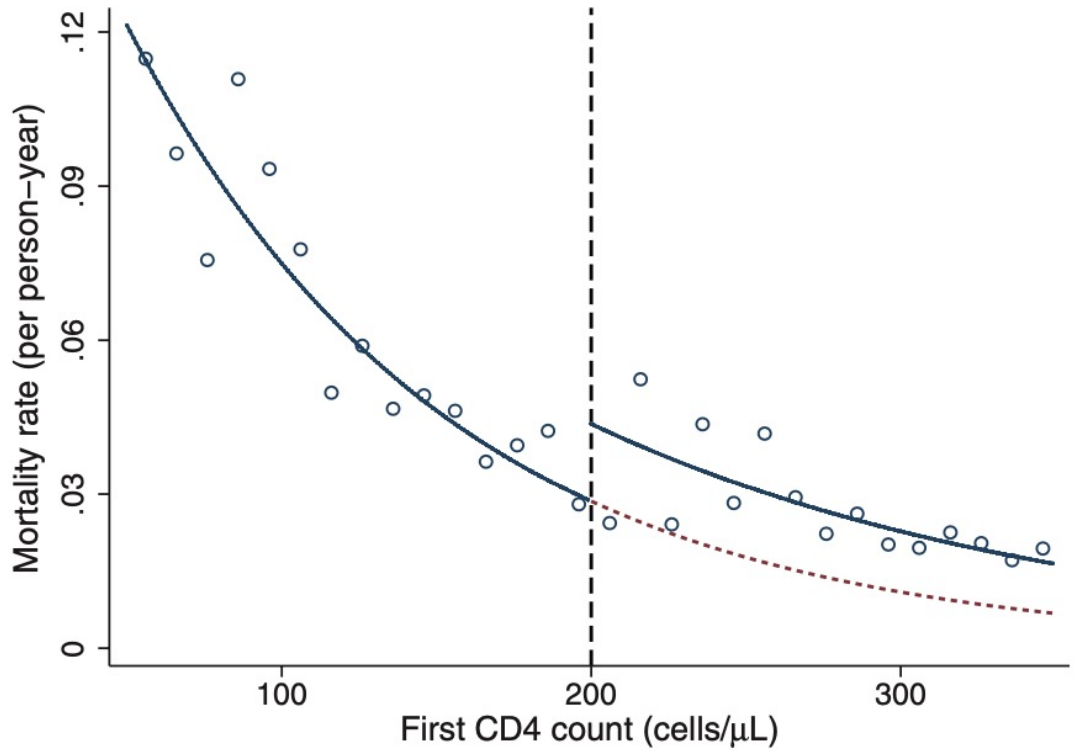


FIGURE 4. First CD4 count and mortality hazard rate. Predicted hazards from the Table, model 2a are displayed as solid lines. Dashed line shows extrapolated prediction if all patients were treatment eligible at first CD4 count. Dots are hazards predicted for CD4 count bins of width 10 cells.

Table 1. PubMed articles with health outcomes using regression discontinuity designs

Authors	Year	Journal	Study topic
Albouy and Lequien [23]	2009	Journal Health Economics	Effect of education on mortality
Almond et al. [15]	2010	Quarterly Journal of Economics	Returns to treatment of low-birth-weight newborns
Andalón [24]	2011	Health Economics	Effect of Oportunidades on obesity
Anderson et al. [25]	2011	Journal of Health Economics	Effect of schooling on children's BMI
Arcand and Wouabe [26]	2010	Health Economics	Effect of teacher training on HIV prevention
Banks and Mazzonna [27]	2012	Economics Journal	Effect of education on old-age cognitive ability
Behrman [28]	2014	Social Science and Medicine	Effect of primary schooling on HIV status
Bor et al. [1]	2014	Epidemiology	Effect of early vs. deferred HIV treatment on mortality
Callaghan et al. [29]	2014	Drug and Alcohol Dependence	Effect of legal drinking age on mortality
Callaghan et al. [30]	2013	American Journal of Public Health	Effect of legal drinking age on alcohol-related morbidity
Callaghan et al. [31]	2013	Addiction	Effect of legal drinking age on inpatient morbidity
Carpenter and Dobkin [16]	2009	AEJ: Applied Economics	Effect of alcohol consumption on mortality
Carpenter and Dobkin [32]	2011	Journal of Economic Perspectives	Minimum legal drinking age and public health
Chen et al. [33]	2013	PNAS	Effect of air pollution on mortality
Conover and Scrimgeour [34]	2013	Journal of Health Economics	Health effects of minimum legal drinking age
De La Mata [35]	2012	Health Economics	Effect of Medicaid eligibility on coverage, utilization, and health
Deza [36]	2014	Health Economics	Effect of alcohol use on drug consumption
Flam-Zalcman et al. [37]	2012	Intl J Psych Research	Effect of criterion-based increase in alcohol treatment
Fletcher [38]	2014	Biodemography and Social Biology	Effect of genetics on stress response
Glance et al. [39]	2014	JAMA Surgery	Effect of hospital report cards on mortality
Gormley et al. [40]	2005	Developmental Psychology	Effect of universal pre-kindergarten on cognitive development
Huang and Zhou [41]	2013	Social Science and Medicine	Effect of education of cognition
Jensen and Wust [42]	2014	Journal of Health Economics	Effect of Caesarean section on maternal and child health
McFarlane et al. [43]	2014	Schizophrenia Bulletin	Effect of treatment program on psychosis onset
Miller et al. [44]	2013	AEJ: Applied Economics	Effect of insurance on health spending, utilization, and health
Nishi et al. [45]	2012	Bulletin of the WHO	Health effects of patient cost-sharing
Pierce et al. [46]	2012	Pers Soc Psych Bulletin	Effect of income disparity in marriage
Sloan and Hanrahan [47]	2014	JAMA Ophthalmology	Effect of new therapies on vision loss among elderly patients
Smith et al. [48]	2014	Canadian Medical Association Journal	Effect of HPV vaccine on sexual behavior
Sood et al. [49]	2014	BMJ	Effect of health insurance on mortality
Weaver et al. [50]	2010	Journal of Traumatic Stress	Effect of cognitive-behavioral therapy on trauma symptoms
Yörük and Yörük [51]	2012	Social Science and Medicine	Effect of alcohol on psychological well-being

Abbreviations: BMI, body mass index; HIV, human immunodeficiency virus; AEJ, American Economic Journal; PNAS, Proceedings of the National Academy of Sciences; Intl J Psych Research, International Journal of Methods in Psychiatric Research; JAMA, Journal of the American Medical Association; WHO, World Health Organization; Pers Soc Psych Bulletin, Personality and Social Psychology Bulletin; HPV, human papillomavirus; BMJ, British Medical Journal.

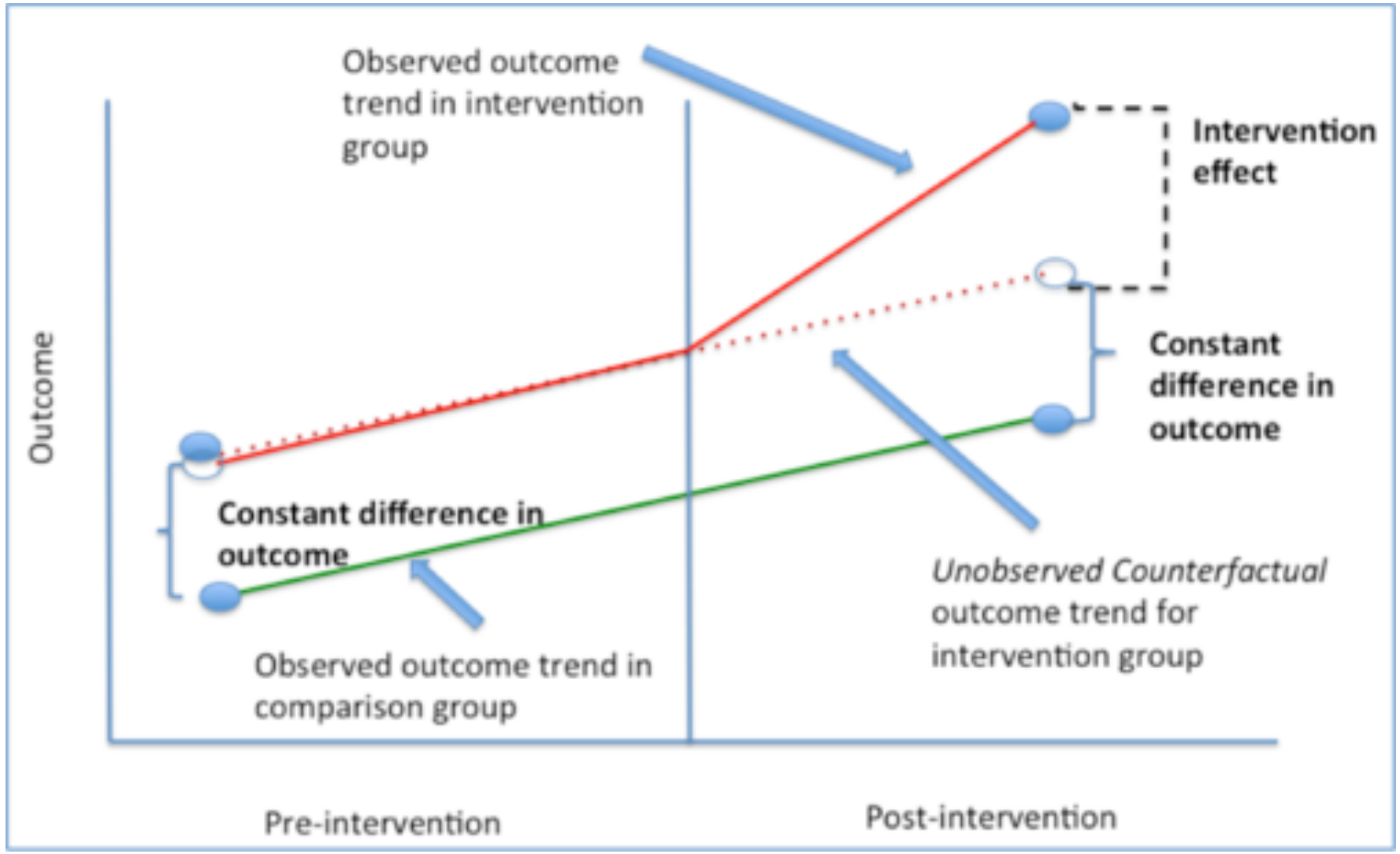
Differences in differences

Differences in differences (DID)

- Uses longitudinal data from ‘treatment’ and ‘control’ groups to obtain an appropriate counterfactual to estimate a causal effect.

- DID is typically used to estimate the effect of a specific intervention or treatment (law, policy intervention) by comparing the changes in outcomes over time between the treatment and control groups

- DID – used by John Snow to investigate the cholera epidemic in London in 1800’s


Observed outcome trend in intervention group
Intervention effect
Constant difference in outcome
Constant difference in outcome
Unobserved Counterfactual outcome trend for intervention group
Observed outcome trend in comparison group
Outcome
Pre-intervention
Post-intervention

Useful links

- Differences in differences
- https://diff.healthpolicydatascience.org

Matching and weighting

Stata: Matching to estimate ‘treatment effects’ for outcomes conditionally independent on treatment

- Implies variables affecting treatment assignment and outcomes are observable
- teffects: continuous, binary, fractional and count outcomes
- stteffects: many functional forms for survival outcomes (parametric)
- telasso: functional forms as for teffects, but with lasso methods for covariate selection
- didregress: difference in differences regression
- xtdidregress: DID for panel (longitudinal) data

- Tools for checking covariate balance, covariate balance summary statistics, overlap plots

Matching methods for treatment effects

- Augmented inverse-probability weighting (AIPW)
- Inverse-probability weighting (IPW)
- Inverse-probability-weighted regression adjustment (IPWRA)
- Nearest neighbour matching (nnmatch)
- Propensity score matching (psmatch)
- Regression adjustment (ra)

The fundamental problem of causal inference

- We only observe one of the potential outcomes in each person
- Instead of assuming that the treatment is randomly assigned, we assume that after conditioning on covariates, the treatment is as good as randomly assigned
  - Conditional mean independence assumption (CMI)
  - After accounting for the covariates, the treatment ‘is as good as random’

Example

- How much is baby birthweight reduced if the mother smokes in pregnancy?

- Model
  - Outcome only (regression adjustment - ra)
  - Treatment only (IPW)
  - Outcome and treatment (Augmented IPW)
  - Outcome and treatment (IPWRA)
  - Outcome (nonparametrically) - nearest neighbour matching
  - Treatment (Propensity score matching)

```
. teffects ra (bweight mmarried prenatal1 fbaby medu) (mbsmoke)

Iteration 0:   EE criterion =  2.336e-23
Iteration 1:   EE criterion =  5.702e-26

Treatment-effects estimation                    Number of obs     =      4,642
Estimator      : regression adjustment
Outcome model  : linear
Treatment model: none
```

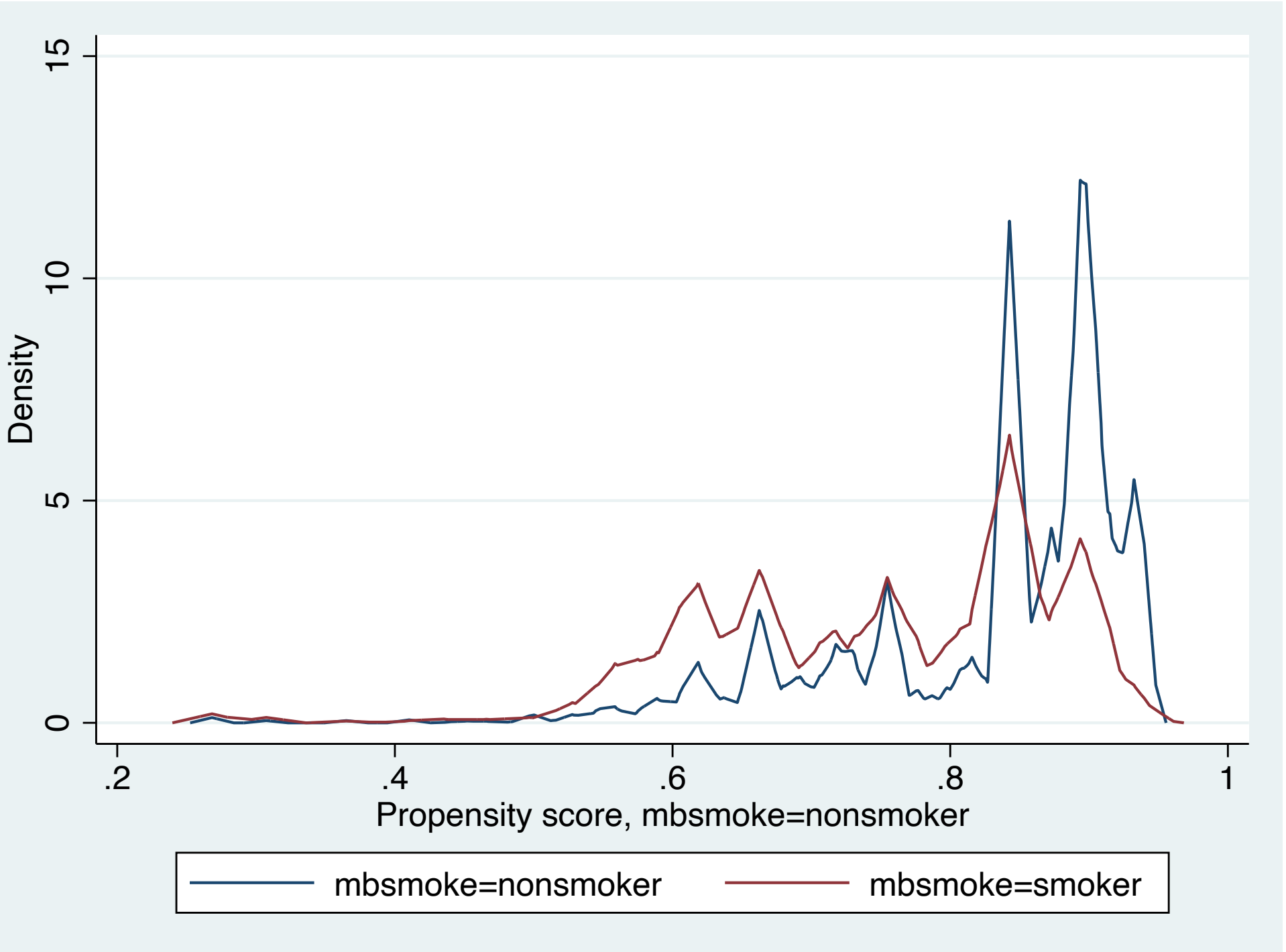
	bweight	Coefficient	Robust std. err.	z	P>|z|	[95% conf.	interval]
ATE							
	mbsmoke						
(smoker vs nonsmoker)		**-230.9541**	**24.34012**	**-9.49**	**0.000**	**-278.6599**	**-183.2484**
POmean							
	mbsmoke						
	nonsmoker	**3402.548**	**9.546721**	**356.41**	**0.000**	**3383.836**	**3421.259**

```
. teffects ipwra (bweight mmarried prenatal1 fbaby medu) (mbsmoke mmarried prenatal1 fbaby medu)

Iteration 0:   EE criterion =  9.630e-22
Iteration 1:   EE criterion =  1.298e-25

Treatment-effects estimation                    Number of obs     =      4,642
Estimator      : IPW regression adjustment
Outcome model  : linear
Treatment model: logit
```

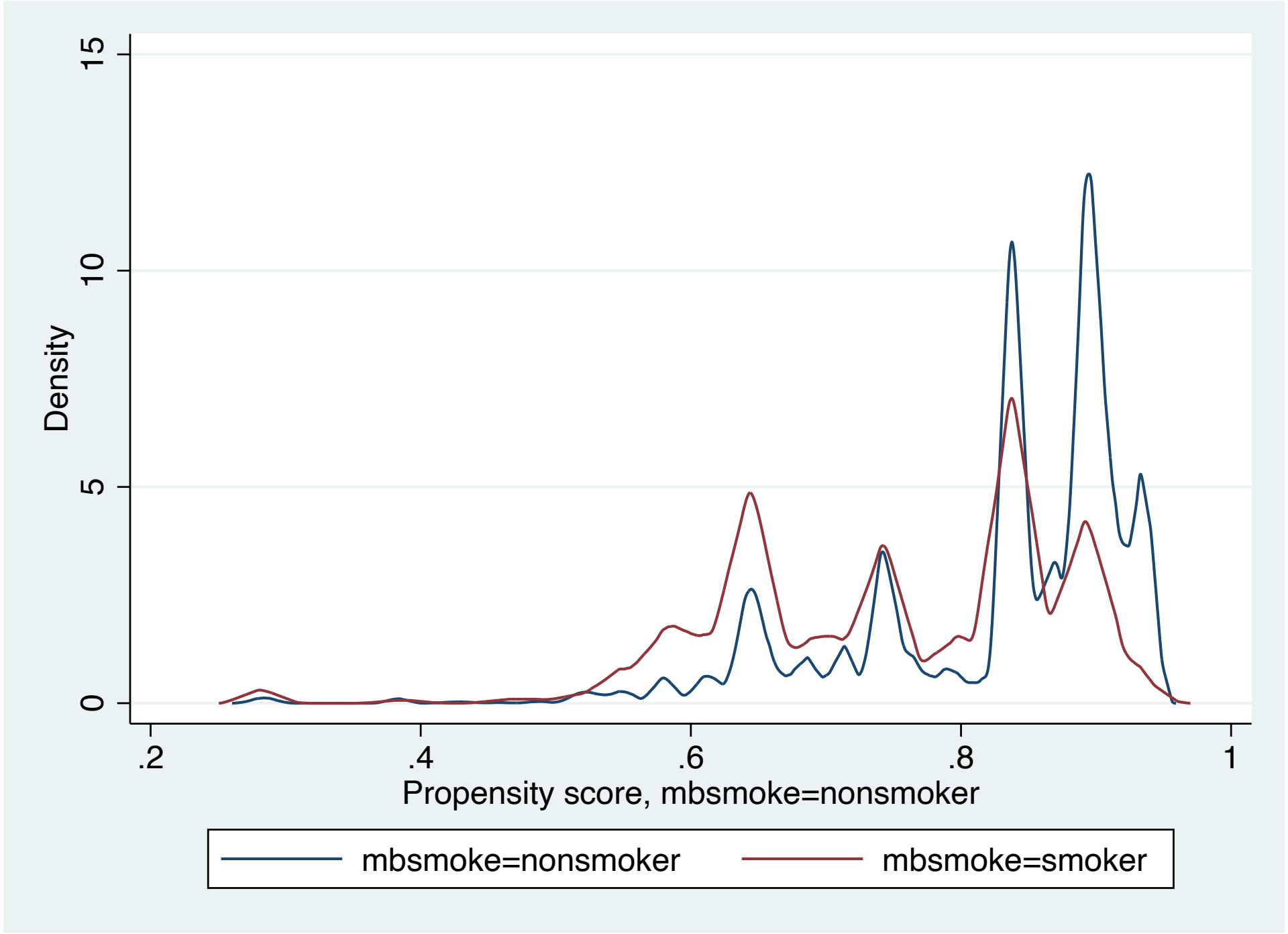
	bweight	Coefficient	Robust std. err.	z	P>|z|	[95% conf.	interval]
ATE							
	mbsmoke (smoker vs nonsmoker)	**-227.4408**	**25.62591**	**-8.88**	**0.000**	**-277.6667**	**-177.215**
POmean							
	mbsmoke nonsmoker	**3403.027**	**9.56025**	**355.96**	**0.000**	**3384.289**	**3421.765**


Density
0
5
10
15
.2
.4
.6
.8
1
Propensity score, mbsmoke=nonsmoker
mbsmoke=nonsmoker
mbsmoke=smoker

```
. teffects psmatch (bweight) (mbsmoke mmarried c.mage fbaby medu), caliper(0.1)

Treatment-effects estimation                   Number of obs      =      4,642
Estimator      : propensity-score matching     Matches: requested =          1
Outcome model  : matching                                  min =          1
Treatment model: logit                                     max =         74
```

	bweight	AI robust Coefficient	std. err.	z	P>|z|	[95% conf.	interval]
ATE							
	mbsmoke (smoker vs nonsmoker)	**-203.9734**	**35.31088**	**-5.78**	**0.000**	**-273.1814**	**-134.7653**


Density
0
5
10
15
.2
.4
.6
.8
1
Propensity score, mbsmoke=nonsmoker
mbsmoke=nonsmoker
mbsmoke=smoker

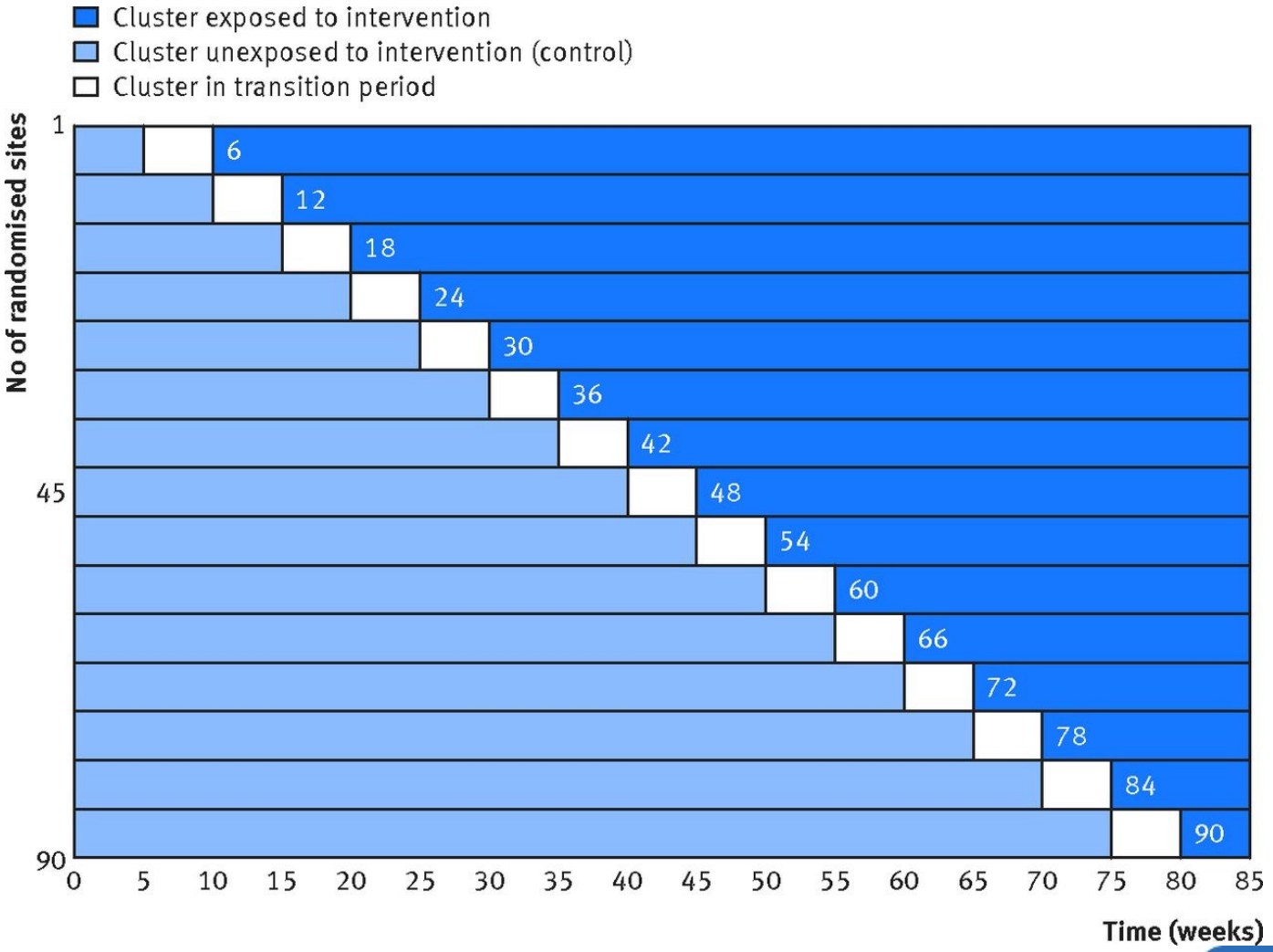
Stepped wedge designs

Stepped wedge (cluster randomised) trials

- Each site serves as their own control (could be a ward in a hospital, or a hospital, or a neighbourhood etc)
- Timing of the crossover is randomised
- Multiple observations per unit (collected at the same time)

- Why?
  - Logistical or financial
  - Systematically evaluate new program shown
  - To study the effect of time on intervention effectiveness (i.e. seasonality, time since introduction)

Fig 2 Schematic representation of the EPOCH stepped wedge study (example 4).



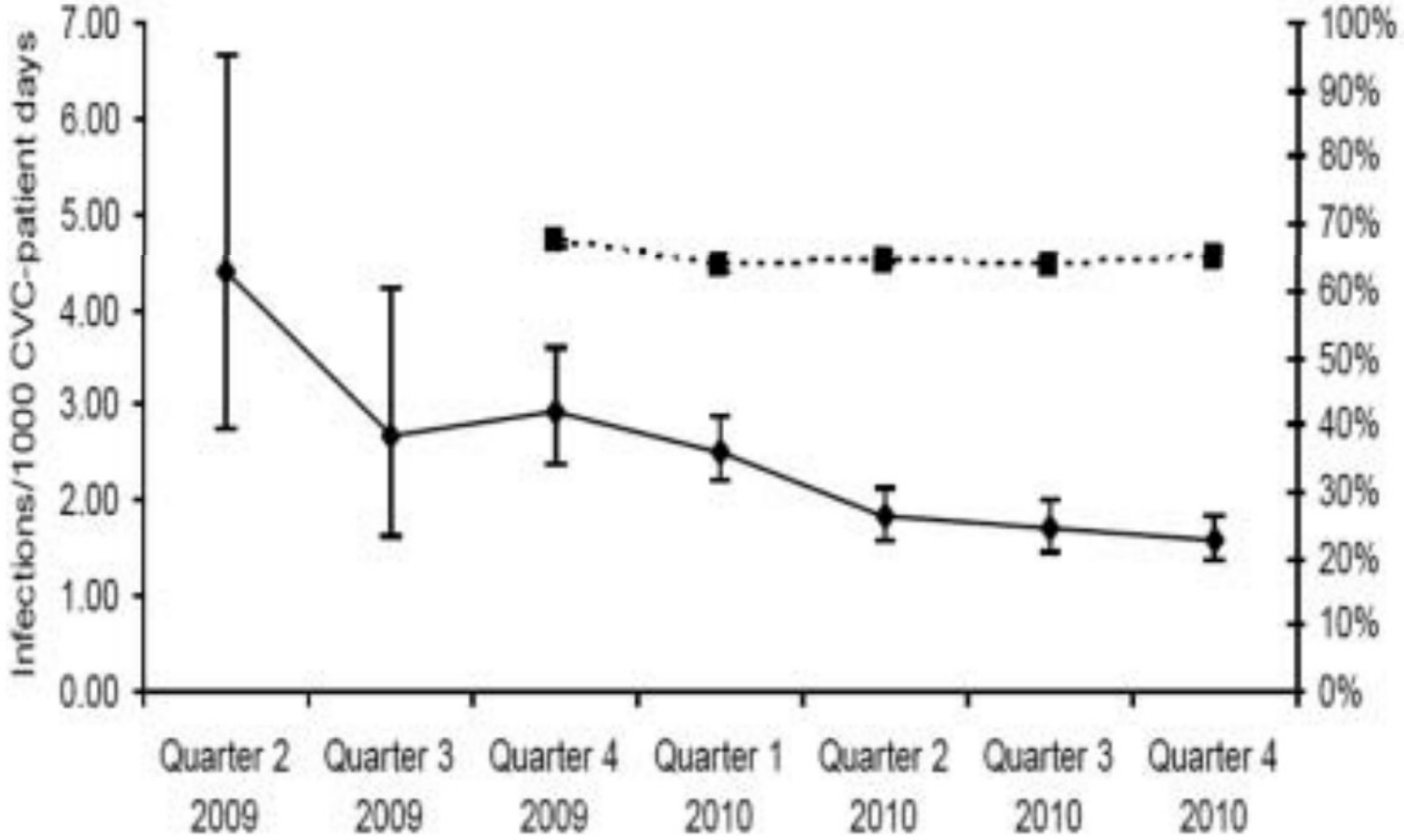
K Hemming et al. BMJ 2015;350:bmj.h391

ORIGINAL RESEARCH

OPEN ACCESS

‘Matching Michigan’: a 2-year stepped interventional programme to minimise central venous catheter-blood stream infections in intensive care units in England

a: Total Adult & Paediatric CVC-BSI Infection Rate (—) and CVC Utilisation ratio % (·····) by Quarter



EQUATOR website

equator network

Enhancing the QUAlity and Transparency Of health Research

Visit the EQUATOR Spanish Website

Home | Library | Toolkits | Courses & events | News | Blog | About us | Contact

The resource centre for good reporting of health research studies

Library for health research reporting

The Library contains a comprehensive searchable database of reporting guidelines and also links to other resources relevant to research reporting.

Search for reporting guidelines

Visit the library for more resources

Key reporting guidelines

CONSORT	Full Record | Checklist | Flow Diagram
STROBE	Full Record | Checklist
PRISMA	Full Record | Checklist | Flow Diagram
STARD	Full Record | Checklist | Flow Diagram
COREQ	Full Record
ENTREQ	Full Record
SQUIRE	Full Record | Checklist
CHEERS	Full Record | Checklist
CARE	Full Record | Checklist
SAMPL	Full Record

Toolkits

The EQUATOR Network works to improve the reliability and value of medical research literature by promoting transparent and accurate reporting of research studies.

Our Toolkits support different user

EQUATOR highlights

3/10/2013 - OPEN: To overcome failure to publish negative findings

The EU-funded OPEN project (Overcome failure to Publish nEgative fiNdings) brought together key opinion leaders from across Europe to address the issue of publications bias. Read More

17/09/2013 - EQUATOR Network at the Peer Review Congress

News

The Checklist Manifesto Meets Clinical Trials
9/10/2013

Publicly available sources provide insufficient information on patient-relevant outcomes of clinical trials
9/10/2013

New EASE Science Editors' Handbook