Introduction to Interrupted Time Series Analysis

Rob Penfold, PhD
April 28, 2023
Washington, DC

SCHOOL OF PUBLIC HEALTH
UNIVERSITY of WASHINGTON

KAISER PERMANENTE®

Agenda

- Introduction to ITS (60 min)
- Questions (10 min) – feel free to interrupt me (ha ha)
- Small Group Exercise: Evaluation of a published study
- Closing Remarks and Wrap-up (10 min)

Overview of Interrupted Time Series Analysis

Basic ITS design

- Why use it?
- How is ITS different from individual-level analyses?
  - When to use individual-level data
- Strengths & weaknesses, threats to validity

Difference-in-Differences (DiD) ITS

- Comparison/control group available
- Appropriate and inappropriate DiD

Overview of Interrupted Time Series Analysis cont'd

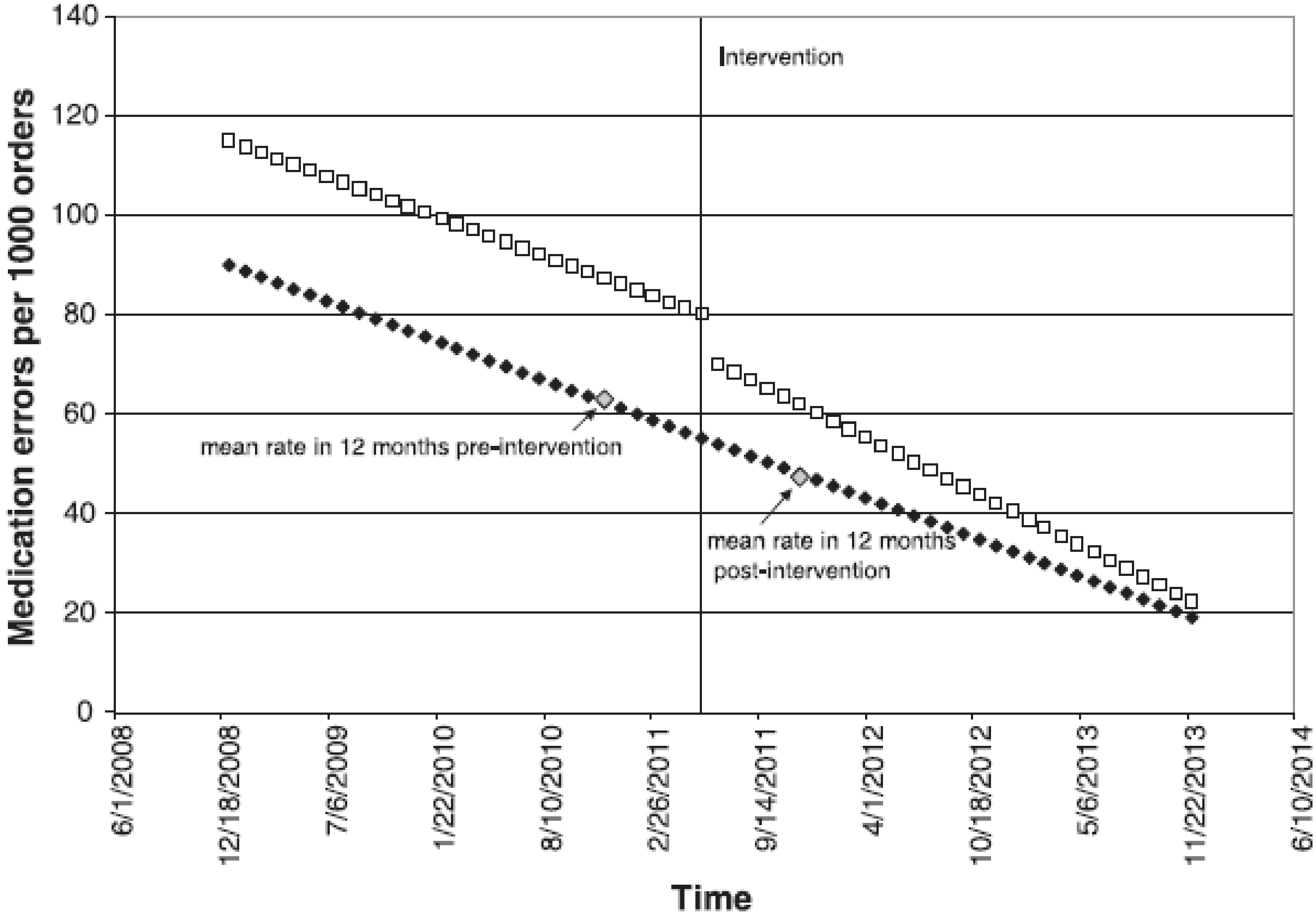
Accommodating "real world" interventions

- Dealing with phased-in interventions/policies
- Multi-component interventions
- Sub-population stratification
- Multi-site interventions with stepped roll-out

Technical issues with time

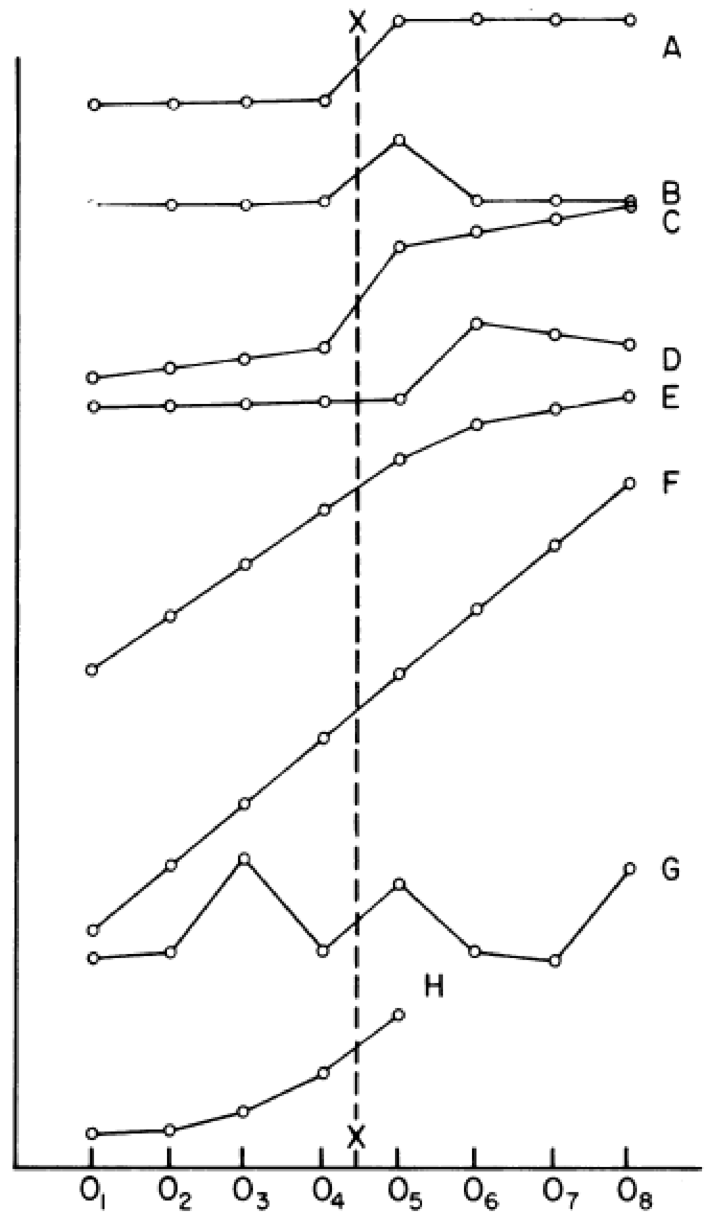
- Unit of time (e.g., month vs. fiscal quarter)
- Seasonality
- Higher order modeling terms

Stronger Design that pre-post diff of means



Gillings (1981) Trend Examples

Be careful about how time is specified in the model

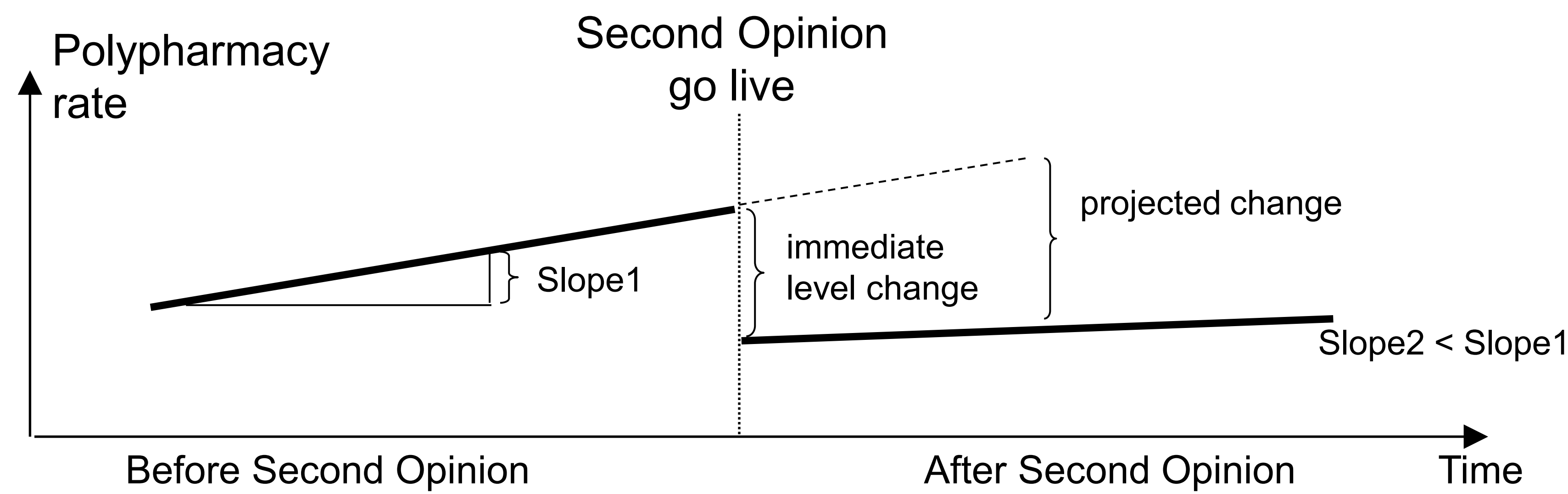


When to use an ITS Design

- Desire to make population-level inferences
  - Mean effect of an intervention or policy
  - Individual-level data are not available or would be difficult/unethical to collect/publish
- Randomization is not feasible, ethical, affordable
  - Take advantage of natural experiments
- Population denominator can be defined

***When it is impossible to obtain/measure individual-level confounders or mediators**

What is an Interrupted Time Series: Changes in Level and Slope



Adapted from Schneeweiss et al, Health Policy 2001

Threats to Validity

Contemporaneous QI programs (competing interventions)

Heterogeneity in the composition of the "population"

Eligibility, attrition, selection

- Maturation
- Mortality

Instrumentation – e.g., new measurement

Speed of implementation and delayed effects

Short time series

Suitability of control population

Competing Interventions (History)

Contemporaneous QI initiatives or programs

- Competing interventions MUST occur at the same time to be confounders.
- Not to be confused with an overall (e.g., national) trend. We can model the secular trend explicitly.

Population Heterogeneity (Selection)

Composition of the population changes over time

- Within-population confounders are no longer constant (controlled)
- Change in risk factors over time may be the cause of increasing/decreasing rates rather than the intervention or policy

Potential solutions

- Standardize the population to the intervention period
  - Similar to standardized mortality ratios
- Limit the population to those individuals observed in all time periods of the analysis (fixed cohort)

Mortality and Attrition (Selection)

Special case of a change in the population over time

- Youth at risk leave the population

Maturation

Particular problem for child and adolescent research

- Youth age out of and into the population at risk

Rolling cohort often better than a fixed cohort

  - Proportion in each age group held constant over time

Stratify by age group

Instrumentation (Measurement)

Ability (effort?) to measure the event changes

- Transition to EHR
- Changes in coding (ICD9 – ICD10)
- Introduction of a new instrument (e.g., screening)
- Change in definition of an event

- Be careful that your intervention does not directly impact your ability to measure (e.g., intervention to reduce unnecessary visits)

- Example: switch to telehealth during COVID
  - Coding of video visits

Multiple sites with Staged Roll-out

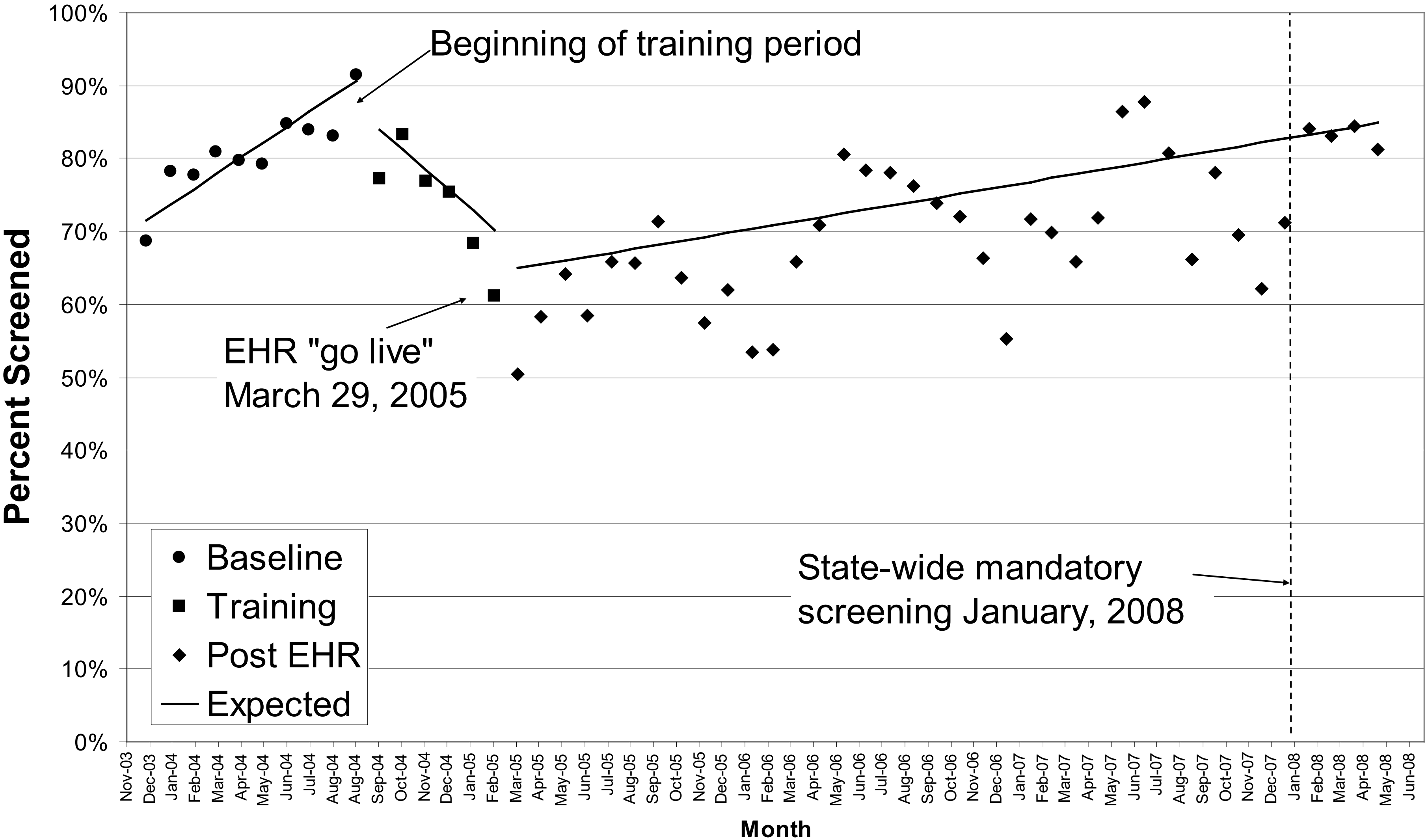
Interventions are often rolled-out sequentially in time across sites

- Changes often occur at "easiest" sites first
- Implementation issues often resolved as additional sites come "online" (i.e., learning occurs)

Address roll-out:

- Censor observations around the go-live date
- Conduct separate ITS by site
- "Center" ITS analyses on the index (implementation) time period even though the calendar dates differ

Hacker and Penfold, 2012



SCHOOL OF PUBLIC HEALTH
UNIVERSITY of WASHINGTON

KAISER PERMANENTE®

Segmented Regression

Calculate crude or adjusted rate by month

Create binary indicator variable for time periods

Create count variable to index time - month 1, month 2 . . .

Interrupted times series model

$$Y_t = \beta_0 + \beta_1 * \text{time}_t + \beta_2 * \text{training}_t + \beta_3 * \text{time after training}_t + \beta_4 * \text{EHR implementation}_t + \beta_5 * \text{time after EHR}_t + e_t$$

Month	Num.	Den.	Rate	time	planning	time after planning	intervention	time after intervention
Dec	55	80	0.688	1	0	0	0	0
Jan	97	124	0.782	2	0	0	0	0
Feb	98	126	0.778	3	0	0	0	0
Mar	127	157	0.809	4	0	0	0	0
Apr	94	118	0.797	5	0	0	0	0
May	130	164	0.793	6	0	0	0	0
Jun	162	191	0.848	7	0	0	0	0
Jul	157	187	0.84	8	0	0	0	0
Aug	250	301	0.831	9	0	0	0	0
Sep	161	176	0.915	10	0	0	0	0
Oct	152	197	0.772	11	1	1	0	0
Nov	140	168	0.833	12	1	2	0	0
Dec	147	191	0.77	13	1	3	0	0
Jan	101	134	0.754	14	1	4	0	0
Feb	80	117	0.684	15	1	5	0	0
Mar	102	167	0.611	16	1	6	0	0
Apr	58	115	0.504	17	1	7	1	1
May	95	163	0.583	18	1	8	1	2
Jun	116	181	0.641	19	1	9	1	3
Jul	83	142	0.585	20	1	10	1	4
Aug	197	299	0.659	21	1	11	1	5
Sep	141	215	0.656	22	1	12	1	6
Oct	122	171	0.713	23	1	13	1	7
Nov	98	154	0.636	24	1	14	1	8

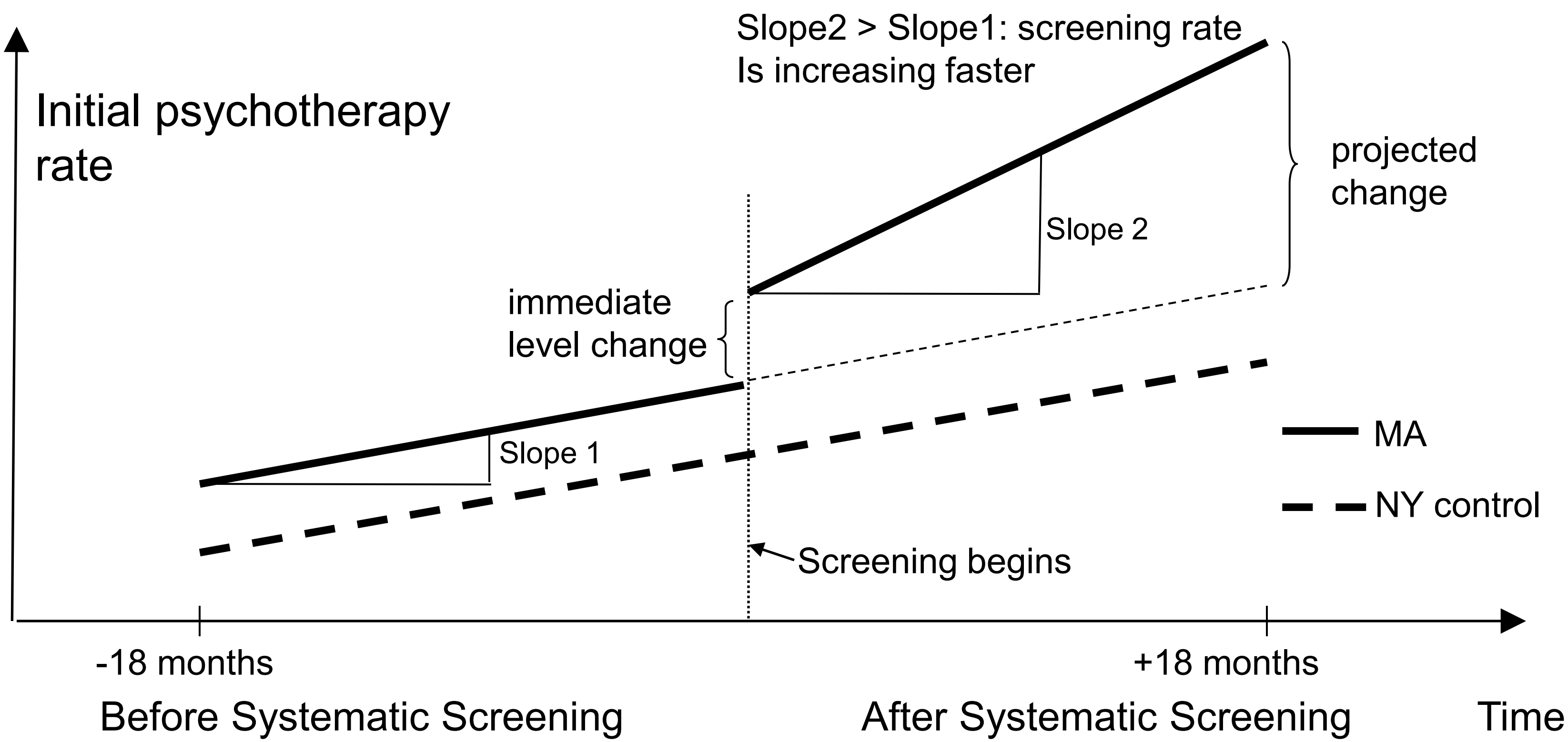
Stronger Design: Difference-in-Differences ITS

Compare the change in the intervention population to the change in an unexposed population

- E.g., intervention clinics to control clinics within Group Health
- Policy state to control state

Segmented regression of the differences in rates rather than the rates/proportions themselves

What is an Interrupted Time Series: Changes in Related Services and Outcomes



Adapted from Schneeweiss et al, 2001

Graph of Differences

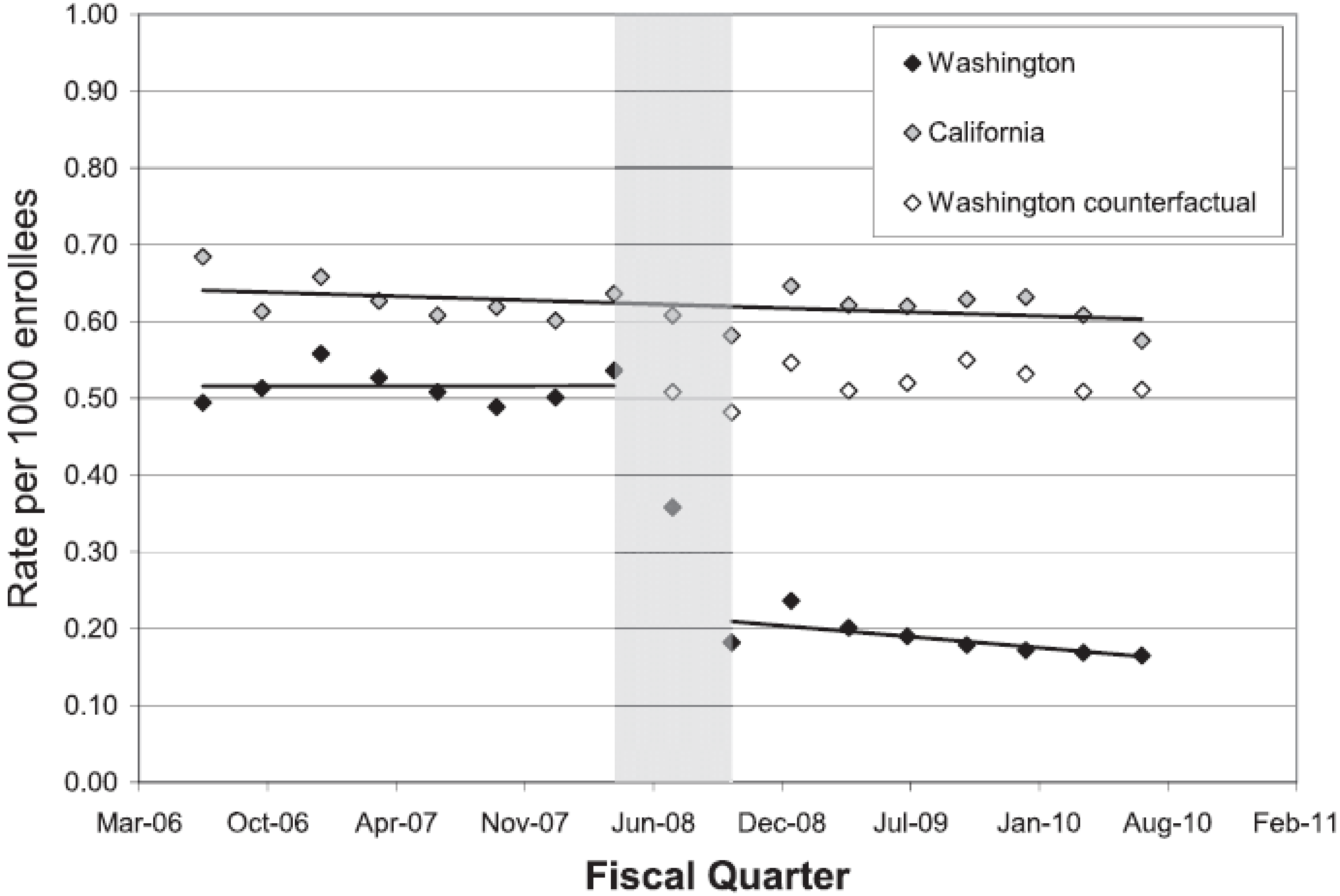
Collapse the intervention and control time series by subtracting the control rate from the intervention rate

- Fit a single model to measure the change in differences

Appropriate

- Intervention and control have similar baseline trend
- Not appropriate when trends converge or diverge prior to the intervention/policy

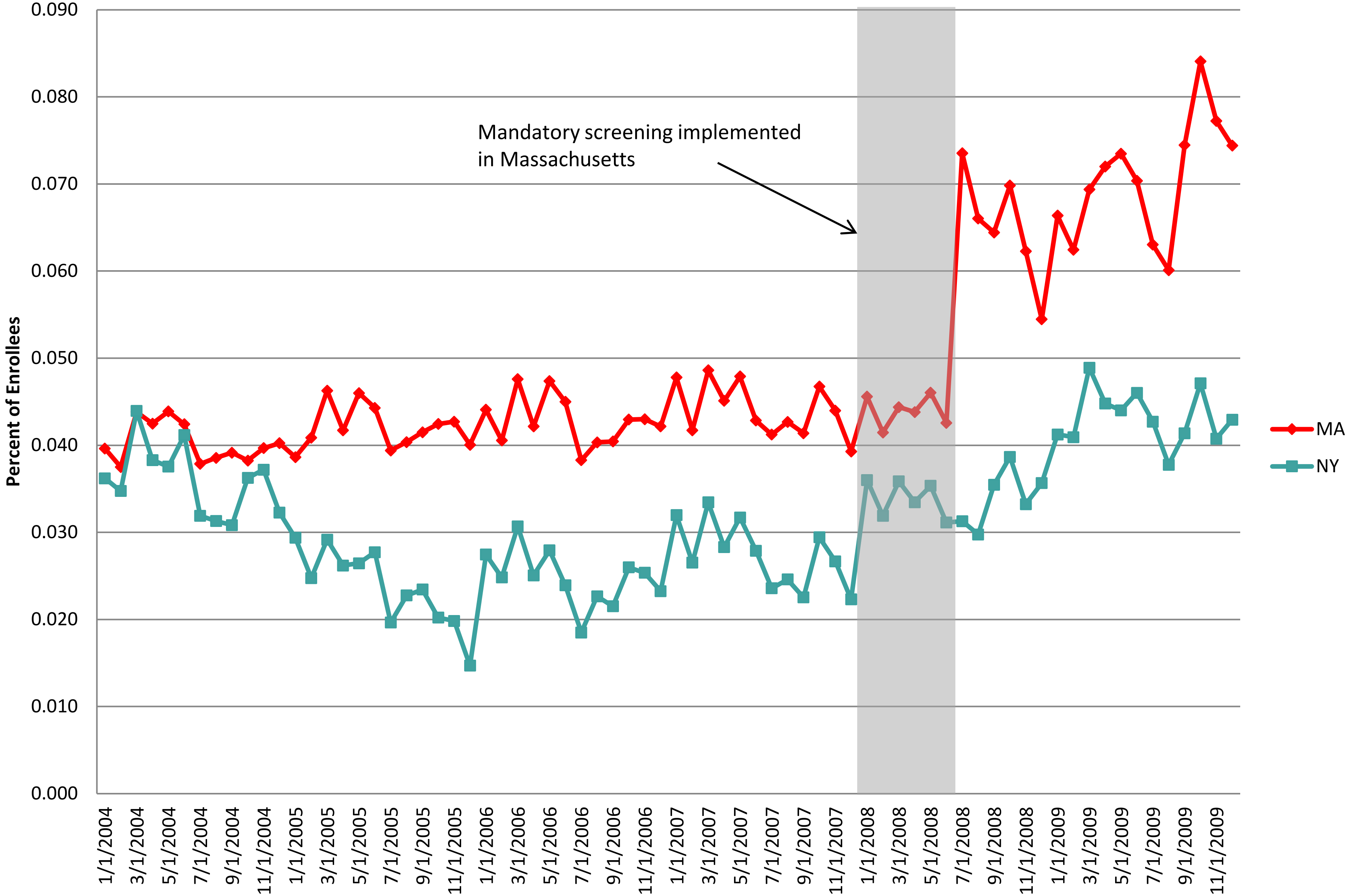
ADHD Hypothetical Example



WA vs. CA Stimulant Initiation

Quarter	WA	CA	WA_CA	Program	Time	Time_After
Jul-06	0.494	0.684	-0.190	0	1	0
Oct-06	0.513	0.613	-0.100	0	2	0
Jan-07	0.558	0.658	-0.100	0	3	0
Apr-07	0.527	0.627	-0.100	0	4	0
Jul-07	0.508	0.608	-0.100	0	5	0
Oct-07	0.489	0.619	-0.130	0	6	0
Jan-08	0.501	0.601	-0.100	0	7	0
Apr-08	0.536	0.636	-0.100	0	8	0
Jul-08	0.358	0.608	-0.250	0	9	0
Oct-08	0.182	0.582	-0.400	1	10	1
Jan-09	0.236	0.646	-0.410	1	11	2
Apr-09	0.201	0.621	-0.420	1	12	3
Jul-09	0.190	0.620	-0.430	1	13	4
Oct-09	0.179	0.629	-0.450	1	14	5
Jan-10	0.172	0.632	-0.460	1	15	6
Apr-10	0.169	0.609	-0.440	1	16	7
Jul-10	0.165	0.575	-0.410	1	17	8

Any Outpatient Mental Health Utilization



SCHOOL OF PUBLIC HEALTH
UNIVERSITY of WASHINGTON

KAISER PERMANENTE®

Other Considerations

Autocorrelation

Seasonality

Power

Binary and count outcomes

Serial Autocorrelation

Observations are correlated over time

- Similarity of observations as a function of the time lag between observations
- Deflates the standard error, inflates the statistical significance of time period as a predictor

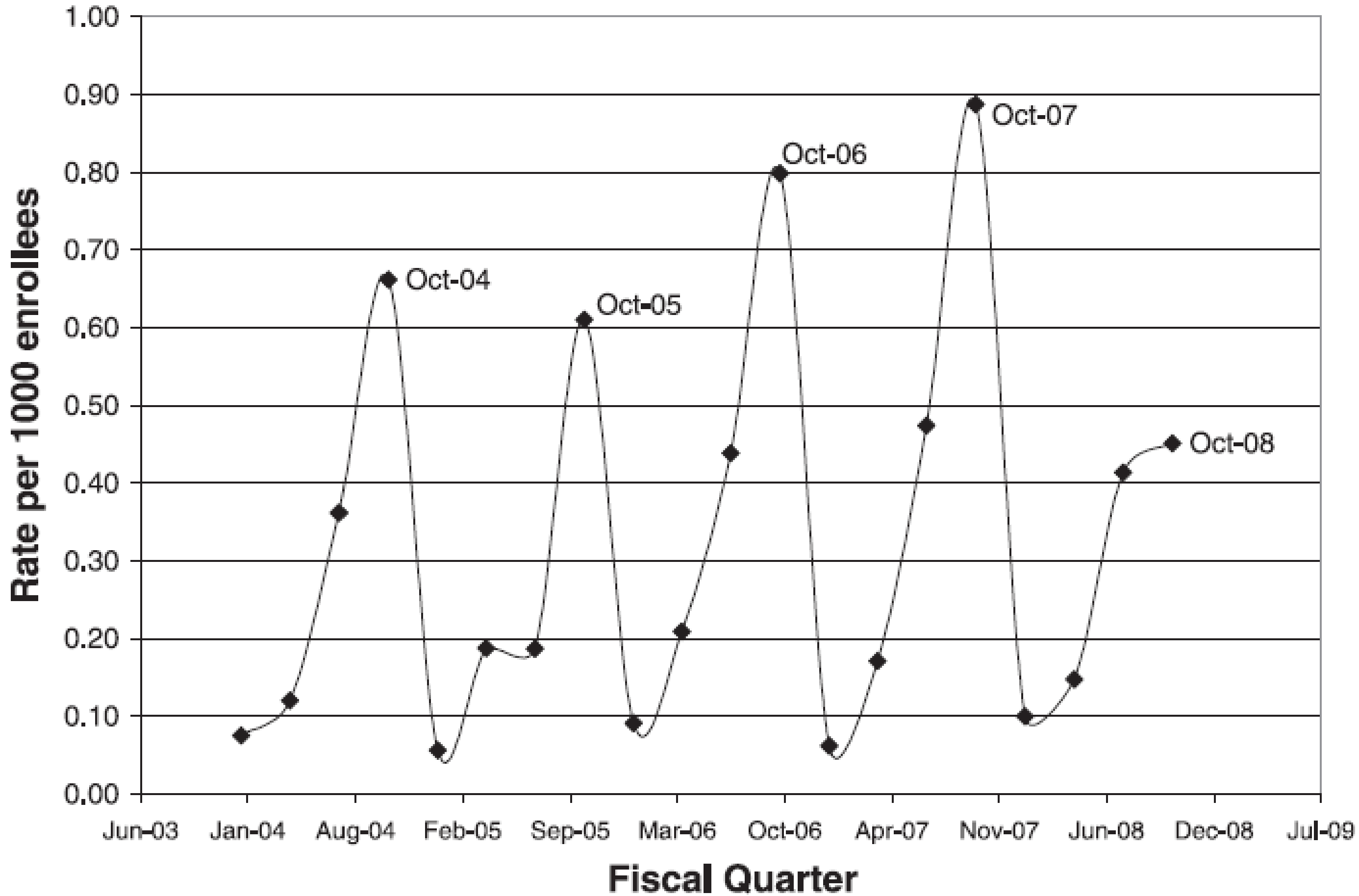
Need to add terms to control for (factor out) autocorrelation

Seasonality

Special case of autocorrelation

- Periodicity in the data associated with the underlying process
  - E.g., school year, flu season, etc

Seasonality in WA ADHD Medication Initiation



SCHOOL OF PUBLIC HEALTH
UNIVERSITY of WASHINGTON

KAISER PERMANENTE®

Smoothing

Aggregate time periods to reduce noise

- E.g., monthly to quarterly

Power

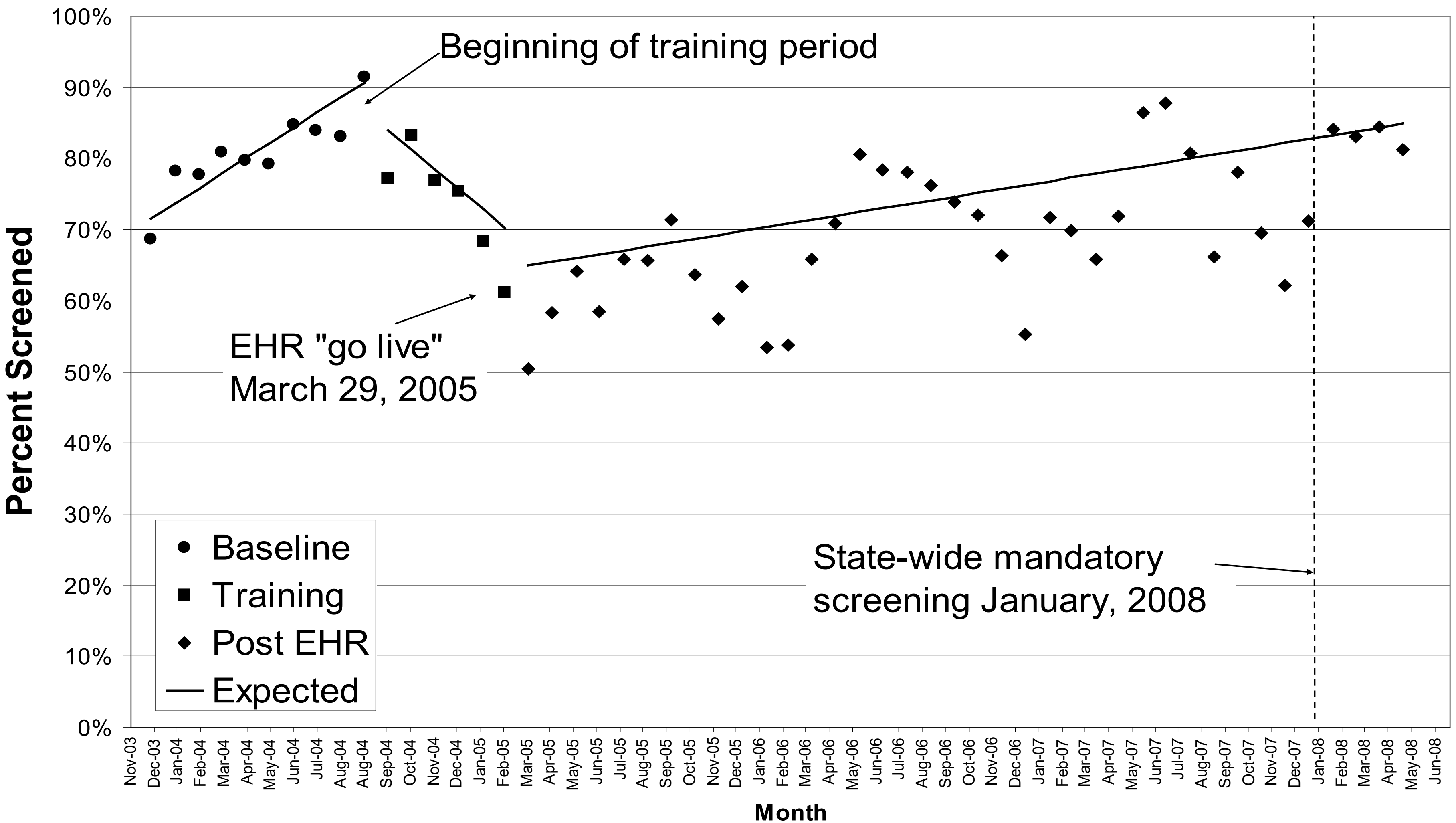
Power to detect a change in level or slope largely determined by:

- Variability over time ("noisy" series)
- Number of time periods

Binary and Count Outcomes

- Individual-level data
- Logistic regression (binary outcome)
- Poisson regression (counts)
- E.g., change in the log odds of the event occurring in the post period compared to the pre-period.

Hacker and Penfold, 2012



SCHOOL OF PUBLIC HEALTH

UNIVERSITY *of* WASHINGTON

KAISER PERMANENTE®

SAS & STATA Programs

SAS

```
PROC AUTOREG DATA=library.data outest=library.est1;
model Rate = time planning  time_after_planning policy  time_after_policy/
    method=ml nlag=6 backstep dwprob loglikl;
      output out=library.output p=predvar r=residvar;
RUN;
```

STATA

```
tsset timevar                    ***Declare a single panel of data***
itsa Rate, single trperiod(11,17) lag(12)
actest, lags(12)                 ***postestimation test for autocorrelation***
```

Example of Segmented Regression Results

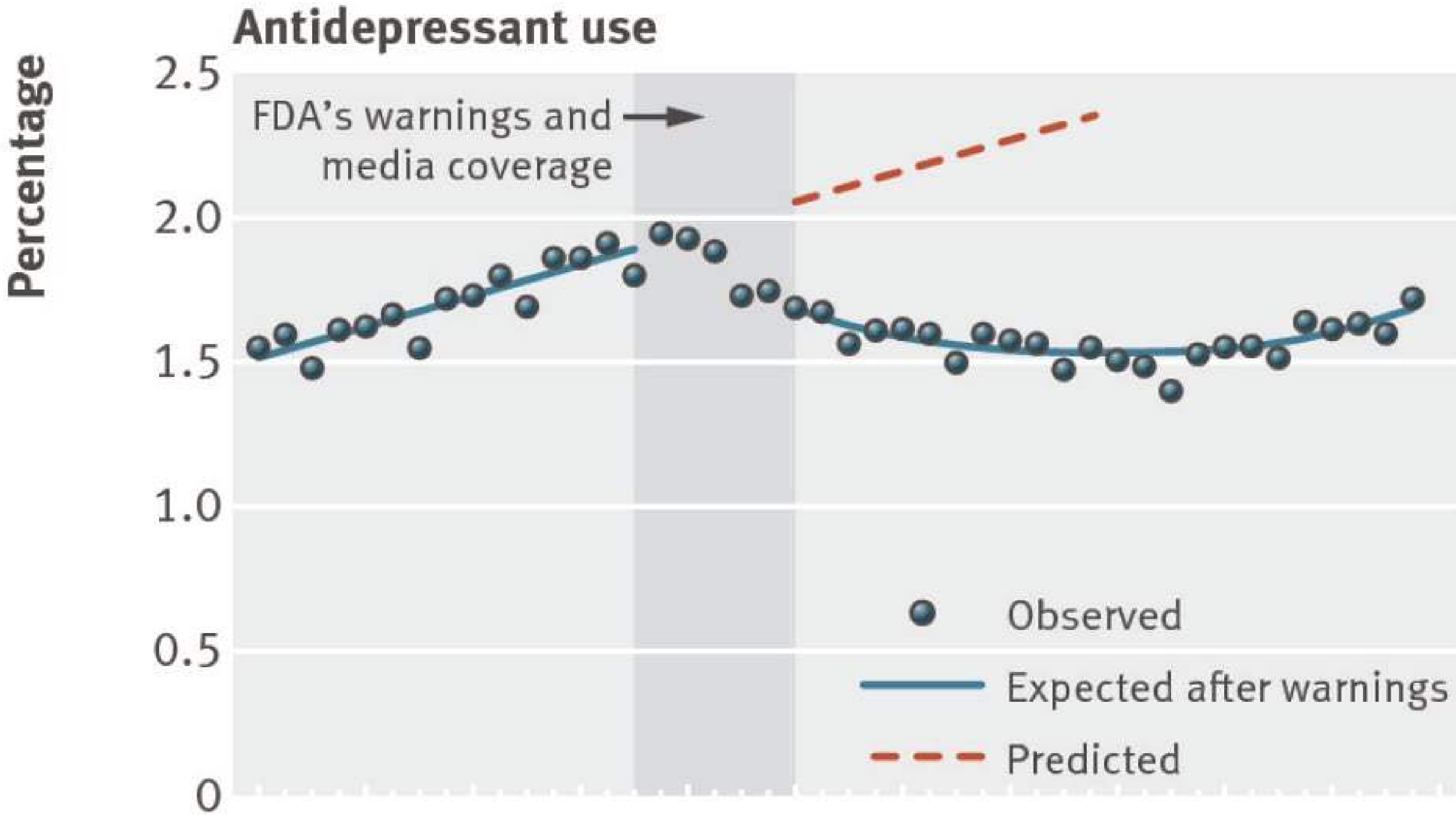
Covariate	Coefficient	Standard Error	T value	alpha
Intercept	0.694	0.044	15.89	<0.0001
Time	0.021	0.007	2.86	0.0063
Training (level)	-0.088	0.056	-1.57	0.1241
Time post training (trend)	-0.049	0.015	-3.26	0.0021
EHR implementation (level)	-0.058	0.040	-1.45	0.1532
Time post-EHR (trend)	0.033	0.011	2.89	0.0058
Autoregressive 5 month	0.358	0.125	2.85	0.0065
Autoregressive 6 month	0.441	0.131	3.37	0.0015

ITS vs. statistical process control (SPC)

- Two advantages of the SPC approach:
  - Managers and clinical staff with little or no statistical knowledge can easily understand and interpret the control charts
  - The charts provide information that supports prompt decision-making.
- Limitations:
  - Requires a stable baseline (no trend)
  - Uses a t-test to compare pre- and post-intervention means
    - Not slopes
  - Cannot control for overall secular trends
  - Control limits may be too narrow - autocorrelation not accounted for
    - May lead to false alerts

Some Examples

FDA Black Box Warning on Antidepressants

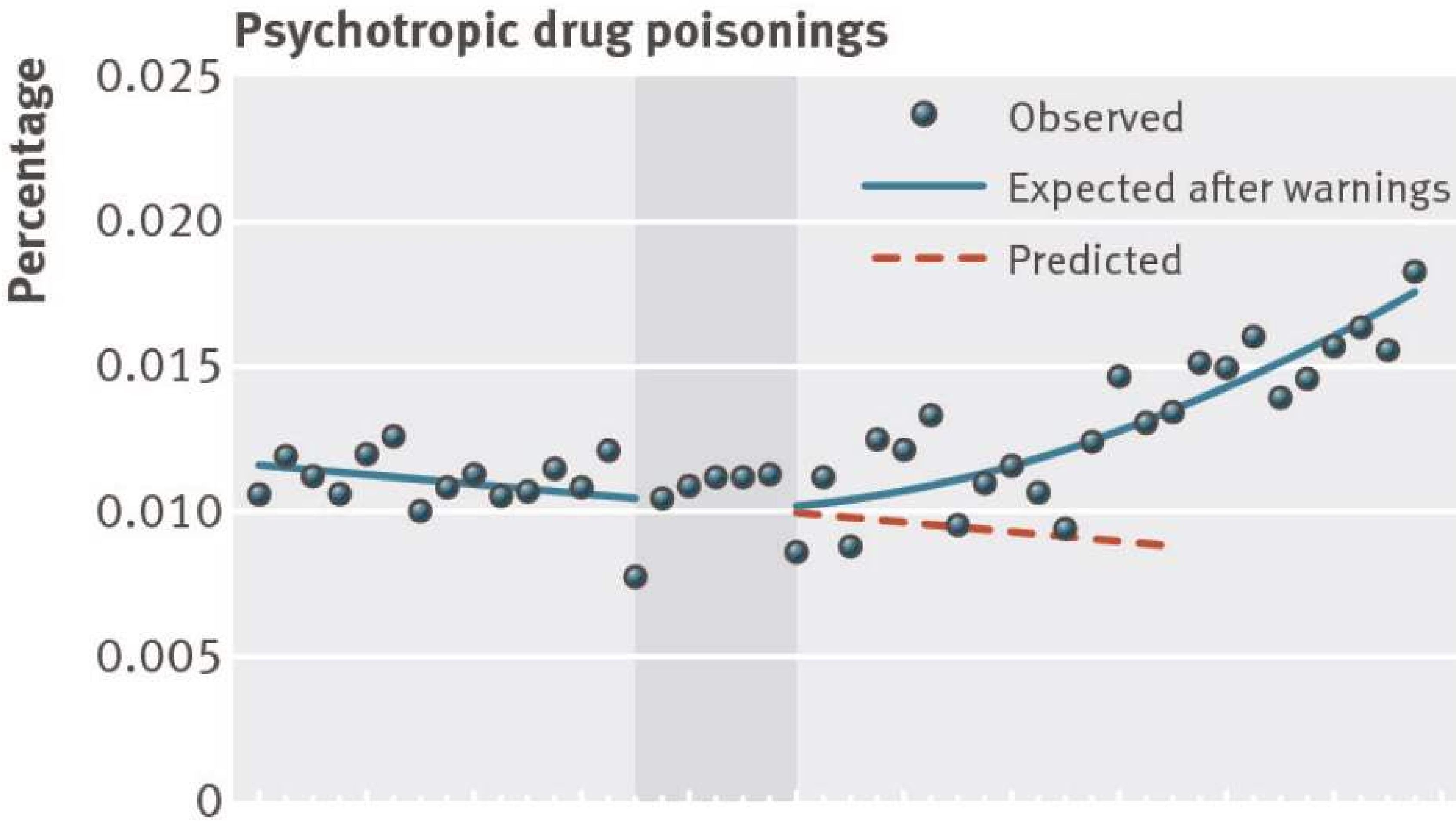


Lu et. al., 2014, BMJ

SCHOOL OF PUBLIC HEALTH
UNIVERSITY of WASHINGTON

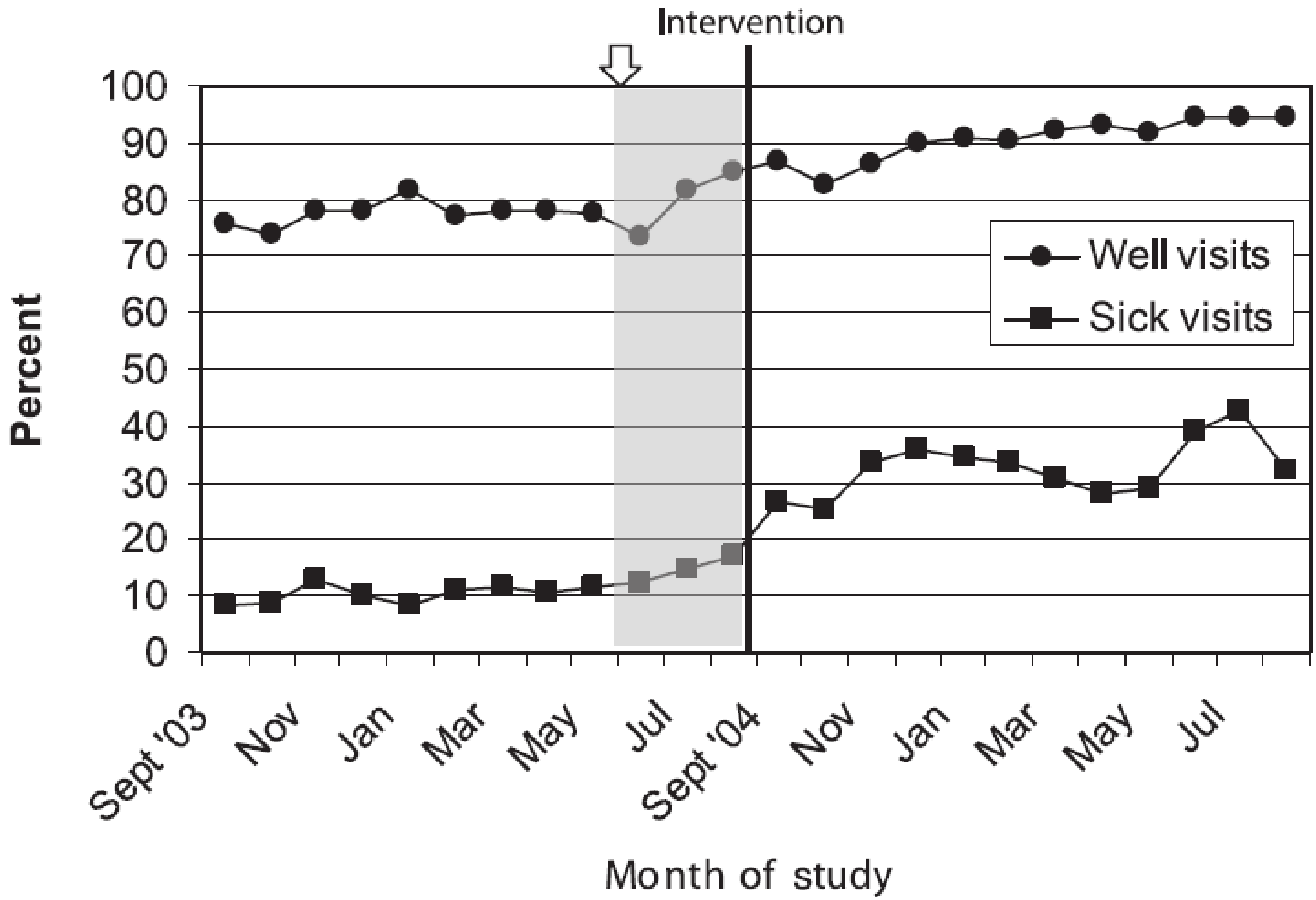
KAISER PERMANENTE

Black Box Warning cont'd

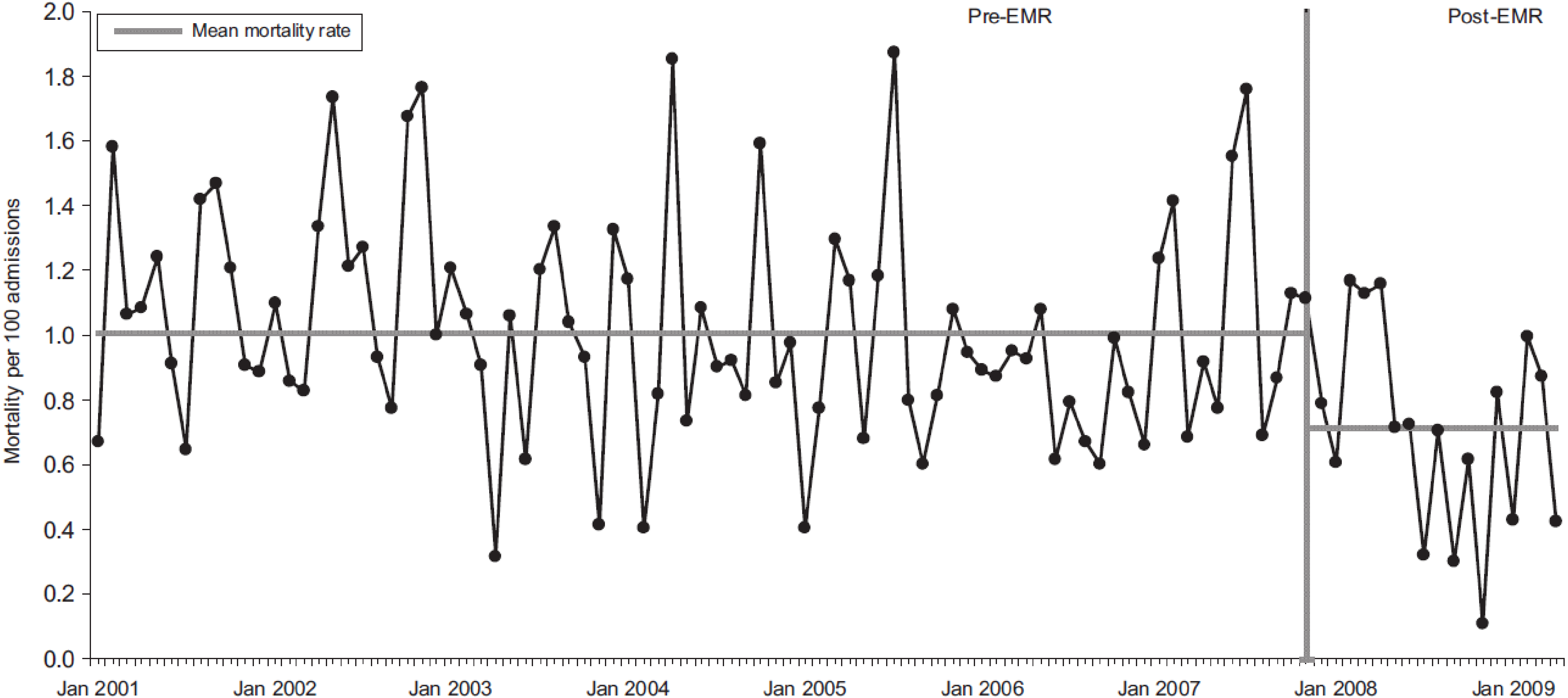


Lu et. al., 2014, BMJ

Fiks et. al., 2007

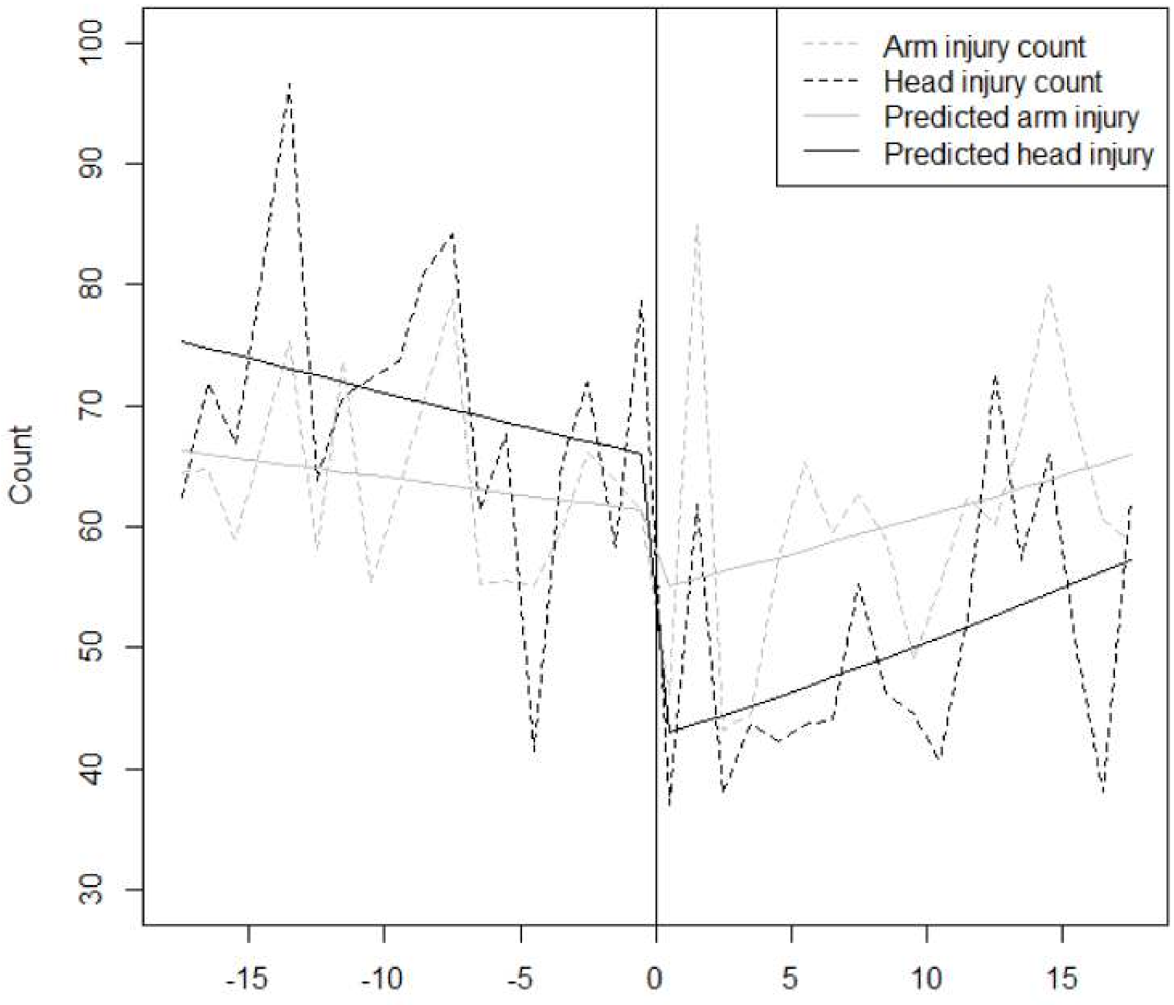


Longhurst et. al., 2010



Do mandatory bicycle helmet laws reduce head injuries?

$$
\begin{aligned}
Y_i \mid \eta_i &\sim Poisson\ (\eta_i \mu_i) \\
\eta_i &\sim Gamma(\alpha, \alpha) \\
\log(\mu_i) &= \beta_0 + \beta_1 TIME + \beta_2 INJURY + \beta_3 LAW + \beta_4 TIME \times INJURY + \beta_5 TIME \times LAW \\
&\quad + \beta_6 INJURY \times LAW + \beta_7 TIME \times INJURY \times LAW + \log(\text{exposure}),
\end{aligned}
$$



Arm injury count
Head injury count
Predicted arm injury
Predicted head injury
Count
100
90
80
70
60
50
40
30
-15
-10
-5
0
5
10
15

ITS Design Standards

Crispness of the policy/intervention implementation

Definition of the population(s)

Composition of the population(s)

Testing the functional form of the statistical model

Inclusion of a graph

Exchangeability of the control population/series

Discussion of competing interventions and other threats to validity

Group Exercise

Read the methods section of the Hacker paper

Use what you know about threats to validity and ITS evaluation standards to review the manuscript

Discuss with your breakout participants

Report back to the larger group on what you think

Group Exercise 2 (time permitting)

Design an ITS evaluation

Evaluate the impact of switching to 100% telehealth service delivery during the COVID-19 pandemic for youth with psychosocial needs at SCH.

- Include any implementation considerations in your design

Design Decisions

Hypotheses

Population inclusion/exclusion criteria

Time periods

QI outcome measurement

Sequential roll-out?

Model specification

Threats to internal/external validity?

How will you select implementation dates (interruptions)?

Stratification?

Sub-population analyses?

Wrap-up

Questions? Comments?

Please fill out an evaluation!

Interrupted
Time
Series
Analysis

https://associationresearch.limequery.com/433562?lang=en