Analysis of time-stratified case-crossover studies in environmental epidemiology using Stata

Aurelio Tobías
Spanish Council for Scientific Research (CSIC), Barcelona, Spain

Ben Armstrong and Antonio Gasparrini
London School of Hygiene and Tropical Medicine, UK

[Updated November 2022]

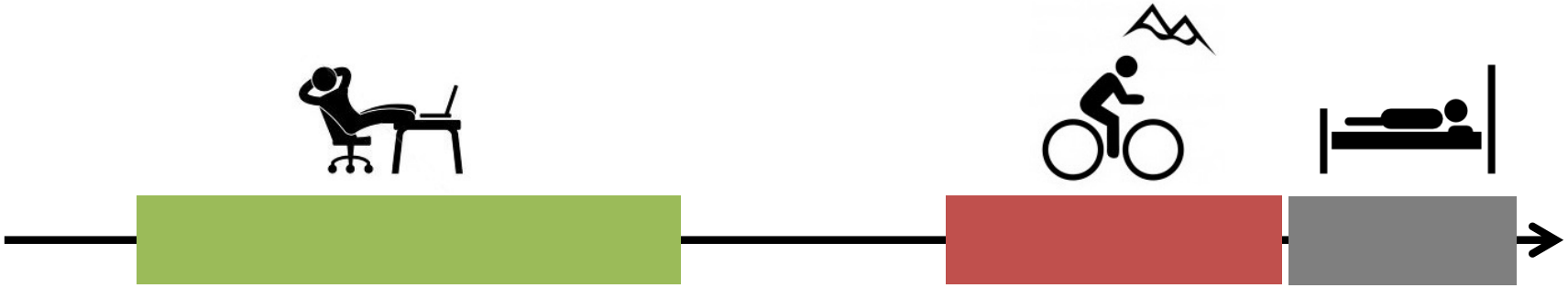
2014 UK Stata Users Group meeting
London, 12th September 2014

Background

- The time stratified case crossover design is a popular alternative to conventional time series regression for analysing associations between time series of environmental exposures (air pollution, weather) and counts of health outcomes

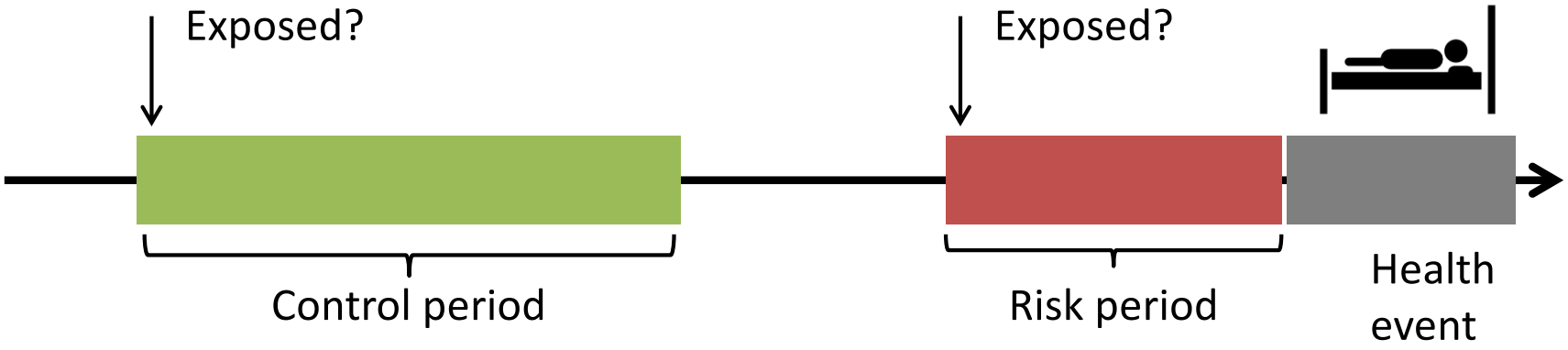
The case-crossover design

- Proposed by *Maclure (1991)* to study transient effects on the risk of acute health events



The case-crossover design

- Proposed by Maclure (1991) to study transient effects on the risk of acute health events



- Analysis likewise a matched case-control

		Control: Exp.	Control: No
Risk period	Exp.	a	b
	No	c	d

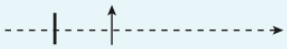
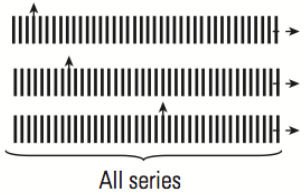
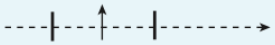
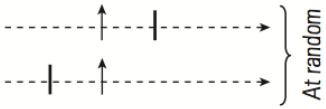
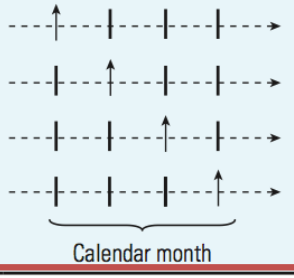
OR = b/c

Application to environmental studies

- Firstly adapted to air pollution studies
  - Philadelphia (Neas et al. 1999) and Barcelona (Sunyer et al. 2000)
- comparing exposure levels for a given day (t) when health event occurs *vs.* levels before and after the health event
- Allows to control for time-trend and seasonality by design, since it compares exposure levels between same weekdays within each month of each year

Time-stratified case-crossover

Table 1. Comparison of different CCO designs.

Reference	Type	Selection of controls	Advantages	Factors that can introduce bias	Selection of controls diagram
Maclure 1991	CCO	One control point before the effect	All possible confounding factors undergoing no change between control periods and effect, automatically controlled for by design	Long-term trends or seasonality	
Navidi 1998	Full-stratum bidirectional	For each case, all the days of the series other than that of the event taken as controls	Provides control for long-term trends	Long-term trends (only partially controlled for) or seasonality	
Bateson and Schwartz 1999	Symmetric bidirectional	Two at equal distance of the event	Provides adequate control for long-term trends and seasonality		
Navidi and Weinhandl 2002	Semisymmetric bidirectional	One chosen at random from the two used for symmetric bidirectional CCO	Provides adequate control for long-term trends and seasonality		
Lumley and Levy 2000	Time stratified	One (or several) within the same time stratum in which the event occurred	Provides adequate control for long-term trends and seasonality		

Arrows pointing up indicate case periods; horizontal arrows represent direction of time within 1 month; dashed lines indicate time periods of 1 day; vertical lines indicate control periods.

(Figueiras et al. 2010)

```
. use madrid, clear
. list date y pm10 temp in 1/31, noobs clean
```

date	y	pm10	temp
01jan2003	64	14	9.6
02jan2003	56	12	11
03jan2003	73	16	9.8
04jan2003	69	14	8.8
05jan2003	71	17	4.3
06jan2003	68	9	7
07jan2003	63	19	3
08jan2003	85	13	6.8
09jan2003	67	13	4.6
10jan2003	80	22	2.7
11jan2003	65	23	1
12jan2003	95	17	.85
13jan2003	60	45	.5
14jan2003	76	60	.45
15jan2003	77	72	1.3
16jan2003	75	54	2.6
17jan2003	76	65	2
18jan2003	75	43	.8
19jan2003	74	12	6.3
20jan2003	79	19	5.8
21jan2003	73	16	8
22jan2003	72	21	6.7
23jan2003	72	25	7
24jan2003	74	46	6.2
25jan2003	59	42	8.1
26jan2003	77	26	10
27jan2003	64	22	15
28jan2003	65	41	9.7
29jan2003	74	18	5.7
30jan2003	77	17	4.8
31jan2003	55	17	2.3

Sun	Mon	Tue	Wed	Thu	Fri	Sat
			1	2	3	4
5	6	7	8	9	10	11
12	13	14	15	16	17	18
19	20	21	22	23	24	25
26	27	28	29	30	31	

```
. egen set = group(yy mm dow)

. list set date y pm10 temp if yy==2003 & mm==2 & dow==1, noobs clean

     set         date      y   pm10    temp
 2003021   06jan2003     68      9       7
 2003021   13jan2003     60     45      .5
 2003021   20jan2003     79     19     5.8
 2003021   27jan2003     64     22    14.8
```

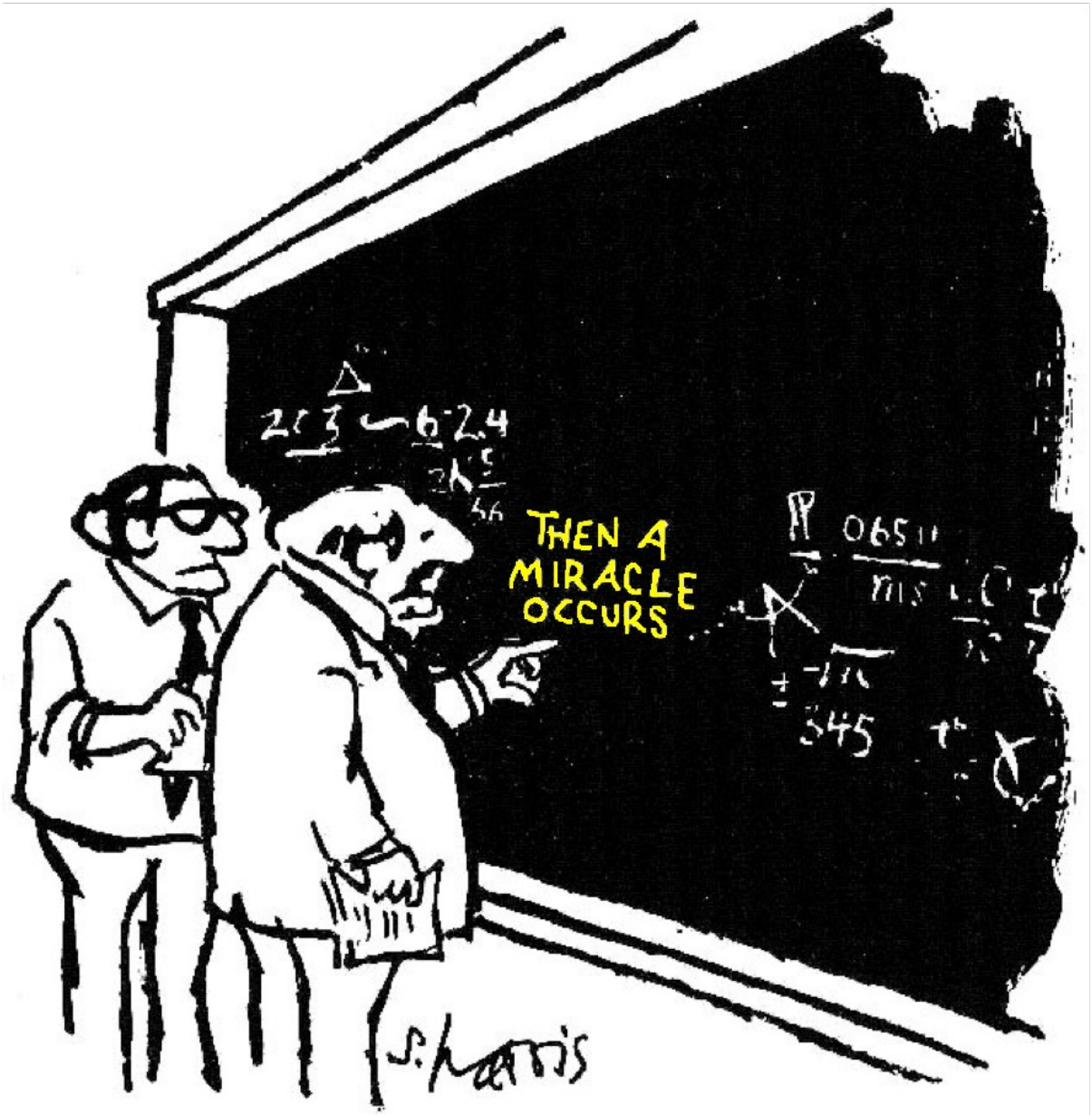
Conditional Poisson regression

- Armstrong et al. *(BMC Med Res Methods 2014)*
  - *" … Poisson models with large numbers of indicator variables can alternatively be fit with conditional Poisson models, conditioning on numbers of events in the time stratum"*
  - Firstly proposed by *Farrington (1995)* for the self-controlled case-series design for vaccine safety studies
- Strata defined by grouping same weekday within each month of each year, and then use Stata's **`xtpoisson`** command jointly with the **`fe`** option
- But overdispersion needs to accounted for with the *new* written-user **`xtodp`** command

```
. egen set = group(yy mm dow)
. xtset set date
...

. xtpoisson y pm10, fe
...

Conditional fixed-effects Poisson regression    Number of obs      =      1096
Group variable: set                             Number of groups   =       252

                                                Obs per group: min =         4
                                                               avg =       4.3
                                                               max =         5

                                                Wald chi2(1)       =      8.42
Log likelihood  = -2854.4979                    Prob > chi2        =    0.0037

------------------------------------------------------------------------------
           y |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        pm10 |   .0007406   .0002552     2.90   0.004     .0002404    .0012408
------------------------------------------------------------------------------

. xtodp

Estimate and standard errors corrected for over-dipersion
df: 843 ; pearson x2:  1065.6 ; dispersion:     1.26
------------------------------------------------------------------------------
           y  |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        pm10 |   .0007406   .0002869     2.58   0.010     .0001782     .001303
------------------------------------------------------------------------------
```

Summary

- Case-crossover studies can easily be analysed using **Stata** with conditional Poisson regression
- Avoids unfriendly data management and large (huge) data-sets at individual level keeping the original time-series format
- Easy fit of interaction terms to investigate effect modifiers (e.g., sex, age)
- Can account for overdispersion installing the user-written command **`xtodp`**, and residual autocorrelation



"I think you should be more explicit here in step two."

Thanks for your attention!