Acute effects of ambient exposures

Time series and case-crossover studies

Brooke Anderson

February 3, 2016

Example data: Chicago NMMAPS

chicagoNMMAPS data

For the examples in this lecture, I'll use some data from Chicago on mortality, temperature, and air pollution. These data are available as part of the `dlnm` package. You can load them in R using the following code:

```
library(dlnm)
```

```
## This is dlnm 2.1.3. For details: help(dlnm) and vignette
## Important changes since version 2.0.0
## See: 'file.show(system.file('Changesince200',package='dl
```

```
data("chicagoNMMAPS")
```

chicagoNMMAPS data

To make the data a little easier to use, I'll rename the data frame as `chic`:

```
chic <- chicagoNMMAPS
chic[1:3, c("date", "death", "cvd", "temp", "dptp", "pm10",
```

```
##         date death cvd       temp   dptp     pm10
## 1 1987-01-01   130  65 -0.2777778 31.500 26.95607 4.3760
## 2 1987-01-02   150  73  0.5555556 29.875       NA 4.9298
## 3 1987-01-03   101  43  0.5555556 27.375 32.83869 3.7510
```

chicagoNMMAPS data

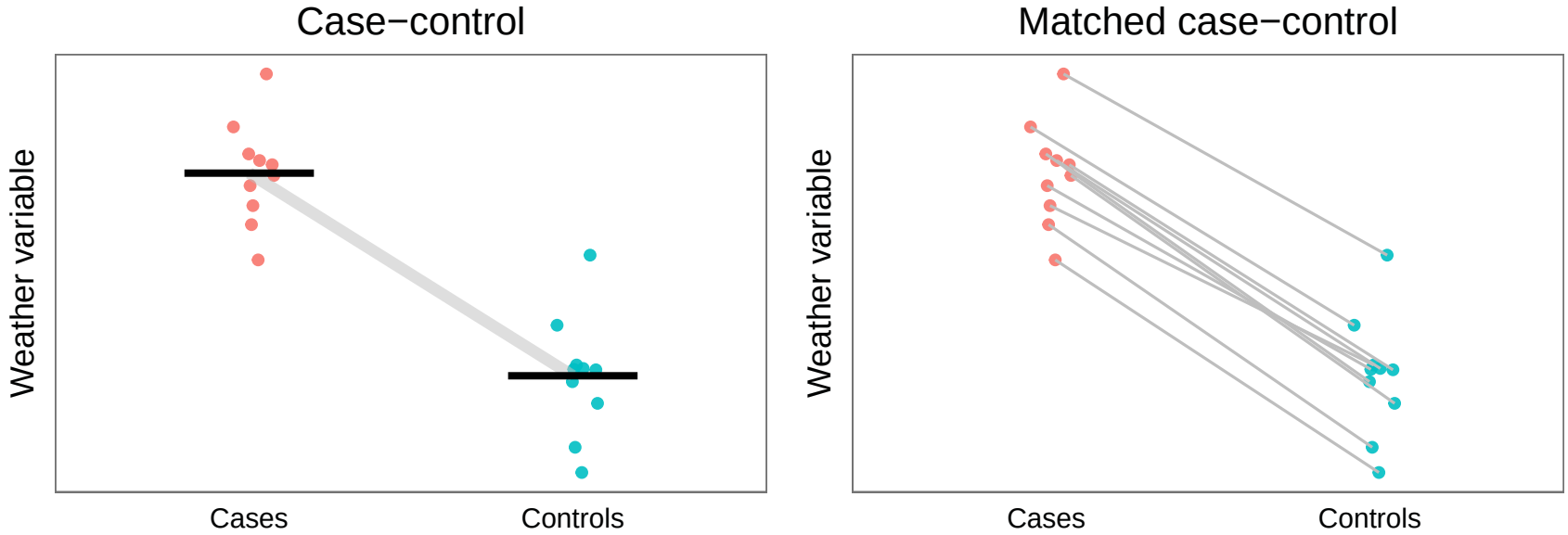
To find out more about this data, you can look at its help file:

```
?chicagoNMMAPS
```

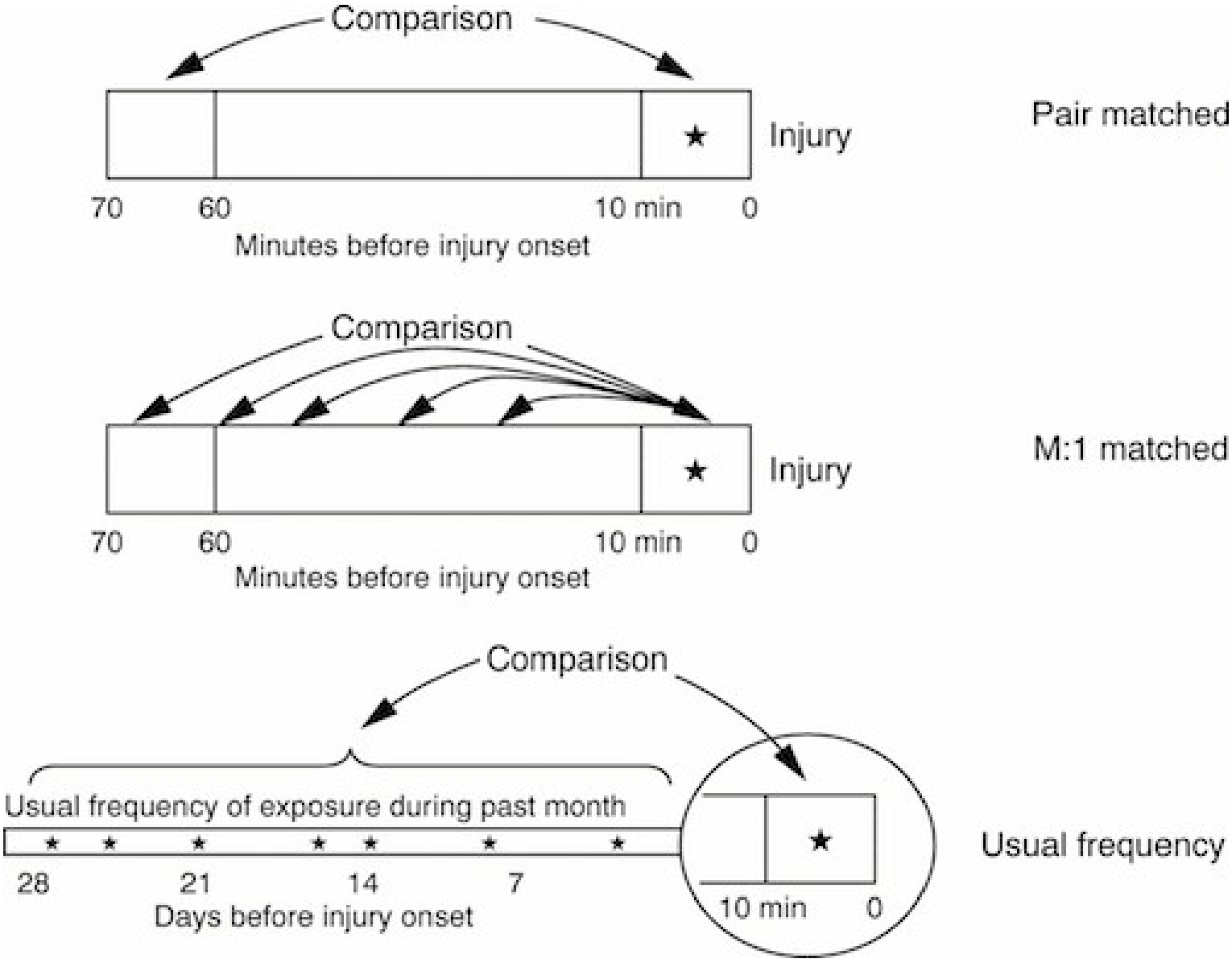
Concept: Case-crossover studies

Case-crossover models

Case-crossover model designs are based on the idea of matched case-control studies. For these, instead of comparing averages of exposure for cases versus controls, you compare the average difference across each matched set of case and control(s).

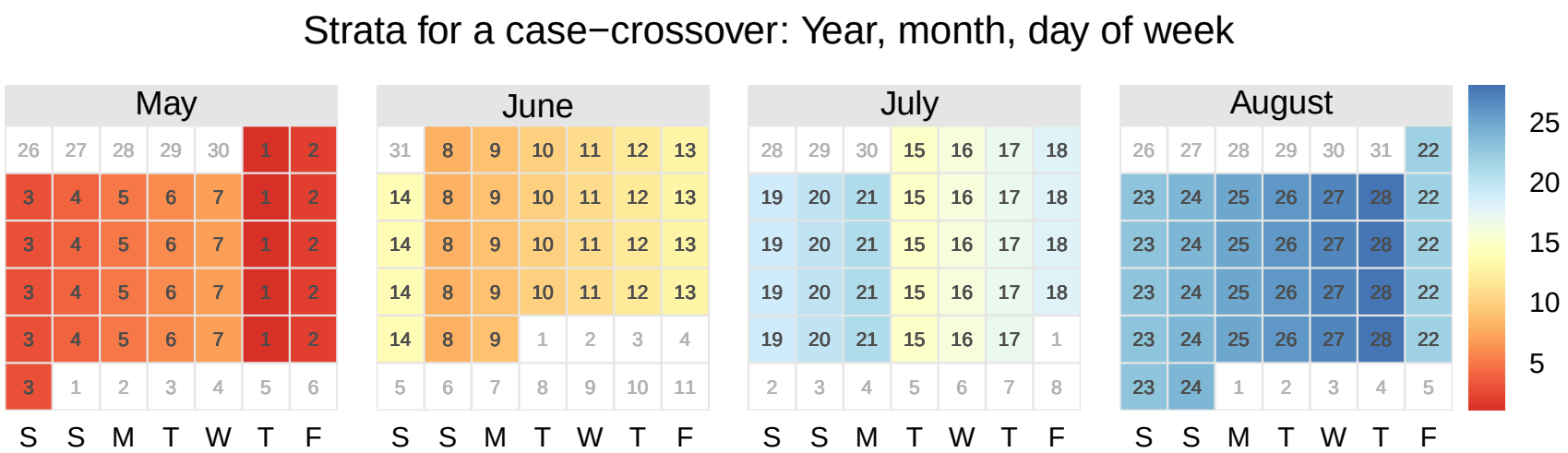


Types of case-crossover designs



Source: Sorock et al. 2001, Injury Prevention

Strata for case-crossover



Implementation: Case-crossover studies

GLM method

One way to fit this type of model is using a GLM, but with control by strata of year-month-day of week.
Case-crossover fit using a GLM:

$$E(log(Y_t)) \sim \beta_0 + \beta_1 PM_t + \beta_2 Stratum_t$$

GLM method

To code this, first you need to create a column with the stratum. In R, you can use `format` with the date to do this easily, and then convert the formatted date for a `factor` class:

```
chic$casecross_stratum <- format(chic$date, "%Y-%m-%a")
chic$casecross_stratum <- factor(chic$casecross_stratum)
head(chic$casecross_stratum, 3)
```

```
## [1] 1987-01-Thu 1987-01-Fri 1987-01-Sat
## 1176 Levels: 1987-01-Fri 1987-01-Mon 1987-01-Sat 1987-01
```

Case-crossover

Now you can include this factor in your model (note: this takes the place of model control for time trends and day of week in a typical time series model):

```
mod_f <- glm(cvd ~ pm10 + ns(temp, 4) + casecross_stratum,
             data = chic,
             family = quasipoisson())
summary(mod_f)$coef[c(1:2, 7:10), ]
```

```
##                                      Estimate   Std. Error
## (Intercept)                      4.0482946294 0.0855215089 4
## pm10                             0.0001909843 0.0001680322
## casecross_stratum1987-01-Mon     0.1907876393 0.1137590495
## casecross_stratum1987-01-Sat     0.0855529446 0.1168756412
## casecross_stratum1987-01-Sun     0.3300835895 0.1099033832
## casecross_stratum1987-01-Thu     0.0462517003 0.1043859066
##                                     Pr(>|t|)
## (Intercept)                     0.000000000
## pm10                            0.255782198
```

Case-crossover

You can interpret the coefficients now in the same way as with the time series model:

```
pm_coef <- summary(mod_f)$coefficients["pm10", ]
100 * (exp(10 * pm_coef[1]) - 1)
```

```
##  Estimate
## 0.1911668
```

Therefore, for this model, there is a 0.191% increase in mortality for an increase of 10 $\mu g/m^3$ PM10.

Case-crossover

There are also other methods for fitting case-crossover models:

- Armstrong et al. (Conditional Poisson models: a flexible alternative to conditional logistic case cross-over analysis) suggest using a conditional Poisson regression model (`gnm()`) to speed up computational time.
- The `casecross` function in the `season` package by Adrian Barnett uses 28-day strata (rather than by month) and a Cox proportional hazards regression model to fit the model.

If you are using this method for a paper, it is worthwhile testing the different methods to see if you get similar results.

Case-crossover

Using a conditional Poisson model:

```
library(gnm)
mod_g <- gnm(cvd ~ pm10 + ns(temp, 4),
             eliminate = casecross_stratum,
             data = chic,
             family = quasipoisson())
pm_coef <- summary(mod_g)$coefficients["pm10", ]
100 * (exp(10 * pm_coef[1]) - 1)
```

```
##  Estimate
## 0.1911668
```

Case-crossover

Using a Cox proportional hazards regression model:

```
library(season)
mod_h <- casecross(cvd ~ pm10 + temp,
                   matchdow = TRUE,
                   data = chic)
```

```
## Note, irregularly spaced data...
## ...check your data for missing days
```

```
pm_coef <- mod_h$c.model$coefficients[1]
100 * (exp(10 * pm_coef[1]) - 1)
```

```
##       pm10
## 0.4320228
```

Multi-city studies

NMMAPS

Source: www.ihapss.jhsph.edu

NMMAPS package

NMMAPSdata package

Data

- akr
- albu
- Anch

and 105 other US cities

- *Meta-data on cities (population, location, counties, Census variables)*

Functions

- readCity
- getMetaData

and various other functions for different versions of the package

Documentation

- PDF users' manual
- Instructions for each function within R
- Examples for each function within R
- Website

Impact of NMMAPS

Research impacts of NMMAPS package

- As of November 2011, 67 publications had been published using this data, with 1,781 citations to these papers
- Research using NMMAPS has been used by the US EPA in creating regulatory impact statements for air pollution (particulates and ozone)
- "Thanks to NMMAPS, there is probably no other country in the world with a greater understanding of the health effects of air pollution and heat waves in its population.""

Source: Barnett, Huang, and Turner, "Benefits of Publicly Available Data", Epidemiology 2012