CoCare Society

東京大学
THE UNIVERSITY OF TOKYO

Analysing seasonality and the role of environmental exposure

Lina Madaniyazi

Shool of Tropical Medicine and Global Health

Nagasaki University

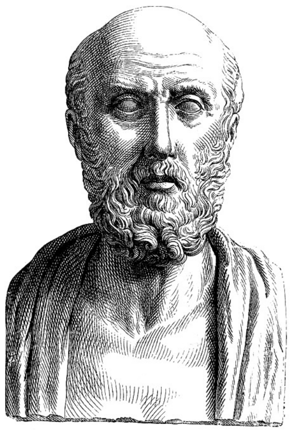
JST
Japan Science and Technology Agency

Outline

To introduce a set of statistical procedures to assess the seasonal variation of a health outcome and its changes after accounting for the potential environmental drivers.

- Seasonality (What & Why)
- Assessing seasonality and the role of environmental exposure
- Hands-on exercise

Seasonality



"All diseases occur at all seasons of the year, but certain of them are more apt to occur and be exacerbated at certain seasons."

- Hippocrates in around 400BC

"A pattern in a health outcome or exposure that increases and then decreases with some regularity"

Source: Analysing Seasonal Health Data

- Adrian G. Barnett and Annette J. Dobson

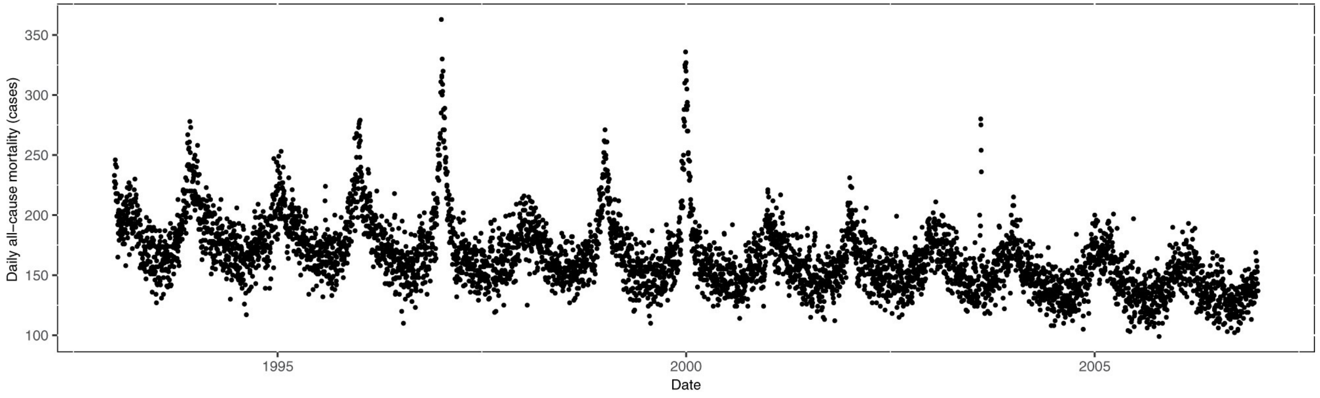


Figure 1. The time series of daily all-cause mortality in London from 1993 to 2006

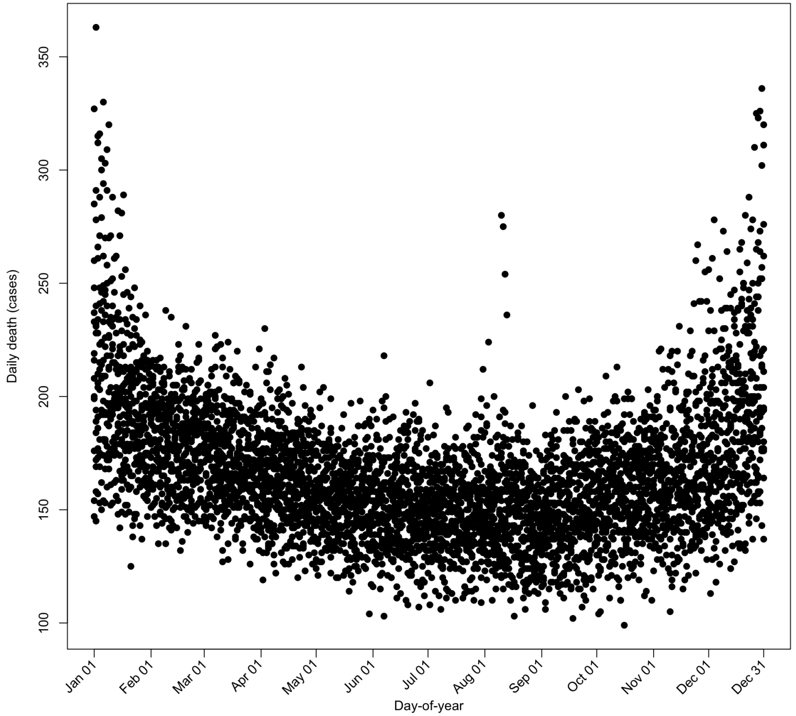
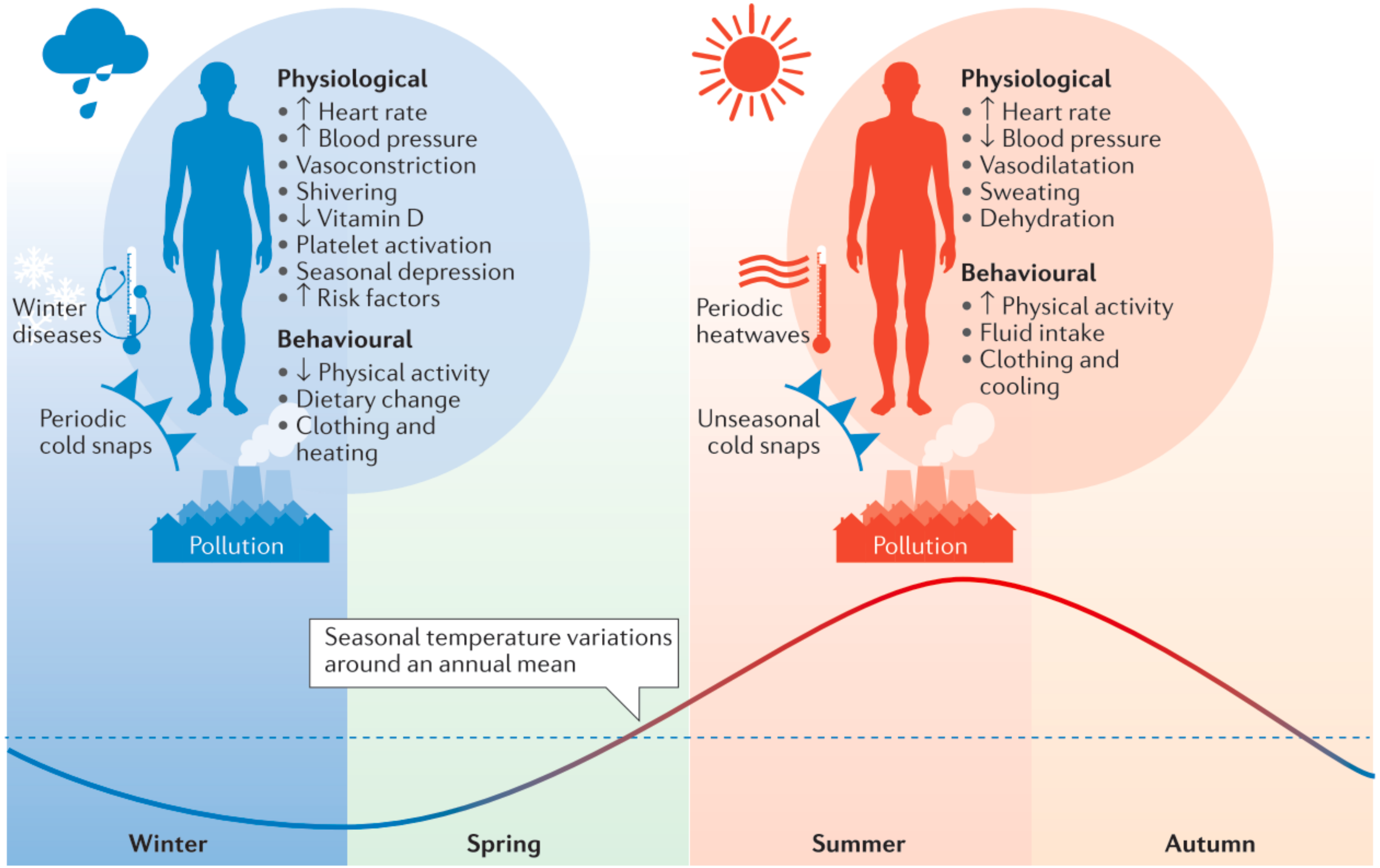


Figure 2. Scatter plot of daily all-cause mortality in London from 1993 to 2006 by the day-of-year



The factors contributing to the seasonality are considered as **seasonal factors.**

Stewart, Simon, et al. "Seasonal variations in cardiovascular disease." Nature Reviews Cardiology 14.11 (2017): 654.

- Seasonality is usually considered as one of the main confounders when examining the health impact of environmental factors.

- Assessing seasonality and the role of environmental exposure is of interest because it provides an insight into the etiology of the disease:

The changes in seasonality after adjustment will help us understand these factors' contributions to seasonality. In addition, the periodic and generally regular patterns in the residual time-series may provide clues as to the presence and importance of other currently unknown or unmeasured causes, such as human behavior.

Time series regression to assess seasonality

$$y_t = m_t + s_t + w_t$$

y_t can be described by a sum of series:
m and s as long-term trend and seasonality; residualsual series w_t

When assessing seasonality, we usually use month-of-year (1 to 12), week-of-year (1 to 52), and day-of-year (1 to 365) as an exposure indicator for seasonality

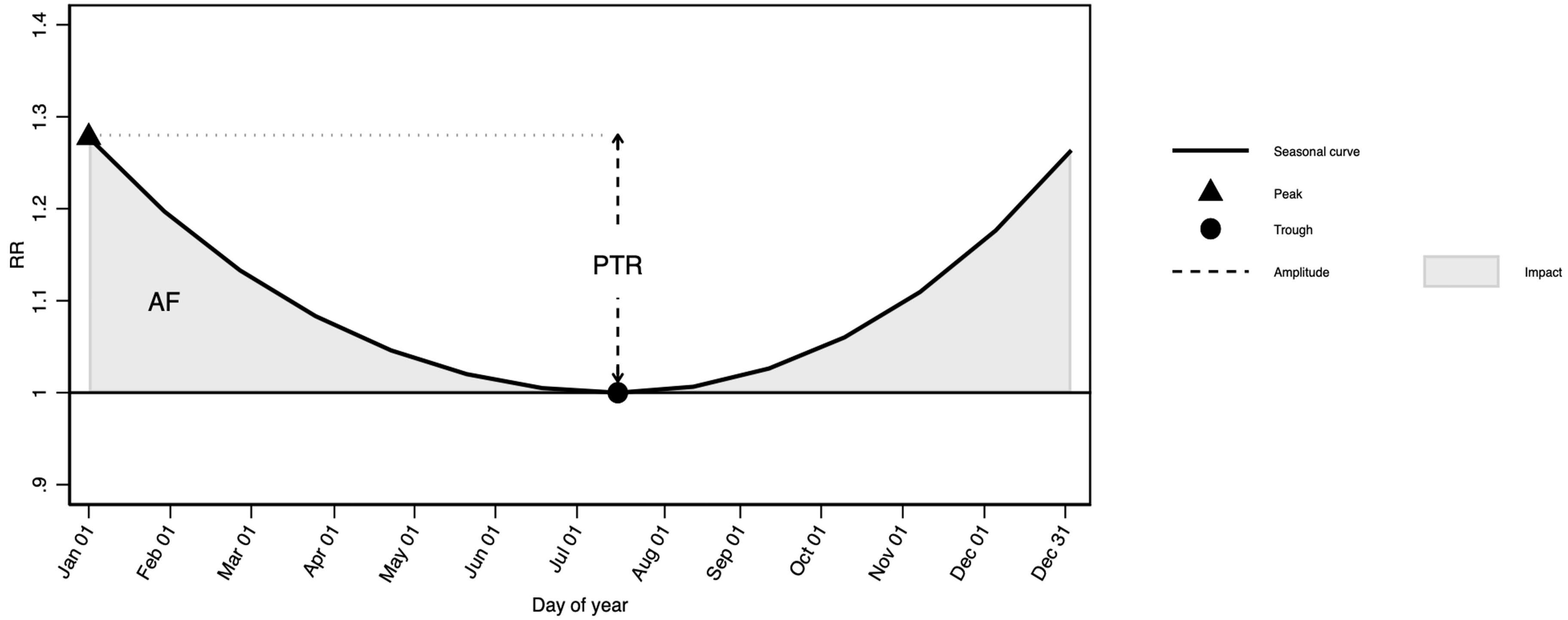
Approach (S_t)

- Strata indicators for months
- Harmonic (Fourier) terms of month or day-of-year
- Cyclic splines of month or day-of-year

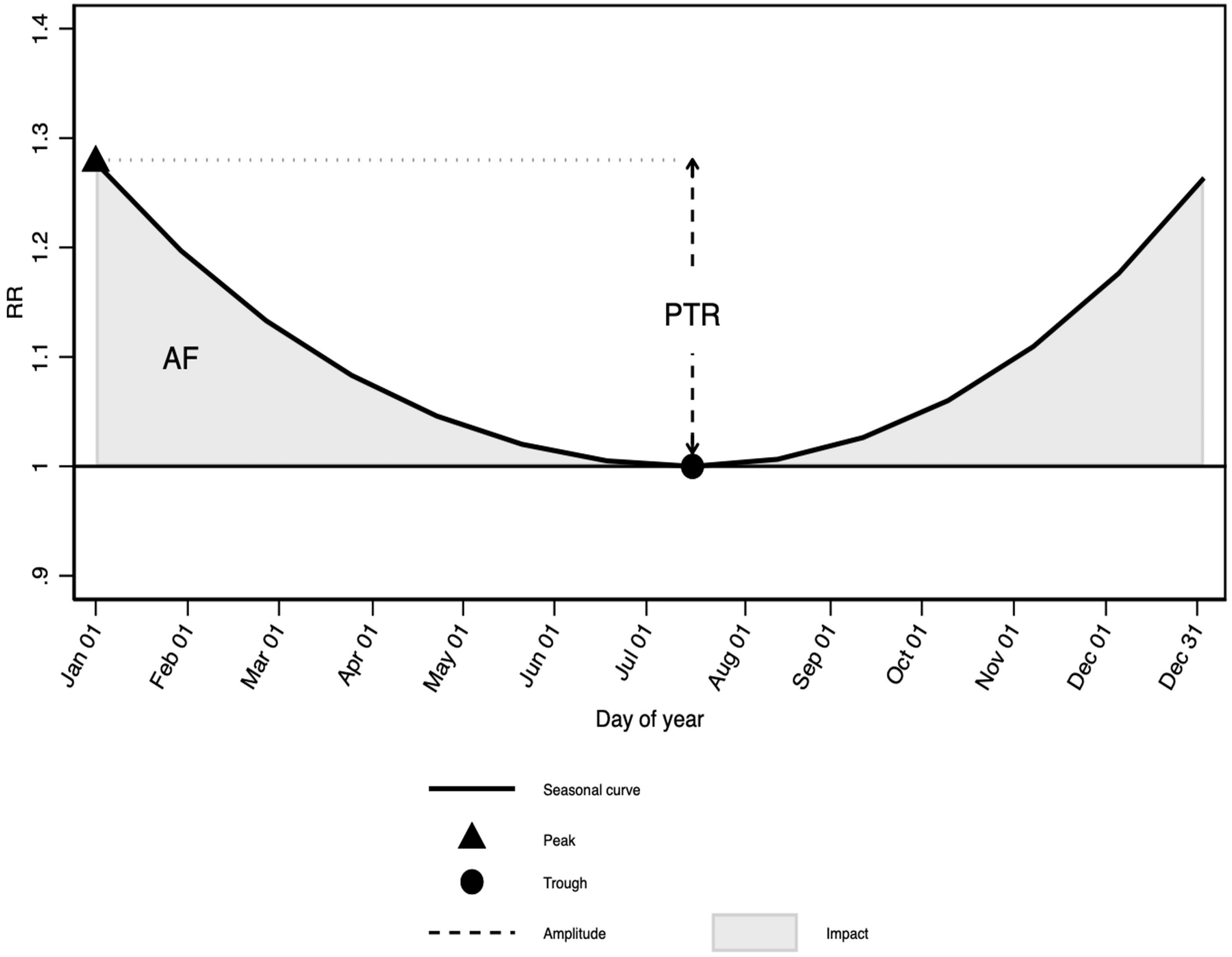
Each method has entails different assumptions and degree of flexibility

You should select the appropriate model based on your data, research question, and model fit.

Summarise **seasonality**



$$\boldsymbol{Relative\ risk} = \frac{(outcome)\ Estimate\ on\ the\ season\_indicator\ x}{(outcome)\ Estimate\ at\ the\ trough}$$

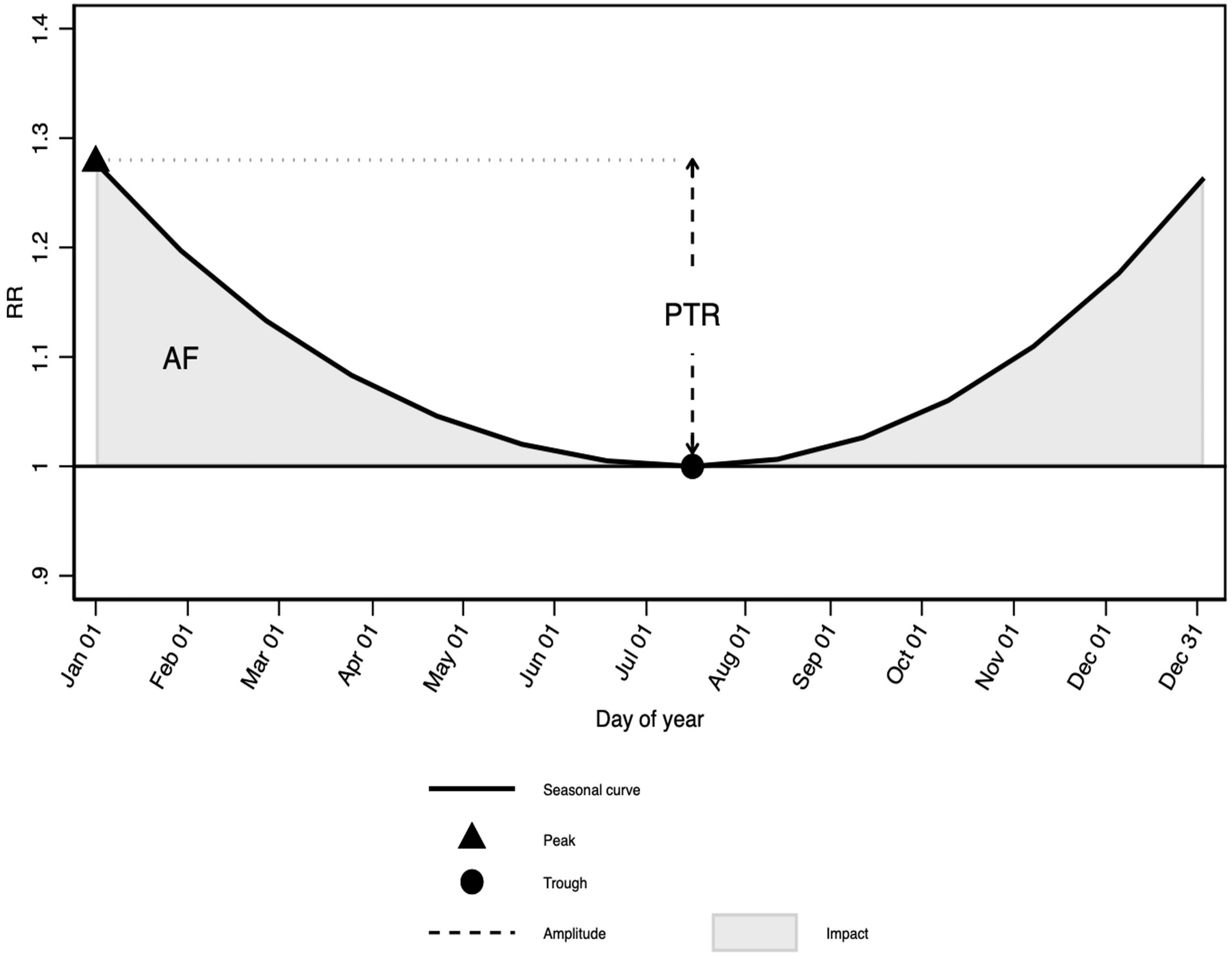


Shape: the fitted seasonal curve

Timings: peak and trough

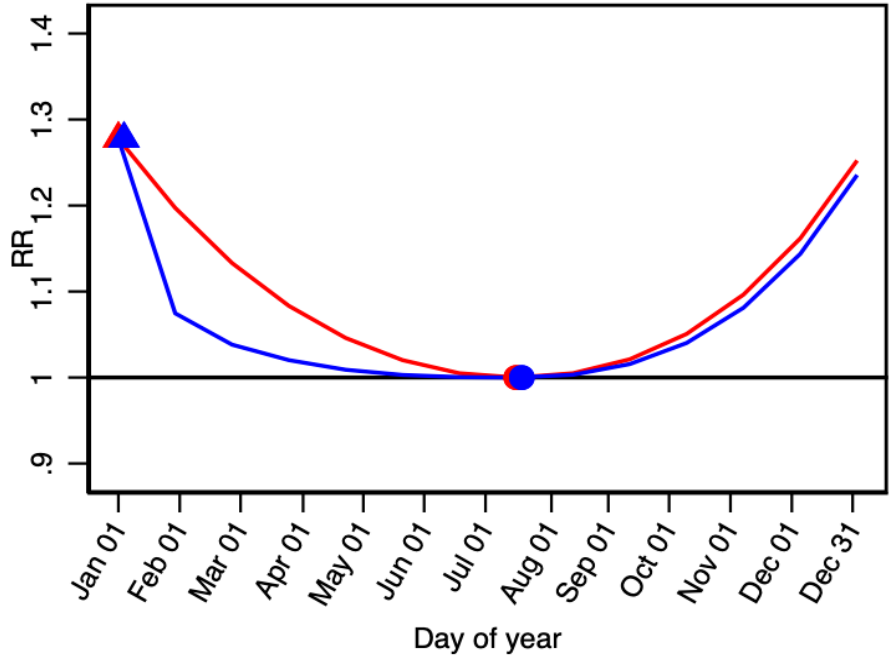
The day-of-year with maximum and minimum outcome estimates from the fitted seasonal curve as the peak and trough, respectively

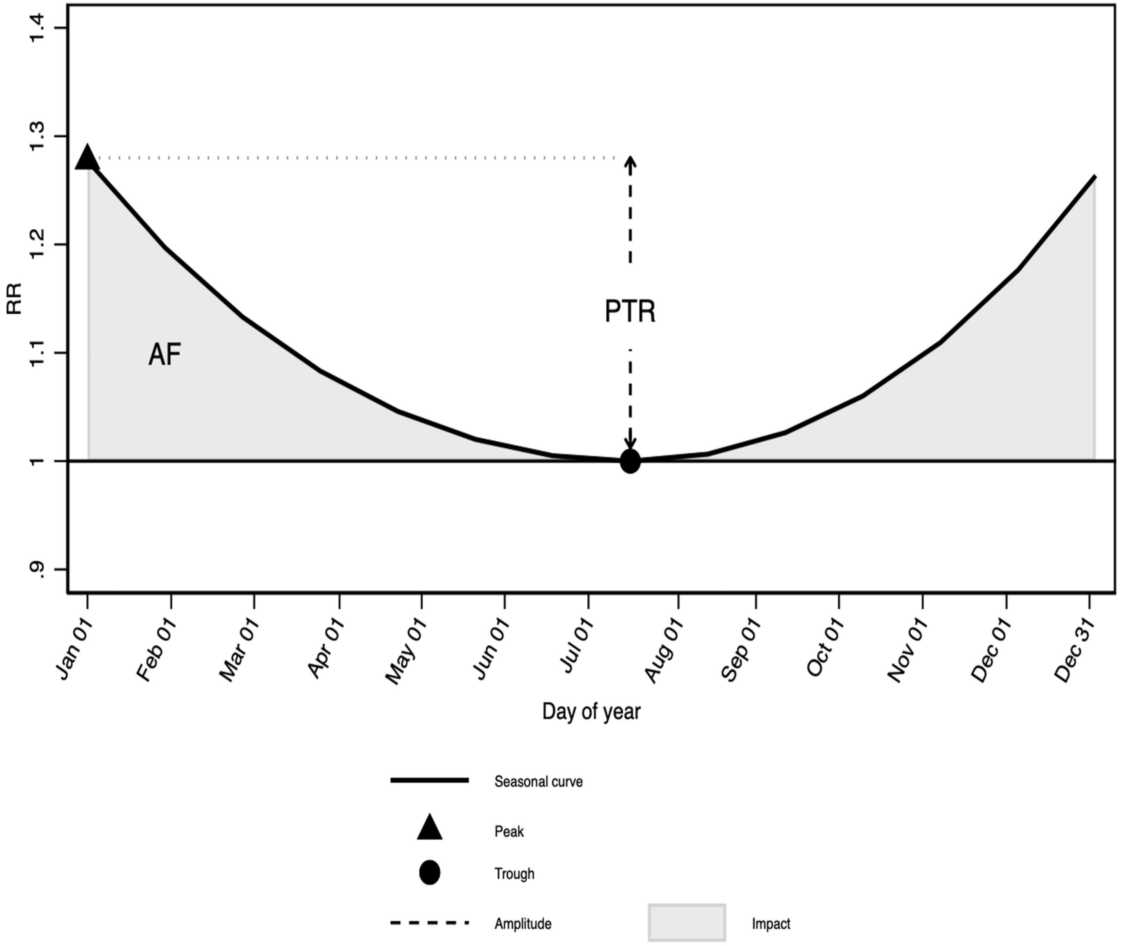
Size: amplitude and impact



Peak-to-trough ratio (PTR)

$$= \frac{Maximum\ (outcome)\ estimate\ at\ the\ peak}{Minimum\ (outcome)\ estimate\ at\ the\ trough}$$





AF estimates the fraction by which the outcome would be reduced in a counterfactual scenario where mortality risk never rose above its seasonal trough.

$$\boldsymbol{Attributable\ fraction} = \sum p_x[1 - \exp(-\beta_x)]$$

x represents the season_indicator,

β_x is the log relative risk of outcome on the season_indicator x compared with the minimum outcome estimated at the trough,

p_x is the percentage of cases on the season_indicator x.

Assess the role of its potential drivers

$$y_t = m_t + s_t + z_t + w_t$$

y_t can be described by a sum of series:
m and s as long-term trend and seasonality; environmental exposure z_t
residualsual series w_t

The difference in seasonality (shape, timings, and size) before and after the adjustment of z_t can be interpreted as the contribution of environmental exposure z to the seasonality.

Comparison of seasonality:

- Shape: Visual inspection of the shape before and after adjustment.
- Timings and size: Calculate their difference before and after adjustment, and interpret the changes in the context of epidemiology and clinical research.

$d = \hat{E}_{after} - \hat{E}_{before}$ and $SE(d) = \sqrt{\widehat{SE}^2_{after} + \widehat{SE}^2_{before}}$,

where $\hat{E}$ are the estimates of the size before and after the adjustment,
$\widehat{SE}$ are their respective standard errors.

∆PTR: (absolute change in PTR after adjustment) = $\log(PTR_{after}) - \log(PTR_{before})$;
∆AF (Change in AF after adjustment) $= AF_{after} - AF_{before}$

Choosing the most appropriate model

- Model choice may be based on model fit criteria, e.g., Akaike's information criterion.
- Regression diagnostics (e.g., residuals)
- Since the modelling process involves many decisions, multiple sensitivity analyses are recommended to check the robustness of the main conclusions

References

Tutorial paper:

- Lina Madaniyazi, Aurelio Tobias, Yoonhee Kim, Yeonseung Chung, Ben Armstrong, Masahiro Hashizume. Assessing seasonality and the role of its potential drivers in environmental epidemiology: a tutorial, International Journal of Epidemiology, Volume 51, Issue 5, October 2022, Pages 1677–1686, https://doi.org/10.1093/ije/dyac115

Other articles and books:

- Barnett AG, Dobson AJ. Analysing seasonal health data. Berlin, Germany:: Springer, 2010 Jan 1.
- Christiansen CF, Pedersen L, Sørensen HT, Rothman KJ. Methods to assess seasonal effects in epidemiological studies of infectious diseases—exemplified by application to the occurrence of meningococcal disease. Clin Microbiol Infect. 2012 Oct 1;18(10):963-9.
- Fisman DN. Seasonality of infectious diseases. Annu Rev Public Health. 2007 Apr 21;28:127-43.

R codes and extensions:
https://github.com/LinaMadaniyazi
https://researchmap.jp/lina

Hands-on exercise

- **Example data:**

Daily time series data on mean PM_{10}, mean temperature, and all-cause mortality from 2002 to 2006 in London

- Seasonal variation in all-cause mortality in London
- Contribution of temperature to seasonal variation in all-cause mortality in London

Real-data practical (Optional)

- **Dataset:**

Daily time series data on daily average WetBulb Global Temperature (WBGT) and heat-stroke (cases) in warm seasons (May- September) from 2015 to 2020 in Tokyo.

**WBGT: a measure of the heat stress in direct sunlight, which takes into account: temperature, humidity, wind speed, sun angle and cloud cover (solar radiation).*

- **Questions:**

- Is there any seasonal variation in heat-stroke?
- Assess its seasonal variation
- Assess the role of WBGT in its seasonal variation