REVIEW 10.1111/j.1469-0691.2012.03966.x

Methods to assess seasonal effects in epidemiological studies of infectious diseases—exemplified by application to the occurrence of meningococcal disease

C. F. Christiansen[1], L. Pedersen[1], H. T. Sørensen[1] and K. J. Rothman[2]

1) *Department of Clinical Epidemiology, Aarhus University Hospital, Aarhus, Denmark* and 2) *RTI Health Solutions, Research Triangle Park, NC, USA*

Abstract

Seasonal variation in occurrence is a common feature of many diseases, especially those of infectious origin. Studies of seasonal variation contribute to healthcare planning and to the understanding of the aetiology of infections. In this article, we provide an overview of statistical methods for the assessment and quantification of seasonality of infectious diseases, as exemplified by their application to meningococcal disease in Denmark in 1995–2011. Additionally, we discuss the conditions under which seasonality should be considered as a covariate in studies of infectious diseases. The methods considered range from the simplest comparison of disease occurrence between the extremes of summer and winter, through modelling of the intensity of seasonal patterns by use of a sine curve, to more advanced generalized linear models. All three classes of method have advantages and disadvantages. The choice among analytical approaches should ideally reflect the research question of interest. Simple methods are compelling, but may overlook important seasonal peaks that would have been identified if more advanced methods had been applied. For most studies, we suggest the use of methods that allow estimation of the magnitude and timing of seasonal peaks and valleys, ideally with a measure of the intensity of seasonality, such as the peak-to-low ratio. Seasonality may be a confounder in studies of infectious disease occurrence when it fulfils the three primary criteria for being a confounder, i.e. when both the disease occurrence and the exposure vary seasonally without seasonality being a step in the causal pathway. In these situations, confounding by seasonality should be controlled as for any confounder.

Keywords: communicable diseases, confounding factors, infectious disease medicine, meningococcal infections, seasons, statistics, time.

Clin Microbiol Infect

Corresponding author: C. F. Christiansen, Aarhus University Hospital, Department of Clinical Epidemiology, Olof Palmes Alle 43–45, 8200 Aarhus N, Denmark

E-mail: cc@dce.au.dk

Introduction

Seasonal variation encompasses cyclic change in either disease occurrence or disease severity over the course of a year [1,2]. Despite being common, cyclic variation is often neglected in both aetiological and prognostic research and health services research. Seasonal variation affects major diseases such as myocardial infarction, stroke, atrial fibrillation, fracture, and cancer [3–9]. Month of birth may also influence the occurrence of non-infectious diseases in childhood and adolescence, such as Crohn's disease and leukaemia [10–12]. Seasonal variation commonly affects many community-acquired infectious diseases.

Several mechanisms may contribute to the seasonal variation of infectious disease [1,13]. First, there are annual cycles in pathogen appearance or virulence, alternating between the northern and southern hemispheres for areas sufficiently remote from the equator. Many of these cyclic patterns are secondary to annual climatic cycles, which affect temperature, rainfall, and humidity. The amount of daylight may also influence the host physiology, affecting immune function and, consequently, disease occurrence. Another factor fostering the annual cyclic occurrence of disease is human behaviour. For example, there is greater crowding of people and seasonal vacation travel during cold and rainy periods, and more use of air-conditioning during warm periods, all of which are phenomena that can be considered to be secondary to climatic changes. Social activities, however, such as those related to specific holidays, may be tied to the calendar without being a consequence of climatic cycles [1,13].

Available methods for the study of seasonality range from simple comparisons across discrete calendar time periods, or simple models such as fitting monthly counts to a sine curve, to more complex and flexible statistical models. The need to control for confounding in studies of seasonal variation is limited by the fact that many common confounders, e.g. age, sex, and lifestyle factors, do not change during the seasons and will therefore not be confounders [14].

In this article, we provide: (i) a brief overview of methods to study the seasonality of infectious diseases; (ii) examples of application in the existing literature; (iii) an example of application of the three methods to the occurrence of meningococcal disease in Denmark; and (iv) a discussion about whether seasonality should be considered as a covariate in studies of infectious diseases.

Methods used to Study Seasonal Variation in Infectious Disease

Several methods have been used to examine seasonal variation in disease occurrence, but in this article we will focus on the three most widely used classes of method: comparison of discrete time periods, geometrical models, and generalized linear models (GLMs). The characteristics of these three classes are summarized in Table 1.

Seasonal variation or seasonality is defined as a periodic variation in the occurrence of disease or disease outcome with calendar time. Occurrence can be measured either as a count of cases per unit time, a rate that relates cases to a denominator of person-time, or an incidence proportion that relates cases to the number of persons at risk. With a single annual cycle, there will be a single peak in occurrence during the year, and ordinarily a single trough, or time of low occurrence, often assumed to be 6 months from the peak. The amplitude of the seasonal pattern is defined as the difference

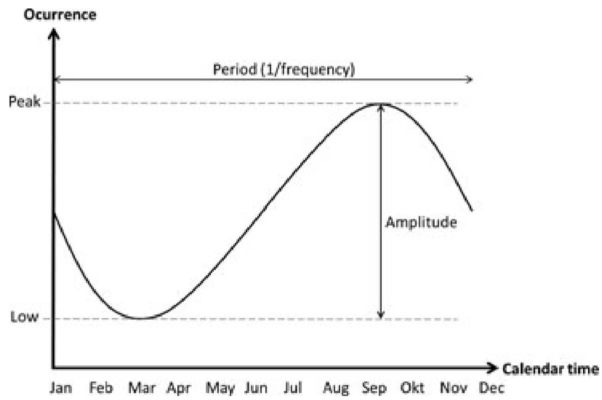

FIG. 1. The terminology of seasonal variation exemplified by a simple sine curve.

in occurrence between the peak and the trough times. The word ‘period’ is used to describe the length of one full cycle, and the frequency is the inverse of the period [13] (Fig. 1).

Direct comparison of discrete time periods

A simple approach to studying seasonal variation is to compare disease occurrence during specific time intervals during a cycle, such as months or quarters during a year. The comparison may involve choosing a reference time during a cycle and comparing the other intervals with the reference. Predefined periods can be compared pairwise by calculating simple risk or incidence rate ratios across time intervals within the cycle. It is common to test whether seasonal variation is present by using statistical significance tests, but such tests are as ill-advised in this situation as they are elsewhere. In brief, statistical significance depends on both the strength of the association and the amount of data, and thus does not measure the strength of seasonal occurrence. Instead, measures that compare estimates of rates, risks or counts of cases should be used. In the rare instances in which there is variation during the cycle in age, sex, or other possible confounders, these may be controlled analytically with traditional methods, such as stratification or regression models. Confounding factors that vary seasonally are unusual, and as a result it is common to see seasonal analyses that involve only crude comparisons, that is, using no adjustment for confounding.

TABLE 1. Characteristics of the three classes of method for the study of seasonal variation

	Comparison of discrete time periods	Geometrical models (e.g. Edwards’ method)	Generalized linear models
Computation	Very simple	Fairly simple	More complicated
Underlying assumption	Departure from equal numbers	Cyclic pattern following a sine curve	Fewer constraining assumptions
Frequency	Predefined	One	Flexible
Secular trend	Normally not addressed. Analyses can be stratified by calendar year	Normally not addressed. Analyses can be stratified by calendar year	Can be included in the model
Identification of time of peak	No	Possible	Possible
Adjustment for covariates	Usually not performed, but possible by stratification or logistic regression analysis	Usually not performed, but possible by stratification or regression analysis	Covariates can easily be included in the model
Examples of test statistics	Chi-squared test	Edwards’ test or recently proposed test statistics	Wald chi-squared test