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.

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
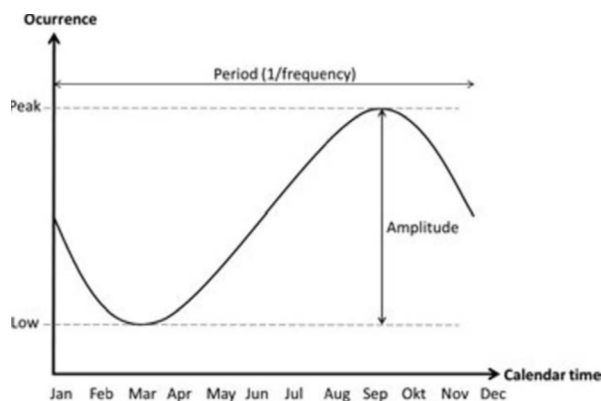

FIG. 1. The terminology of seasonal variation exemplified by a simple sine curve.

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

Using direct comparison of discrete time intervals, although straightforward, is limited by the need for predefined definitions of seasons and by the inability to compare more than two periods at once. This approach also seldom takes into account any secular trend that may be superimposed on the seasonal pattern.

The following are some examples. In a study of antibiotic-resistant *Streptococcus* pneumonia in two populations, Dagan *et al.* [15] found that more prescriptions were written in the cold months than in the warm months (291 vs. 222 prescriptions per 1000 children).

Al-Hasan *et al.* [16] studied the seasonal variation of *Escherichia coli* bloodstream infection (BSI) in 461 patients between 1998 and 2007 in a county in Minnesota. Their hypothesis was that a warm climate would increase the risk of BSI. They simply compared the four warmest months (June–September) with the remainder of the year, and found a 35% increased risk of BSI during the summer (incidence rate ratio 1.35, 95% CI 1.12–1.66) [16]. The time of the peak was not examined, which was also pointed out in the accompanying editorial [17].

In another example, Reddy *et al.* [18] compared the seasonal distribution of cases with microsporidial keratitis in a tertiary centre in India from 2006 to 2008. They identified 30 cases, 20 of which occurred during the monsoon (June–September), six during the winter (October–January), and four during the summer (February–May). Their interpretation was limited by a lack of information about the size of and seasonal changes in the referral source population, and by the fact that the authors compared seasons only by means of statistical tests of significance [18].

Logar *et al.* [19] compared the rate of positive test results for acute toxoplasmosis in pregnant women during the four seasons, and found a lower rate during the summer (0.27% of tested women; 95% CI 0.17–0.37) than during the winter (0.48%; 95% CI 0.34–0.62), but they, too, only compared periods by statistical tests of significance.

Geometrical model assuming a sinusoidal cyclic pattern

The second category of methods is based on harmonic (periodic) regression, an approach that fits a sine curve to a time series of frequencies by the use of ordinary regression methods. The most widely used approach is based on the harmonic technique of Edwards, which assumes that counts of disease are derived from a non-homogeneous Poisson distribution [20].

The outcome measure is usually the peak-to-low ratio, interpreted as a measure of relative risk (RR) that compares the month with the highest incidence (peak) with the month with the lowest incidence (low or trough).

$$
\begin{aligned}
\text{Relative risk (RR)} &= \text{Peak-to-low ratio} \\
&= \frac{1 + \text{hemi amplitude}}{1 - \text{hemi amplitude}} \\
&= \frac{1 + \text{Edwards estimator}}{1 - \text{Edwards estimator}}
\end{aligned}
$$

The method of Edwards uses simple formulas to fit the sine curve; this approach has been modified slightly by Brookhart and Rothman [20,21] to improve the statistical performance of the estimator and derive confidence limits by the use of straightforward formulas.

The main limitation of using a fitted sine curve is the inability to adjust for covariates, including a secular trend in occurrence. As with the direct comparison approach, however, traditional methods to control confounding, such as stratification, can be applied. As a variant of the method of Edwards, the geometrical approach can also be applied in a linear regression model including a sine and a cosine term [22]. Such a model would allow adjustments for covariates.

The geometrical model can easily be applied by using the free programmed spreadsheet, Episheet [23], which provides a graphical presentation of seasonal variation, and estimates the time of the peak, the peak-to-low ratio, and a confidence interval for the latter. The only data entry needed is the set of 12 frequencies measuring the number of cases occurring in each month of the year; optionally, if the denominator is known and varies, 12 denominator frequencies may also be entered [23].

The following are some examples. Akhtar and Mohammad studied the seasonality of 4608 cases of pulmonary tuberculosis among 2.3 million immigrants in Kuwait between 1997 and 2006 [24]. The highest frequency of occurrence was in late April, and the peak-to-low ratio was 1.51 (95% CI 1.39–1.65). Seasonal variation was similar in the first and the second half of the study period [24]. A study by Yamaguchi *et al.* [25] described the epidemiology of measles in a cohort of 674 measles cases occurring in a city in Malawi between 1996 and 1998. A graphical presentation of the number of cases per month showed annual peaks in April 1996, October 1997, and June 1998. The exact date of the peaks was estimated by the method of Edwards. This study underscores the importance of graphical presentation of data, and demonstrates the importance of examining the data in each calendar year when feasible [25].

GLMs, including Poisson regression

GLMs are a group of statistical models that provide a flexible approach in studies of seasonality, because they allow data to be fitted to various underlying mathematical functions [3]. A log-linear Poisson regression model is a commonly used

underlying function [21]. The model may include not only seasonality, but also covariates and secular trends, i.e.

$$\log(\text{disease occurence}) = \text{Seasonal variation} + \text{Secular trend} + \text{Covariates}$$

The terms in this model are flexible; for example, the seasonal variation term can be considered as several overlapping sinusoid functions with different frequencies [26]. It allows several annual peaks and adjustment for covariates [3]. The method also allows computation of less biased peak-to-low-ratios than geometrical models [26]. Despite the advantages of these models as compared with the geometrical models, their application and interpretation are more complex.

The following are some examples. Eber *et al.* examined seasonal variation in the frequencies of Gram-negative BSIs in a study that included >200 000 blood cultures from hospitalized patients in 132 US hospitals. Their Poisson model accounted for long-term trends, but the estimation focused on differences between the four seasons. The most pronounced seasonal variation was found for *Acinetobacter*, which was 52% more frequent during the summer months than during the winter months [27]. This comparison of mean frequencies during seasons will underestimate the peak-to-low ratio.

A study from The Netherlands examined the secular trend and seasonality of pertussis in 1996–2006 in age-stratified models including seasonal (monthly) variation and secular variation, corrected for autocorrelation [28]. They found seasonal variation with a peak incidence in August, except for children aged 13–18 years, whose incidence peaked in November. The peak-to-low ratio of the incidence ranged from 1.36 (95% CI 1.12–1.66) to 2.86 (95% CI 2.30–3.55) in the different age groups under study. There was a slight increase in the occurrence of pertussis during the study period [28].

A study of 2810 *E. coli* BSIs during an 8-year period in northern Israel compared the incidence rate between three predefined periods by using a GLM that also accounted for long-term trends in the study period [29]. The incidence rate was highest during the summer, with an incidence rate ratio of 1.19 (95% CI 1.12–1.26) as compared with the transitional season (March, April, and November). The times of peak and low were not reported, owing to the predefined time periods [29]. This study also included a time-series analysis. This kind of analysis takes into account the fact that adjacent observations may be correlated [30]. A subtype of time-series analysis, the autoregressive integrated moving average (ARIMA) models, includes a component allowing observed outcomes to depend on previous outcomes (the autoregres-

sive component) varying with lag-time (the moving average component), and also allows the examination of long-term trends (the integrative component) [30,31]. Such an ARIMA model was used in a study of the seasonal variation of sepsis [32].

An example of application: seasonal variation of meningococcal disease

In this section, we use meningococcal disease in Denmark to illustrate the application of the three classes of method.

Identification of hospitalizations with meningococcal disease. We extracted the number of patients hospitalized with a diagnosis of meningococcal disease from 1995 to 2011 by using the Danish National Registry of Patients (DNRP), which covers all Danish hospitals. The DNRP has recorded >99% of acute-care hospital admissions in Denmark since 1977 and admissions to outpatient clinic and emergency room visits since 1995 [33,34]. DNRP records include dates of admission and discharge, one primary diagnosis (main reason for hospitalization), and up to 19 secondary diagnoses, treatments, and procedures, including intensive-care observation/therapy. Diagnoses were coded according to the International Classification of Diseases, 8th revision (ICD-8) to 1993, and have been coded according to the 10th revision (ICD-10) since 1994. We used the ICD-10 code A39x to identify meningococcal disease.

We identified 2407 patients hospitalized with meningococcal disease during this 17-year period. Monthly counts are plotted in Fig. 2. We weighted the number of monthly counts to the length of the month by multiplying the count by 30 divided by the length of the month.
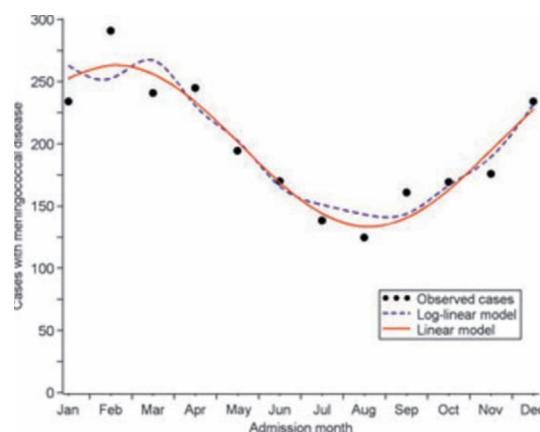

FIG. 2. Example of monthly counts of meningococcal disease weighted by the length of the month (black dots) and fitted to a sine curve (red line/linear model) and to a log-linear Poisson model (blue dashed line/log-linear model).

Direct comparison of discrete time periods. First, we computed rate ratio as the sum of time-weighted counts during the winter (December–February) divided by that for the summer (June–August), and computed 95% CIs [35]. Including the entire period, the overall rate ratio was 1.75 (95% CI 1.56–1.97). There were no major disparities between the included years, although the estimates were imprecise for individual years (Table 2).

Application of a geometrical model. Second, we used the geometrical approach, fitting the weighted monthly counts to a sine curve, using Episheet [23]. This assumes one yearly peak and one low with 6 months in between [20]. The red line in Fig. 2 shows such a fitted sine curve. The peak-to-low ratio was 1.96 (95% CI 1.76–2.18) and the peak was on 18 February. Seasonality was evident for all years, although it was less pronounced in a few years, e.g. 2002 and 2010 (Table 2).

Application of a GLM. Third, we used log-linear Poisson regression (a GLM) [26]. The fitted log-linear function is illustrated by a blue dashed line in Fig. 2. This method revealed a peak-to-low ratio of 1.94 (95% CI 1.72–2.17) with a peak on 17 February. Analyses stratified by calendar year showed a pattern very similar to that obtained with the geometrical model (Table 2).

Summary of the applied methods. In this example, all three classes of method found seasonality with similar estimated amplitudes. Although the simple comparison of winter and summer led to the same overall conclusion, the method overlooked seasonal variation in 2006 and 2007. The estimated peak-to-low ratios, including CIs, were similar for both the geometrical model and the GLM when the entire study period was summed up. However, the geometrical model overestimated the high peak-to-low ratio in 2001 (4.80 vs. 4.10) and underestimated the low peak-to-low ratio in 2010 (1.18 vs. 1.33) as compared with the log-linear Poisson model, which should provide less biased estimates [26] (Table 2).

Controlling for Seasonality in Studies of Infectious Disease

When seasonal occurrence confounds another factor that is being studied, seasonality should be controlled as for any confounder. Like any confounder, seasonality will, broadly speaking, be confounding if it is associated with both the exposure and the outcome without being in the causal pathway [36,37]. Thus, confounding must be evaluated for each separate hypothesis and association studied. For example, in a study of recent influenza infection as a risk factor for meningococcal disease, both influenza and meningococcal disease occurred with seasonal variation, and it would therefore be relevant to control for seasonal variation when studying the effect of influenza infection on meningococcal risk [38]. In contrast, it would not be necessary to control for seasonal variation in a study of diabetes as a risk factor meningococcal disease, because the prevalence of diabetes is not expected to change seasonally and would therefore not be a confounder of a seasonal risk factor.

If a seasonally varying factor is a confounder, it should be controlled for as for any confounder, e.g. by stratification, matching, restriction, or adjustment in regression analyses.

TABLE 2. Output from the three classes of method applied to monthly counts of meningococcal disease in Denmark

Year	Comparison of discrete time periods	Geometrical model (Episheet)		Generalized linear model (log-linear Poisson model)	
	Rate ratio, winter/summer (95% CI)	Peak-to-low ratio (95% CI)	Day of peak	Peak-to low ratio (95% CI)	Day of peak
1995	1.78 (1.23–2.56)	1.82 (1.31–2.53)	23 January	1.88 (1.31–2.71)	22 January
1996	1.36 (0.93–1.98)	1.46 (1.06–2.00)	20 March	1.55 (1.07–2.23)	17 March
1997	1.88 (1.30–2.71)	2.04 (1.46–2.86)	19 February	2.06 (1.44–2.95)	18 February
1998	2.31 (1.42–3.75)	2.11 (1.40–3.18)	6 February	2.18 (1.41–3.37)	5 February
1999	1.70 (1.09–2.64)	2.26 (1.53–3.35)	6 March	2.32 (1.54–3.50)	5 March
2000	1.69 (1.10–2.61)	1.81 (1.21–2.71)	19 February	1.91 (1.22–2.99)	18 February
2001	3.01 (1.87–4.84)	4.80 (2.60–8.85)	2 March	4.10 (2.60–6.46)	1 March
2002	1.98 (1.05–3.71)	1.36 (1.00–2.21)	23 January	1.53 (0.87–2.70)	22 January
2003	1.85 (1.10–3.13)	2.67 (1.53–4.66)	7 February	2.72 (1.57–4.71)	6 February
2004	1.67 (0.95–2.92)	1.59 (1.00–2.57)	1 January	1.75 (1.02–3.02)	1 January
2005	1.47 (0.85–2.54)	1.71 (1.05–2.79)	7 March	1.88 (1.09–3.26)	6 March
2006	1.13 (0.62–2.08)	1.63 (1.00–2.75)	29 March	1.81 (1.00–3.28)	29 March
2007	1.19 (0.60–2.38)	1.93 (1.08–3.47)	25 March	2.11 (1.11–4.00)	26 March
2008	2.26 (1.17–4.36)	2.09 (1.18–3.69)	30 December	2.26 (1.23–4.16)	29 December
2009	1.44 (0.88–2.37)	1.62 (1.00–2.69)	12 February	1.76 (0.99–3.14)	10 February
2010	1.38 (0.77–2.46)	1.18 (1.00–1.91)	12 February	1.33 (0.76–2.36)	9 February
2011	2.09 (1.19–3.69)	2.18 (1.29–3.69)	19 February	2.31 (1.33–4.03)	18 February
All years	1.75 (1.56–1.97)	1.96 (1.76–2.18)	18 February	1.94 (1.72–2.17)	17 February

A simple way to address potential confounding would be to repeat the analyses with stratification by months or seasons.

Conclusion

Seasonal variation commonly affects infectious disease occurrence. Studies with the aim of studying seasonal variation should use appropriate methods to identify and report the seasonal variation. The simplest class of methods compare discrete time periods pairwise, but are limited by the need to predefine seasons. More advanced methods are needed in order to quantify the seasonal variation, e.g. by a peak-to-low ratio. The geometrical models, e.g. fitting monthly frequencies of disease to a sine curve, allow estimation of the peak-to-low ratio and identification of the timing of the peak. GLMs, such as the log-linear Poisson model, are more complicated to apply, but allow the inclusion of covariates in the model, and provide more precise estimates of the peak-to-low ratio and its CI than the geometrical model. In our example of meningococcal disease, the three methods reached almost the same conclusion about seasonality, which may be the case in large studies with moderate seasonal variation. It is seldom necessary to adjust for confounding in studies of seasonal variation, because the prevalence of common confounders rarely changes during the seasons.

Authorship/contribution

H. T. Sørensen and K. J. Rothman contributed to the design of the article. C. F. Christiansen reviewed the literature and wrote the first draft. L. Pedersen conducted the statistical analyses. All authors interpreted data and reviewed the article critically. All authors read and approved the final manuscript.

Transparency Declaration

References

1. Dowell SF. Seasonal variation in host susceptibility and cycles of certain infectious diseases. *Emerg Infect Dis* 2001; 7: 369–374.
2. Grassly NC, Fraser C. Seasonal infectious disease epidemiology. *Proc Biol Sci* 2006; 273: 2541–2550.
3. Lundbye-Christensen S, Dethlefsen C, Gorst-Rasmussen A *et al.* Examining secular trends and seasonality in count data using dynamic generalized linear modelling: a new methodological approach illustrated with hospital discharge data on myocardial infarction. *Eur J Epidemiol* 2009; 24: 225–230.
4. Seretakis D, Lagiou P, Lipworth L, Signorello LB, Rothman KJ, Trichopoulos D. Changing seasonality of mortality from coronary heart disease. *JAMA* 1997; 278: 1012–1014.
5. Fischer T, Johnsen SP, Pedersen L, Gaist D, Sorensen HT, Rothman KJ. Seasonal variation in hospitalization and case fatality of subarachnoid hemorrhage—a nationwide Danish study on 9,367 patients. *Neuroepidemiology* 2005; 24: 32–37.
6. Bischoff-Ferrari HA, Orav JE, Barrett JA, Baron JA. Effect of seasonality and weather on fracture risk in individuals 65 years and older. *Osteoporos Int* 2007; 18: 1225–1233.
7. Lambe M, Blomqvist P, Bellocco R. Seasonal variation in the diagnosis of cancer: a study based on national cancer registration in Sweden. *Br J Cancer* 2003; 88: 1358–1360.
8. Oberg AL, Ferguson JA, McIntyre LM, Horner RD. Incidence of stroke and season of the year: evidence of an association. *Am J Epidemiol* 2000; 152: 558–564.
9. Frost L, Johnsen SP, Pedersen L *et al.* Seasonal variation in hospital discharge diagnosis of atrial fibrillation: a population-based study. *Epidemiology* 2002; 13: 211–215.
10. Sorensen HT, Pedersen L, Norgard B, Fonager K, Rothman KJ. Does month of birth affect risk of Crohn's disease in childhood and adolescence? *BMJ* 2001; 323: 907.
11. Sorensen HT, Pedersen L, Olsen J, Rothman K. Seasonal variation in month of birth and diagnosis of early childhood acute lymphoblastic leukemia. *JAMA* 2001; 285: 168–169.
12. Van RM, Joossens M, Joossens S *et al.* Crohn's disease and month of birth. *Inflamm Bowel Dis* 2005; 11: 597–599.
13. Naumova EN. Mystery of seasonality: getting the rhythm of nature. *J Public Health Policy* 2006; 27: 2–12.
14. Rothman KJ, Greenland S, Lash TL. *Modern epidemiology*, 3rd edn. Philadelphia, PA: Lippincott William & Wilkins, 2008; 598–619.
15. Dagan R, Barkai G, Givon-Lavi N *et al.* Seasonality of antibiotic-resistant *Streptococcus pneumoniae* that causes acute otitis media: a clue for an antibiotic-restriction policy? *J Infect Dis* 2008; 197: 1094–1102.
16. Al-Hasan MN, Lahr BD, Eckel-Passow JE, Baddour LM. Seasonal variation in *Escherichia coli* bloodstream infection: a population-based study. *Clin Microbiol Infect* 2009; 15: 947–950.
17. Freeman JT, Anderson DJ, Sexton DJ. Seasonal peaks in *Escherichia coli* infections: possible explanations and implications. *Clin Microbiol Infect* 2009; 15: 951–953.
18. Reddy AK, Balne PK, Garg P, Krishnaiah S. Is microsporidial keratitis a seasonal infection in India? *Clin Microbiol Infect* 2011; 17: 1114–1116.
19. Logar J, Soba B, Premru-Srsen T, Novak-Antolic Z. Seasonal variations in acute toxoplasmosis in pregnant women in Slovenia. *Clin Microbiol Infect* 2005; 11: 852–855.
20. Edwards JH. The recognition and estimation of cyclic trends. *Ann Hum Genet* 1961; 25: 83–87.
21. Brookhart MA, Rothman KJ. Simple estimators of the intensity of seasonal occurrence. *BMC Med Res Methodol* 2008; 8: 67.
22. Stolwijk AM, Straatman H, Zielhuis GA. Studying seasonality by using sine and cosine functions in regression analysis. *J Epidemiol Community Health* 1999; 53: 235–238.
23. Rothman KJ. Episheet. Available at: http://krothman.hostbyet2.com/Episheet.xls (last accessed 30 April 2012).
24. Akhtar S, Mohammad HG. Seasonality in pulmonary tuberculosis among migrant workers entering Kuwait. *BMC Infect Dis* 2008; 8: 3.

25. Yamaguchi S, Dunga A, Broadhead RL, Brabin BJ. Epidemiology of measles in Blantyre, Malawi: analyses of passive surveillance data from 1996 to 1998. *Epidemiol Infect* 2002; 129: 361–369.

26. Christensen AL, Lundbye-Christensen S, Dethlefsen C. Poisson regression models outperform the geometrical model in estimating the peak-to-trough ratio of seasonal variation: a simulation study. *Comput Methods Programs Biomed* 2011; 104: 333–340.

27. Eber MR, Shardell M, Schweizer ML, Laxminarayan R, Perencevich EN. Seasonal and temperature-associated increases in gram-negative bacterial bloodstream infections among hospitalized patients. *PLoS One* 2011; 6: e25298.

28. De Greeff SC, Dekkers AL, Teunis P, Rahamat-Langendoen JC, Mooi FR, De Melker HE. Seasonal patterns in time series of pertussis. *Epidemiol Infect* 2009; 137: 1388–1395.

29. Chazan B, Colodner R, Edelstein H, Raz R. Seasonal variation in *Escherichia coli* bloodstream infections in northern Israel. *Clin Microbiol Infect* 2011; 17: 851–854.

30. Zeger SL, Irizarry R, Peng RD. On time series analysis of public health and biomedical data. *Annu Rev Public Health* 2006; 27: 57–79.

31. Morgenstern H. Ecologic studies. In: Rothman KJ, Greenland S, Lash TL, eds. *Modern epidemiology*. Philadelphia, PA: Lippincott William & Wilkins, 2008; 511–531.

32. Danai PA, Sinha S, Moss M, Haber MJ, Martin GS. Seasonal variation in the epidemiology of sepsis. *Crit Care Med* 2007; 35: 410–415.

33. Andersen TF, Madsen M, Jorgensen J, Mellemkjoer L, Olsen JH. The Danish National Hospital Register. A valuable source of data for modern health sciences. *Dan Med Bull* 1999; 46: 263–268.

34. Lynge E, Sandegaard JL, Rebolj M. The Danish National Patient Register. *Scand J Public Health* 2011; 39: 30–33.

35. Kirkwood BR, Sterne JAC. Comparing rates. In: Kirkwood BR, Sterne JAC, eds. *Essential medical statistics*. Malden, MA: Blackwell Science, 2003; 240–248.

36. Rothman KJ. *Epidemiology—an introduction*, 1st edn. New York: Oxford University Press, 2002.

37. Greenland S, Pearl J, Robins JM. Causal diagrams for epidemiologic research. *Epidemiology* 1999; 10: 37–48.

38. Jensen ES, Lundbye-Christensen S, Samuelsson S, Sorensen HT, Schonheyder HC. A 20-year ecological study of the temporal association between influenza and meningococcal disease. *Eur J Epidemiol* 2004; 19: 181–187.