REVIEW 10.1111/j.1469-0691.2012.03968.x

Seasonality of viral infections: mechanisms and unknowns

D. Fisman

The Division of Epidemiology, Dalla Lana School of Public Health; The Institute for Health Policy, Management, and Evaluation; The Department of Medicine, Faculty of Medicine, University of Toronto, Toronto, ON, Canada

Abstract

Seasonality is a long-recognized attribute of many viral infections of humans, but the mechanisms underlying seasonality, particularly for person-to-person communicable diseases, remain poorly understood. Better understanding of drivers of seasonality could provide insights into the relationship between the physical environment and infection risk, which is particularly important in the context of global ecological change in general, and climate change in particular. In broad terms, seasonality represents oscillation in pathogens' effective reproductive number, which, in turn, must reflect oscillatory changes in infectiousness, contact patterns, pathogen survival, or host susceptibility. Epidemiological challenges to correct identification of seasonal drivers of risk include failure to adjust for predictable correlation between disease incidence and seasonal exposures, and unmeasured confounding. The existing evidence suggests that the seasonality of some enteric and respiratory viral pathogens may be driven by enhanced wintertime survival of pathogens, and also by increased host susceptibility resulting from relative 'wintertime immune suppression'. For vector-borne diseases and zoonoses, environmental influences on vector or reservoir abundance, and vector biting rates, are probably more important. However, numerous areas of uncertainty exist, making this an exciting area for future research.

Keywords: Climate change, ecosystem, epidemiology, seasonality, virus

Clin Microbiol Infect

Corresponding author: D. Fisman, 155 College Street, Room 547, Toronto, ON, Canada M5T 3M7
E-mail: david.fisman@utoronto.ca

Introduction

The seasonality of infectious diseases is a phenomenon so widespread, and so familiar, that it has worked its way into the English language vernacular. We speak of 'flu season' in winter, usually without stopping to wonder why it might be that influenza viruses, which circulate year-round, appear to have a greater reproductive number when it is cold outside [1–3]. More recently, concerns around climate change, environmental degradation and an apparent surge in infectious disease emergence have led some investigators to rigorously research environmental, behavioural and immunological factors that could be responsible for the seasonality of viral infectious diseases [4–7].

Before discussing possible mechanisms underlying seasonality in greater depth, it is important to introduce a working definition of the concept of seasonality. Whereas a dictionary definition might simply suggest that a disease is seasonal if peak incidence is correlated with a particular period of the calendar year, an alternative definition is 'the state of recurring at regular intervals'[8], and, indeed, the seasonality of infectious diseases usually represents a periodic process: one that has sinusoidal peaks and troughs at regular intervals. An infectious disease may have several different periods: for example, in the pre-vaccination era, measles outbreaks typically had autumn seasonality (usually ascribed to enhanced transmission among schoolchildren with the start of the school year), but 3-year periodicity (typically attributed to the time period necessary to re-accumulate a critical fraction of susceptible children following a large epidemic) [9,10]. For the purposes of this review, we regard seasonal processes as those with an incidence associated with a particular calendar period, and which have periodicity, although this is not limited to annual periodicity.

Seasonal patterns of disease risk have been sufficiently obvious to observers to be accurately described in texts

from the pre-microbial era (see, for example, the Hippocratic 'Epidemics', or William Cullen's description of the wintertime 'bastard peripneumony'[11,12]). Nonetheless, rigorous epidemiological evaluation of environmental, immunological and behavioural drivers of infectious disease risk poses several methodological challenges [2]. I begin by describing such challenges, as well as methodologies for their remediation, in the epidemiological study of seasonal viral infectious diseases (or any seasonal phenomenon, for that matter). I then focus on what is known, and unknown, about environmental drivers, and drivers of seasonality of occurrence of three broad classes of viral infectious disease: viral gastrointestinal infections; viral respiratory infections; and vector-borne and zoonotic viral diseases. I close by noting potential avenues for future research.

Challenges to Study, and Useful Epidemiological Tools

A number of epidemiological factors complicate the study of seasonal drivers of infectious diseases. It is my opinion that principal among these is the issue of co-seasonality of numerous environmental, social and behavioural phenomena that could influence the reproductive number of an infectious disease (the number of incident cases of infection created by a prevalent case). This creates significant potential for numerous, predictable but fallacious associations between exposures and disease outcomes: for example, if two diseases (e.g. influenza and invasive pneumococcal disease) have shared seasonality [13], casual correlations will be observed whether or not one increases vulnerability to the other [2]. There is also the potential for epidemiological confounding by failure to adjust for seasonal behavioural phenomena that are correlated with environmental exposures (as, for example, when attempting to identify the role that summertime school closure may play in dampening influenza epidemic 'waves' [14]).

The relative rarity of some infectious disease outcomes of interest means that disease outcomes (and exposures) need to be aggregated at weekly, monthly or yearly levels for study. This creates the potential for a temporal version of 'ecological fallacy', with changes in disease risk being attributed to average exposure levels during a given time period, when in fact they are correlated with a subset of exposures that are effaced because of aggregation [15,16]. As a more concrete example, if the cause of a surge in disease was a rainy day that occurred during an unusually dry month, we might, at the monthly level, attribute that surge to drought rather than rain.

Finally, Rothman [17] noted that there may be between-population variability in the roles that different component causes play in creating a 'sufficient cause' for disease occurrence, and that such geographical variability might result in geographical variation in the magnitude of (true) relative risks associated with a given disease outcome. In this respect, it must be noted that geographical variation in the strength of risk factors for disease occurrence, and drivers of seasonality, may be real. The identification of sunlight, temperature and humidity as distinct drivers of influenza seasonality in various jurisdictions [14,18,19] does not necessarily mean that any of these findings are incorrect, athough, of course, they may be, for the reasons described above.

This author believes that the major threat to the internal validity of epidemiological evaluations relates to failure to adjust for underlying and predictable seasonality of candidate exposures and disease outcomes. Epidemiological studies that utilize regression models need to adjust for predictable underlying seasonality by using 'smoothers', such as fast Fourier transforms (sinusoidal time functions) [20,21], or spline smoothers [22]. The fact that many infectious diseases are also communicable from person to person needs to be addressed as well; often, the use of an autocorrelation structure in models can adjust for such effects [23]. When diseases of interest are rare, factors contributing to the seasonality of case occurrence can be studied with 'case-only' methods, such as case-crossover designs [20,24,25]; however, it has been noted previously that control periods for case-crossover designs of seasonal phenomena need to be selected with random directionality, to avoid biased estimation [26].

Drivers of Seasonality

Numerous environmental, behavioural and immunological mechanisms have been proposed for the distinct seasonality of many viral pathogens [1,27]. These mechanisms are outlined more fully in Table 1. Simply considered, seasonal oscillation in incidence for diseases that are communicable (directly or indirectly via the environment) requires oscillation of the 'reproductive number' of the disease: the average number of secondary cases of disease produced by a primary infective case [28]. The effective reproductive number is a function of the 'basic' reproductive number (often denoted as R_0, the reproductive number extant when all in the population are susceptible to infection) multiplied by the fraction of the population susceptible to infection [29]. Consequently, it can be seen that seasonality depends on either periodic fluctuation in R_0, or periodic fluctuation in population-level susceptibility to infection. The latter may result from either host effects (e.g. seasonal surges in births [30] or seasonal 'immune deficiency' caused by vitamin D deficiency [27]), or