The Art of Modeling the Mortality Impact of Winter-Seasonal Pathogens

Lone Simonsen1,2 and Cécile Viboud2

1Department of Global Health, George Washington University, Washington, DC; and 2Fogarty International Center, National Institutes of Health, Bethesda, Maryland

(See the article by van Asten et al, on pages 628–39.)

Estimation of the mortality burden of winter-seasonal pathogens is notoriously difficult. In this issue of the journal, van Asten et al estimate the mortality burden of 9 viral and bacterial pathogens in Dutch persons aged ≥65 years using statistical time series models [1]. They report that influenza A and respiratory syncytial virus (RSV) each cause approximately 1.5% of total deaths in this age group, whereas parainfluenza, influenza B, and norovirus account for respectively 0.9%, 0.6%, and 0.2% of deaths. In contrast, enterovirus, rotavirus, Campylobacter, and Salmonella were not significantly associated with mortality. The authors use multivariate regression models to link mortality statistics with pathogen-specific laboratory surveillance data.

The mortality burden of acute infections has traditionally been indirectly assessed using statistical time series models because specific pathogens triggering death are rarely laboratory confirmed and identified as such. For example, only a small subset of influenza-related deaths receive an influenza code on the death certificate; rather, influenza is a “hit-and-run” pathogen that leads to secondary bacterial coinfection or exacerbation of underlying chronic diseases. Such influenza-related deaths, especially among older patients, will be captured in broader respiratory or chronic diseases death outcomes categories. Therefore, the total influenza-related mortality burden-defined as the number of deaths that are triggered by influenza and are, therefore, in principle preventable with influenza vaccine—is estimated based on winter-seasonal elevation in mortality temporally coinciding with influenza virus activity.

Recent statistical modeling approaches have estimated the relative contribution of cocirculating pathogens, and have been applied to various disease outcomes including mortality, hospitalization, and physicians visits [1–5]. In such models, weekly mortality data are regressed against weekly proxies of pathogen activity, secular time trends, and terms adjusting for seasonality and noninfectious factors. There is debate about which proxy for viral activity is most appropriate, but it generally includes some component of laboratory-surveillance [2, 3], syndromic surveillance [6], or their combination [7]. A multipathogen modeling study thus requires the availability of robust weekly national laboratory surveillance data for multiple years-data which have been available since 1997 in the Netherlands.

Modeling the burden of influenza and cocirculating pathogens is an art. The choices of mortality outcomes endpoints, adjustment covariates included in the model, and model assumptions leave room for interpretation. Some prefer using cause-specific mortality, such as deaths from respiratory, gastroenteritic, or cardiac causes, because these causes of deaths are related to—yet may not capture all—deaths from common acute infections. In contrast, using all-cause mortality data ensures that all deaths are captured; however, specificity (and therefore precision of the estimate) is sacrificed: Influenza, for example, contributes only marginally to total mortality [1, 8]. Whenever possible, it is therefore advisable to model a variety of mortality outcomes endpoints.

Another tricky modeling issue is how to capture background seasonal patterns in mortality data. In theory, the cyclical seasonal baseline represents the effect of unknown factors causing seasonal elevation in mortality, which are unrelated to the pathogens of interest. In practice, seasonal terms can compete with the impact of the pathogens included in the model. For example, attempts to include an RSV term in monthly disease burden models were unsuccessful due to high correlation between RSV activity and background seasonal terms in 1 US study [9]; then a more detailed model relying on weekly data solved this issue [10]. A related issue is whether (and how) if and how to adjust for extreme temperatures, especially cold spells, which appear to cause substantial mortality in the United Kingdom [11, 12].
but not in the Netherlands [1]. Theoretically, it is expected that with sufficient pathogen and temperature terms included in the model, background cyclical terms representing unknown seasonal mortality triggers will no longer be needed.

Van Asten et al’s interesting modeling strategy not only introduces 9 pathogen terms and temperature in the model but also uses innovative approaches to adjust for time trends in surveillance intensity and variation in virulence of influenza virus [1]. Each influenza season is modeled separately and the impact of lagging laboratory surveillance with respect to mortality data is explored. The inclusion of lags reflects both the time interval from infection to death but also potential asynchrony in epidemic dynamics between the population captured by laboratory surveillance (eg, children for RSV [2]), and the population studied for disease burden (eg, elderly persons [1]). In the Dutch study, the best mortality model had no infection-to-death lag for influenza A and a 0–3 week lag for RSV increasing with older age, which produced similar burden estimates for these viruses. The proposed RSV burden “delay” among elderly persons has previously been championed by UK studies in which the attributable mortality fraction was higher for RSV than influenza [13] and is consistent with a 3–5 week lag between RSV laboratory reports from children and elderly persons [2]. In contrast with the United Kingdom and the Netherlands, lag-free models have attributed 3–6-fold more elderly person deaths to influenza than RSV in an average season [2, 3]. Taken together, these conflicting findings on the absolute and relative mortality burden of RSV and influenza cannot easily be reconciled.

In addition to influenza and RSV, van Asten et al report a significant mortality impact for norovirus among older elderly persons and no significant effect for enterovirus, rotavirus, Campylobacter, and Salmonella [1]. In line with these findings, norovirus was responsible for 16% of deaths from intestinal diseases in people aged ≥65 years in a UK modeling study, whereas 10 other bacterial and viral enteric pathogens included in the model were not associated with mortality [4]. But the most surprising finding of the van Asten study is perhaps the sizable mortality estimate attributable to parainfluenza in older seniors. Further research should focus on evaluating the burden of parainfluenza in other countries while accounting for the intriguing differences in seasonality and periodicity of the main parainfluenza serotypes [14].

Although the work of van Asten et al is innovative, some issues are worth noting [1]. In the Netherlands, the RSV season systematically preceded the influenza season by several weeks. Thus, introducing a 3-week lag for RSV may accidentally attribute some of the influenza-related deaths to RSV. Another weakness is the insistence on using all-cause mortality and the omission of confidence intervals on the attribution estimates—which are likely broad because the studied pathogens accounted for <5% of all elderly person deaths. Hence, it will be particularly important to confirm the estimates using cause-specific mortality data, as done previously by the same team [6, 15]. Furthermore, it would be informative to include the contribution of bacterial pathogens that have a significant impact on mortality and potentially interact with influenza and RSV, such as pneumococcus [2, 16, 17]. Taken together, these sensitivity analyses would add to the robustness of the relative and absolute mortality burden estimates associated with common respiratory and enteric infections in Dutch elderly persons.

Going forward, setting the debate regarding the magnitude of RSV impact in elderly persons relative to influenza is an important agenda item for future research. Although laboratory-based prospective studies of elderly patients suggest a substantial RSV burden relative to influenza [18], time series modeling studies have reached various conclusions. In our opinion, analysis of unusual winter seasons is a unique opportunity to resolve this question from a modeling perspective: for instance, in the 2009–2010 respiratory season, pandemic influenza A/H1N1 activity occurred unusually early, predating RSV activity by several weeks in most Northern Hemisphere temperate countries. A recently published companion study on the burden of the 2009 influenza pandemic in the Netherlands does not discuss the consequences of including this ideal season for RSV estimation unfortunately [6]. It would be easy to check whether lagging RSV activity 3 weeks into an influenza-free time period coincides with a mortality peak among elderly persons in 2009–Winter 2010.

Another important area for improvement of disease burden models relates to the availability of laboratory surveillance data. In the modern era of rapid multiplex polymerase chain reaction testing, a definite (and cost-saving) improvement would come from integrated surveillance for multiple respiratory and enteric pathogens in patients seen with disease in ambulatory outpatient and hospital settings. This is routinely done in Denmark for respiratory infections [19] and, amazingly, such a system has recently been put in place in Lao PDR, a resource-poor setting [20]. Sadly, in most other countries, surveillance systems for different pathogens remain separate.

Lastly, we strongly believe that timeliness of national mortality vital statistics release needs to be improved and should be made a priority for public health monitoring. Critically, disease burden models require mortality and hospital discharge data with sufficient temporal, cause, and age detail. Why must such critical data lag by several years? There are great disparities between countries in this respect. For Mexico, detailed and timely online mortality data allowed rapid assessment of the 2009 pandemic impact [21], whereas US mortality statistics lag by 3–4 years, and pandemic burden estimates are
still preliminary in this country [22]. In the age of electronic health records and death certificates, there is really no reason why a representative sample cannot be made available rapidly to allow timely public health research. Such improvement, which is critically needed to prioritize research for new vaccines, monitor the effectiveness of existing vaccine and other intervention programs, and evaluate the severity of pandemics and other emerging threats.

**Notes**

**Acknowledgements.** This work was supported by the in-house influenza research program of the Division of International Epidemiology and Population Studies, Fogarty International Center, National Institutes of Health, which is funded by the International Influenza Unit, Office of Global Affairs, Department of Health and Human Services. LS acknowledges support from the RAPIDD program of the Science and Technology Directorate, Department of Homeland Security. Potential conflicts of interest. All authors: No reported conflicts. All authors have submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest. Conflicts that the editors consider relevant to the content of the manuscript have been disclosed.

**References**