Late season interim estimates of influenza vaccine effectiveness reliably predict end of season estimates in Victoria, Australia, 2007 to 2012

S G Sullivan (Sheena.Sullivan@influenzacentre.org), H Kelly²,³
¹. WHO Collaborating Centre for Reference and Research on Influenza, Melbourne, Australia
². Victorian Infectious Diseases Reference Laboratory, Melbourne, Australia
³. Australian National University, Canberra, Australia

Citation style for this article:

Article submitted on 24 September 2013 / published on 10 October 2013

Twice each year the World Health Organization makes a recommendation for the composition of the influenza vaccine, based on circulating strains of influenza A(H3N2), A(H1N1) and B. Strain selection has always been based on immunogenicity studies with limited human data. Immunogenicity can be considered as a proxy for vaccine effectiveness (VE). However, only interim VE estimates for the target hemisphere can be considered in time for the strain selection meeting. Using surveillance data from Victoria, Australia, we retrospectively estimated and compared interim and final VE estimates for 2007 to 2012. In general, interim estimates were within five percentage points of final estimates. However, estimates made too early or in years of low influenza activity may be unreliable.

Introduction
Twice every year, the World Health Organization (WHO) hosts an influenza vaccine strain selection meeting where data gathered by members of the Global Influenza Surveillance and Response System (GISRS) are reviewed and used to generate formal recommendations for the composition of seasonal influenza vaccines [1,2]. Recommendations for the northern hemisphere vaccine are made in February and for the southern hemisphere in September. Strain selection is based on serological data, with human data used to estimate immunogenicity, generally considered as a proxy for vaccine effectiveness (VE). At the February 2013 meeting, epidemiological data were submitted reporting interim VE estimates from surveillance systems in Canada, Europe and the United States (US). These estimates were published [3-8] and included for the first time with the package reviewed by GISRS meeting members. Final season estimates for the northern hemisphere were recently presented at the September 2013 meeting, as well as interim estimates for the southern hemisphere.

Interim estimates are vulnerable to change. First, as the season reaches its peak more data become available. For example, early US interim estimates released for the period from 3 December to 2 January 2013 [9] suggested a crude VE of 62% (95% confidence interval (CI): 51 to 71), while a second, lower interim estimate of 51% (95% CI: 43 to 58) was released later in February, with an adjusted estimate of 56% (95% CI: 47 to 63) [6]. The second interim estimate was made after the peak of influenza circulation had been reached, included more than double the sample size and had more complete information on covariates for adjustment and/or exclusion. Interim and final season estimates might also be expected to differ when the predominant type or subtype shifts within a season. For example, the early adjusted VE estimate from pooled European Influenza Monitoring Vaccine Effectiveness (I-MOVE) data in 2010/11 was 42% (95% CI: −7 to 69) [10] while the final estimate was 52% (95% CI: 30 to 67) [11]. In the interim analysis, 77% of viruses were influenza A(H1)pdm09 and 21% were influenza B, whereas the final analysis included 58% influenza A(H1)pdm09 and 38% influenza B viruses [10,11]. In the 2011/12 season final European estimates against influenza A(H3) were all revised down [12-14]. This was attributed to waning immunity against A(H3), a phenomenon that is being further investigated.

In the 2007/08 season, Belongia and colleagues compared interim and final estimates from US data, observing a difference of about seven percentage points (44%; 95% CI: 11 to 65 versus 37%; 95% CI: 22 to 49) [15]. They concluded that interim estimates were a useful indicator of VE mid-season. Systematic comparison of interim and final estimates has not been done since and has not been done for multiple seasons. To assess whether interim estimates reliably predicted final estimates, we compared retrospective interim and final VE estimates against influenza A and B for six influenza seasons, from 2007 to 2012, in Victoria, Australia.
Methods
We used data collected as part of the Victorian General Practice Sentinel Surveillance network for the years 2007 to 2012 to calculate retrospective interim and final estimates. This network has been described in detail elsewhere [16]. Briefly, recruitment follows the case test-negative design [17-19]: a subset of patients seeing their general practitioner (GP) for influenza-like-illness (ILI; combination fever (measured or history of), cough and fatigue [20]) during the southern hemisphere influenza surveillance period are recruited at the GP’s discretion, swabbed and tested for influenza by real-time reverse transcription-polymerase chain reaction (RT-PCR). Those testing positive for influenza A or B are cases and those testing negative are non-cases. The GPs collect demographic data (age, sex), symptom onset date, vaccination status, vaccination date and, since 2011, the presence of conditions predisposing the patient to severe influenza (chronic heart disease, chronic respiratory disease, diabetes, impaired immunity, obesity, pregnancy). Surveillance generally begins in epidemiological week 18 in April/May and ends in week 44 in October/November.

VE estimates from the sentinel network have been reported for all years from 2007 to 2012 [21-24].
For all years, the peak of the season preceded the end of the interim period. The characteristics of participants were not different for interim and final estimates (Table).

VE was estimated as \((1 - OR)\) using logistic regression. Patients were considered vaccinated if they had received the vaccine \(\geq 14\) days prior to the onset of symptoms and excluded if vaccination took place \(\leq 14\) days. Patients were considered influenza-positive if they tested positive to any of influenza A(H1), A(H3) or influenza B viruses by real-time RT-PCR, but no separate analyses were conducted for type or subtype. Models were adjusted for age group (\(\geq 18, 18–64, \geq 65\) years) and week of presentation. The sensitivity of the estimates was tested in four ways: (i) the end of the interim period was brought forward by one and two weeks; (ii) final estimates excluded patients presenting more than eight days after symptom onset to reduce the possibility of false negative results, an exclusion which may not be possible in an interim analysis; (iii) estimates were restricted to people in a target group for vaccination (people with predisposing conditions or aged \(\geq 65\) years); and (iv) different variables for adjustment were used in the interim and final models. The fourth sensitivity analysis was based on the likely scenario where some information would be missing for the interim analysis, such as complete data on the presence of a condition predisposing to severe influenza and the date of onset, and where a decision was made to change the age groups used to increase comparability with other studies. Thus, the model for the final VE estimate included a variable representing the presence of at least one comorbid condition and a variable indicating the time between onset and consultation, an additional age group was added (\(\geq 18, 18–44, 45–64, \geq 65\) years) and month was used instead of week to denote calendar time. The third and fourth sensitivity analyses were restricted to the years 2011 and 2012 because data on predisposing conditions were only available for these two years.

**Results**

The data available for each year are shown in Figure 1 with lines indicating the end of the interim period. For all years, the peak of the season preceded the end of the interim period. The characteristics of participants were not different for interim and final estimates (Table).

Interim and final estimates were determined using the same model adjusting only for age group and week. There were no statistical differences between estimates, with point estimates varying by up to five percentage points (Figure 2). For 2007, 2008, 2010 and 2011 interim point estimates were lower than final estimates, while for other years interim estimates were higher. In the first sensitivity analysis, when the interim period was shortened by one week, there continued to be little difference in estimates for all years except 2008, where estimates differed by more than 10 percentage points. Shortened by a further week, estimates for 2008 continued to show great variability, and the direction of effect was reversed.

The second sensitivity analysis excluded people presenting more than eight days since symptoms onset (or for whom the onset date was not recorded), resulting in the exclusion of 437 people from the analysis (Table). VE estimates with this exclusion criterion were 61% (95% CI: 30 to 79) for 2007, −7% (95% CI: −123 to 49) for 2008, −2% (95% CI: −49 to 30) for 2009, −1% (95% CI: −41 to 84) for 2010, 48% (95% CI: −2 to 74) for 2011, and 44% (95% CI: 12 to 64) for 2012. VE estimates were within four percentage points of those made without the exclusion criterion, with the exception of 2009; The 2009 estimates differed by nine percentage points and this year had the most exclusions (n=197).

The third and fourth sensitivity analyses were restricted to 2011 and 2012, the only two years for which data on comorbidity were collected. When interim and final estimates were compared for only those people in a target group for vaccination, the interim estimate was 61% (95% CI: −149 to 94) for 2011, while the final estimate was 48% (95% CI: −110 to 87). For 2012, the interim estimate was 30% (95% CI: −60 to 69), while the final estimate was 32% (95% CI: −52 to 70). Interim and final estimates were also compared when using different models for the estimates. The interim model adjusted for age group (\(\geq 18, 18–64, \geq 65\) years) and week, while the final model included presence of a predisposing condition, days between symptom onset and consultation, an additional age group (\(\geq 18, 18–44, 45–64, \geq 65\) years) and month. The interim and final estimates for 2011 were 44% (95% CI: −26 to 75) and 43% (95% CI: −20 to 73) respectively, and for 2012 were 49% (95% CI: 21 to 68) and 44% (95% CI: 13 to 69).

**Discussion**

We found that interim VE estimates over six influenza seasons closely approximated final estimates when the interim period was limited to week 36. When the interim period was shortened, estimates for 2008 were different by more than ten percentage points. Estimates for 2008 showed the greatest instability, which may be explained by that year’s smaller sample size and the timing of the season, the peak of which fell later than in other years in week 35 (Figure 1). Only one interim estimate has previously been reported for this surveillance network, for the 2009 pandemic [26], a season
### Characteristics of patients with influenza-like illness, Victoria, Australia, 2007–2012

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>2007</th>
<th>2008</th>
<th>2009</th>
<th>2010</th>
<th>2011</th>
<th>2012</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total swabbed</td>
<td>392</td>
<td>466</td>
<td>ND</td>
<td>310</td>
<td>404</td>
<td>ND</td>
</tr>
<tr>
<td>Total included</td>
<td>386</td>
<td>458</td>
<td>ND</td>
<td>304</td>
<td>396</td>
<td>ND</td>
</tr>
<tr>
<td>Age group</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;18 years</td>
<td>64 (17)</td>
<td>71 (17)</td>
<td>44 (14)</td>
<td>91 (15)</td>
<td>75 (14)</td>
<td>90 (15)</td>
</tr>
<tr>
<td>≥18–64 years</td>
<td>298 (77)</td>
<td>307 (62)</td>
<td>235 (77)</td>
<td>306 (77)</td>
<td>653 (68)</td>
<td>691 (68)</td>
</tr>
<tr>
<td>≥65 years</td>
<td>24 (6)</td>
<td>28 (6)</td>
<td>25 (8)</td>
<td>28 (7)</td>
<td>21 (4)</td>
<td>12 (3)</td>
</tr>
<tr>
<td>Vaccination status</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unvaccinated</td>
<td>316 (82)</td>
<td>370 (81)</td>
<td>244 (80)</td>
<td>326 (80)</td>
<td>760 (80)</td>
<td>798 (80)</td>
</tr>
<tr>
<td>Vaccinated</td>
<td>70 (18)</td>
<td>88 (19)</td>
<td>60 (20)</td>
<td>70 (18)</td>
<td>195 (20)</td>
<td>205 (20)</td>
</tr>
<tr>
<td>Influenza</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A(H1)</td>
<td>46 (24)</td>
<td>48 (22)</td>
<td>343 (89)</td>
<td>343 (89)</td>
<td>121 (86)</td>
<td>146 (87)</td>
</tr>
<tr>
<td>A(H3)</td>
<td>115 (61)</td>
<td>127 (58)</td>
<td>22 (31)</td>
<td>41 (36)</td>
<td>7 (2)</td>
<td>7 (2)</td>
</tr>
<tr>
<td>A(mixed)</td>
<td>1 (1)</td>
<td>1 (0)</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>A(NS)</td>
<td>7 (4)</td>
<td>7 (3)</td>
<td>1 (1)</td>
<td>6 (5)</td>
<td>35 (9)</td>
<td>36 (9)</td>
</tr>
<tr>
<td>B</td>
<td>19 (10)</td>
<td>35 (16)</td>
<td>45 (63)</td>
<td>63 (55)</td>
<td>0 (0)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>Presentation after time of symptoms onset (exclusion of &gt; 8 days or date of symptom unknown made for final estimates only in sensitivity analysis 2)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>≤8 days</td>
<td>412 (86)</td>
<td>550 (87)</td>
<td>359 (91)</td>
<td>37 (9)</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>&gt;8 days or onset date unknown</td>
<td>498 (79)</td>
<td>542 (80)</td>
<td>473 (79)</td>
<td>64 (9)</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Has a condition predisposing to severe influenza (for sensitivity analysis 3 and 4)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td>Yes</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
</tr>
</tbody>
</table>
which started and peaked earlier than in other years. The final estimate for that year used a more complete model with data collected over a much longer period but the VE point estimate was the same, and differed by only four percentage points when the analysis was restricted to the weeks of maximum influenza activity [22]. Thus, interim estimates may be their most reliable when made after the peak, which is more likely in seasons which start early. For weeks 18 to 36 of the 2013 season, our interim estimate at the time of writing was 43% (95% CI: −30 to 75). However, like 2008, the 2013 season has been characterised by a late start and relatively low activity. In addition the VE estimate was unusually sensitive to the model used. Consequently, we expect the interim estimate for 2013 to be less reliable than in other years.

We expected interim estimates would be their least reliable when made for specific, smaller groups, such as people in a target group for vaccination. Moreover in the presence of waning VE within a season, it would be expected that the distribution of vaccinated cases would be skewed towards the end of the season, resulting in a lower final VE estimate. For example, in the 2011/12 European season, pooled estimates against A(H3) for people in a target group for vaccination were 43% (95% CI: −0.4 to 68) in the interim [27] and 25% (95% CI: −6 to 46) at the end of the season [12]. Similarly, our final estimates for people in a target group for vaccination in 2011 declined 13 percentage points, from 61% to 48%. Conversely, estimates increased two percentage points in 2012 but the sample size was larger in 2012 and the season peaked earlier. Meaningful differences between interim and final estimates might also be expected when making estimates separately for each type/subtype, as the distribution of cases may be skewed or the number of exposed cases may be small.
Estimates varied little when the model used for final estimates was altered, as was done in the second and fourth sensitivity analyses. The restriction of patients to include only those presenting within eight days was done to reduce the possibility of false negatives. However, this modification to the model may not be expected to alter estimates, as imperfect, nondifferential sensitivity in the presence of perfect specificity will not usually bias estimates [19, 28]. In contrast, modifying the covariates in the model might be expected to have a greater impact on both point estimates (by removing or introducing bias) and precision. By comparison with the principle analysis, we observed modest changes using a larger model, suggesting the more parsimonious model would have sufficed for these data. However, this may not always be the case. In some circumstances, larger models can reduce the effective sample size used due to complete or quasi-complete separation (also known as perfect prediction), which can inflate estimates and reduce precision [29]. For the data reported here, perfect prediction led to the loss of the oldest age group in the 2010 analyses. This problem has also been reported by the I-MOVE investigators when including calendar time (week) as a categorical variable, due to perfect prediction within a week or weeks [11]. In such cases, it may be preferable to employ methods such as exact logistic regression or penalised likelihood estimation to avoid generating biased estimates [29].

The relatively recent adoption of the test-negative design study [17-19] has permitted rapid dissemination of VE estimates on a yearly and interim basis, and consideration of these estimates in vaccine strain selection meetings has been suggested for some time [30]. However at this early stage, VE estimates are unlikely to influence strain selection; VE studies are less developed than the immunogenicity studies that have been used for decades to guide strain selection and should not yet be expected to provide reliable estimates of VE by type, sub-type, age-group and target group. However, our results illustrate the likely range of protection afforded by trivalent influenza vaccines (the only vaccines licensed in Australia) and support the use of late interim estimates as a proxy for final estimates. As VE studies evolve, their usefulness for strain selection should improve. The results presented here are hypothetical comparisons using cleaned data. It would be instructive to see a similar retrospective comparison for countries in the northern hemisphere.

Acknowledgements

The General Practice Sentinel Surveillance network is partly funded by the Victorian Government Department of Health. The Melbourne WHO Collaborating Centre for Reference and Research on Influenza is supported by the Australian Government Department of Health.

Conflict of interest

None declared.

Authors’ contributions

SGS conceptualised the study, undertook analysis and interpretation of the data, and participated in writing the manuscript. HK conceptualised the study, undertook interpretation of the data and participated in writing the manuscript.

References


