Information and Communication Technologies

EPIWORK

Developing the Framework for an Epidemic Forecast Infrastructure

http://www.epiwork.eu

Project no. 231807

D6.3 Establishment of a New Cohort for the Population-based Approach (PBA) to influenza surveillance

Period covered: Month 13th-24th
Start date of project: February 1st, 2009
Due date of deliverable: Month 24
Distribution: public

Date of preparation: 02-15-11
Duration: 12 Months
Actual submission date: 03-30-11
Status:

Project Coordinator: Alessandro Vespignani
Project Coordinator Organization Name: ISI Foundation
Lead contractor for this deliverable: SMI

Deliverable 6.3

The overall objective of the work produced within Workpackage 6 is to provide a comparative analysis of three systems for disease surveillance in Sweden: (1) the existing General Practice-based sentinel system of influenza surveillance with the (2) new Internet monitoring system (IMS) that is implemented within the framework of Workpackage 5, and (3) a new population-based approach (PBA), recently developed and evaluated at the Swedish Institute for Communicable Disease Control (Smittskyddsinstitutet – SMI). The first deliverable (6.1), due on February 1, 2010, was a fully functioning technical infrastructure for the PBA surveillance. This was accomplished with considerable margin. Since then, we have continued to run the population-based surveillance system on a full scale. We have also finished the analyses of the validation effort and a manuscript is ready for submission.

Deliverable 6.3 is a new cohort for the PBA ("Sjukrapport"). This cohort was established during August through September 2010 in close collaboration with Statistics Sweden, which holds the continuously updated computerized Swedish population register.

In order to draw a representative sample from the population, a sampling frame was first created. The sampling frame demarcates, identifies and enables linkage to the objects in the population. The sampling frame in this study was created proceeding from the Swedish Register of the Total Population (RTB) and consisted of all individuals registered as residents of Stockholm County, aged 0-95 years. The number of individuals in the sampling frame was 2,028,172 on May 31st 2010. The age distribution in the Stockholm County population is exhibited in Figure 1 below.

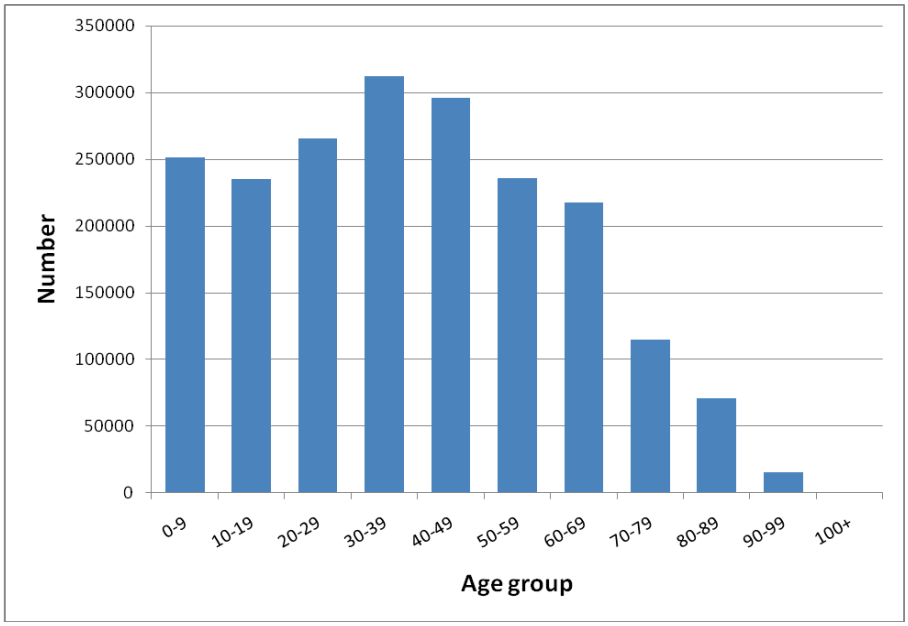

Figure 1: The age distribution (in 10-year age-bands) among people registered as residents of Stockholm county on May 31st 2010.

Since we strived for a high resolution among children, who are the most important vehicles for the spread of influenza in the population, and also wished to compensate

2011 © Copyright lies with the respective authors and their institutions

for an expected low participation of young adults (based on experiences from previous years), we aimed to over-sample among children and young adults. We therefore decided to perform a stratified random sample. The population described above was divided into four strata (0-14, 15-39, 40-64, and 65-95 years). The sizes of the strata are shown in Figure 2 below.

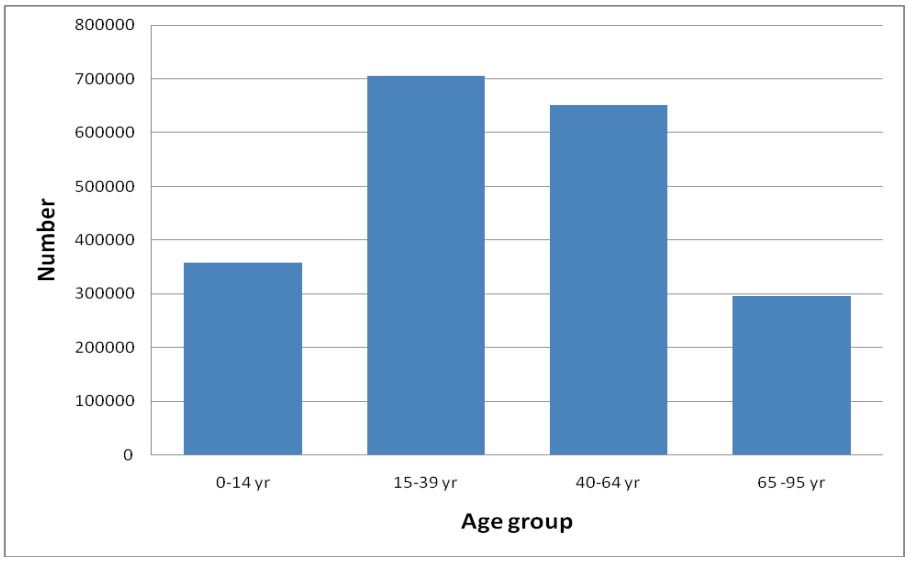

Figure 2: The 2,028,172 eligible persons, aged 0-95 years and registered as residents of Stockholm county on May 31st 2010, divided into four age strata.

Accordingly, we drew a stratified independent random sample of 14,000 individuals from the sampling frame, using an in-house selection program. A stratified independent random sample means that every object within each stratum has equal probability of being selected. The sizes of the stratum-specific samples are shown in Table 1.

Table 1: Description of strata and sample size

Strata	Description	Number
1	Stockholm County random sample born 1996-2010	4,490
2	Stockholm County random sample born 1971-1995	5,738
3	Stockholm County random sample born 1946-1970	2,259
4	Stockholm County random sample born 1915-1945	1,513
Total		**14,000**

As can be seen in Figure 3, which shows the relative distribution of the age strata in the population (blue bars) and the corresponding distribution in the sample (red bars), there was a deliberate and considerable over-representation of children 0-14 years of age, a less marked but still substantial over-representation of young adults (15-39 years of age), and consequently a corresponding under-representation of the age groups 40-64 years and 65-95 years (where the participation was expected to be higher than among young adults).

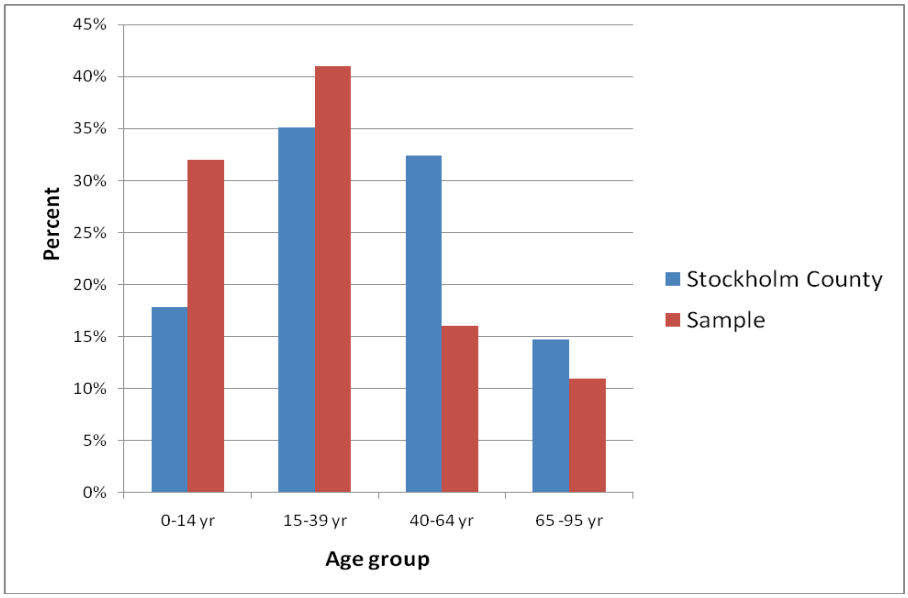

Figure 3: Relative distributions of age strata in the population (blue bars) and in the sample (red bars).

The selected individuals were contacted through regular mail. Letters were dispatched on 3 occasions during August and September and consisted of:

- First round: Two single-sided A4 which constituted the invitation and instructions, along with a fridge magnet.
- Second round: Reminder postcard, A5 format
- Third round: Reminder postcard, A5 format

Before the first invitations were sent out, an identification check of the selected persons was performed against population data to get the current address information. Upon this check, it was found that 72 persons no longer belonged to the population and thus constituted so called over-coverage. The most common reasons for the over-coverage was that the individuals had emigrated (n=16), moved out of the study area (n=55) or died (n=1) after the sample had been drawn.

A special contract (personbiträdesavtal) was drawn up between SCB and SMI, at SMI's initiative. SCB was advised by its juridical office to remove the link to the personal data three months after the last mailing round.

The first mailing round contained an information letter, in which the invitees could read about the background and objectives of the investigation, along with an outline of their commitment. Those who accepted to participate were exhorted to send a notification via Sjukrapport's website or interactive voice response telephone service (Figure 4).

2011 © Copyright lies with the respective authors and their institutions

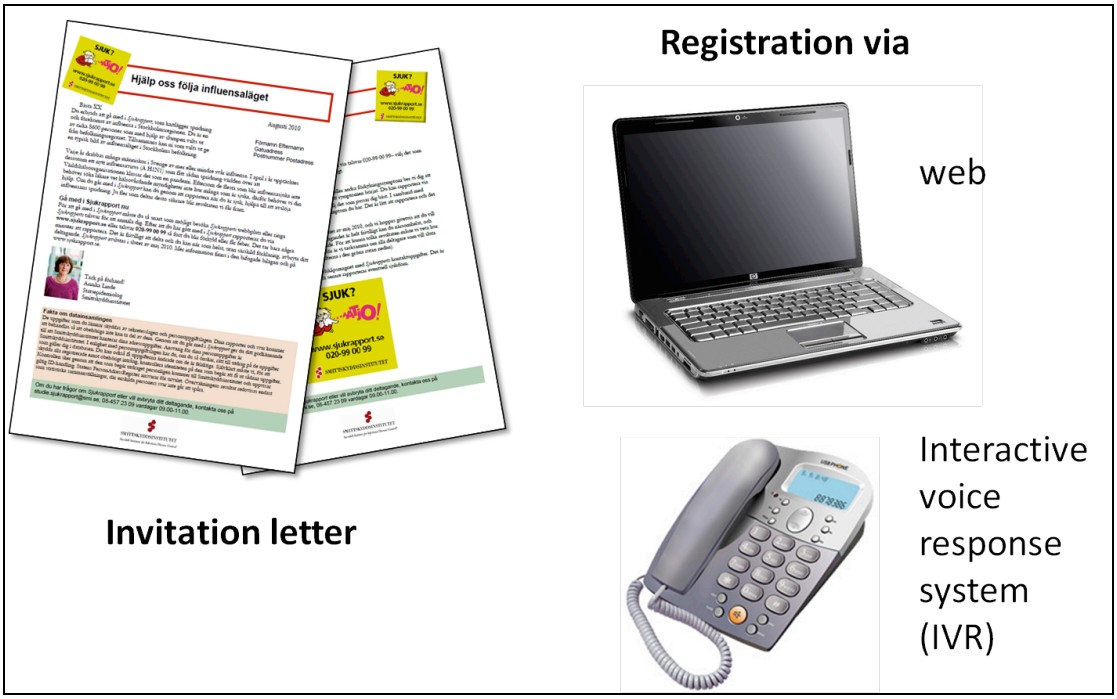

Figure 4: The structure of the invitation to the PBA cohort.

The first mailing round was dispatched on August 12, 2010. Thereafter, reminders were sent to those who had not responded to the previous mailing. Both reminders were postcards of an A5 format. The first reminder was sent out on August 27. The second reminder was divided into two mailings, one regular and one extra mailing, dispatched on September 9th and 10th, respectively.

A few days prior to the scheduled mailing date, SMI delivered an Excel file that contained the National Registration Number/running ID number of persons who had answered and thus were to be excluded from the reminder file.

In anticipation of reminder 1, such a file was delivered on August 23 and contained 924 individuals. Upon an identity check, it was found that 17 of the National Registration Numbers were not part of the original sample. Thus, 907 persons in the sample were excluded from the first reminder. Of these 907 individuals in file 1, the following were registered:

- 795 (of whom 3 were without names) were either participants in the study or invitees who had actively declined participation;
- 109 running ID numbers representing either invitations that were returned to sender or invitees who had actively declined participation.

In anticipation of reminder 2, a new file was delivered by SMI on September 7, containing 1233 new National Registration Numbers/running ID numbers. Upon an identity check, it was found that 45 of the National Registration Numbers were not part of the original sample. Thus, 1188 additional persons were excluded from the original selection file. Of these 1188 individuals in file 2, the following were registered:

- 1117 were either participants in the study or invitees who had actively declined participation;
- 71 running ID numbers representing either invitations that were returned to sender or invitees who had actively declined participation.

After the last reminder, the final total number of participants amounted to 2701 (19.29%). The day-by-day inclusion of participants is depicted in Figure 5 below. The effect of the 2 reminders can be clearly seen.

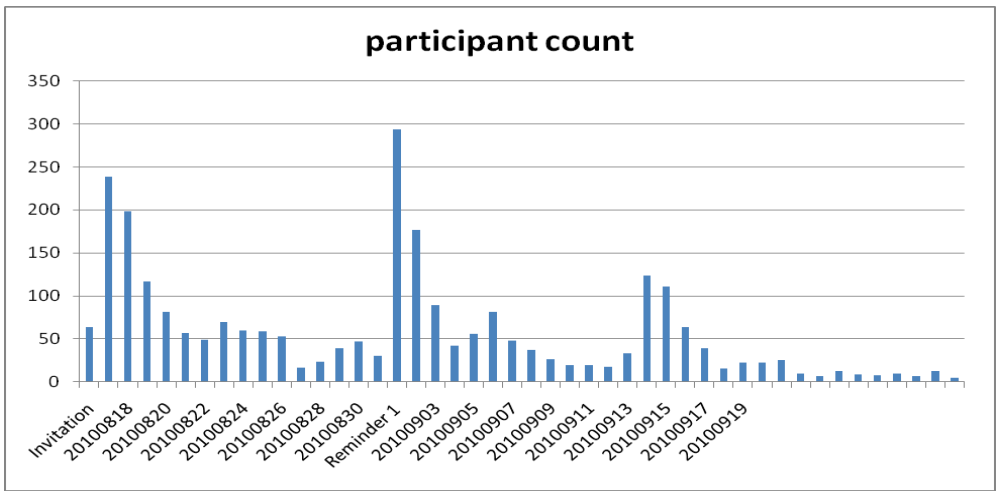

Figure 5: Number of participants included on each day of the recruitment period.

Of the 2,701 participants, approximately 75% registered via the web, while 25% registered via our interactive voice response telephone service (Figure 6).

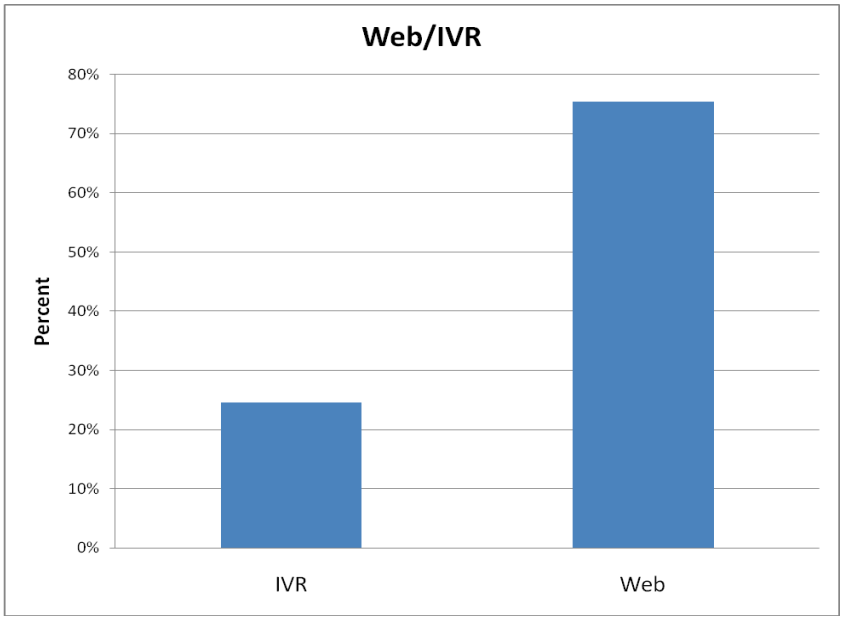

Figure 6: The proportions of 2,701 participants who registered their participation through the Interactive Voice Response (IVR) telephone and the PBA website, respectively.

2011 © Copyright lies with the respective authors and their institutions

The final relative distribution of age strata among participants, along with the corresponding distributions in the sampling frame and the sample, is depicted in Figure 7. The deliberate over-representation among children 0-14 years of age was further enhanced by a high participation rate so that they constituted 39% instead of 18% in the population. The over-sampling among young adults, on the other hand, did not fully compensate for the low participation rate; they constituted 28% of the participants while the percentage in the population was 35%. The consequential under-sampling among 40-64 and 65-95 year olds was to some extent counter-balanced by higher participation rates. It should be emphasized that exact representation of the age structure in the population is not required. When combining data from different age strata, the imbalances are easily fixed through appropriate weighting. Low participation among young adults could, however, introduce selection bias.

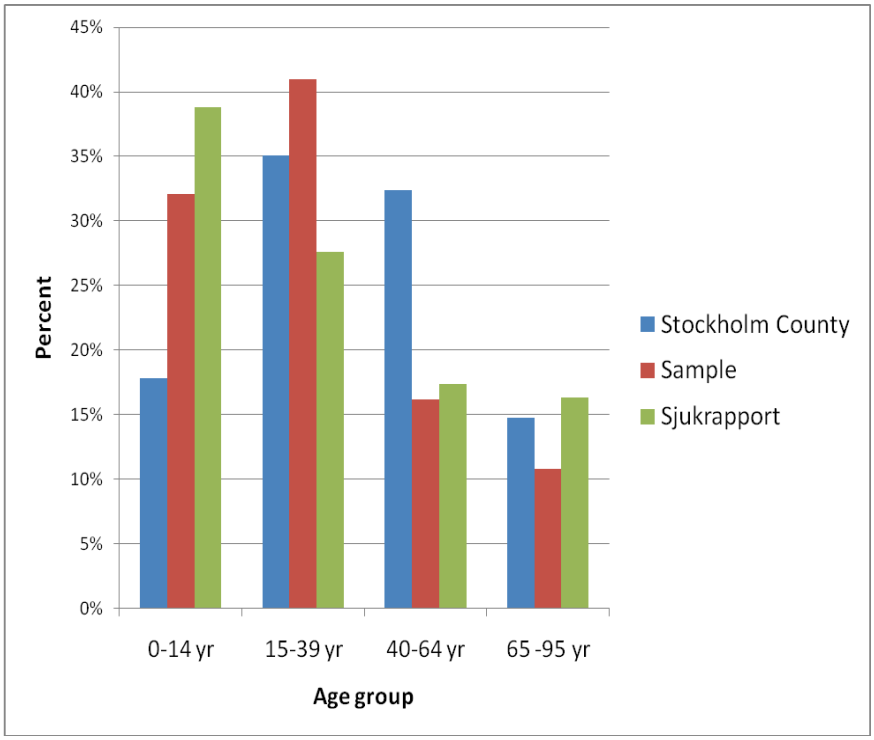

Figure 7: *Relative distributions of age strata in the population (blue bars), in the sample (red bars), and in the final cohort of participants (green bars).*

Not unexpectedly, there was a slight over-representation of women and a corresponding under-representation of men (Figure 8).

The participation rate was much lower than during previous seasons. This could be partly explained by an unusually early start; the first invitation was sent out when many people were still on summer vacation, reflected by a low yield from the initial mailing round and a richer yield from the first reminder. Another explanation is that this invitation coincided with the aftermath of the swine flu pandemic, during which Swedish authorities, among them SMI, were subject to critique and discontent from the public. In particular, the adequacy of the mass vaccination program was questioned when it was realized that the pandemic was much less severe than had

been feared. So the Swedish population was (and probably still is) somewhat influenza-tired, and the motivation to contribute to influenza surveillance is generally low at present.

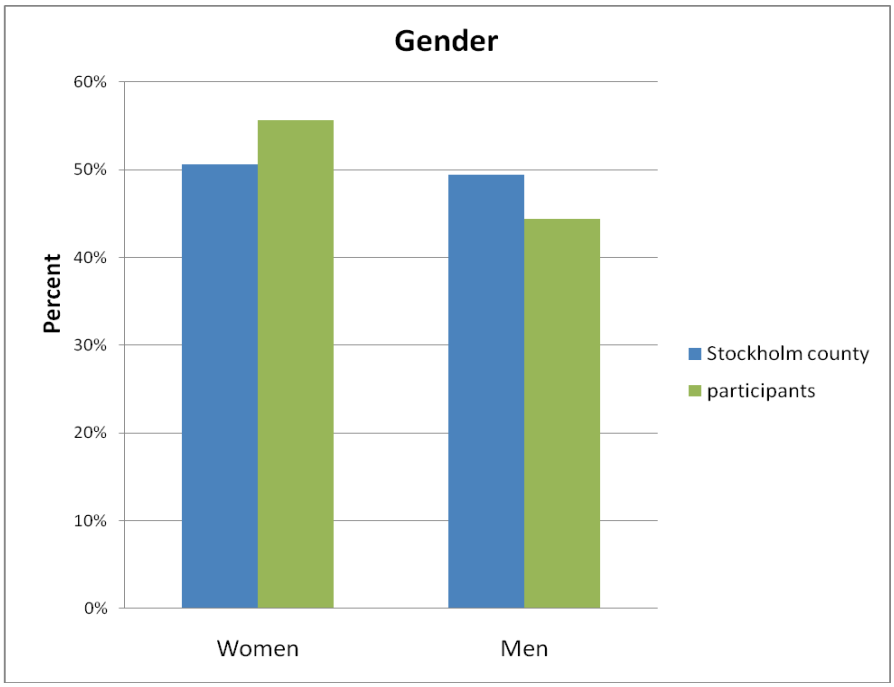

Figure 8: The gender distribution in the sampling frame (blue bars) and among the participants (green bars).

The disease reporting is ongoing. Below are two snapshots of the epidemic curves based on this year's registration (purple curve), compared to the two previous seasons. The first one concerns all acute upper respiratory tract infections (ARI). As can be seen, the weekly prevalence of ARI (around 3%) is slightly higher than during last year (when extraordinary precautions were taken by the general population to prevent transmission of the swine flu), but still considerably lower than in the season 2008/2009.

2011 © Copyright lies with the respective authors and their institutions

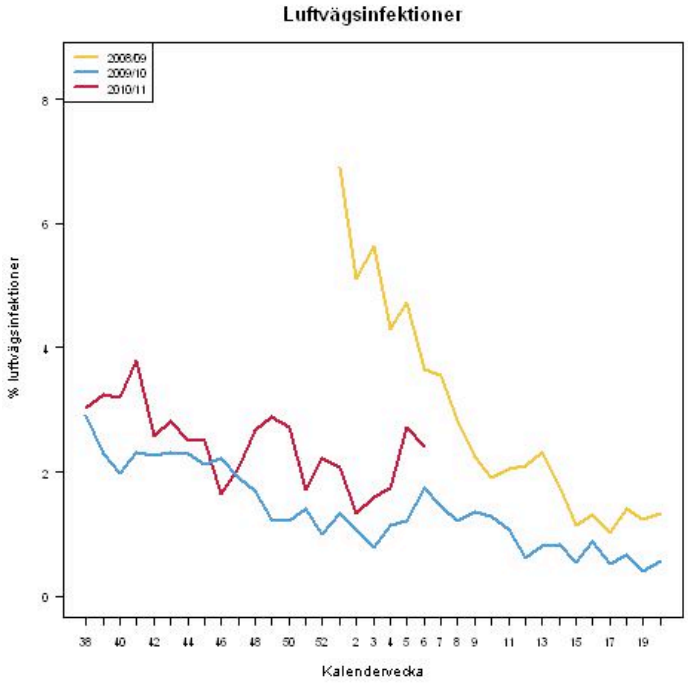

Figure 9: Epidemic curves for total acute upper respiratory tract infections for the seasons 2008/2009 (yellow curve, starting on January 1, 2009), 2009/2010 (blue curve), and 2010/2011 (purple curve). Week 38 is third week of September, week 1 first week of January, and Week 19 third week of May.

The second graph concerns influenza-like illness (ILI):

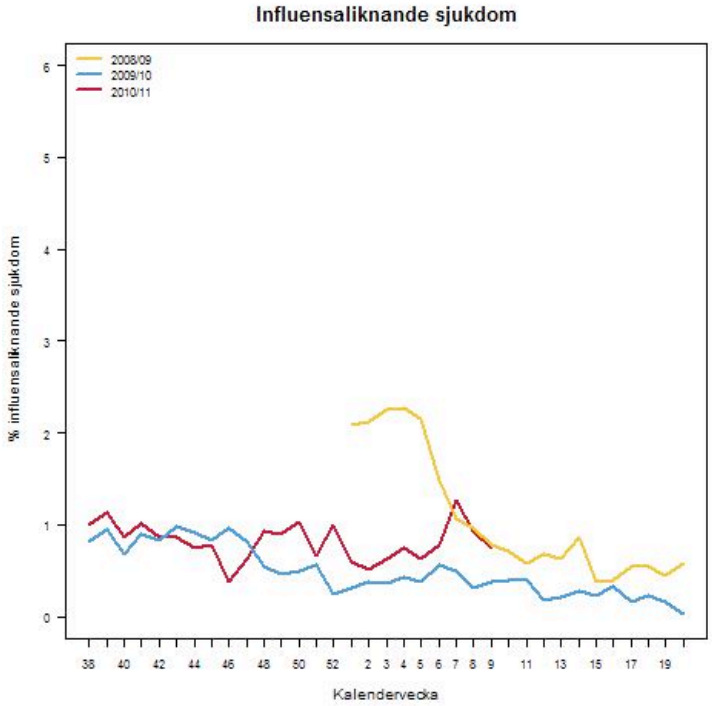

Figure 10: Epidemic curves for influenza-like illness (ILI) for the seasons 2008/2009 (yellow curve, starting on January 1, 2009), 2009/2010 (blue curve), and 2010/2011 (purple curve).

In view of a cluster of severe or deadly cases of the swine flu in the first weeks of this year, the moderate weekly prevalence of ILI in the general population is of considerable interest. Admittedly, there was a moderately high peak between weeks 5 and 9, during which time the rates were considerably higher than during the same period last year (but substantially lower than during the seasonal influenza season 2008-2009). The discrepancy between the rate in the population and the number of registered cases with severe influenza seen in health care is not at all unexpected; due to the precautions in the last season, the elderly high-risk population was spared, but this year, the old high-risk individuals are more exposed. But most high-risk individuals are vaccinated. However, a small proportion remains unvaccinated, and it is from this very subpopulation that the recent severe cases emanate.

The continuously updated incidence curves can also be followed on SMI's website http://www.smittskyddsinstitutet.se/publikationer/smis-nyhetsbrev/sjukrapport/sasongen-20102011/

2011 © Copyright lies with the respective authors and their institutions