Appendix S1: Test statistics for comparisons of morbidity ratios

Let $N_M$ and $N_F$ be the subpopulations of males and females in an age group in Japanese population, i.e. fixed values, and let $n_M$ and $n_F$ be the random variables that describe the numbers of male and female patients with pdmH1N1 in the age group identified through ~5000 sentinel points in Japan, respectively. Since the present sampling is based on the data reported from the sentinel points, the morbidities at the current time $p_M$ and $p_F$ cannot be estimated from the patient data. Let $\pi_M$ and $\pi_F$ be the probabilities that male and female patients in the age group visit the sentinel points, respectively. Then, $\pi_M p_M$ and $\pi_F p_F$ are estimated by $\hat{\pi}_M p_M = \frac{n_M}{N_M}$ and $\hat{\pi}_F p_F = \frac{n_F}{N_F}$, respectively. In this paper, $\gamma = \frac{\pi_M p_M}{\pi_F p_F}$ is referred to as the M/F ratio. The ratio is the apparent M/F ratio, if $\pi_M \neq \pi_F$, and it is the true M/F ratio if $\pi_M = \pi_F$. The M/F ratio is estimated by $\hat{\gamma} = \frac{\pi_M p_M}{\pi_F p_F} = \frac{n_M N_F}{n_F N_M}$. Since populations $N_M$ and $N_F$ are very large, random variables $n_M$ and $n_F$ can be considered to be Poisson-distributed with means $N_M \pi_M p_M$ and $N_F \pi_F p_F$, respectively. In the surveillance data, the means are sufficiently large, so from the central limit theorem, $n_M$ and $n_F$ are asymptotically normally distributed with variances $N_M \pi_M p_M$ and $N_F \pi_F p_F$, respectively. In order to test hypotheses with respect to the M/F ratio, the following statistic can be used:

$$Z = \frac{\frac{\pi_M p_M}{N_M} - \gamma \frac{\pi_F p_F}{N_F}}{\sqrt{\frac{\pi_M p_M}{N_M} + \gamma^2 \frac{\pi_F p_F}{N_F}}} \approx \frac{\frac{\pi_M p_M}{N_M} - \gamma \frac{\pi_F p_F}{N_F}}{\sqrt{\frac{\pi_M p_M}{N_M} + \gamma^2 \frac{\pi_F p_F}{N_F}}}.$$  \hfill (A1)

The above statistic is asymptotically distributed according to the standard normal distribution. Under the null hypothesis $H_0$, i.e. $\gamma = 1$, the above statistics becomes
\[ Z = \frac{\pi_{MPM} - \pi_{FPF}}{\sqrt{\frac{\pi_{MPM}}{N_M} + \frac{\pi_{FPF}}{N_F}}} \approx \frac{\tilde{\pi}_{MPM} - \tilde{\pi}_{FPF}}{\sqrt{\frac{\tilde{\pi}_{MPM}}{N_M} + \frac{\tilde{\pi}_{FPF}}{N_F}}} \]

Based on the above test statistic, we can test \( H_0: \pi_{MPM} = \pi_{FPF}, \) i.e. \( \gamma = 1 \) vs. \( H_1: \pi_{MPM} \neq \pi_{FPF}, \) i.e. \( \gamma \neq 1. \)

Next, the asymptotic confidence interval of the M/F ratio \( \gamma = \frac{\pi_{MPM}}{\pi_{FPF}} \) is derived. For large sample, (A1) is asymptotically equal to

\[ Z = \frac{\tilde{\gamma} - \gamma}{\sqrt{\frac{\tilde{\pi}_{MPM}}{N_M} + \frac{\tilde{\pi}_{FPF}}{N_F}}} = \frac{\tilde{\gamma} - \gamma}{\sqrt{\gamma^2 \left( \frac{1}{n_M} + \frac{1}{n_F} \right)}}. \]  

(A2)

From the above standardized statistic, we have the 95% confidence interval of the M/F,

\[ \tilde{\gamma} \left( 1 \pm 1.96 \sqrt{\frac{1}{n_M} + \frac{1}{n_F}} \right). \]

Finally, we can test the M/F ratio differences between pdmH1N1 and Year 2005 influenza infections using the following statistic. Let \( (p_{iM}, p_{iF}, \pi_{iM}, \pi_{iF}; \gamma_i) \) be parameters in an age group for pdmH1N1 \((i = 1)\) and those for Year 2005 influenza \((i = 2)\), which are defined above; and let \( (n_{iM}, n_{iF}) \) and \( (N_{iM}, N_{iF}) \) be the numbers of patients and subpopulations of males and females. From (A2), the estimators of M/F ratios, \( \hat{\gamma}_i = \frac{n_{iM}N_{iF}}{n_{iF}N_{iM}} \) \((i = 1, 2)\), are asymptotically normally distributed with means \( \gamma_i \) and variances \( \frac{\gamma_i^2}{n_{iM} + n_{iF}}, \) respectively. Since \( \hat{\gamma}_i \) \((i = 1, 2)\) are independent, the following statistic is distributed according to an asymptotic normal distribution with variance 1:

\[ Z_{12} = \frac{\hat{\gamma}_1 - \hat{\gamma}_2}{\sqrt{\frac{\gamma_1^2}{n_{1M} + n_{1F}} + \frac{\gamma_2^2}{n_{2M} + n_{2F}}}}. \]  

(A3)

By using the above statistic, we can test hypotheses, \( H_0: \gamma_1 = \gamma_2 \) vs. \( H_1: \gamma_1 \neq \gamma_2. \) Under the null hypothesis, the above statistic is also asymptotically normally distributed with mean 0 and variance 1.