A Practical Guide for Designing and Conducting Influenza Disease Burden Studies

World Health Organization
Western Pacific Region
November 2008
Table of Contents

Abbreviations and Acronyms .................................................................................................................... iii

Introduction .................................................................................................................................................. 1

Section 1: Studies Directly Measuring The Disease Burden Of Influenza .......................................................... 3
  1.1. Indicator-based Disease Burden Studies .......................................................................................... 3
    1.1.1 Terms and Definitions .......................................................................................................... 3
    1.1.2 Case Definitions .................................................................................................................. 4
    1.1.3 Study Population.................................................................................................................... 4
      1.1.3.1 Hospitalized Acute Respiratory Infections (sARI)...................................................... 4
      1.1.3.2 Outpatients With ILI .................................................................................................. 4
    1.1.4 Specimen Collection and Handling ...................................................................................... 5
    1.1.5 Laboratory Diagnosis ......................................................................................................... 6
    1.1.6 Data Analysis: Simple Proportions and Incidence Rates ...................................................... 6
      1.1.6.1 Proportions .............................................................................................................. 6
      1.1.6.2 Incidence Rates ......................................................................................................... 7
      1.1.6.3 Individual Level Versus Population Level Data ........................................................ 7
    1.1.7 Essential Indicators For Influenza Disease Burden Measurement ......................................... 7
      1.1.7.1 Hospitalized Sari Due To Laboratory-confirmed Influenza Infection ....................... 7
      1.1.7.2 Influenza Managed In The Outpatient Setting ......................................................... 8
  1.2. Absenteeism As A Measure Of Disease Burden ............................................................................. 8
    1.2.1 Absenteeism In Schools ........................................................................................................ 8
      1.2.1.1 Methods and Materials ............................................................................................ 8
      1.2.1.2 Data Storage and Analysis ....................................................................................... 9
      1.2.1.3 Monitoring ................................................................................................................ 10
      1.2.1.4 Ethical Considerations .............................................................................................. 10
    1.2.2 Absenteeism In Workplaces .................................................................................................. 10
      1.2.2.1 Methods and Materials ............................................................................................ 10
      1.2.2.2 Data Storage and Analysis ....................................................................................... 11
      1.2.2.3 Monitoring ................................................................................................................ 12
      1.2.2.4 Ethical Considerations .............................................................................................. 12

Section 2: Studies Indirectly Measuring The Disease Burden Of Influenza ......................................................... 13
  2.1. Estimating Disease Burden Of Influenza Through Mathematical Modelling ............................... 13
    2.1.1 Introduction ...................................................................................................................... 13
    2.1.2 Data Preparations ............................................................................................................ 14
      2.1.2.1 Health Outcome Data .............................................................................................. 14
      2.1.2.2 Influenza Virus Laboratory Surveillance Data .......................................................... 14
      2.1.2.3 Data On Other Covariates ....................................................................................... 15
### Abbreviations and Acronyms

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>APACI</td>
<td>Asia-Pacific Advisory Committee on Influenza</td>
</tr>
<tr>
<td>APSED</td>
<td>Asia Pacific Strategy for Emerging Diseases</td>
</tr>
<tr>
<td>CDC</td>
<td>Centers for Disease Control and Prevention</td>
</tr>
<tr>
<td>IHR</td>
<td>International Health Regulations</td>
</tr>
<tr>
<td>ILI</td>
<td>Influenza-like Illness</td>
</tr>
<tr>
<td>MOH</td>
<td>Ministry of Health</td>
</tr>
<tr>
<td>NIC</td>
<td>National Influenza Center</td>
</tr>
<tr>
<td>NIID</td>
<td>National Institute of Infectious Diseases</td>
</tr>
<tr>
<td>RSV</td>
<td>Respiratory Syncytial Virus</td>
</tr>
<tr>
<td>RT-PCR</td>
<td>Reverse transcriptase polymerase chain reaction</td>
</tr>
<tr>
<td>sARI</td>
<td>Severe Acute Respiratory Infection</td>
</tr>
<tr>
<td>SEARO</td>
<td>South-East Asia Regional Office of WHO</td>
</tr>
<tr>
<td>WHO</td>
<td>World Health Organization</td>
</tr>
<tr>
<td>WPRO</td>
<td>Western Pacific Regional Office of WHO</td>
</tr>
</tbody>
</table>
Introduction

The epidemiology and disease burden of human influenza in most tropical and subtropical developing countries has not been adequately described. Consequently, no consensus on the public health importance of human influenza has been reached and national influenza vaccination programmes are uncommon. However, virological surveillance for influenza is well established in many countries in the Western Pacific and South-East Asia Regions. At present, there are 19 National Influenza Centres (NIC) in 14 countries in the Western Pacific Region and eight National Influenza Centres in six countries in the South-East Asia Region (1, 2). Historically, these NICs have focused on collecting clinical specimens, isolating and characterizing circulating influenza viruses, and monitoring for novel strains with pandemic potential. Despite recent efforts by Member States and the World Health Organization (WHO) to increase the scope of influenza surveillance to include disease burden data, the quality and capacity of many surveillance systems needs improvement.

While such virological surveillance is essential, in the absence of robust data describing the human health and socioeconomic burden of disease, influenza is seldom considered to be an important cause of illness. In the United States and Europe, many studies have documented a substantial disease burden in terms of morbidity, mortality and socioeconomic costs (3-5). These data have supported the development and expansion of influenza control programmes, primarily by vaccinating groups at highest risk for serious complications from influenza infection (6-8). The WHO Global Agenda on Influenza Surveillance and Control (9) calls for efforts to strengthen capacity to conduct epidemiological studies on influenza disease burden and to evaluate the clinical and economic burden of disease in countries where there is little recognition of influenza and no control policies are in place. Therefore, high quality, multi-year influenza disease burden studies from the Western Pacific and South-East Asia Regions are a priority.

During the 1st Meeting of The National Influenza Centres in the Western Pacific and South-East Asia Regions in Melbourne, Australia in May 2007, a biregional four-year work plan for strengthening national influenza surveillance capacity was developed. The work plan also requires WHO to facilitate the development of protocols to assist countries to conduct influenza disease burden research. The work plan was subsequently endorsed by the bi-regional Technical Advisory Group as a sub-work plan of the Asia Pacific Strategy for Emerging Diseases (APSED) (10).

On November 12-13, 2007 the Communicable Disease Surveillance and Response Unit of the WHO Regional Office for the Western Pacific (WPRO) organized an Informal Consultation on Protocols for Influenza Disease Burden Studies in Manila, Philippines. The objectives of this
consultation were to: 1) develop generic protocols for influenza disease burden studies; 2) review and introduce mathematical modelling methods for influenza disease burden studies; and 3) draft a guide for data collection and preparation for influenza disease burden studies through mathematical modelling. Eleven experts from eight countries participated in this consultation. The Asia-Pacific Advisory Committee on Influenza (APACI) was also represented during this meeting. This document was developed by the expert consultants of this meeting to provide specific guidance for the planning and operation of influenza disease burden studies.
Section 1: Studies Directly Measuring The Disease Burden of Influenza

1.1. INDICATOR-BASED DISEASE BURDEN STUDIES

1.1.1 Terms and Definitions

*Disease burden:* Morbidity and mortality due to influenza infection and the related costs to the health care system and to society (e.g. loss of employment, missed schooling).

*Eligible patient:* A patient who meets the ILI or sARI case definition and can be enrolled in the study.

*Enrolled patient:* A patient who meets the ILI or sARI case definition and has provided informed consent and respiratory specimens.

*Informed consent:* Agreement by the patient to participate in the study after receiving a thorough explanation of the study’s purpose, risks and benefits communicated in a manner understandable to the patient.

*Influenza patient:* A person with laboratory-confirmed influenza infection.

*Incidence rate:* The number of new influenza cases per unit of person-time at risk, often expressed as \( n \) cases / 100 000 persons/year

*Respiratory specimen:* Respiratory secretions obtained specifically for virus detection.

1.1.2 Case Definitions

*Influenza-like Illness (ILI):* A person with sudden onset of fever of > 38°C and cough or sore throat in the absence of other diagnoses.

*Severe Acute Respiratory Infection (sARI)*\(^1\): Meet ILI case definition (Sudden onset of fever over 38°C and cough or sore throat in the absence of other diagnosis),

**AND**

- Shortness of breath or difficulty breathing, **and**
- Requiring hospital admission

---

\(^1\) Also referred to as Acute Lower Respiratory Tract Infection (aLRTI)
1.1.3 Study Population

The study population should be representative of the wider population at risk for influenza infection. On balance, patients meeting the study case definition who are enrolled in the study and provide informed consent and respiratory specimens should reflect their communities in terms of age, gender, ethnicity and socioeconomic status.

1.1.3.1 Hospitalized sARI patients

A retrospective analysis of patients admitted to hospitals in the previous year with severe acute respiratory infection should be carried out in order to estimate the number of patients that could meet the case definition during the prospective study. Depending on funding and staffing, efforts should be made to obtain informed consent and respiratory specimens from all patients who meet the sARI case definition. If this is not feasible, either due to large numbers of eligible patients or a limited study budget, a systematic sampling strategy should be instituted allowing for selection of every \( n \)th eligible patient for taking respiratory specimens, but all of the patients who meet the sARI case definition in a specified period (week or month) should be recorded. This sampling strategy will reduce the potential for bias in the selection of study patients.

Example:

Step 1: Review a random subset of hospital discharge diagnoses from various months during the preceding year.

Step 2: Determine the proportion of hospital patients from the previous year that would have met the sARI case definition.

Step 3: Determine the number of patients and laboratory tests that the study budget will allow. Recognize that not all patients meeting the sARI case definition will be enrolled for various reasons including refusal to provide informed consent.

Step 4: If the budget does not allow for study participation from at least 50% of the estimated number of potentially eligible patients, design a random sampling plan that will attempt to enrol one out of every two eligible patients. Similarly, if the budget and staffing will only allow for testing of specimens from 25% of the expected eligible patient population, plan to systematically enrol every fourth hospitalized patient meeting the sARI case definition.

1.1.3.2 Outpatients with ILI

For studies in outpatients, general practice or “polyclinics” are preferable as they are more likely to represent the wider community in terms of age, gender and socioeconomic status than patients visiting emergency departments or other specialty clinics. Ideally, all patients who meet the case definition should be enrolled during randomly selected, complete/entire clinic days to allow for the collection of the total number of outpatient visits (denominator) as well as the total number meeting the ILI case definition. The proportion of patients meeting the WHO case definition for ILI will vary according to local seasonal patterns of influenza activity.

Example:

Step 1: Begin by randomly selecting one (or more) polyclinic days per week as study budget and staffing permits.
Step 2: Assign a research nurse or assistant to monitor the outpatient department intake desk to record the age, gender and reason for visit of all patients attending the clinic from opening to closing time.

Step 3: Self-reporting of the presence of a fever must be verified with measurement.

Step 4: If the patient has a measured fever over 38°C, determine if the patient also has either a cough or a sore throat. If yes, the patient meets the WHO case definition for ILI. Attempt to obtain informed consent and respiratory swab specimens from the patient.

Step 5: At the end of the clinic day, record: (1) the total number of all outpatient department patient visits, (2) the number of patients that met the WHO ILI case definition, (3) the number of patients that enrolled in the study. The laboratory results of this cohort should be followed, recording the number of patients with positive test results.

1.1.4 Specimen Collection and Handling

Diagnosis of infection with influenza relies on the collection of high quality specimens, their rapid transport to the laboratory and appropriate storage before laboratory testing (11). Influenza is best detected in specimens containing infected cells and secretions collected during the first three days after the onset of clinical symptoms. Therefore, every effort should be made to collect specimens as close to the onset of illness as possible. The time between the onset of illness and specimen collection should be recorded.

A variety of specimens are suitable, including:

- Nasal swab
- Throat swab
- Nasopharyngeal swab
- Nasopharyngeal aspirates or washes
- Nasal wash

Nasopharyngeal swabs, aspirates and washes are the best specimens for cell culture and PCR (12). However, these can be technically difficult, require careful training to ensure proper technique, and may be unpleasant for the patient. An acceptable alternative is to collect a nasal and a throat swab from the same patient and then combine these swabs into a single vial of transport medium.

Failure to carefully handle respiratory swab specimens can significantly reduce their usefulness. It is essential that swab specimens be carefully and consistently managed. Respiratory specimens should be collected, transported and stored in virus transport medium. Swab specimens should be promptly transported to the laboratory and not exposed to heat or allowed to dry. The specimens can be stored at 2º-8ºC for up to 72 hours prior to processing by PCR or cell culture. Specimens for direct detection of viral antigens by immunofluorescence staining of infected cells should be refrigerated and processed within two days. Specimens for use with commercial rapid influenza testing kits should be handled and stored in accordance with the manufacturer’s instructions.

Each specimen should be divided into aliquots for additional testing, re-testing or archiving prior to freezing at -70°C for long-term storage. Heating and repeat freeze-thawing of specimens greatly affect the quality and diagnostic yield, therefore should be avoided. Regular training and careful supervision of staff responsible for specimen collection and storage is essential to ensure high-quality specimens. Appropriate infection control and biosafety precautions (13, 14) must be
implemented from the point of specimen collection through to completion of laboratory testing and specimen disposal. The details of specimen collection procedures are provided in Annex 2.

### 1.1.5 Laboratory Diagnosis

RT-PCR and real time RT-PCR have been demonstrated to be more sensitive and timely than cell culture methods (15), more resilient to variations in specimen quality, and are the methods of choice for influenza disease burden studies. Cell culture provides viral isolates useful for vaccine strain selection but is relatively slow and is very sensitive to inappropriate specimen handling. However, it can be used as a quality control method by testing a subset of RT-PCR positive and negative specimens. Antigen detection by direct or indirect immunofluorescence is also a timely and sensitive diagnostic method but ideally requires a nasopharyngeal swab or aspirate specimen and should be conducted within two days of specimen collection. While the sensitivity of immunofluorescence is lower than both RT-PCR and cell culture, rapid antigen diagnostic tests can be useful. These tests are best suited for use in outpatient settings where patients are in the early stages of illness. Rapid antigen tests can facilitate the collection of cost-of-illness data since influenza patients may be followed from early in the course of their illness. Because of their reduced sensitivity in patients who are several days from the onset of symptoms, rapid antigen tests are not recommended for use with hospitalized sARI patients (15).

### 1.1.6 Data Analysis:

#### 1.1.6.1 Proportions

Simple proportions can provide meaningful information on disease burden. Examples include: the proportion of all outpatients that meet the case definition of ILI, the proportion of all outpatients with ILI that have influenza infection, and the proportion of all hospitalized inpatients with sARI caused by influenza virus infection. To calculate proportions, denominator data must be collected. This means that an accurate record must be kept of the total number of patients by age groups that meet specific criteria during a specified time period (clinic day, week, month and year).

**Examples of useful proportions:**

\[
\frac{\text{Number of influenza positive ILI outpatients}}{\text{Total number of ILI patients tested}}
\]

\[
\frac{\text{Number of sARI inpatients with influenza}}{\text{Total number of sARI patients}}
\]

#### 1.1.6.2 Incidence rates

An incidence rate is the number of new influenza cases per unit of person-time at risk. While incidence rates can be expressed in many ways, it is helpful to report incidence rates in a consistent manner to facilitate comparisons between studies and over time.

Example: \( n \) influenza cases/ 100 000 persons/ year
To calculate an incidence rate, the size of the population at risk must be known. This can be difficult when studying patient populations that can access health care services at multiple sites. Studies that are designed to examine specific populations whose census is known are referred to as population-based studies. In the absence of such information, certain adjustments and assumptions are necessary in order to estimate incidence rates. In such cases, advanced epidemiological assistance should be sought and the methods used to estimate incidence should be clearly and completely described in subsequent reports or publications.

1.1.6.3 Individual level versus population level data

Disease burden studies typically report data that describe the population at risk. For example, the age distribution, socioeconomic status, gender and ethnicities of the population under study. The WHO Global Agenda on Influenza Surveillance and Control (10) also calls for evaluation of the clinical aspects of influenza infection. Reporting such patient-level clinical information can be useful for clinicians providing care for influenza patients and may also help to improve outcomes for patients with serious complications. Therefore, wherever possible, disease burden researchers should also collect and report individual level data describing the clinical presentation of influenza infection, underlying risk factors, medical interventions and the outcome of influenza infection. Examples of patient level data that could be collected are provided in Annex 3.

1.1.7 Indicators for Influenza Disease Burden Measurement

1.1.7.1 Hospitalized sARI due to laboratory-confirmed influenza infection

Disease burden:

**Essential data:**
- Number of sARI cases per week, month, year
- Proportion of all sARI cases caused by influenza per week, month, year
- Distribution by age groups (0-2 years, 3-4 years, 5-17 years, 18-49 years, 50-64 years, ≥ 65 years)
- Demographic data of study population (gender, residence, ethnicity)
- Clinical data (history, symptom presentation, discharge status)

**Desirable data:**
- If catchment population is known or can be estimated, incidence per 100,000 persons per year
- Indicators of severity and outcome such as need for oxygen therapy, treatment in Intensive Care Unit and endotracheal intubation
- Mean, median and range of length of hospital stay
- Clinical status 21 days post-discharge from hospital (dead or alive)

**Socioeconomic burden data:**
- Direct treatment costs
- Lost work and lost school days
- Out-of-pocket costs to patient and families
Section 1: Studies Directly Measuring The Disease Burden of Influenza

1.1.7.2 Influenza managed in the outpatient setting

Disease burden:

*Essential data:*
- Total number of outpatient visits per week, month, year
- Proportion of outpatients with ILI per week, month, year
- Proportion of ILI patients with laboratory-confirmed influenza infection
- Age distribution of ILI patients and laboratory-confirmed influenza infections
- Seasonality: proportion of ILI cases with laboratory-confirmed influenza infection reported by week and month

*Desirable data:*
- If catchment population is known or can be estimated, incidence per 100,000 persons per year
- Distribution by age groups (0-2 years, 3-4 years, 5-17 years, 18-49 years, 50-64 years, ≥ 65 years)
- Clinical data, including history, symptoms at presentation, medical intervention

*Socioeconomic burden data:*
- Direct treatment costs
- Lost work/school days due to laboratory-confirmed influenza infection
- Out-of-pocket costs to patient and families

1.2. ABSENTEEISM AS A MEASURE OF DISEASE BURDEN

Absenteeism at schools and workplaces can provide important information on the social and economic costs of influenza. Research has shown that trends in absenteeism often correspond to seasonal variations in influenza activity (16).

1.2.1 Absenteeism in Schools

As school-aged children are at high risk for influenza infection, data on absenteeism due to influenza-like illness (ILI) from schools may provide useful information on the seasonality, burden and social costs of influenza. Individual schools may be selected to act as sentinels and followed for one or more school years to determine the frequency of ILI-associated absenteeism. If laboratory diagnosis is available, an improved understanding of the burden and seasonality of laboratory-confirmed influenza infection in the general community may also be gained by the study of school-age populations.

1.2.1.1 Methods and Materials

*Study design:*
Prospective, longitudinal, observational study

*Study population:*
All children attending selected schools in the survey area.

*Selection of schools:*
Schools should be randomly selected from among all schools in a specified administrative area.

*Data to be collected:*
Number of children under surveillance
Number of absence events/episodes
Number of missed school days
Number of absence events associated with an ILI
Number of missed school days associated with an ILI
Number of absence events associated with laboratory-confirmed influenza
Number of lost school days due to laboratory-confirmed influenza infection

Survey procedure:
1. Selected schools keep daily attendance books throughout the survey time period.
2. All school absences are reported to the study coordinator on a daily basis (Annex 4).
3. The school nurse or designated study staff then contacts the child’s family to inquire if the absence is related to illness. If yes, the student’s parents will complete a reporting form (Annex 5). The reporting form details symptoms of influenza-like illness (ILI), such as fever, cough or sore throat as well as any visits to a medical provider. ILI-associated absences are recorded and tallied.
4. If the study has sufficient financial resources and access to a qualified diagnostic laboratory, children with an ILI-associated absence may be tested for the presence of an influenza virus infection. In this case, a health care worker from the study will visit the ill child in their home to collect clinical swab specimens for laboratory diagnosis. If available, a rapid influenza test may be used to diagnose influenza infection during the home visit.

1.2.1.2 Data Storage and Analysis

An efficient system for storage and management of data should be established prior to the commencement of the survey. Weekly tallies of the number of ILI-associated absences can be used as an indicator of increased influenza activity.

There are three levels of potential data:
1. Total (All-cause) Absence Episodes
2. Total ILI-Associated Absence Episodes
3. Total Laboratory-confirmed Influenza Infection Absence Episodes

Data analysis:

1. Calculate the proportion of absentees by:

\[
\frac{\text{Total number of students absent at the school during the week}}{\text{Total number of students attending the school on the first day of the week}}
\]

If a student is absent across more than one week, he/she contributes to the numerator for both weeks

2. Attempt to identify the commencement of the influenza season by increases in absenteeism.

3. Estimate influenza burden by comparing the proportion of absentees during the influenza season with those during the non-influenza season (excess influenza absenteeism).

4. Estimate the relative influenza burden by calculating the proportion of absenteeism due to influenza, based on results of influenza diagnostic laboratory tests.
1.2.1.3 Monitoring

An ongoing comparison of school ILI-associated absenteeism with the reported incidence of influenza from the national surveillance system will allow an assessment of the reliability of the survey results.

1.2.1.4 Ethical Considerations

This type of study will, in most cases, require a formal protocol and clearance from the relevant Institutional Review Board(s). The coordinators of the survey must protect the privacy of individuals when reporting and publishing data. Before commencing such a survey, the coordinators must thoroughly explain the purpose, protocols, and material of the survey to the school administration and to parents. Written consent should be obtained, and signed and dated by both the school caretaker and the study coordinator. The titles, names, and contact information of the signatories should be clearly provided. The original copies of the document are filed at the school and at the local health care centre. Informed consent must also be obtained from a parent or legal guardian of the absentee.

1.2.2 Absenteeism in Workplaces

Selected workplaces are identified to act as sentinels and monitored prospectively for absences due to ILI. People of working age are considered to be at lower risk for influenza morbidity and its complications than other age groups. Workers may continue to work while ill or return to work before they are fully recovered. Consequently, workplace absenteeism survey data may underestimate the true burden.

1.2.2.1 Methods and Materials

Study design: Prospective, longitudinal, observational study

Study population: All employees working in selected workplaces

Selection of workplaces: When possible, workplaces should be randomly selected from all workplaces within a given region.

Data to be collected:
- Number of all-cause absentee episodes
- Number of ILI-associated absentee episodes
- Number of absentees episodes with laboratory-confirmed influenza
- Costs incurred due to medical assessment/intervention

Survey procedure:
1. Selected workplaces keep daily attendance books throughout the survey year. The number of absentees is recorded and reported to the study coordinating centre every day (Annex 6).
Section 1: Studies Directly Measuring The Disease Burden of Influenza

Employees absent from work provide a completed reporting form (Annex 7) at the time of return. The reporting form details symptoms of influenza-like illness (ILI), such as fever, cough or sore throat as well as any visits to a medical provider. ILI-associated absences are recorded and tallied.

2. If the study has sufficient financial resources and access to a qualified diagnostic laboratory, workers with an ILI-associated absence may be tested for the presence of an influenza virus infection. In this case, a health care worker from the study will visit the worker in their home to collect clinical swab specimens for laboratory diagnosis. If available, a rapid influenza test may be used to diagnose influenza infection.

3. Workplaces should calculate employees’ medical costs associated with these illnesses.

1.2.2.2 Data Storage and Analysis

An efficient system for storage and management of data should be established prior to the commencement of the survey. Weekly tallies of the number of ILI-associated absences can be used as an indicator of increased influenza activity.

Data analysis:
1. Calculate the proportion of absentees by:

\[
\frac{\text{Total number of employees absent at the workplace during the week}}{\text{Total number of employees attending the workplace on the first day of the week}}
\]

If an employee is absent across more than one week, he/she contributes to the numerator for both weeks.

2. Attempt to identify the commencement of the influenza season by increases in absenteeism.

3. Estimate influenza burden by comparing the proportion of absentees during the influenza season with those during the non-influenza season (excess influenza absenteeism).

4. Estimate the relative influenza burden by calculating the proportion of absenteeism due to influenza, based on results of influenza diagnostic tests.

5. Calculate excess medical expenses due to influenza by comparing absentees’ medical expenses with non-absentees’ medical expenses during influenza season. It is important to recognize that the proportion of people with underlying disease may be larger among absentees than among non-absentees. Thus, excess medical expenses due to influenza can also be estimated by comparing medical costs of absentees during the influenza season with that during the non-influenza season.

1.2.2.3 Monitoring

An ongoing comparison of ILI-associated workplace absenteeism with the reported incidence of influenza from the national surveillance system will allow an assessment of the reliability of the survey results.
1.2.2.4 Ethical Considerations

This type of study will require a formal protocol and clearance from the relevant Institutional Review Board(s). The coordinators of the survey must protect the privacy of individuals when reporting and publishing data. Before commencing such a survey, the coordinators must thoroughly explain the purpose, protocols, and material of the survey to the business administration and to the workers. Written consent should be obtained, and signed and dated by both all workers. The titles, names, and contact information of the signatories should be clearly provided. The original copies of the documents are filed at the study coordinating centre.
Section 2: Studies Indirectly Measuring The Disease Burden Of Influenza

2.1. Estimating the Disease Burden through Mathematical Modelling

2.1.1 Introduction

Influenza has a long history with documented epidemics and pandemics of disease in humans dating back to more than a century ago. In the last century there have been three pandemics, the most severe of which was the Spanish influenza pandemic in 1918 which led to more than 40-50 million deaths worldwide (17). After 1918, there were two other milder influenza pandemics which spread globally in 1957 and 1968, however still with significant disease burden for mortality and morbidity (17).

During the inter-pandemic periods, seasonal influenza epidemics have long been recognized as major sources of disease burden in temperate climates (18-21). But in the tropics and subtropics, the health impacts of seasonal influenza have not been recognized. The major problems in the assessment outside the temperate climates lie in the lack of a well-defined seasonality of influenza outbreaks since seasonality has been a necessary tool in such assessment. Following a better understanding of seasonal occurrence in the tropics and subtropics and with the establishment of new approaches in the assessments, influenza has been also found associated with mortality and morbidity with the impacts similar to those identified in the temperature climates (22-25). However the disease burden assessments in the tropics and subtropics have only been carried out in some well developed countries or cities and the estimates are far from representative and complete for the tropical and subtropical regions.

In this document we outline the essential features of several state-of-the-art approaches including risk difference method (details of the statistical model shown in Annex 8); Serfling...
regression method (Annex 9); robust regression method (Annex 9); ARIMA method (Annex 10); stochastic frontier estimation method (Annex 11); and Poisson regression method (Annexes 12 - 13). Several of these methods can be applied in countries with both temperate and tropical/subtropical climates, irrespective of the seasonality patterns.

Based on these approaches we propose a generic statistical approach in quantifying the health burden of seasonal influenza for member countries, and list out the epidemiological and viral surveillance data which are normally required for such approaches.

2.1.2 Data Preparations

2.1.2.1 Health outcome data

The time series of the health outcomes are either weekly (more preferable) or monthly counts of all-cause deaths, or deaths from specific underlying causes (e.g. due to cardio-respiratory diseases, and pneumonia and influenza), or the weekly or monthly counts of hospital discharges with the main diagnoses of a certain underlying causes (e.g. due to cardio-respiratory diseases, and pneumonia and influenza). The specific causes are categorized using either the International Classification of Disease, 9th Revision (ICD-9) or the International Classification of Disease, 10th Revision (ICD-10) as follows:

<table>
<thead>
<tr>
<th>Underlying cause of death/hospitalization</th>
<th>ICD-9</th>
<th>ICD-10</th>
</tr>
</thead>
<tbody>
<tr>
<td>All-cause</td>
<td>All ICD codes</td>
<td>All ICD codes</td>
</tr>
<tr>
<td>Cardio-respiratory</td>
<td>390-519</td>
<td>I00-I99, J00-J99</td>
</tr>
<tr>
<td>Pneumonia and influenza</td>
<td>480-487</td>
<td>J10–J18</td>
</tr>
</tbody>
</table>

The sources of the mortality data are most likely derived from the national death registration systems for most member countries. There are usually less issues on completeness and comparability of the mortality data among member countries.

However, for hospital data, representative and nation-wide data are usually not available or would be problematic. In some member countries only government operated national registers are accessible but there may be substantial hospitalization undertaking in private healthcare institutions. In addition, the diagnosis may not be completely specified and coded for every hospital discharge.

The analysis is based on all ages. But if age information is available from the death/hospitalization records, it would be desirable also to perform the analysis for stratified sub-group who are at high risk for influenza infection, such as the 65 years of age or older group.

2.1.2.2 Influenza virus laboratory surveillance data

The best set of influenza virus surveillance data would be obtained from national influenza centres or national epidemiology surveillance systems. It is because specimens from these centres or systems are usually better representative and good coverage of the countries or cities, and are consistently collected and tested for influenza viruses over time. The data aggregated into weekly
or monthly numbers of total respiratory specimens tested for influenza virus, and numbers tested positive for influenza virus isolates are used to measure influenza activities in defining epidemic and non-epidemic periods as well as proportion of positive results (influenza A or B types and subtypes) in specimens tested for influenza viruses.

### 2.1.2.3 Data on other covariates

Time series data on weekly or monthly mean temperature (°C) and relative humidity (%) are used to control for the potential confounding effects of weather. Information of unusual events (e.g. heat or cold wave, strike in the health services, flood, earthquake, infectious disease outbreak e.g. SARS, etc) will be collected whenever available. But for the control of other confounding factors such as seasonality and long-term time trends, and calendar and holiday effects, no special collection of the data is necessary. Since the respiratory syncytial virus (RSV) is known to co-circulate with influenza viruses and affect morbidity and mortality among both young children and the elderly (26-29). RSV data in numbers and proportion of specimens tested positive for RSV will also be collected and included in the statistical models.

### 2.1.3 Statistical methods in disease burden of influenza estimation

The quantification for the disease burden of seasonal influenza has long been regarded as an important issue but has not been fully developed particular for the tropical and subtropical regions. There are many approaches to disease burden modelling for influenza. The first method is the “risk difference method” which relies on definition of epidemic and non-epidemic (assumed to have low or no influenza viruses in the circulation as a baseline or control) periods so that the difference in observed mortality and morbidity between the two periods can be assessed to estimate the excess impacts (27, 30). The second approach is the “regression method”, including the Serfling method (20) and several adaptations of the Serfling method (31-34). The third approach is the robust regression method (35). The fourth approach is the classical ARIMA method (18, 36). The fifth approach is the stochastic frontier estimation method (37). Finally the sixth approach is the Poisson regression method (23-24, 28, 38-39).

The regression methods, robust regression method, ARIMA method, and stochastic frontier estimation method require well-defined seasonal patterns of influenza activities. However, in order to overcome the problems arising from a lack of well-defined seasonality patterns in the tropics and subtropics, several alternative approaches have been identified. One approach is the simple risk difference method which defines periods based on available influenza viral surveillance data (27, 40). Another approach is the Poisson regression model using viral surveillance data as the predictors in the model and also including the seasonal terms in the model. In the temperate climates, influenza activities usually follow well-defined seasonal patterns so that the epidemic periods with excess mortality and morbidity outcomes can be recognized. But in the tropics and subtropics, estimation approaches become difficult for the following reasons:

i) There are no well-defined seasonal influenza patterns to allow estimation of the epidemic and non-epidemic baseline periods because the influenza virus can be circulating year round.

ii) A surveillance system with laboratory defined influenza activities data is either not available or not complete.

iii) Population level mortality and morbidity databases are not available or not complete.
2.1.3.1 Summary for different approaches and data requirement for disease burden estimations

The data requirement and the statistical methods for each of the approaches 1–6 are summarized in the following table.

<table>
<thead>
<tr>
<th>Approach</th>
<th>Data requirement</th>
<th>Statistical methods</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Risk Difference Method</td>
<td>1) Weekly or monthly morbidity or mortality data 2) Weekly or monthly viral surveillance data to define epidemic periods</td>
<td>Defined epidemic/ nonepидemic periods.</td>
</tr>
<tr>
<td>2. Serfling Regression Method</td>
<td>1) Weekly or monthly morbidity or mortality data 2) Clear seasonal component</td>
<td>Serfling’s method seasonal and expected occurrence during baseline period and to identify the threshold for influenza epidemic.</td>
</tr>
<tr>
<td>3. Robust Regression Method</td>
<td>1) Weekly or monthly morbidity or mortality data 2) Clear seasonal component</td>
<td>Robust regression models which do not require Gaussian distribution and equal variance assumptions, and are robust to the presence of outliers in the health outcomes.</td>
</tr>
<tr>
<td>4. ARIMA Method</td>
<td>1) Weekly or monthly morbidity or mortality data 2) Clear seasonal component</td>
<td>ARIMA models which are liable to over-fitting the health outcomes.</td>
</tr>
<tr>
<td>5. Stochastic Frontier Estimation Method</td>
<td>1) Weekly or monthly morbidity or mortality data 2) Clear seasonal component</td>
<td>Stochastic frontier estimation model to identify excess health outcomes during periods with unexpected high level of health outcomes (influenza data needed to validate the periods).</td>
</tr>
<tr>
<td>6. Poisson Regression Method</td>
<td>1) Weekly or monthly morbidity or mortality data 2) Weekly or monthly viral surveillance data</td>
<td>Poisson regression to model the time series patterns of the health outcomes due to seasonality and time varying confounding. Then the numbers of health outcomes when influenza activities are assumed to be zero (E_0). The excess number = (\text{observed number} - E_0).</td>
</tr>
</tbody>
</table>

The above Table 1 illustrates the statistical models and the underlying statistical methods available to us for estimation of disease burden of influenza. In Table 1 the data required in carrying out the estimation using different approaches are also specified. However in order to guide in the selection of approach, a flow chart (Figure 1) is also provided. It is important to know whether the required data for the estimation are available. If we have weekly or monthly data for the health
outcomes and for the virology activity measures, we can apply Poisson regression (Approach 6). If we only have virology data (to measure influenza activities) and do not have long series of the health outcome data, we can define influenza epidemics and baseline periods in each year and then assess the rate difference between the two periods in each year, providing that we are accessible to the disease incidence rates. But if we have long series of the health outcome data but not virology data, we can apply either the Robust Regression model or the Stochastic Frontier Estimation model, depending on whether the data are from baseline years or not. If we know which are the baseline years, we apply Robust Regression model or if not we apply the Stochastic Frontier Estimation model.

Other approaches based on population surveys or based on intensive case finding for influenza diagnoses in health care institution can be applied. But this is beyond the scope of this guideline. Also to be stringent, validation of the Robust Regression model and Stochastic Frontier Estimation model could be performed first before their recommendation for application in tropical and subtropical regions. Finally disease burden may vary from place to place depending on the susceptibility of the population, and also from year to year depending on emerging of new or recurring strains. Disease burden due to seasonal influenza in the tropics and subtropics is largely unexplored, which warrants urgent investigation in order to raise the awareness and also preparedness of the public, the health professions and the vaccine production capacities during the inter-pandemic periods.

Figure 1: Flow chart of statistical approaches to be chosen for estimating disease burden of influenza

![Flow chart of statistical approaches](image-url)
Section 2: Studies Indirectly Measuring The Disease Burden Of Influenza
References

1. WHO National Influenza Centres.  

2. WHO Global Influenza Programme. Terms of Reference for National Influenza Centres.  

3. Ryan J, et al., Establishing the health and economic impact of influenza vaccination within the  


5. O'Brien M, et al., Incidence of outpatient visits and hospitalizations related to influenza in  


7. Fredrickson K., et al., Influenza vaccination coverage among children aged 6-23 months- six  
   immunization information system sentinel sites, United States, 2005-06 influenza season.  


9. WHO Global Agenda on Influenza Surveillance and Control. Recommended Surveillance  
   Standards  
   http://www.who.int/csr/resources/publications/surveillance/WHO_CDS_CSR_ISR_99_2_EN/  
   en/.

10. WHO Asia Pacific Strategy for Emerging Diseases – South-East Asia Region, Western Pacific  
    Region, 2005.  
    http://www.wpro.who.int/sites/csr/documents/APSED.htm

11. WHO guidelines for collecting, preserving and shipping specimens for the diagnosis of avian  
    influenza A(H5N1) virus infection  
    http://www.who.int/csr/resources/publications/surveillance/CDS_EPR_ARO_2006_1.pdf


ANNEX 1

Informal Consultation on Protocols of Influenza Disease Burden Studies
Manila, Philippines
12-13 November, 2007

Temporary Advisors

Professor Robert BOOY
National Centre for Immunization Research and Surveillance
The University of Sydney and The Children’s Hospital at Westmead
Kerry Packer Institute for Child Health Research
Locked Bag 4001
Westmead, NSW 2145, Australia

Professor Paul CHAN
Department of Microbiology
The Chinese University of Hong Kong
1/F Clinical Science Building
Prince of Wales Hospital
Shatin, New Territories
Hong Kong

Dr Yoshio HIROTA
Professor and Chairman
Department of Public Health
Osaka City University
Graduate School of Medicine
Osaka, 545 8585 Japan

Dr Lance JENNINGS
Clinical Virologist
Canterbury Health Laboratories
P.O. Box 151
Christchurch, New Zealand

Dr Stefan MA
Acting Deputy Director
(Biostatistics and Research)
Epidemiology and Disease Control Division
Ministry of Health, Singapore

Dr Satoko OHFUJI
Research Associate
Department of Public Health
Osaka City University
Graduate School of Medicine
Osaka, Japan
Dr Yasushi OHKUSA
Infectious Disease Surveillance Center
National Institute of Infectious Diseases
1-23-1 Toyama Shinjuku
Tokyo, Japan

Dr David SHAY
Epidemiologist and Team Leader
Epidemiology and surveillance Branch
Influenza Division
Centers for Disease Control and Prevention
1600 Clifton Road
Atlanta, GA 30333, USA

Dr Mark SIMMERMAN
Epidemiologist, Influenza Division
Coordinating Centre for Infectious Diseases
Centers for Disease Control and Prevention
Bangkok, Thailand
msimmerman@cdc.gov

Dr Sirenda VONG
Chief, Epidemiology and Public Health Unit
Pasteur Institute
P.O. Box 983, 5 Bldg Monivong
Phnom Penh, Cambodia

Dr Chit-Ming WONG
Associate Professor
The University of Hong Kong
Department of Community Medicine
5/F William MW Mong Block
21 Sassoon Road
Pokfulam, Hong Kong

Dr Takeshi KASAI
Regional Adviser
Communicable Disease Surveillance and Response
World Health Organization
Regional Office for the Western Pacific
P.O. Box 2932
1000 Manila, Philippines

Dr Weigong ZHOU
Medical Officer
Communicable Disease Surveillance and Response
World Health Organization
Regional Office for the Western Pacific
P.O. Box 2932
1000 Manila, Philippines
ANNEX 2

Procedures for specimen collection

Materials required
- Sputum/mucus trap
- Polyester fibre-tipped applicator (Swabs with a wooden shaft should not be used as they may carry inhibitors for PCR and cell culture)
- Plastic vials
- Tongue depressor
- 15mL conical centrifuge tubes
- Specimen collection cup or Petri dishes
- Transfer pipettes
- Virus transport medium

Virus transportation medium for use in collecting throat and nasal swabs
1. Add 10g veal infusion broth and 2g bovine albumin fraction V to sterile distilled water (to make a total volume 400mL)
2. Add 0.8mL gentamicin sulfate solution (50mg/ml) and 3.2mL amphotericin B (250 μg/ml)
3. Sterilize by filtration

Methods of collection

Clinical specimens should be collected as described below and added to transport medium. Nasal or nasopharyngeal swabs can be combined in the same vial of virus transport medium. The following information should be recorded on the Field Data Collection Form: general patient information, type of specimen(s), date of collection, contact information of person completing the form.

Standard precautions including barrier protection should always be followed during sampling.

Nasal swab
A dry polyester swab is inserted into the nostril, parallel to the palate, and left in place for a few seconds. It is then slowly withdrawn with a rotating motion. Specimens from both nostrils are obtained with the same swab. The tip of the swab is put into a plastic vial containing 2-3mL virus transport medium and the applicator stick is broken off.

Nasopharyngeal swab
A flexible, fine-shafted polyester swab is inserted via the nostril into the nasopharynx and left in place for a few seconds. It is then slowly withdrawn with a rotating motion. A second swab should be used for the second nostril. The tip of the swab is put into a vial containing 2-3mL virus transport medium and the shaft cut.

Nasopharyngeal aspirate
Nasopharyngeal secretions are aspirated through a catheter connected to a mucus trap and fitted to a vacuum source. The catheter is inserted into the nostril parallel to the palate. The vacuum is applied and the catheter is slowly withdrawn with a rotating motion. Mucus from the second nostril is collected with the same catheter in a similar manner. After mucus has been collected from both nostrils, the catheter is flushed with 3mL virus transport medium.
Nasal wash
The patient sits in a comfortable position with the head slightly tilted backward and is advised to keep the pharynx closed by saying “K” while the washing fluid (usually 0.9% sterile saline) is applied into the nostril. With a transfer pipette, 1-1.5mL washing fluid is instilled into one nostril at a time. The patient then tilts the head forward and lets the washing fluid flow into a specimen cup or a Petri dish. The process is repeated with alternate nostrils until a total of 10-15mL washing fluid has been used. Dilute approximately 3mL washing fluid 1:2 in virus transport medium.

Throat swab
The posterior pharynx is swabbed vigorously, and the swab is placed in virus transport medium as described for nasal swabs.
Annex 3:

Key Clinical Data and Investigational Findings

Age
Gender
Admission from home / long-term care facility

Co-existing illness:
- Congestive heart failure
- Chronic obstructive airway disease
- Bronchiectasis
- Asthma
- Pneumoconiosis
- Old pulmonary tuberculosis
- Cerebrovascular disease
- Neoplastic disease
- Renal disease
- Liver disease

Presenting symptoms:
- Fever
- Cough
- Presence of sputum (clear / yellow / blood stained)
- Sore throat
- Coryza

Clinical parameters:
- Blood pressure
- Pulse rate
- Respiratory rate
- Percutaneous oxygen saturation (SaO$_2$ %)
- Temperature (tympanic / oral)
- Changes in mental status or other evidence of neurological complications

Laboratory results
- Complete blood count
- Arterial blood gases
- Blood glucose
- Electrolytes
- Urea

Chest radiographic findings (lobe involvement / effusion)
Requirement for Intensive Care
Need for artificial ventilation, number of days requiring artificial ventilation
Length of stay in hospital
Annex 4:

Health Care Center Recording Form (School Absenteeism)

Date of Information : ___/___/_______ (dd/mm/yy)

Name of School:

School Close : Yes _____  No _____

<table>
<thead>
<tr>
<th></th>
<th>Grade 1</th>
<th>Grade 2</th>
<th>Grade 3</th>
<th>Grade 4</th>
<th>Grade 5</th>
<th>Grade 6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of absentees</td>
<td>Male</td>
<td>Female</td>
<td>Male</td>
<td>Female</td>
<td>Male</td>
<td>Female</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of absentees from febrile disease</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of laboratory confirmed influenza</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of AH1 influenza</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of AH3 influenza</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of B influenza</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Annex 5:

Student Absentee Reporting Form

Name of school:

Grade :

Name :

Periods of absenteeism: from __/__/__ (dd/mm/yy) to __/__/__ (dd/mm/yy)

Consultation in a clinic (yes • no)

Symptoms during the period of absenteeism.

1. Max temperature : degrees

2. Nasal discharge ; (presence • absence)

3. Cough : (presence • absence)

4. Sputum : (presence • absence)

5. Sore throat : (presence • absence)

6. Joint pain : (presence • absence)

<table>
<thead>
<tr>
<th>Diagnosis</th>
<th>Influenza : others ( )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rapid diagnostic test</td>
<td>done ( A • B • negative ) • not done</td>
</tr>
<tr>
<td>Pneumonia</td>
<td>presence • absence</td>
</tr>
<tr>
<td>Encephalopathy</td>
<td>presence • absence</td>
</tr>
<tr>
<td>Otitis Media</td>
<td>presence • absence</td>
</tr>
</tbody>
</table>
Annex 6:

Health Care Center Recording Form (Workplace Absenteeism)

Date of Information : ___/___/_______ (dd/mm/yy)

Name of work place :

<table>
<thead>
<tr>
<th>Age</th>
<th>&lt; 20 years old</th>
<th>20 - 29</th>
<th>30 - 39</th>
<th>40 - 49</th>
<th>50 - 64</th>
<th>65+</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sex</td>
<td>Male</td>
<td>Female</td>
<td>Male</td>
<td>Female</td>
<td>Male</td>
<td>Female</td>
</tr>
<tr>
<td>Number of absentees</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Annex 7:

Employee Absentee Reporting Form

Name of workplace : 

Unit / Department : 

Name : 

Periods of absenteeism : from __/__/__ (dd/mm/yy) to __/__/__ (dd/mm/yy)

Consultation in a clinic ( yes • no )

Symptoms during the period of absenteeism.

1. Max temperature : degrees
2. Nasal discharge : ( presence • absence )
3. Cough : ( presence • absence )
4. Sputum : ( presence • absence )
5. Sore throat : ( presence • absence )
6. Joint pain : ( presence • absence )

<table>
<thead>
<tr>
<th>Diagnosis</th>
<th>Influenza : others ( )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rapid diagnostic test</td>
<td>done ( A • B • negative ) • not done</td>
</tr>
<tr>
<td>Pneumonia</td>
<td>presence • absence</td>
</tr>
</tbody>
</table>
Annex 8: 

Approach 1 - Risk Difference Method

Risk Difference method which is simple and require fewer assumptions, has been used for many years to estimate influenza-associated morbidity and mortality. In this model, the average numbers of deaths or hospital admissions during the months assumed to have low or no influenza virus circulation (baseline periods) are defined, followed by calculation of the excess mortality or hospitalization by subtracting this the numbers during the baseline periods from the observed numbers of deaths or hospital admissions during influenza epidemics. These methods often make use of viral surveillance data, but only to establish periods of influenza epidemic or baseline periods.

In one innovative application of this method in the subtropical region (Hong Kong) has been adopted by Chiu et al (26), in estimation for the burden of influenza on hospitalizations among children 15 years of age or younger. In her method she defined influenza predominance periods to be the influenza epidemic periods in which RSV activities are low, and special baseline to be periods which both influenza and RSV activities are low. She then demonstrated that there was a significant excess rate of hospitalization for acute respiratory disease during the weeks when influenza was predominant, as compared with weeks when neither influenza nor RSV had substantial activities.
Annex 9:

Approach 2 and 3 - Serfling’s Method and Robust Regression Method

The Serfling’s method uses cyclic regression to model the weekly proportion of deaths from pneumonia and influenza and to define an epidemic threshold that is adjusted for seasonal effects. The cyclic regression equation can be written as:

\[ Y_t = \alpha + \gamma T + \beta_1 \sin\left(\frac{2\pi t}{52}\right) + \beta_2 \cos\left(\frac{2\pi t}{52}\right) + \varepsilon_t \]

where \( Y_t \) is the number of deaths from pneumonia and influenza (or some other health outcomes) in week \( t \), during the period when there is no epidemic. The parameter \( \alpha \) is the constant term (representing the overall weekly number of deaths), and \( \varepsilon_t \) is an error term which is assumed to be normally distributed with a mean 0 and variance \( \sigma^2 \). The component \( \gamma T \) specifies the secular trend, and \( \beta_1 \sin\left(\frac{2\pi t}{52}\right) + \beta_2 \cos\left(\frac{2\pi t}{52}\right) \) the sinusoidal trend for the model of the annual recurrence of influenza epidemics. The ordinary least squares (OLS) or maximum likelihood estimation method can be used to estimate the model parameters \( \alpha, \gamma, \beta_1, \beta_2, \sigma^2 \) (in data for non-epidemic periods), and to compute confidence bounds about the predicted values. The predicted values are given by

\[ \hat{y}_t = \hat{\alpha} + \hat{\gamma} t + \hat{\beta}_1 \sin\left(\frac{2\pi t}{52}\right) + \hat{\beta}_2 \cos\left(\frac{2\pi t}{52}\right) \]

The confidence bounds are then computed as

\[ \hat{y}_t \pm t_{\alpha/2} SE(\hat{y}_t) \]

where \( SE(\hat{y}_t) \) is the estimated standard error of the prediction, and \( t_{\alpha/2} \) is the \((1-\alpha/2)\) percentile of a Student’s \( t \) distribution. The confidence bounds are used to define a time varying epidemic threshold that is adjusted for trend and seasonal effects.

As an example, the following figure described the epidemic threshold computed by using Serfling’s method on the proportion of death for pneumonia and influenza record in the US between 1999 to 2002.

![Epidemic threshold graph](image_url)
The method is well implemented in temperate countries where there are well-established and clear seasonal patterns of influenza, however it is not the case in sub-tropic and tropic countries where the seasonality of influenza is not that well defined.

Nevertheless, this model has been extended by US CDC to a robust estimation method (namely robust regression method) using iteration of weighted residuals as follows: at first, it applies

\[ f(Y_t) = \alpha_0^0 + \alpha_1^0 + \alpha_2^0 \sin\left(\frac{2\pi t}{52}\right) + \beta_1^0 \cos\left(\frac{2\pi t}{52}\right) + \epsilon_t^0 \]  

(1)

for all data by OLS where \( f(.) \) stands for some transformation if as necessary. By using the estimated residual \( \hat{\epsilon}_t^0 \), the weighted function in the next step is defined as

\[ W_t^i = \begin{cases} \frac{\sin \hat{\epsilon}_t^0 / (c\sigma)}{\hat{\epsilon}_t^0 / (c\sigma)} & f\left|\hat{\epsilon}_t^0 / (c\sigma)\right| < \pi c \\ 0 & \text{otherwise} \end{cases} \]  

(2)

where \( c \) is some positive constant. The second step and following step are weighted regression for

\[ f(Y_t) = \alpha_0^i + \alpha_1^i + \alpha_2^i T^2 + \beta_1^i \sin\left(\frac{2\pi t}{52}\right) + \beta_2^i \cos\left(\frac{2\pi t}{52}\right) + \epsilon_t^i \]  

(3)

using weights \( W_t^i \). The weight for the next step \( i+1 \) is

\[ W_{t+1}^{i+1} = \begin{cases} \frac{\sin \hat{\epsilon}_t^{i+1} / (c\sigma)}{\hat{\epsilon}_t^{i+1} / (c\sigma)} & f\left|\hat{\epsilon}_t^{i+1} / (c\sigma)\right| < \pi c \\ 0 & \text{otherwise} \end{cases} \]  

(4)

where \( \hat{\epsilon}_t^{i+1} \) is the estimated residual in the \( i \)th step. Then we continue this process until convergence. The procedure is known as andrews weighting (see reference 35 for detail).

\( \beta_1 \sin\left(\frac{2\pi t}{52}\right) + \beta_1 \cos\left(\frac{2\pi t}{52}\right) \) represents clear regular seasonality, and thus it may be inappropriate for tropical and sub-tropical areas. Additional higher term of trigonometric curve,

\( \sin\left(\frac{4\pi t}{52}\right) + \cos\left(\frac{4\pi t}{52}\right) \) i.e. or some other weekly dummies like those specified under stochastic frontier estimation model as in the Annex 11 below may be necessary.
1.1. Autoregressive Integrated Moving Average (ARIMA) Model:
The ARIMA model is used to forecast using a time series data. A time series is a group of observations indexed by time $t$. For example, weekly deaths from influenza is a time series. The ARIMA model uses past values of a time series to forecast future values. This is the main benefit of using ARIMA as we do not need information on any other covariates. For example, to forecast deaths from influenza, we just need data on weekly deaths from influenza. We do not need any additional data such as viral surveillance data. The main components of an ARIMA model are the autoregressive model, moving average model, and the level of differencing needed to make the time series stationary (to be described in 1.1c).

1.1. a. Autoregressive (AR) Model:
In the autoregressive model, the value of the time series in the present period is a proportion of values of the time series in the past. The simplest AR model is the first order autoregressive process, denoted as AR(1). In AR(1) model, number of deaths in the current period are modelled as a proportion of deaths in the immediately prior time period plus a random error. We fit the model on the weekly deaths data to estimate this proportion. Once we estimate this proportion, the forecasted deaths are a proportion of deaths in the previous time period.

We can also increase the number of past time periods used in the AR model. In the second order autoregressive model (AR(2) model), the number of deaths in the current time period are modelled as a proportion of deaths in the two previous time periods plus a random error. We can estimate these proportions by fitting the model on the data (the proportions can be different for the two time periods). The forecasted deaths, from the AR(2) model, are a sum of proportion of deaths in the two last periods. The order of an AR model corresponds to the number of previous time periods used in the model. For example, an autoregressive model of order $p$, written as AR(p), contains $p$ previous time periods.

1.1. b. Moving Average (MA) Model:
In a moving average model, we base the deaths in the current period on the random error (shock) in the past. In the first order moving average model, denoted as MA(1), number of deaths in the current period are modeled as a proportion of random error in the last period plus the random error in the current period. The forecasted deaths in the MA(1) model are a proportion of the random error in the last period. The order of a moving average model corresponds to the number of previous random errors used in the model. For example, in a moving average of second order, denoted as MA(2), we use random errors in the last two periods to model deaths in the current period. In a MA(2) model, deaths in the current period are modelled as a proportion of random errors in the last two time periods plus a random error in the current period. A moving average model of order $q$, models deaths in the current period as a proportion of random errors in the $q$ previous time periods and the random error in the current period.

1.1. c. Autoregressive Moving Average (ARMA) Model:
The autoregressive moving average model, written as ARMA ($p$, $q$), combines both the autoregressive model and the moving average model, where $p$ is the order of autoregressive
part and \( q \) is the order of moving average part. For example, in a ARMA\((1,1)\) model deaths in the current time period are modelled as a proportion of deaths and random error in the last time period plus a random error in the current period. ARMA model can be used only for time series that are stationary. A time series is stationary if it has no trend and the variance is constant across time. If the time series is not stationary we use the ARIMA model.

1.1. d. Autoregressive Integrated Moving Average (ARIMA) Model:

If the time series is non-stationary, we can take first differences to make it stationary. Differencing a series means subtracting each observation from the subsequent observation. If the series is still not stationary we can difference it once more (take second differences). Once we have a stationary series we can apply the ARMA\((p, q)\) model. The ARIMA model is written as ARIMA\((p, d, q)\), where \( p \) is the order of autoregressive part, \( d \) is the level of differencing needed to make the series stationary, and \( q \) is the order of moving average part.

1.2. Seasonality:

Some time series show a seasonal pattern. For example, a time series of weekly deaths from influenza might show a seasonal pattern, where every year the deaths peak in winter and are at a minimum in summer. There are two ways to model seasonality in the ARIMA process. If the time series shows a strong seasonal pattern, we can difference the series with respect to the seasonal lag. For example, if we use the time series of weekly influenza deaths, we can take a 52 week difference and then apply ARMA model. We can also model seasonality as a multiplicative term in the ARIMA model if the series shows significant correlation at seasonal lags.

1.3. Box and Jenkins’ (1976) (1) four step procedure to forecast using the ARIMA process:

1) Transform the data if necessary to meet the assumption of covariance stationarity. If the series is non stationary, we can difference it to make it stationary. We may also need to take a seasonal difference.

2) Once we have a stationary time series we use the autocorrelation function (ACF), inverse autocorrelation function (IACF), and the partial autocorrelation function (PACF) to identify the time series. Identification of a time series means looking at the sample ACF, IACF, and PACF to make an initial guess about the \( p \) and \( q \) values of the ARMA \((p, q)\) model.

3) Once we have identified the model, we estimate the parameters. Mostly we use a maximum likelihood estimation procedure.

4) After estimating the model, we should also perform diagnostic checking to confirm the appropriateness of the model selected. We should check that the parameters are significant and the errors do not have a pattern. We can either plot the errors to look for a pattern or use a test statistic. The test statistic we used was the Ljung modification of the Box-Pierce \( Q \) statistic (ref Ljung and Box, 1978). The model is a good fit if the errors have no correlation and the \( p \) value of the test statistic is large.

1.4. Calculating Excess Deaths:

The analysis consists of two steps. We first calculated excess deaths for the first two seasons by constructing a Fourier series equation. We used Fourier series for the first two seasons because of two reasons. Firstly, to model a seasonal time series we need at least two seasons of data. Secondly, Choi and Thacker (2) also used a Fourier series to model first two seasons of data.

1.4.1 Fourier Method Steps:

1) To identify epidemic weeks for the first influenza season, we calculated mean and standard deviation for the time series from week 27 of the first season to week 42 of the next season. The epidemic weeks for the first season were defined as any two consecutive weeks with deaths greater than two standard deviations from the mean.
2) **Model fitting:** We used Fourier analysis to produce an equation statistically describing the number of deaths during non-epidemic weeks of the first year of data.

3) **Forecasting:** We then used the Fourier equation to predict deaths during the first influenza epidemic.

4) **Excess Deaths:** The actual deaths during these epidemic weeks were then compared with the forecasted deaths, with the difference between the actual recorded deaths and the predicted deaths being the excess deaths attributed to influenza.

5) The deaths during the epidemic weeks during the first season were replaced with forecasted deaths to predict excess deaths for the second season.

6) **Second season:** We calculated mean and standard deviation of deaths from week 42 of the second season until week 42 of the subsequent season. The epidemic weeks for the second season were defined as any two consecutive weeks with deaths greater than two standard deviations from the mean for week 42 of the second season until week 42 of the third season. The method required the Fourier series equation to be updated/re-calibrated, using data from the first season (which included the forecasted, not actual, deaths during the epidemic period) to the beginning of the epidemic period of the second season. The recalibrated equation was then to be used to forecast deaths for the second season epidemic period. As before, the difference between the actual recorded deaths and the predicted deaths would be the excess deaths attributed to influenza.

1.4.2. **ARIMA Method Steps:**

1) **Fitting the model on non-epidemic data from previous years:** We used the data set from week 27 of 1972 to week 42 of 1974, consisting of actual, non-epidemic deaths and predicted (from the Fourier series equations) epidemic period deaths (for 1972-73 season), as inputs into an ARIMA model.

2) **Checking for goodness of fit:** After fitting the model, we checked for goodness-of-fit by testing for autocorrelation in the residuals. The test statistic we used was the Ljung modification of the Box-Pierce $Q$ statistic (refer to Ljung and Box, 1978). The model is a good fit if the errors have no correlation and the p value of the test statistic is large. We changed the model if the p value of the test statistic was smaller than 5% for many lags.

3) **Forecast forward:** We used ARIMA methodology to forecast the next 52 weeks of deaths.

4) **Epidemic Weeks:** The epidemic weeks during the next season (3rd season) were defined as two or more consecutive weeks when actual mortality was greater than the 95 percent confidence interval of the forecasted deaths.

5) **Excess Deaths:** As before, the difference between the actual recorded deaths and the forecasted deaths during the defined epidemic weeks was the excess deaths attributed to influenza.

6) We then replaced the actual deaths for the epidemic weeks of the 3rd season with the predicted deaths.

7) We repeated the process for each subsequent season. For each season we forecasted deaths (for next 52 weeks) from week 42 onwards using the data for all previous weeks. After predicting deaths for a given year, that year (with predicted deaths) was included in the data set used to predict subsequent years (i.e., the data set used to predict grew). When predicting deaths for each year, we re-estimated the coefficients of the ARIMA equation.

1.5. **Additional Technical Details:**

1.5.1. **Lag Operator:**

The lag operator is a highly useful operator in time series analysis. Sometimes it is easier to write the time series model using the lag operator. Suppose we have a series $x_t$, then the application of lag operator on $x_t$ will generate a new series $y_t$ where the value of $y$ on date $t$ is equal to the value $x$ took on at date $t-1$:

$$y_t = Lx_t \equiv x_{t-1}$$
The lag operator satisfies the following properties:

\[ L(Lx_t) = L(x_{t-1}) = x_{t-2} \]

\[ L^2x_t = L(x_{t-1}) = x_{t-2} \]

\[ L^kx_t = x_{t-k} \]

\[ L(\beta x_t) = \beta Lx_t = \beta x_{t-1}, \text{ where } \beta \text{ is a constant} \]

\[ L(x_t + w_t) = LX_t + Lw_t = x_{t-1} + w_{t-1} \]

Following is an example of the use of lag operator

\[ y_t = (\alpha + \beta L)Lx_t = (\alpha L + \beta L^2)x_t = \alpha x_{t-1} + \beta x_{t-2} \]

1.5.2. Autoregressive Integrated Moving Average (ARIMA) Process:

The ARIMA model combines the moving average and the autoregressive models with a differencing term to remove trend in the data. A moving average model of order \( q \), MA(\( q \)), is written as

\[ X_t = \mu + \varepsilon_t - \theta_1\varepsilon_{t-1} - \theta_2\varepsilon_{t-2} - \cdots - \theta_q\varepsilon_{t-q} \]

We can also write the above MA(\( q \)) model using the lag operator

\[ X_t = \mu + \varepsilon_t - \theta_1L\varepsilon_t - \theta_2L^2\varepsilon_t - \cdots - \theta_qL^q\varepsilon_t \]

or

\[ X_t = \mu + \left( 1 - \sum_{i=1}^{q} \theta_i L^i \right) \varepsilon_t \]

where \( \mu \) is the mean of the series, \( \theta_i \) are the parameters of the moving average part, \( q \) is the number of moving average terms, and \( \varepsilon_t \) is the error term (or random shock) with zero mean, constant variance and \( \varepsilon_t \) are uncorrelated over time. We call it moving average since it is an average of random shocks in the past, where \( \varepsilon_t \) is the random shock in period \( t \), \( \varepsilon_{t-1} \) is the random shock in period \( t-1 \), and so on. Thus the value of series in any time period is expressed as a sum of current random shock and \( q \) previous period random shocks.

The autoregressive model of order \( p \), AR(\( p \)), is written as

\[ X_t = c + \phi_1X_{t-1} + \phi_2X_{t-2} + \cdots + \phi_pX_{t-p} + \varepsilon_t \]

Using the lag operator, we can also write the above equation as

\[ X_t = c + \phi_1LX_t + \phi_2L^2X_t + \cdots + \phi_pL^pX_t + \varepsilon_t \]

or

\[ X_t = c + \left( \sum_{i=1}^{p} \phi_i L^i \right) X_t + \varepsilon_t \]
where \( c \) is the mean of the series, \( \theta_j \) are the parameters of the autoregressive part, and \( p \) is the number of autoregressive terms. In the autoregressive model, past values of the series influence the current value of the series, where \( X \) is the value of series in time period \( t \), \( X_{t-1} \) is the value of series in time period \( t-1 \), and so on. Thus, in the AR\((p)\) model, the value of series in any time period is modelled as a sum of the random shock in the current period and a portion of \( p \) previous values of the series.

An autoregressive moving average model (ARMA) combines both moving average and autoregressive models. An ARMA \((p, q)\) model contains \( p \) autoregressive terms and \( q \) moving average terms and is written as

\[
X_t = \mu + \theta_1 X_{t-1} + \theta_2 X_{t-2} + \ldots + \theta_p X_{t-p} + \epsilon_t - \theta_{p+1} \epsilon_{t-1} - \theta_{p+2} \epsilon_{t-2} - \ldots - \theta_q \epsilon_{t-q}
\]

or in lag operator terms as

\[
(1 - \sum_{i=1}^{q} \theta_i L^i) X_t = \mu + (1 - \sum_{i=1}^{q} \theta_i L^i) \epsilon_t
\]

Thus, in an ARMA \((p, q)\) model, the current value of the series is influenced by the past values of the series and also by the past random shocks.

If the time series data is not stationary, i.e. it shows a trend, we can first take a one period difference of the series to remove the trend before we do the analysis. After differencing we get a new series \( W_t \) such that

\[
W_t = X_t - X_{t-1}
\]

Since the differenced series \( W_t \) is now stationary (no trend), we can use the ARMA\((p, q)\) model on \( W_t \) for forecasting. If the first differenced series \( W_t \) is also not stationary, we can do a second differencing to make the series stationary (we rarely go beyond second differencing) and use ARMA\((p, q)\) model on the stationary \( Z_t \) series

\[
Z_t = W_t - W_{t-1}
\]

Thus if we need to difference a series to make it stationary, we use the ARIMA \((p, d, q)\) model, where \( d \) is level of differencing needed to make it stationary. The ARIMA model is similar to an ARMA model, except that we first difference the series and then apply the ARMA model on the differenced series.

An ARIMA \((p, d, q)\) model can be written as,

\[
(1 - \sum_{i=1}^{q} \theta_i L^i) (1-L)^d X_t = (1 - \sum_{i=1}^{q} \theta_i L^i) \epsilon_t
\]

Where \( L \) is the lag operator, and \( d \) is level of differencing needed to make the series stationary.

### 1.5.3. Seasonality:

Some time series show a seasonal pattern. For example, a time series of weekly deaths from influenza might show a seasonal pattern, where every year the deaths peak in winter and are at a minimum in summer. There are two ways to model seasonality in the ARIMA process. If the time series shows a strong seasonal pattern, we can difference the series with respect to the seasonal lag. For example, if we use the time series of weekly influenza deaths and we find that
the autocorrelations have strong wave like shape with spikes at 52 weeks interval, we can take a 52 week difference and then apply ARMA model. An ARIMA model with 52 week difference can be written as

\[
(1 - \sum_{i=1}^{q} \phi_i L^i)(1-L)^d X_t = \left( 1 - \sum_{i=1}^{q} \phi_i L^i \right) \epsilon_t
\]

Note that in the above ARIMA model we also need a d period differencing to make the time series stationary. We may also need to incorporate seasonality as multiplicative term in the ARIMA model if the autocorrelations of the differenced series have spikes at the seasonal lag. In the following equations we have modelled seasonality as a multiplicative term in the AR and MA parts respectively.

\[
(1 - \sum_{i=1}^{q} \phi_i L^i)(1-L)^d X_t = \left( \sum_{i=1}^{q} \phi_i L^i \right) \epsilon_t
\]

\[
(1 - \sum_{i=1}^{q} \phi_i L^i)(1-L)^d X_t = \left( 1 - \sum_{i=1}^{q} \phi_i L^i \right) \epsilon_t
\]

References;


Annex 11: Approach 5 - Stochastic Frontier Estimation Method

The Stochastic Frontier Estimation method is the official procedure for the estimation of excess mortality in Japan. It applies the stochastic frontier estimation to mortality data. Excess mortality is defined as the difference in the total number of deaths and the hypothetical number of deaths if there is no influenza epidemic. Thus the baseline must indicate the number of deaths if there is no influenza epidemic. This stochastic frontier estimation can represent the baseline which is the minimum number of deaths in the month. The estimation equation for mortality in all causes at year $t$, $Y_t$, is

$$\log Y_t = \alpha + \gamma_1 T + \gamma_2 T^2 + \sum \beta_i M_i + \epsilon_i$$

(5)

where $T$ is the linear time trend and $M_i$ are dummy variables for month. The third degree of polynomial equation is not significant, therefore time trend variables for the third and higher degree polynomial are not included in the model.

A disturbance term is defined as

$$\epsilon_i = v_i + |\omega_i|$$

(6)

where $v_i$ and $\omega_i$ are mutually independent random variables, and $v_i \sim N(0, \mu^2)$ and $\omega_i \sim N(0, \xi^2)$. A random variable $v_i$ is a purely stochastic disturbance and $\omega_i$ represents the magnitude of the epidemic which is a non-negative deviation term from the expected number of deaths. The equation form (5) is selected by statistical inference. The probability density function of this model is

$$f(\epsilon_i) = \frac{2}{\sqrt{\mu^2 + \xi^2}} \frac{1}{\sqrt{\mu^2 + \xi^2}} \Phi\left(\frac{\epsilon_i - \xi}{\mu}\right)$$

(7)

where $\Phi$ is the probability density function for standard normal distribution and $\Phi$ is its cumulative distribution function. The likelihood function is defined as a product of the probability density function over time: $L(\epsilon_1, \epsilon_2, \ldots, \epsilon_t) = \prod f(\epsilon_i)$.

Estimators for $\alpha$, $\gamma_1$, $\gamma_2$, $\beta$s, $\xi / \mu$ and $\sqrt{\mu^2 + \xi^2}$ are chosen so as to maximize the likelihood function. The estimators are therefore maximum likelihood estimators. This model is known in economics as the stochastic frontier estimation. In economics, $\omega_i$ is often defined as inefficiency, and stochastic frontier estimation is used to measure this inefficiency. For example, deviation from the most efficient production is defined as inefficiency in estimation of production function, and deviation from the least cost is defined as inefficiency in estimation of cost function. In this context, we defined $\omega_i$ as the magnitude of the epidemic. In this way we can estimate the excess mortality without making any arbitrary assumptions of what epidemic or non-epidemic weeks are. For example, the referred models arbitrarily define a non-epidemic season and estimate the excess mortality during the epidemic season using this “non-epidemic season” as a baseline. The epidemic threshold is usually defined as an upper limit of 95% confidence bound for stochastic disturbance terms, $v_i$. Therefore, the epidemic threshold should be

$$\hat{\alpha} + \hat{\gamma}_1 T + \hat{\gamma}_2 T^2 + \sum \hat{\beta}_i M_i + 1.96\hat{\mu}$$

(8)
where \* indicates the estimated parameters and \( \mu \) is defined as

\[
\sqrt{\frac{(\mu^2 + \beta^2)^{1/2}}{1 + (\xi/\mu)^2}}
\]

Comparing the stochastic frontier estimation model with the robust regression model, these are similar in the sense of all information other than mortality. However, robust regression model has less or no weight on the lower outlier as well as the upper outlier, which is the epidemic. Conversely, the stochastic frontier estimation model put some weight on the lower outlier because the estimated baseline in the model represents the minimum mortality in the month. In this sense, we may sometimes find a lower outlier in the robust regression model, and it may contradict the definition of baseline, which is the minimum number of deaths in the month if there is no influenza outbreak.

The choice of such a seasonality or function \( f(.) \) may be different from country to country, based on statistical inferences.
Annex 12: Approach 6 - Poisson Regression Method

(a) Poisson regression (assumed weekly data to be used)

$$\log E(Y_t) = \alpha + \beta_1 Z_{1t} + \ldots + \beta_q Z_{qt} + \theta X_t$$  \hspace{1cm} (9)

where for each week index $t \,(t = 1, 2, \ldots, n)$, observation $Y_t$ follows Poisson distribution, $E(Y)$ is the expected numbers of health outcome; $Z_{1t}, Z_{2t}, \ldots, Z_{qt}$ are the covariates for trends, seasonality, temperature, relative humidity and proportion of positive results tested for RSV; $X_t$ is the weekly proportion of positive results tested for influenza; $\beta_1, \beta_2, \ldots, \beta_q$ and $\theta$ are the estimated parameters for respective covariates and $\alpha$ is an intercept of the model.

Note: Here we use dummy variable to define each year and all the year dummy variable in model to control for trend. We also use

$$\sin\left(\frac{2k\pi t}{52}\right) \quad \text{and} \quad \cos\left(\frac{2k\pi t}{52}\right)$$

where $k=1$ and 2 cycles a year, $(t=1, 2, \ldots, n)$ are indices for weeks, to define seasonality and product terms of year-by-seasonality variables to specify seasonality of each of the years. The year-by-seasonality variables are put into the model to control for seasonality with an amplitude specific for each year.

Quasi-likelihood estimation method is used to adjust for over dispersion of weekly death counts. In addition, adjustment for serial correlation may be performed whenever necessary.

The effects of influenza on health may not be appeared at the same week, in order words observed delay effects of influenza could come in effects on following weeks or so. Lags 0 up to 3 week of influenza data could be considered and tested in the model.

The modelling approach of this kind is to model the health outcome which can be explained by covariates as much as possible. The model adequacy should be carefully examined by plots of residuals.

(b) Avoidable fraction ($x$ AF) for elimination of weekly influenza proportion from $x$ down to zero:

$$AF_x = \frac{\sum_{t=1}^{n} (E_{X_t=x} - E_{X_t=0})}{\sum_{t=1}^{n} (E_{X_t=x})}$$  \hspace{1cm} (10)

where $E_{X_t=x}$ and $E_{X_t=0}$ are weekly expected numbers, derived from (9), of health outcome for influenza A+B proportion equals to $x_t$ and 0 respectively.


\[ AF_x = 1 - \frac{\sum_{t=1}^{n} (E_{X_t=0})}{\sum_{t=1}^{n} (E_{X_t=x_t})} \]

\[ = 1 - \frac{\sum_{t=1}^{n} \exp (\alpha + \beta_1 Z_{1t} + \ldots + \beta_q Z_{qt})}{\sum_{t=1}^{n} \exp (\alpha + \beta_1 Z_{1t} + \ldots + \beta_q Z_{qt} + \theta X_t)} \]

\[ = 1 - \frac{1}{\sum_{t=1}^{n} \exp (\alpha + \beta_1 Z_{1t} + \ldots + \beta_q Z_{qt} + \theta X_t)} \]

\[ = 1 - \frac{1}{\sum_{t=1}^{n} \exp (\alpha + \beta_1 Z_{1t} + \ldots + \beta_q Z_{qt}) \exp(\theta X_t)} \]

(11)

Note: For computation we can make use of total expected equals total observed in (10). Therefore, \( AF_x \) is total observed minus total expected when \( X \) is set to zero, and then divided by total observed.
Annex 13: Data Template to be collected for estimation of influenza burden using Approach 6

<table>
<thead>
<tr>
<th>Week / Month, $t$</th>
<th>No. of Deaths</th>
<th>No. of Hospitalizations</th>
<th>No. of positive test results</th>
<th>No. of Specimens tested</th>
<th>No. of positive test results</th>
<th>No. of specimens tested</th>
<th>Mean Temperature (°C)</th>
<th>Mean relative humidity</th>
<th>Unusual event (coded 1 if yes, 0 otherwise)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* Respiratory syncytial virus