Testimony to the National Committee on Vital and Health Statistics

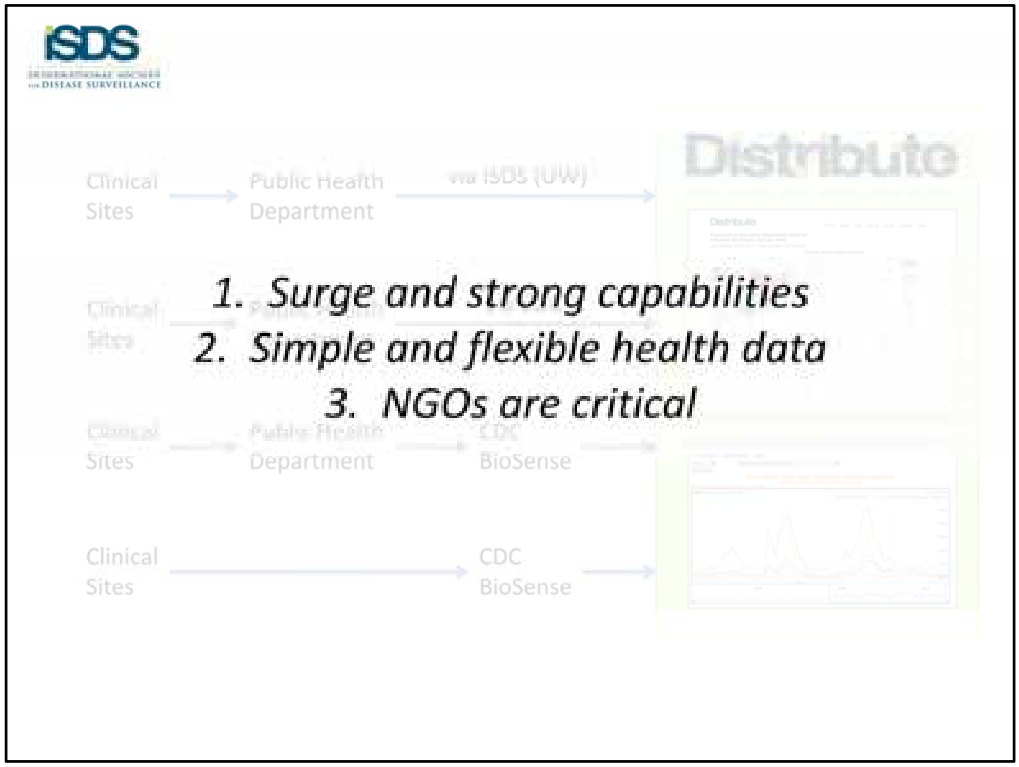
The Current State of Health Data Standards for Syndromic Surveillance Among U.S. Governmental Public Health Agencies
November 12, 2013

Charles K.T. Ishikawa, MSPH
Associate Director of Public Health Programs, ISDS
On behalf of the International Society for Disease Surveillance (ISDS)
ISDS is a national, 501C(3) non-profit organization based in Boston, Massachusetts. We are a society of innovators and early adopters of information technology for better disease surveillance practice and more effective public health action. Our community spans disciplines, market sectors, governmental agencies and geopolitical boundaries [1].

In recent years, we have become the *de facto* organization for developing syndromic surveillance data standards in the United States. The Markle Foundation, CDC, ONC, NIST, our JPHIT partner associations and many others institutions and companies are our valued partners in this work--work that we all do to help local and state public health agencies meet contemporary demands for relevant and timely health information.

To speak about the drivers and facilitators of health data standards for syndromic surveillance, I will share with you ISDS's experience in developing national standards in two areas:
Aggregate health data exchange; and
EHR messaging for meaningful use.

In 2009, we leveraged an aggregate health data standard to establish a national, near-real time influenza-like-illness, or ILI, surveillance system for the H1N1 pandemic response. The system we created was called Distribute [2].

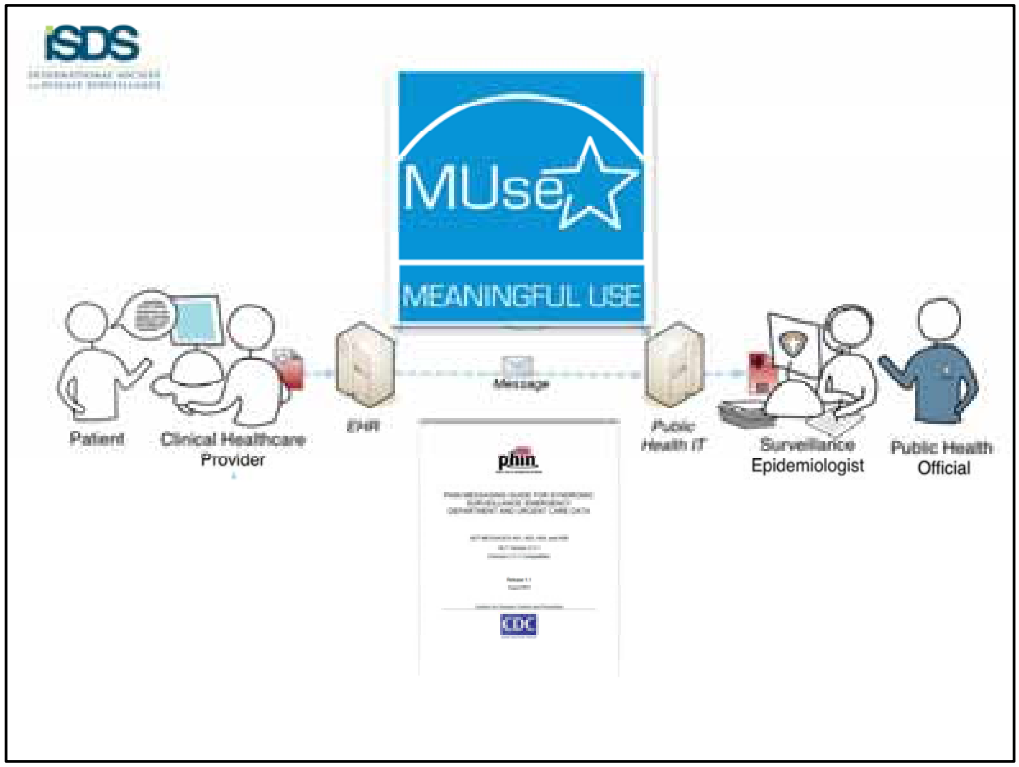


Distribute was a system of 43 local and state jurisdictions sharing their daily counts of ILI emergency department visits with one another and the general public. We worked with the CDC and PHII to build-up this system of systems within three (3) months, and at its peak, Distribute data covered more than 40% of all emergency departments in the United States.

There are three things that we learned about adopting a public health, aggregate health data standard from Distribute.

- First, adoption of an aggregate health data standard requires surge capacity and strong level local and state surveillance system capabilities. Driven by an emergency need for national situational awareness, federal resources enabled us to stand-up a national system a remarkably small amount of time. Although that infusion of surge capacity was critical, it was really a small part of the Distribute infrastructure; local and state syndromic surveillance systems were its foundation. Notably, many of those systems were built and maintained with Public Health Emergency Preparedness and Response grants.
- Second, aggregate health data standards must be simple and flexible for successful adoption. Distribute data were messaged using a comma separated value file sent over the internet via sFTP [3]. Only four data elements were required:
    - Encounter date;
    - Three digit Zip Code
    - Age Group; and
    - A percentage or count of ILI visits using the agency's preferred ILI syndrome definition.

  This simplicity of Distribute's technical requirements provided a low bar of entry for participation. The allowance for site specific ILI case definitions provided the

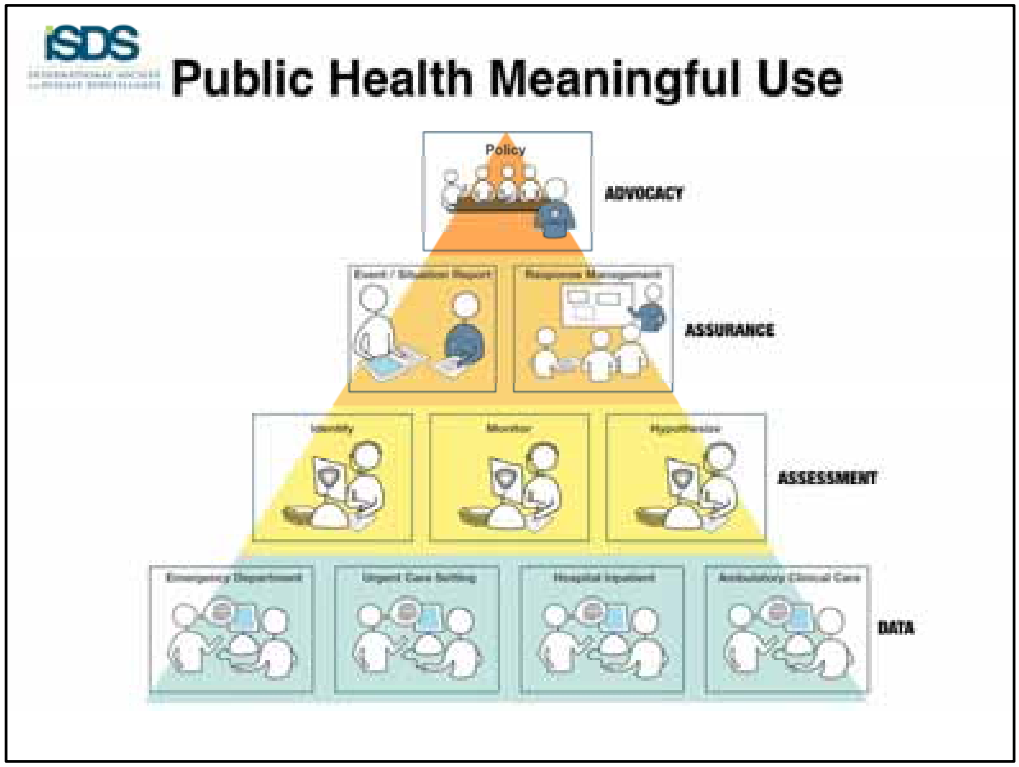


They are also used to help monitor, gauge and assess the population health impacts of extreme cold weather events, hurricanes, terrorist attacks and other natural and man-made disasters. In all these situations, patient encounter data from healthcare settings are critical inputs.

Mainly because they are used for similar purposes, there is some degree of standardization among different syndromic surveillance systems. Generally, health data for syndromic surveillance represent all the patient encounters of a particular type within a treatment facility, not a subset a specific patient health criteria. Examples include: All emergency department registrations, inpatient admissions, or office visits. Most systems collect these data at least once every 24 hours, and often more frequently, to maximize signal sensitivity with an acceptable loss in specificity. And since syndromic analyses, by-in-large, do not require the identification of individuals, many systems do not collect limited personal identifiable information.

Currently, Meaningful Use and federal assists from the CDC are driving greater standardization in how hospitals provide these health data.

The PHIN Messaging Guide for Syndromic Surveillance Emergency Department and Urgent Care data is now the standard to which EHR technology are certified, and hospitals attest for the syndromic surveillance Meaningful Use measure [4]. It is a relatively simple standard. Ubiquitous admission, discharge and transfer, or ADT HL7 messages are used for packaging data that support common and more advanced syndromic applications. It is also a flexible data standard that health agencies can tailor to suit their jurisdictional and system needs. This standard was developed, in 2010, with support from the CDC to ISDS, to ensure that Meaningful Use will maintain the integrity of established syndromic surveillance systems. Now, as Meaningful Use Stage 2 begins and on-going syndromic surveillance data transfer is core for eligible hospitals, BioSense 2.0 is providing added capacity for some public health agencies to receive these data and develop core syndromic surveillance capabilities for the first time [5].



Theoretically, standardizing syndromic surveillance data messaging nationwide is beneficial. With every health department receiving a core set syndromic surveillance data, communities at every level could gain access to more relevant and more local health data. Reaching this dream, however, will take far more than a messaging standard. Those messages must carry quality data, and public health agencies must understand those data and be able to use and share them effectively.

To help minimize problems and ensure implementation success, public health agencies and national associations need targeted funds that support communities of practice for front line workers to rapidly develop and share practical solutions. Syndromic surveillance also needs resources to further develop the PHIN Guide for messaging de-identified laboratory result data, write a Guide for messaging ambulatory clinical care data, and assess the meaningful impacts of health data standards on population health.

This is the state of syndromic surveillance health data standards in the United States. Please be aware that there are also many health-related data that public health agencies collect for syndromic surveillance [6]. Even with those data we generally find that:

1. Data standards are used because they build a critical capability or solve a common problem;
2. Surge capacity and baseline capabilities are necessary for standard development and adoption; and
3. Non-governmental organizations and institutions are critical for successful data standards at a national level.

I would like to leave you with one closing thought: We need to know when NOT to impose standards.

> One of the strengths of this disease surveillance methodology is in its ability to identify and track something that you never thought of--something that you could not have built into a standard case definition. Thus the critical importance of using free-text to collect chief complaints, reasons for admission, and causes of death--not just pre-coded, structured or standardized representations of them. Sometimes a data standard does not help pubic health.