International Society for Disease Surveillance
Technical Conventions Committee

Date: January 21, 2014; 1:00 – 2:00pm ET
Description: Committee Meeting

Notes:

• Review of ISDS Technical Conventions Committee: http://www.syndromic.org/communities/technical-conventions
• Project Tycho – Freely available national reportable disease data (See presentation slides for details)
  ▪ Rationale
    § Make PH data more available for decision making
  ▪ US Surveillance data
    § Digitization
      • US weekly disease reports from 1888-2013
      • Cities: 1888 – 1948; States: 1928 – 2013
    § Sources
      • Archived in PubMed Central (freely available)
      • MMWR
      • Hathi Trust Digital Library
      • University of Pittsburgh Library -> Supplement incomplete reports
  ▪ Distinctions
    • Geography (e.g., city level vs. state level)
    • Morbidity vs. Mortality
  ▪ Visualization
    • Can explore diseases reported over time
      ▪ Number of reportable diseases has increased over time
    • Can visualize impact of vaccination
  ▪ NEJM paper (PDF): http://www.febrilnotropeni.net/newsfiles/3777NEJMms1215400.pdf
  ▪ Levels of data redistribution
    ▪ Level 1: Data used in the NEJM article - various types of counts standardized for 8 diseases, most clean, complete, and standardized
    ▪ Level 2: Data reported in a consistent format (filter applied): ready for analysis, but gaps remain due to incomplete data standardization
    ▪ Level 3: Raw data
  ▪ Future direction:
    ▪ Expand data
    ▪ Classify, normalize and standardize
    ▪ Annotation and metadata
    ▪ Improved visualization and new demonstration analysis
  ▪ Discussion
    ▪ Data:
• Fairly easy to access on http://www.tycho.pitt.edu/
• Interested in working with people, especially on Level 3 data
  o Clean data and fill in gaps
• Is there a certain field using other than weekly? Found gaps in the weekly data (specifically looking at Pertussis)
  o For Pertussis, there were no weekly reports between 1954 ~ 1976
  o Transition from weekly reports to cumulative reports
    ▪ Did not include cumulative reports in Level 2 data
• Over last 20 years, CDC published weekly totals in MMWR
  o CDC has collected case-level data, and so could include additional information in correct week if they wish
  o When comparing cumulative to weekly totals, cumulative may include updates from states
• Have you considered contacting states and asking them to review data and compare with their data in terms of reporting lag and completeness?
  o Have started and are very interested in doing more
    ▪ Contacted CSTE
  o Want to look at how well: County level to state level data match and state level to national level published in MMWR match
  o More difficult for historical data

• Current Use Cases
  o NC DETECT Use Case
    ▪ Close to 200,000 records in data set
    ▪ E-mail Amy if interested in accessing. Essentially,
      • Involves completing online form
      • Will then receive DUA specific to this project from NC Division of PH
      • Complete and return DUA to Division of PH
  o Data Quality Use Case
    ▪ 3 individuals have signed DUAs
    ▪ If others are interested in accessing data, contact Marcus and/or Atar (will need to sign a DUA for each dataset)
  o ED-based surveillance during large-scale events
    ▪ Have worked on defining
      • Considering narrowing the Use Case rather than more widely applicable (in other jurisdictions)
      • Mass events vary widely – difficult to create a use case that will capture the variability
      • Use case specifies data fields that would be required
    ▪ Looking into options for sharing data

• What happens once solution has been developed?
  ▪ Depends on the solutions that are developed, and how easy it is to implement
  ▪ Use case developers are generally willing to share solutions (if solution developer is willing) and make implementation documentation available
  ▪ When it makes sense, ISDS will help in dissemination of solutions and associated documentation
o Publishing articles on solution
  ▪ Primary purpose is to empower PH departments (not publish papers)
  ▪ Many DUAs include language about publishing
  ▪ Want to be clear about expectations regarding publication when working with solution developers
o Different use cases have different data solutions (based on data owners’ individual data sharing restrictions)

• Additional Use Case Ideas
  o Focused anomaly detection
  o Alert investigation
  o Outbreak tracking
  o Multiple data types
    ▪ Antimicrobial drug resistance
    ▪ Severe weather effect surveillance
  o Combining multiple data sources
    ▪ Syndromic counts or ratios
    ▪ Environmental measurements
    ▪ Absenteeism rates
    ▪ Over-the-counter product sales
• Bringing Use Cases to the attention of solution developers
  o Presentations at conferences (upcoming: JSM)
  o Other groups could be reaching – How?
    ▪ Sylvia has helped: Presented at Machine Learning meetings
    ▪ Try to find similar meetings
  o Are there options for funding?

Next steps:

• Make committee aware of the availability of Project Tycho data
• Consider ways to partner with Project Tycho, and possibly CDC
• Schedule a call on ED-based surveillance during large-scale events to further discuss Use Case (Goal: Meeting in 1-3 weeks)
  o Is it worth the time? Will it be applicable to other jurisdictions?