ISDS Public Health Practice Problem Definition

USE- CASE/PROBLEM TITLE

Reportable disease cluster detection in the context of sporadic adoption of PCR-based diagnostic tests

CONTACT INFORMATION

Submitter name: Sharon K. Greene, PhD, MPH

Jurisdiction or affiliation: Bureau of Communicable Disease, New York City Department of Health and Mental Hygiene

Phone: (347) 396-2679

Email: sgreene4@health.nyc.gov

Co-submitters and affiliations: Eric Peterson, MPH; Kristen Lee, MPH; Ana Maria Fireteanu, MPH; Annie Fine, MD

PROBLEM DESCRIPTION

Summarize the problem:

For timely reportable disease cluster detection, we conduct weekly analyses using the refined historical limits method [1] and daily analyses using the prospective space-time permutation scan statistic in SaTScan [2]. Both methods compare counts of recent cases with a historical baseline, accounting for seasonality and secular trends.

The sporadic adoption of PCR-based tests (e.g., the BioFire FilmArray gastrointestinal panel [3]) at different times by different hospital laboratories serving different parts of our catchment area threatens the validity of the comparison of current vs. historical data, biasing toward increased signaling. PCR-based tests are more sensitive than traditional, culture-based methods and are easier to perform, leading to increased case ascertainment from more people being tested, more true positives, and more false positives [4-6]. It is currently challenging to distinguish signals attributable to changing testing practices and improved case ascertainment vs. true excess disease activity.

Would you please provide input on the below proposal to make recent and historical data more comparable, and/or provide alternative suggestions to suppress signals attributable to improved case ascertainment? Separately for each affected disease:

1. Using passive surveillance data, identify laboratories (by Lab CLIA) performing these tests in NYC (a non-trivial problem, as reported test types can be unclear).
2. For each lab performing PCR-based tests separately, plot reported test types over time. Identify a cut-off date for “routine” PCR-based test use (pre/post), e.g., ≥50% of confirmed/probable/suspected cases of a disease reported by a laboratory have a PCR-based test.
    a. “Routine” use is not strictly defined; if, on visual inspection, a time period cannot be clearly classified as either pre- or post-routine use (e.g., due to very recent PCR-based test adoption, or these tests constituting a very small proportion of the overall test types for a disease from a lab), then it will be considered a washout period, with its data excluded from step 3.
3. Estimate the magnitude of increased case ascertainment by performing a Poisson regression (or negative binomial regression if the data are overdispersed). Prepare a dataset with the following columns: month of event date (1–12), year of event date (2013–present), lab indicator, pre/post indicator (0=lab not routinely using PCR-based test, 1=lab routinely using PCR-based test), and number of events (all case statuses except unresolved) of disease X. Include rows with 0 events. Model the number of disease events (outcome), with pre/post indicator as the independent variable of interest, adjusting for month and year, and accounting for multiple observations per lab (in proc genmod, use repeated subject=lab). Exponentiate the parameter estimate for the pre/post indicator to determine the risk ratio (RR) of increased case ascertainment with PCR-based tests, adjusting for seasonality and secular trends.
4. Modify the data input into cluster detection algorithms:
    a. For historical limits method: For cases reported from labs using PCR-based tests, pre-routine adoption, multiply case count in baseline by the RR determined in step #3 (i.e., increase historical counts so more comparable with recent data).
    b. For the prospective space-time permutation scan statistic: shorten the study period (currently 1 year, 1.5 years, or 2 years, disease-depending) to 90 days (to reduce the amount of a time a transitioning lab presents a problem).
        i. if labs DO NOT transition to routine PCR-based testing at the same time: For simplicity, exclude from the case input file all events from labs that transitioned to routine PCR-based testing within the prior 90 days.
        ii. If labs DO transition to routine PCR-based testing at the same time (such that we cannot effectively detect clusters while excluding data from these labs): for cases reported from labs using PCR-based tests, post-routine adoption: include in the case input file with a probability of 1/RR (i.e., decrease recent counts so more comparable with historical data).
5. Maintenance: repeat steps #1-3 on a regular basis (e.g., monthly) to update the RR applied in step #4. Continue maintenance until no labs have newly adopted routine PCR-based test use since the start of the baseline period.

This plan is subject to at least four limitations. First, we assume no outbreaks exist in the baseline. Second, we assume the risk ratio applied in step 4 is known without error. Third, by temporarily excluding (for simplicity) in step 4.b.i. any events from labs that transitioned to routine PCR-based testing within the prior 90 days, we might miss a cluster in an area primarily serviced by that lab. Fourth, by shortening the study period in step 4.b., we might reduce power to detect a cluster, especially for rare diseases, for which longer baselines are preferable; we will monitor performance and might extend the baseline if warranted.

[1] Levin-Rector A, et al. Refining historical limits method to improve disease cluster detection, New York City, New York, USA. Emerg Infect Dis. 2015;21(2):265-72.
[2] Kulldorff M, et al. A space-time permutation scan statistic for disease outbreak detection. PLoS Med. 2005;2(3):e59.
[3] Buss SN, et al. Multicenter evaluation of the BioFire FilmArray gastrointestinal panel for etiologic diagnosis of infectious gastroenteritis. J Clin Microbiol. 2015;53(3):915-25.
[4] Iwamoto M, et al. Bacterial enteric infections detected by culture-independent diagnostic tests--FoodNet, United States, 2012-2014. MMWR Morb Mortal Wkly Rep. 2015;64(9):252-7
[5] Langley G, et al. Effect of culture-independent diagnostic tests on future Emerging Infections Program surveillance. Emerg Infect Dis. 2015; 21(9): 1582–1588.
[6] Cronquist A, et al. Impacts of culture-independent diagnostic practices on public health surveillance for bacterial enteric pathogens. Clin Infect Dis. (2012) 54 (suppl 5): S432-S439.

SOLUTION REQUIREMENTS

Describe the type of solution you are seeking (e.g., anomaly detection, signal validation, data quality characterization):

We are seeking input and suggestions on our proposal to modify input data for anomaly detection.

Describe what type of solution would enable you to implement it in your practice setting (e.g., Do you need an algorithm? Do you need code? If you need code, does it have to be written in any particular programming language?).

We are seeking only guidance at this time, e.g., a conference call convening biostatisticians with expertise in prospective cluster detection.

Describe who will use the solution. For example, how many users will there be and what level of skill do the users have? Are the users all within a single jurisdiction/organization?

We are asking on behalf of the Bureau of Communicable Disease, New York City Department of Health and Mental Hygiene, but health departments in many jurisdictions are currently (or will soon be) facing the same problem. This problem seems analogous to a problem faced by syndromic surveillance: chief complaint (CC) text is changing to a more standardized clinical description. Because CC keyword searches and ICD-10 discharge codes are used to define syndromes, recent changes to the CC/discharge diagnosis fields affect how an ED visit is defined. If the change is substantial, using a longer baseline becomes problematic. Discharge diagnosis is doubly problematic because of the switch to ICD10, which is more specific than ICD9, and reduced missingness with Meaningful Use-compliant EDs.

Note any other constraints:

We prefer the simplest/most practical/valid solution possible. For example, for the daily SaTScan analysis, we would prefer to continue to run the analysis once daily for each disease (vs. multiple times daily using different versions of input data and having to reconcile results across runs).

VALIDATION

Does a gold standard exist with which to validate the proposed solutions?

- ☐ Gold standard exists within the provided data set (e.g., an outbreak signal nested within baseline data)
- ☐ Gold standard exists in a separate data set, which can be provided to the workgroup (e.g., laboratory data to validate ED data)
- ☐ Gold standard exists but cannot be furnished
- ☒ Gold standard does not exist

INPUT DATA

List the minimum data elements that can be provided to address the problem:

N/A (thought experiment)

How much historical data can be provided?

N/A

Describe any restrictions for sharing the data:

N/A

Note any other relevant data characteristics:
For a description of using our reportable disease data for cluster detection, please see: Levin-Rector A, et al. Refining historical limits method to improve disease cluster detection, New York City, New York, USA. Emerg Infect Dis. 2015;21(2):265-72.

Disease-specific considerations:

Disease	If the only test result is a positive PCR-based test from a hospital or commercial lab, does the case meet the CDC/CSTE case definition?	If the test results are a positive PCR-based test from a hospital or commercial lab but a negative confirmatory test from a city or state public health lab, does the case meet the CDC/CSTE case definition?*	Comments
Amebiasis	Yes (case status=suspected)	No. "Not a case"	These are NYC-specific classifications, since not nationally notifiable since 1994.
Campylobacteriosis	Yes (case status=probable)	No. No effect on case definition, remains "Probable"	
Cryptosporidiosis	Yes (case status=confirmed)		Increased case ascertainment a substantial problem.
Cyclosporiasis	Yes (case status=confirmed)		
Giardiasis	Yes (case status=confirmed)		
Salmonellosis	Yes (case status=suspected)	No. No effect on case definition, remains "Suspected"	The Bureau of Communicable Disease performs *Salmonella* cluster detection only at the level of serotype, which requires a culture. So, we are not trying to determine how to incorporate PCR-based *Salmonella* results in cluster detection. Peterson ER, et al. Prospective spatio-temporal and temporal cluster detection by *Salmonella* serotype (oral presentation). In: Abstracts of the 14th Annual International Society for Disease Surveillance Conference; Denver; 2015 Dec 9-10.
Shiga toxin-producing *E. coli*	"Unresolved" case status	No. No effect on case definition, remains "Unresolved"	We have already excluded cases where the only laboratory report is from a PCR-based test. However, increased case ascertainment is still a concern, because a positive PCR-based test leads to more cases undergoing confirmatory testing and reporting.
Shigellosis	Yes (case status=suspected)	No. No effect on case definition, remains "Suspected"	The BioFire gastrointestinal panel cannot discriminate between *Shigella* (which is reportable) and enteroinvasive *E. coli* (which is not reportable), so we classify such cases as suspected *Shigella*. Increased ascertainment of confirmed cases is also an issue, since a positive *Shigella*/EIEC result might prompt follow-up cultures that would not otherwise have been performed.
Vibriosis	"Unresolved" case status	No. No effect on case definition, remains "Unresolved"	
Yersiniosis	"Unresolved" case status	Not nationally notifiable. No effect on case definition, remains "Unresolved"	

*Updating from a case to "not a case" based on additional laboratory information is NYC-specific; state health departments participating in FoodCORE indicate they do not downgrade cases to "not a case" based on a negative confirmatory test.

If you have remaining questions about the characteristics of these data, please ask.

OUTPUT DATA

NOTES