A Novel Syndrome Definition Validation Approach for Rarely Occurring Diseases

Julio C. Silva[1], Shital C. Shah[1], Dino P. Rumoro[1], Marilyn M. Hallock[1], Gillian S. Gibbs*[1] and Michael J. Waddell[2]

[1]Rush University, Chicago, IL, USA; [2]Pangaea Information Technologies, Chicago, IL, USA

Objective

To develop and test a novel syndrome definition validation approach for rarely occurring diseases.

Introduction

Early detection of rarely occurring but potentially harmful diseases such as bio-threat agents (e.g., anthrax), chemical agents (e.g., sarin), and naturally occurring diseases (e.g., meningitis) is critical for rapid initiation of treatment, infection control measures, and emergency response plans. To facilitate clinicians' ability to detect these diseases, various syndrome definitions have been developed. Due to the rarity of these diseases, standard statistical methodologies for validating syndrome definitions are not applicable.

Methods

Syndrome definitions were developed by researchers for the Geographic Utilization of Artificial Intelligence in Real-Time for Disease Identification and Alert Notification (GUARDIAN) surveillance system (1). The main steps for validation of the syndrome definitions were:

1) Partition of literature articles: Literature articles that described positive cases were randomly divided to generate detection (75% of articles) and testing (25% of articles) syndrome definitions.

2) Synthetic case generation: Syndrome definitions and associated statistical measures were reverse engineered using probability of occurrence and inverse Gaussian function to generate potentially infinite positive artificial cases.

3) Clinical filter application: To avoid clinically incompatible combinations of newly generated symptoms, rules based on clinically guided knowledge from emergency department (ED) physicians were applied. Steps 2 and 3 were repeated for both detection and testing syndrome definitions.

4) Detection phase validation:

a. ED negative case sample: Detection syndrome definitions were tested using a random sample of negative ED cases. Knowledge gained through false positive cases was utilized to modify the surveillance algorithms and system thresholds.

b. 10-fold cross-validation: Standard 10-fold cross-validation on detection articles of positive cases and ED negative cases was utilized to generate performance metrics. Suspected cases were reviewed by ED clinicians for threshold enhancement.

c. Literature articles (n=1): The ability of syndrome definitions to correctly flag literature articles with n=1 case was documented.

5) Testing phase validation:

a. Testing sample: Synthetic positive cases generated from the testing articles along with another set of ED negative cases were evaluated by the respective syndrome definition. Suspected cases were clinically evaluated.

b. Literature articles (n=1): Similar to detection step 4c, articles with n=1 were tested using syndrome definitions.

c. True positive samples: When available, true positive cases from an ED were identified and sent through the GUARDIAN system.

6) Multi-syndrome validation: A combined sample of positive cases of multiple syndromes and ED negative cases were evaluated for detection of individual syndromes among other similar syndromes.

Results

To demonstrate the validation approach, the anthrax syndrome definition was utilized. This syndrome definition was developed with 25 articles containing positive anthrax cases used for detection, and the remaining 11 articles used for testing. With a 10-fold cross validation of the detection phase, the initial results showed accuracy was 99.4% (false positive rate of 0.65% and false negative rate of 0.00%). The testing phase initial validation revealed 99.2% accuracy for the anthrax syndrome definition.

Conclusions

Syndrome specific synthetic samples that are validated through clinical filters allowed the generation of an unlimited number of positive cases. Correct identification by GUARDIAN of these cases indicates robust and reliable syndrome definitions. Utilization of these cases, in conjunction with adherence to a methodological process, was the cornerstone of the GUARDIAN syndrome definition validation approach. The validation approach was successfully demonstrated on anthrax and can be applied to other bio-threat agents, chemical agents, and naturally occurring diseases.

Keywords

syndromic surveillance; bioterrorism; infectious diseases

References

1. J. Silva, D. Rumoro, M. Hallock, S. Shah, G. Gibbs, M. Waddell, K. Thomas, Disease profile development methodology for syndromic surveillance of biological threat agents, Emerging Health Threats Journal, 2011, 4:11129.

*Gillian S. Gibbs
E-mail: Gillian_Gibbs@rush.edu