User Centred Networked Health Care
A. Moen et al. (Eds.)
IOS Press, 2011

doi:10.3233/978-1-60750-806-9-564

Network Analysis of Possible Anaphylaxis Cases Reported to the US Vaccine Adverse Event Reporting System after H1N1 Influenza Vaccine

Taxiarchis BOTSIS[a,b,1], Robert BALL[a]
[a] *Office of Biostatistics and Epidemiology, Center for Biologics Evaluation and Research (CBER), Food and Drug Administration (FDA), Rockville, MD, USA*
[b] *Department of Computer Science, University of Tromsø, Tromsø, Norway*

Abstract. The identification of signals from spontaneous reporting systems plays an important role in monitoring the safety of medical products. Network analysis (NA) allows the representation of complex interactions among the key elements of such systems. We developed a network for a subset of the US Vaccine Adverse Event Reporting System (VAERS) by representing the vaccines/adverse events (AEs) and their interconnections as the nodes and the edges, respectively; this subset we focused upon included possible anaphylaxis reports that were submitted for the H1N1 influenza vaccine. Subsequently, we calculated the main metrics that characterize the connectivity of the nodes and applied the island algorithm to identify the densest region in the network and, thus, identify potential safety signals. AEs associated with anaphylaxis formed a dense region in the 'anaphylaxis' network demonstrating the strength of NA techniques for pattern recognition. Additional validation and development of this approach is needed to improve future pharmacovigilance efforts.

Keywords. Spontaneous Reporting System, Network Analysis, VAERS, H1N1.

1. Introduction

More than 10,000 reports of adverse events following more than 82.4 million doses of the H1N1 2009 monovalent vaccine were submitted to the United States (US) Vaccine Adverse Event Reporting System (VAERS) [1]. VAERS is the repository for adverse events (AEs) that are reported after vaccinations by health care providers, vaccine recipients and other interested parties, and by manufacturers as required by regulation. Well-trained nurses code these reports using the Medical Dictionary for Regulatory Activities (MedDRA) and assign preferred terms (PTs) that represent the AEs described in the narratives. Data collected in VAERS is analyzed to identify safety signals [2]. The traditional approach combines the review of individual reports by Medical Officers (MOs) and statistical data mining algorithms (DMAs), that are scientifically based on the detection of disproportionality of reporting [3]. Current

[1] Corresponding Author: Taxiarchis Botsis, OBE|CBER|FDA, Woodmont Office Complex 1, Room 306N, 1401 Rockville Pike, Rockville, MD 20852, USA, e-mail: Taxiarchis.Botsis@fda.hhs.gov.

DMAs are generally limited in their ability to evaluate the multiple interactions among all the vaccines and AEs in the database. Thus, the methodologies to identify patterns of AEs related to the administration of a vaccine or the co-administration of multiple agents need to be improved.

If a safety concern is identified, MOs follow up with more detailed analysis, including an evaluation of a series of cases with subsequent classification according to a predefined case definition. Responding to a safety signal for anaphylaxis after H1N1 influenza vaccine that was received from the Canadian Ministry of Health in mid-November 2009 [1], FDA systematically reviewed all (N=6034) case reports submitted to VAERS related to the H1N1 vaccine from November 22, 2009 through January 31, 2010 to evaluate whether a similar safety signal for anaphylaxis existed in VAERS. Although there was not a safety signal for anaphylaxis after H1N1 influenza vaccine in VAERS, the dataset generated by this review provided an opportunity to investigate whether applying the principles of network analysis (NA) would allow us to identify a pattern of PTs within the network that had performance characteristics nearly equal to manual case classification. The VAERS subset was viewed and analyzed as a network with the vaccines/PTs and their interconnections being the nodes and the edges, respectively.

2. Methods

MOs manually screened the narratives and PTs of all reports of possible anaplylaxis (N=237). The possible 'anaphylaxis' subset was preprocessed to facilitate the subsequent NA. Each report was first represented as a vector (Rx) consisted of vaccines and PTs; then, the vectors were decomposed into pairs of vaccine (Vax) or PT and report ID. For example, the vector Rx= [IDx Vax_1 Vax_2 PT_1 PT_2 PT_3] was decomposed to Vax_1-IDx, Vax_2-IDx, PT_1-IDx, PT_2-IDx, PT_3-IDx. The vaccines/PTs were tied by their co-occurrence in an individual report that is being part of pairs with the same IDx. The number of reports containing a particular tie was the weight for each element that was included in an adjacency matrix; this matrix facilitated the construction of the 'anaphylaxis' network.

We focused on identifying patterns among the PTs consistent with anaphylaxis in this network. In terms of topology a dense region within the network structure would represent a pattern. NA offers the possibility for the qualitative evaluation and quantification of these areas. Particularly, we used certain node centrality metrics: hub centrality, which measures the degree of connectivity of a node to other important nodes in the network [4]; betweenness centrality, which measures the extent to which each node acts as a 'bridge' between other nodes [5]; and, inverse closeness centrality, which measures the average distance from a node to the other nodes [4]. We calculated these metrics for the anaphylaxis network and scaled them according to the top value, i.e. all values in each metric were divided by that top value. Subsequently, we selected the top 20 nodes according to hub centrality and constructed betweenness vs. inverse closeness centrality diagram to illustrate the connectivity of these nodes.

Further evaluation included the visual representation of the densest area of the network that might hide the pattern of interest. To reduce the full network, we selected the 'islands' algorithm that identifies all the maximal islands within a predefined node interval for an edge weight threshold [6] and combined it with triangular weight (TW) that is equal to the number of triangles each line of the original network is

contained [7]. We hypothesized that the use of TWs instead of the original weights would emphasize multiple interactions, filter out weak connections and reveal the patterns; thus, we applied it to both networks.

While no clear "gold standard" exists for pattern recognition, we compared the PTs identified by the above network analysis with criteria in the Brighton Collaboration (BC) case definition for anaphylaxis [8]. Based on these criteria (Table 1), the patterns related to anaphylaxis are defined. Here, we were interested in finding these criteria through both qualitative and quantitative NA of the 'anaphylaxis' subset. MedDRA does not include all the appropriate PTs to fully represent the BC criteria; however, BC case definition was a guide for recognizing the PTs that describe these criteria in the identified patterns. Pajek 2.01 and ORA 2.2.5 were the tools used for the network analysis.

3. Results

The original 'anaphylaxis' network included 301 nodes. The diagrams in Figure 1A present the metrics for the 20 top nodes according to hub centrality. The original network was reduced to include a community of 30 nodes by combining the TW (TW threshold equal to 70) with the 'island' algorithms (Figure 1B). Network analysis showed a clear pattern for anaphylaxis syndrome with all the PTs (shown in red crosses in the betweenness vs. inverse closeness diagrams) that characterize this condition being part of the 'anaphylaxis' island and among the top nodes in terms of all centrality metrics (Figure 1A). In line with the Brighton Collaboration criteria the symptoms for the four organ systems (dermatological/mucosal, cardiovascular, respiratory and gastrointestinal) were represented in the network image as well as in the top 20 nodes. As expected FLU(H1N1) node was the most central; the other two influenza vaccines were also among the top nodes.

Table 1. Summarized criteria for the Brighton Collaboration case definition of anaphylaxis.

Organ Systems	Major Criteria	Minor Criteria
Dermatologic or mucosal	urticaria (hives) or erythema, generalized	generalized pruritus without skin rash
	angioedema, localized or generalized	generalized prickle sensation
	generalized pruritus with skin rash	localized injection site urticaria
		red and itchy eyes
Cardiovascular	measured hypotension	reduced peripheral circulation (tachycardia, a capillary refill time of >3 sec without hypotension, a decreased level of consciousness)
	uncompensated shock (tachycardia, capillary refill time >3 sec, reduced central pulse volume, decreased level or loss of consciousness)	
Respiratory	bilateral wheeze (bronchospasm), stridor	persistent dry cough, hoarse voice
	upper airway swelling (lip, tongue, throat, uvula, or larynx)	difficulty breathing (no wheeze or stridor)
	respiratory distress (tachypnoea, increased use of accessory respiratory muscles, recession, cyanosis, grunting)	sensation of throat closure
		sneezing, rhinorrhea
Laboratory		Mast cell tryptase elevation > upper normal limit
Gastrointestinal		Diarrhoea, abdominal pain, nausea, vomiting

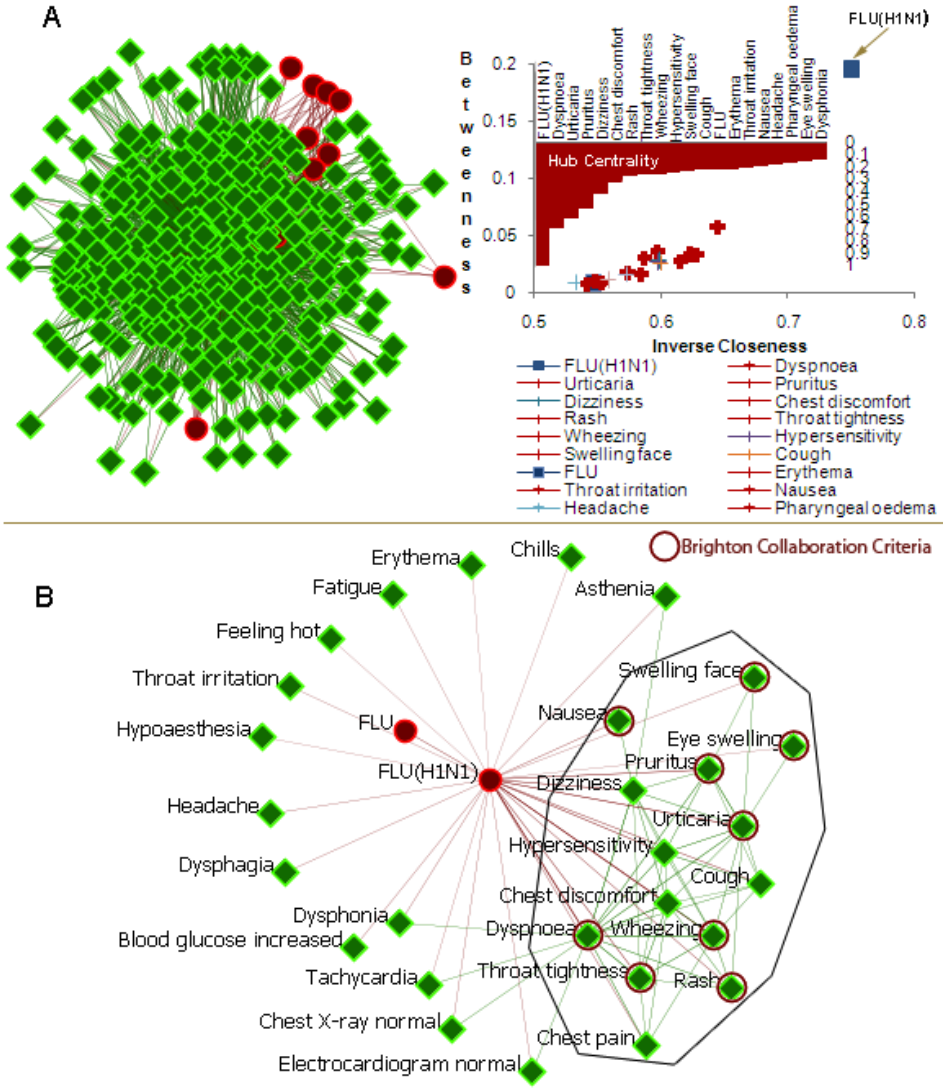

Figure 1. A. 'Anaphylaxis' network and centrality metrics for the top 20 nodes; for illustration purposes, node labels are not presented and hub centrality diagram is reversed. B. 'Anaphylaxis' island and pattern.

4. Discussion

This work demonstrates the potential use of NA for pattern identification in VAERS as this was discussed in our first study (also the first in the area) that dealt with the same issue [9]. Filling the gap of traditional approaches, we analyzed the multiple interactions of the critical terms (vaccines and PTs) in VAERS reports using a dataset related to adverse events reported after H1N1 vaccination. Through the anaphylaxis example, we showed that it is possible to isolate the densest region in a network using certain metrics and algorithms. Using a certain standard (e.g. BC criteria) this region

could be characterized as a pattern that deserves further investigation. While not the focus of our study, NA might serve as an efficient way to begin development of Standardized MedDRA Queries [10].

This study has some limitations. First, we did not apply a statistical framework for identifying the anaphylaxis pattern but empirically evaluated the results of NA. Second, we did not follow a validated rule for selecting the node interval in the 'islands' algorithm; it was considered that this number should be adequate to reveal a strong pattern. It could be also argued that our sample included retrospectively classified reports and this might reduce the value of our analysis; however, our main scope was the investigation of the possible benefits from applying NA to VAERS data.

Various algorithms have been applied before for the detection of clustered regions in a network. For example, Newman described the identification of communities based on the concept of modularity [11]. The evaluation of other approaches in addition to the 'islands' algorithm should be included in the next steps of our work. The evaluation framework should be extended to include a statistical aspect e.g. a thorough analysis of the centrality metrics. The current study is one step in evaluating the NA potential to recognize safety patterns in VAERS. We plan to further study this approach by addressing the aforementioned limitations and the application of our ideas to prospectively collected data for prediction purposes.

Acknowledgements: This project was supported in part by an appointment to the Research Participation Program at the Center for Biologics Evaluation and Research administered by the Oak Ridge Institute for Science and Education through an interagency agreement between the U.S. Department of Energy and the U.S. Food and Drug Administration. We thank the Medical Officers at FDA who evaluated the reports and those who reported them.

References

[1] Reblin T. AREPANRIX™ H1N1 Vaccine Authorization for Sale and Post-Market Activities. 11-12-2009. Canadian Ministry of Health.
[2] Varricchio F, Iskander J, Destefano F, et al. Understanding vaccine safety information from the vaccine adverse event reporting system, The Pediatric Infectious Disease Journal, 23 (2004) 287-294.
[3] Stephenson WP, Hauben M. Data mining for signals in spontaneous reporting databases: proceed with caution, Pharmacoepidemiology and Drug Safety, 16 (2007) 359-365.
[4] Newman MEJ. Networks: an introduction., Oxford University Press, New York 2010.
[5] Freeman LC. Set of Measures of Centrality Based on Betweenness, Sociometry, 40 (1977) 35-41.
[6] Zaversnik M, Batagelj V. Islands, Sunbelt XXIV, 2004.
[7] Batagelj V, Mrvar M. Analysis of Large Networks with Pajek, Sunbelt XXIX, 2009.
[8] Ruggeberg JU, Gold MS, Bayas JM, et a. Anaphylaxis: Case definition and guidelines for data collection, analysis, and presentation of immunization safety data, Vaccine, 25 (2007) 5675-5684.
[9] Ball R, Botsis T. Can network analysis improve pattern recognition among adverse events following immunization reported to VAERS? Clinical Pharmacology & Therapeutics (in press).
[10] Bate A, Evans SJW. Quantitative signal detection using spontaneous ADR reporting. Pharmacoepidemiology and Drug Safety, 18 (2009) 427-436.
[11] Newman MEJ. Fast algorithm for detecting community structure in networks, Physical Review, 69 (2004).