

Monitoring biomedical literature for post-market safety purposes by analyzing networks of text-based coded information

Taxiarchis Botsis, MSc, PhD¹, Matthew Foster, MSc¹, Kory Kreimeyer, MSc¹, Abhishek Pandey, MSc¹, Richard Forshee, PhD¹

¹Office of Biostatistics and Epidemiology, Center for Biologics Evaluation & Research, US Food and Drug Administration, Silver Spring, MD

Abstract

Literature review is critical but time-consuming in the post-market surveillance of medical products. We focused on the safety signal of intussusception after the vaccination of infants with the Rotashield Vaccine in 1999 and retrieved all PubMed abstracts for rotavirus vaccines published after January 1, 1998. We used the Event-based Text-mining of Health Electronic Records system, the MetaMap tool, and the National Center for Biomedical Ontologies Annotator to process the abstracts and generate coded terms stamped with the date of publication. Data were analyzed in the Pattern-based and Advanced Network Analyzer for Clinical Evaluation and Assessment to evaluate the intussusception-related findings before and after the release of the new rotavirus vaccines in 2006. The tight connection of intussusception with the historical signal in the first period and the absence of any safety concern for the new vaccines in the second period were verified. We demonstrated the feasibility for semi-automated solutions that may assist medical reviewers in monitoring biomedical literature.

Introduction

Reports for adverse events (AEs) related to the use of licensed medical products in the US are submitted to the Food and Drug Administration (FDA) for safety monitoring by health care providers, manufacturers, patients or their legal representatives, and patients' relatives. The Vaccine Adverse Event Reporting System (VAERS) and FDA Adverse Event Reporting System (FAERS) are repositories that store these safety reports for vaccines, drugs and biologics. Each report contains certain structured fields for patient information (e.g. demographics) and other reported information (e.g. dates), as well as free-text fields to describe the AE(s). Medical codes (known as Preferred Terms or PTs) from the Medical Dictionary for Regulatory Activities (MedDRA) are also assigned to each report by either specialized personnel or the product manufacturer to represent key medical information included in the narrative description.

Medical experts at the FDA review reports submitted to VAERS and FAERS to monitor for unusual patterns in the data and verify safety signals. Their review includes the analysis of report-specific features, such as time to onset, symptoms, and diagnoses; as well as the evaluation of other data sources such as Electronic Health Record (EHR) data and biomedical literature. Figure 1 illustrates the submission of safety reports to the VAERS and FAERS systems and the review process centered around the medical experts at the FDA. The full review process involves numerous data sources and may be expanded in the future to include data pools such as new reports and social media. While this process is currently supported by some automated tools, it is largely manual and time intensive and would benefit from automated methodologies.

Making this process more efficient, effective, and rigorous is a high priority for the FDA. Most of the greatest challenges to overcome are related to: (i) the automation of currently manual steps, such as the retrieval of key information from the free text (report narratives, EHR data, and biomedical literature); and (ii) the use of more advanced techniques for pattern evaluation, such as the recognition and analysis of the complex AE patterns in sets of reports. We previously focused on the automated classification of cases for particular conditions (1, 2) and the text mining of safety reports (3-5) aimed at automating these major steps in the review process. Inspired by the principles of graph theory and social network analysis, we developed a network analysis (NA) tool to support evaluation of the complex relationships in our data. This line of research started a few years ago with the development of the US FDA Adverse Event Network Analyzer that handled our 3-mode data (reports, medical products, AEs) and presented them in a dynamic environment (6). Building on this work and the investigation of various NA techniques (7-9), we further developed a complete NA platform tailored to our needs, the Pattern-based and Advanced Network Analyzer for Clinical Evaluation and Assessment (PANACEA).

PANACEA is part of a decision support environment that also includes the Event-based Text-mining of Health Electronic Records (ETHER) system (10). ETHER processes the narratives of post-market reports and extracts key

clinical features (primary, secondary and rule-out diagnosis, cause of death, medical and family history, symptoms, and medical product exposures) along with timestamps for these features. PANACEA and ETHER are tightly integrated, i.e. it is possible to use the ETHER PTs to create networks in PANACEA or to isolate subnetworks in PANACEA and view the reports in ETHER. While both tools have been developed to process post-market reports, it is possible to apply them to other types of data. We previously demonstrated the automated development of case definitions from medical literature by processing PubMed abstracts with our text mining algorithm and analyzing the information associations with a widely used network analysis tool (11).

The starting point for this study was the known safety signal of intussusception (IS) following the administration of Rotashield Vaccine (RV) found in 1999 (12). Intussusception is a bowel obstruction in which one segment of bowel becomes enfolded within another segment. RV was licensed in August 1998 for the vaccination of infants and was withdrawn from the market in October 1999 following the identification of its association with IS (13). The next two rotavirus vaccines, RotaTeq and RotaRix, were not approved until February 3, 2006 and April 3, 2008, respectively (14, 15). Although no association of RotaTeq or Rotarix with IS or any other AE has been detected since the licensure of the two vaccine products, the past association between RV and IS justifies the continued monitoring of this adverse event. This includes the review and screening of biomedical literature regarding these products. We here evaluate how ETHER and PANACEA might efficiently automate the information retrieval and processing for relevant publications.

This paper is organized as follows. In the methodology section, we describe the retrieval of publications from PubMed and the processing of their abstracts with ETHER, the National Library of Medicine MetaMap tool (16) and the National Center for Biomedical Ontologies (NCBO) Annotator (17) to generate network data for PANACEA. Since the main analysis was performed in PANACEA, we focus on PANACEA functionalities used in our analysis and describe the corresponding steps. We then present the results and discuss the main aspects of our work.

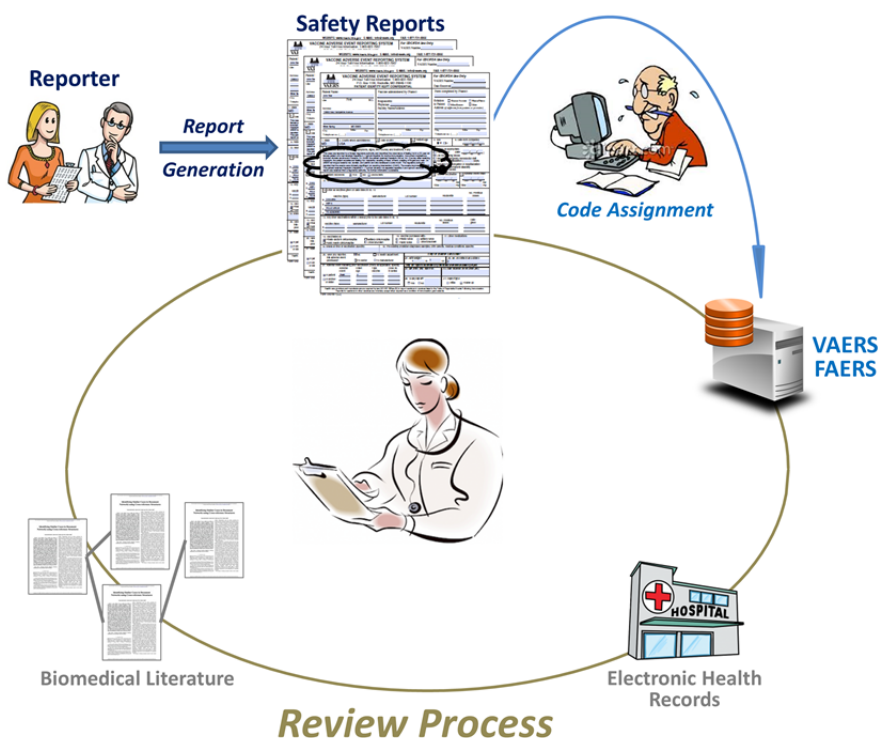


Figure 1. The submission of safety reports and the review process includes the evaluation of various data sources. Medical Dictionary for Regulatory Activities (on MedDRA) Codes are assigned to the reports by either the reporters or well-trained personnel to represent the key medical information included in the narrative descriptions. Medical experts and epidemiologists utilize these codes in the review process. VAERS: Vaccine Adverse Event Reporting System; FAERS: FDA Vaccine Adverse Event Reporting System.

Methods

We hypothesized that: (i) all the papers published around the licensure of RV (in 1998 and the next few years) would discuss the association of this product with IS; and, (ii) anything published in recent years and especially after the release of the two new vaccines (2006 and onwards) would potentially describe analyses on the topic. To evaluate our hypothesis, we ran the following query in PubMed on May 10, 2016 and also applied a publication date filter to it:

Search: (rotavirus vaccine) and ((adverse events) or (side effects) or (adverse effects) or (toxicities) or (adverse event) or (side effect) or (adverse effect) or (toxicity))

Filter: Publication date from January 1, 1998

This query supported the retrieval of all articles that described the side effects of rotavirus vaccines published between January 1, 1998 and the date of the query. Articles that did not contain an abstract were excluded from the next steps.

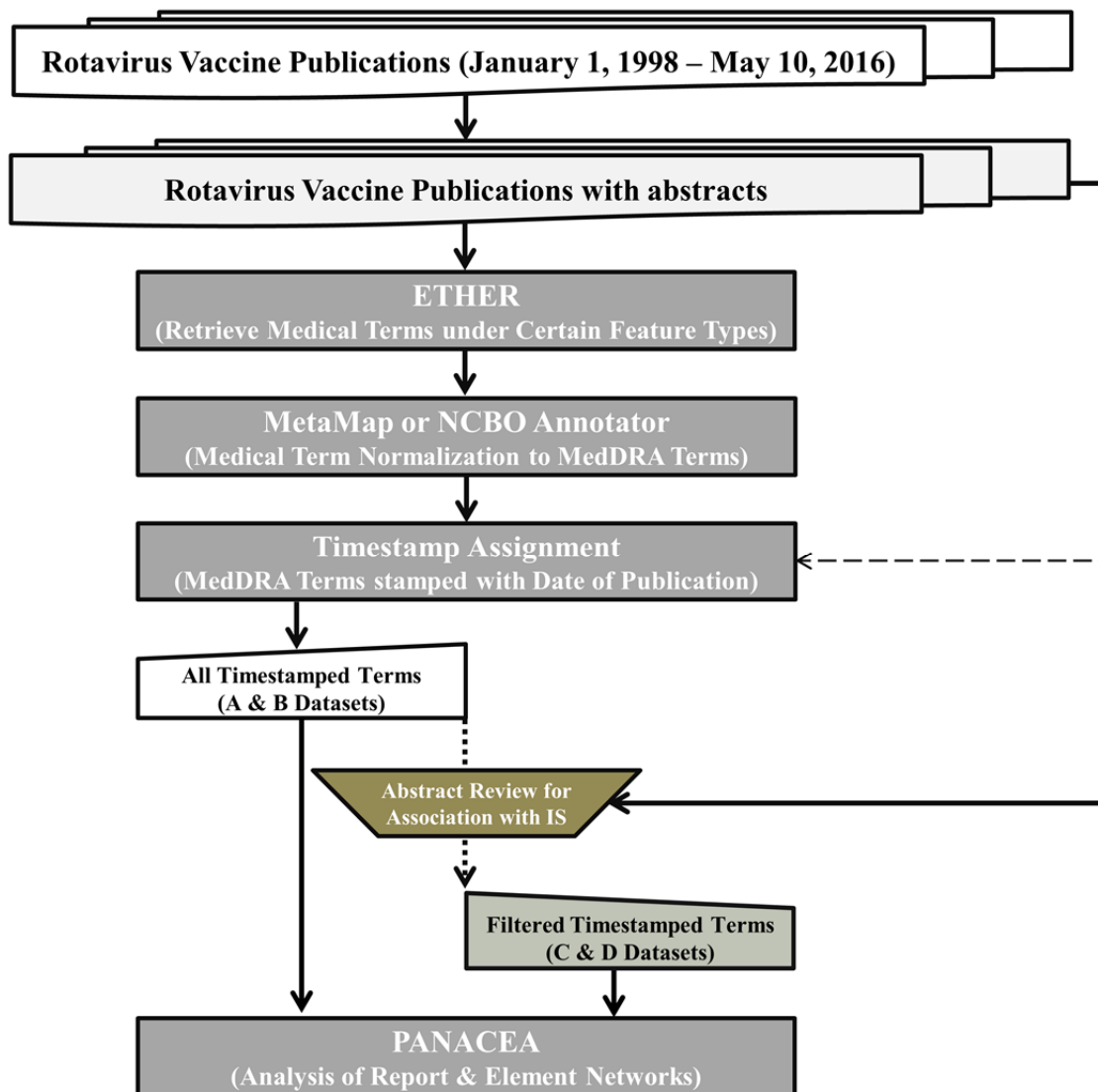


Figure 2. The multiple steps followed for the generation of the coded datasets processed in the Pattern-based and Advanced Network Analyzer for Clinical Evaluation and Assessment (PANACEA). ETHER: Event-based Text-mining of Health Electronic Records; NCBO: National Center for Biomedical Ontology.

As shown in Figure 2, the final set of abstracts was processed by ETHER and the medical terms extracted as primary and secondary diagnosis, cause of death, or plain symptoms were normalized to MedDRA terms. This was done with two separate tools, so the results could be compared. First, the MetaMap tool was used through its online interface. A batch processing job was submitted on January 3, 2017 containing all of the extracted snippets of text and using the 2016AA knowledge source, the NLM data version and the relaxed data model. The word sense disambiguation and all derivational variants were used ignoring the word order and MetaMap score. The second tool was the NCBO Annotator, which was accessed through its Web Service API. For each text snippet, one annotation request was submitted with the *Match Longest Only* and *Match Partial Words* settings turned off. The NCBO process was performed on December 20, 2016 when the currently uploaded MedDRA version was the 2016AA submission from September 06, 2016. The terms identified by each tool were timestamped with the date of publication of the articles they belonged to. All articles, along with their MetaMap- and NCBO-based terms and associated time information, formed the first (A) and second (B) datasets, respectively, for our analysis. We also took an additional step and manually reviewed each abstract where IS was assigned for the nature of association between the rotavirus vaccine and intussusception. When no real association was reported, we removed IS from the list of MedDRA terms for that article in the A and B datasets. This process resulted in the creation of a third (C) and a fourth (D) dataset, respectively, with the same articles but the “noisy” information removed. Table 1 summarizes the four datasets that were subsequently imported into PANACEA.

Table 1. Brief description of the four datasets.

Dataset	MedDRA Term Source	Intussusception Mentions
A	MetaMap	All
B	NCBO Annotator	All
C	MetaMap	Only Associated
D	NCBO Annotator	Only Associated

PANACEA uses the medical terms and product names to create two types of networks for the post-market data: a Report Network and an Element Network. In the Report Network each node represents a post-market report and the edges between nodes represent terms or products that appear in both reports. In the Element Network, each node is a report element (product or term) and each edge represents at least one post-market report containing both elements. A report network may support the identification of reports including the same or similar information as well as the retrieval of reports associated with particular condition(s). The element network focuses on the review and analysis of the actual information in the reports. Assume a network of reports for a particular medical product is constructed to support the identification of reports associated with anaphylaxis. In that network we can either identify an existing anaphylaxis-related report node or synthesize and import a new anaphylaxis-related report node to find the tightly connected nodes that represent reports potentially associated with this AE. Following report retrieval through the identification of the corresponding region in the network, we can examine the particular AE information by constructing an element network. The recognition of an anaphylaxis pattern in the element network confirms the findings in the report network and may result in additional explorations.

The current standard version of PANACEA supports many traditional and other more advanced network analysis techniques. In this study we particularly used the principal component and the island height layout. The principal components layout is based on the idea of using the first two eigenvectors of the weighted adjacency matrix to create a useful and visually appealing 2-dimensional layout. The two eigenvectors with the largest eigenvalues are estimated sequentially using the Power method, and each node’s position along these vectors is used to map it to an onscreen location with the x-axis corresponding to the largest eigenvalue. In practice, the principal component on the x-axis corresponds with node centrality (mainly strength), and the component on the y-axis represents structural differences within the network. The island height layout is based on the “island” idea previously described by the Pajek developers (18). PANACEA uses the heaviest edge weight of each node to assign the maximum height to that node, with the expectation that nodes with large edge weights will be more important within the network. The y-position of each node in the layout corresponds to its island height, while the x-position is determined by the number

of nodes co-appearing at the same height as well as their number of connections to the remaining nodes; the most connected nodes will hold the most central position at the corresponding height.

We used the MedDRA terms of each dataset to construct a report network (hereafter, article network) where each node represented an article from the dataset and each edge the connection between two articles that contained at least one term in common. The edges in the article network were weighted based on the number of the same terms that appeared in two articles. We then switched to the element network where the nodes represented the terms and the rotavirus vaccine. The latter could have been removed from the element network since it appeared in all articles and did not add any real value to the analysis. However, this would also result in the removal of nodes for terms that appeared in one abstract only with no other coded AEs and, thus, affect the network topology. It was therefore considered important to include the vaccine node in our networks. The date of publication stamped each article and its terms. This time information was used in PANACEA and supported the calculation of the node metrics in the four element networks (one for each dataset) over time as well as the creation of subnetworks by using the date of publication as a filter.

We further evaluated the strength evolution of the IS node in the networks. Strength was preferred over the other metrics since it not only illustrates the number of connections between IS and other nodes (i.e. the degree centrality) but also indicates the number of articles where IS was co-reported with the rotavirus vaccine and, potentially, other AEs. We then examined the corresponding temporal evolution plots and identified the time point(s) at which IS strength reached its plateau. We assumed that the period following that time point would not include any other publications for either the co-occurrence of IS and rotavirus vaccines (A and B datasets) or the real association between any rotavirus vaccine and IS (C and D datasets). This time point was considered to be the cutting point for creating two subnetworks, one before and one after it. Each network was visualized using the island height layout that supported the qualitative evaluation of IS' position and importance in the network.

Results

The PubMed query returned 563 articles. Abstracts were available for 420 publications. The first paper with an abstract discussing rotavirus vaccination was published in April 1998 and a total of 7 papers followed within the same year. As expected, IS was not reported until July 1999, the month the Centers for Disease Control and Prevention released the Morbidity and Mortality Weekly Report for the association between RV and IS (13). A total of 166 articles mentioned IS in the next few years with an expected peak in the first years following the IS safety signal. Further peaks were observed in 2006 when the next rotavirus vaccine was licensed and in the more recent years from 2012 to 2015 (Figure 3). Based on the manual review of all abstracts reporting IS, twenty-six abstracts for papers that were published until the end of 2006 described the confirmed association between RV and IS. No association was found between the new vaccines and IS with the exception of one paper published 2012 (19). It should be noted though that the claims of the authors in that paper were not supported by either the FDA safety review or the large number of other articles published for the same topic. One other paper published in 2010 referred to the old 1999 signal.

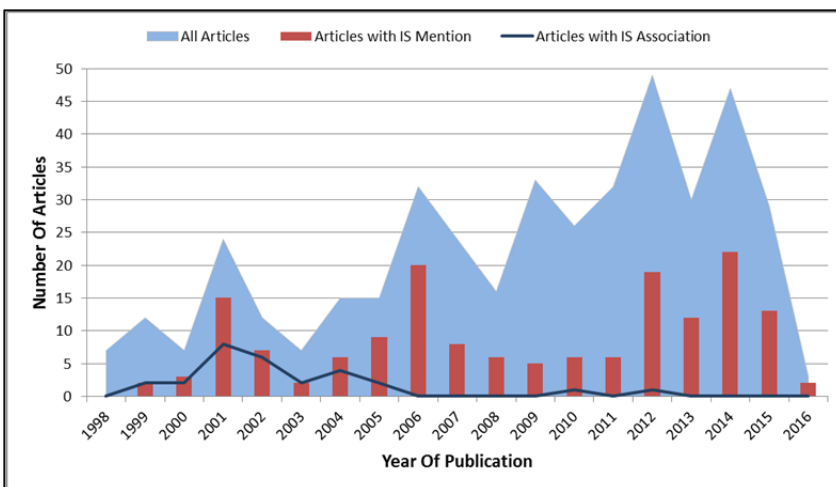


Figure 3. The articles for rotavirus vaccine published since January 1, 1998 with a closer look at the abstracts with either an intussusception (IS) mention or a clear association between the vaccine and IS.

We processed the 420 abstracts by ETHER and MetaMap or NCBO and created the four datasets that were then imported in PANACEA. NCBO annotator generated fewer MedDRA terms than MetaMap (1941 vs. 2589) and also encoded the ETHER output for fewer abstracts (409 vs. 418). The article networks and the subnetworks for the four datasets are shown in Figure 4 and 5. We set a threshold at 2006 to split the article networks into two subnetworks each with the term information included in the abstracts published until December 31, 2005 and after January 1, 2016. We also synthesized an article node that contained the IS term only and imported it into each subnetwork to identify the nodes that represented articles containing the IS term. All abstracts describing the real association between rotavirus vaccines and IS (26 and 2 before and after January 1, 2016, respectively) were identified in the second subnetworks with the synthesized article node. This finding indicated that both MetaMap and NCBO generated the IS term for those abstracts.

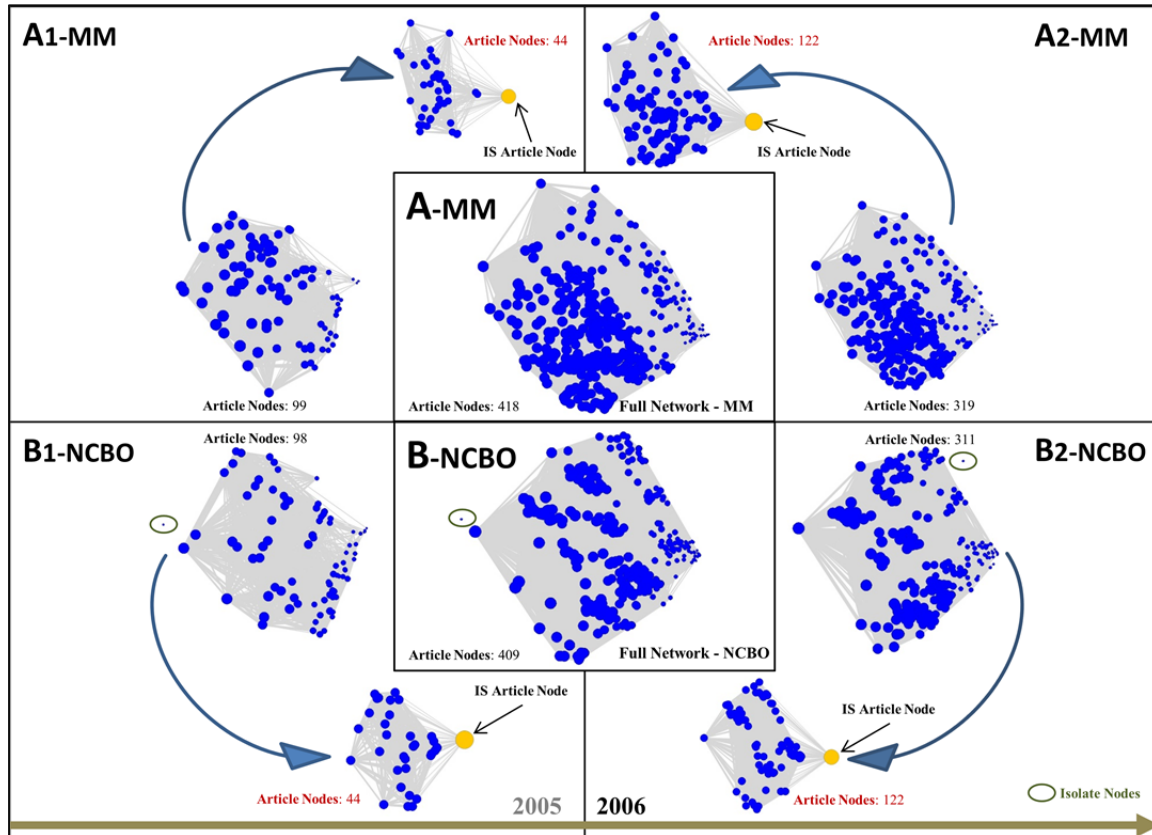


Figure 4. The article networks based on the full list of terms generated by either the MetaMap tool (MM) or the National Center for Biomedical Ontologies (NCBO) Annotator are shown in the center of the image (A-MM and B-NCBO, respectively). A1-MM and A2-MM subnetworks were built based on the A dataset and the 2006 time threshold (licensure of the new rotavirus vaccine). Similarly, the B1-NCBO and B2-NCBO subnetworks were constructed using the B dataset and the 2006 time threshold. In each panel a second subnetwork is shown after the import of a synthesized article node that contains the Intussusception (IS) term only and supports the identification of the IS-related abstracts. Small differences are observed between the MetaMap and NCBO subnetworks in terms of the network size. The isolate nodes in the NCBO (sub)networks contain the rotavirus vaccine and only one term that does not appear in any other article; it is therefore impossible to establish any connections with other nodes. The principal components layout was used for the network visualizations.

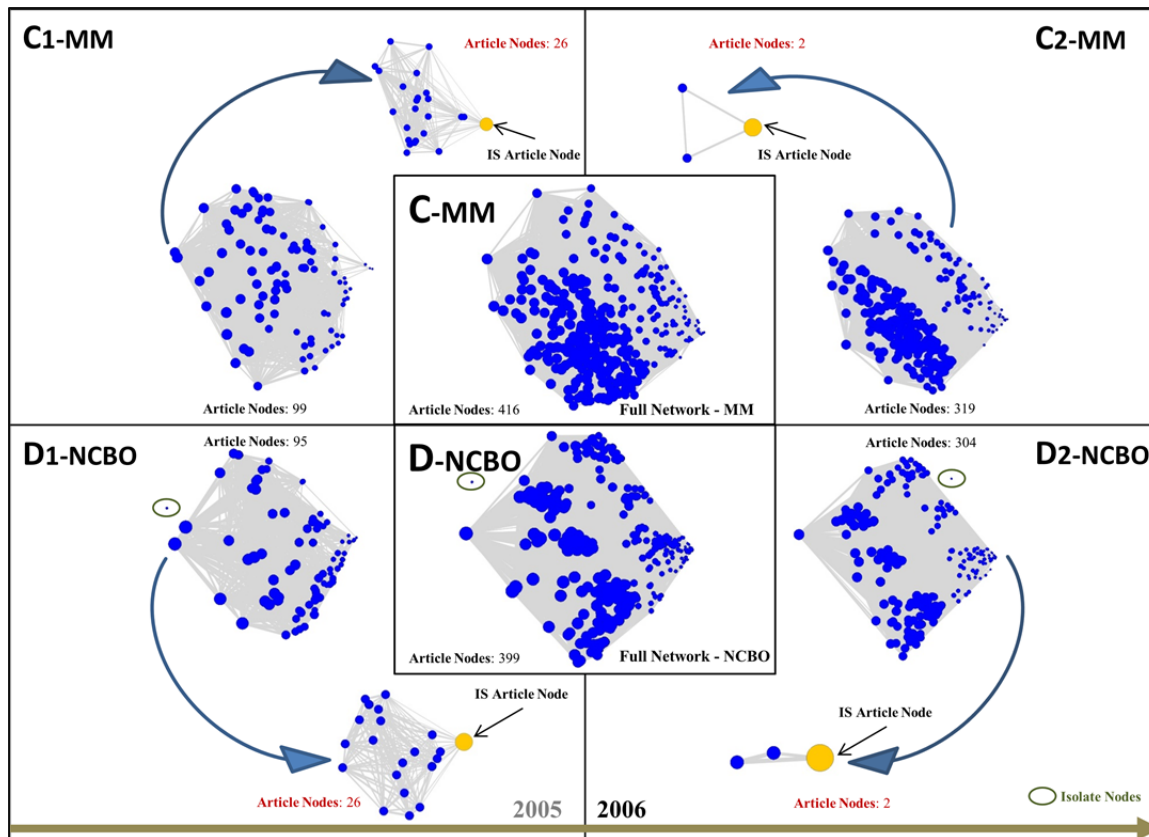


Figure 5. The article networks based on the full list of terms generated by either the MetaMap tool (MM) or the National Center for Biomedical Ontologies (NCBO) Annotator are shown in the center of the image (C-MM and D-NCBO, respectively). C1-MM and C2-MM subnetworks were built based on the C dataset and the 2006 time threshold (licensure of the new rotavirus vaccine). Similarly, the D1-NCBO and D2-NCBO subnetworks were constructed using the D dataset and the 2006 time threshold. In each panel a second subnetwork is shown after the import of a synthesized article node that contains the Intussusception (IS) term only and supports the identification of the IS-related abstracts. Some differences exist between the MetaMap and NCBO subnetworks in terms of the network size. The isolate nodes in the NCBO (sub)networks contain the rotavirus vaccine and only one term that does not appear in any other article; it is therefore impossible to establish any connections with other nodes. The Principal Components Layout was used for the network visualizations.

The evaluation of the actual information contained in the four datasets was performed in the element networks shown in Figures 6 and 7. We set again our threshold at 2006 to split the element networks into subnetworks each with the term information included in the abstracts published until December 31, 2005 and after January 1, 2006. The island height layout was used for the network visualizations and the side-by-side comparisons. The IS node appears as one of the top nodes in the subnetworks representing the A and B datasets, which include all IS mentions. For the C and D datasets, IS was one of the top nodes in the subnetworks representing the abstract information from articles published before 2006. In the subnetworks of the second period (Figure 7, Panels C2 and D2), the IS node is at the bottom of the visualization hidden within other nodes, an indication of the lack of real association between the new rotavirus vaccines and IS.

The element networks further supported the creation of the temporal evolution plot for the IS strength versus the year of publication. As shown in Figure 8, the IS node strength in the A and B datasets never reached a plateau, which is explained by the continuous mentions of IS in many articles over the years. For the C and D datasets though, there was a clear change in 2005 when the IS node strength reached its maximum value. The strength-based findings confirm the visual observations shown in Figures 6 and 7.

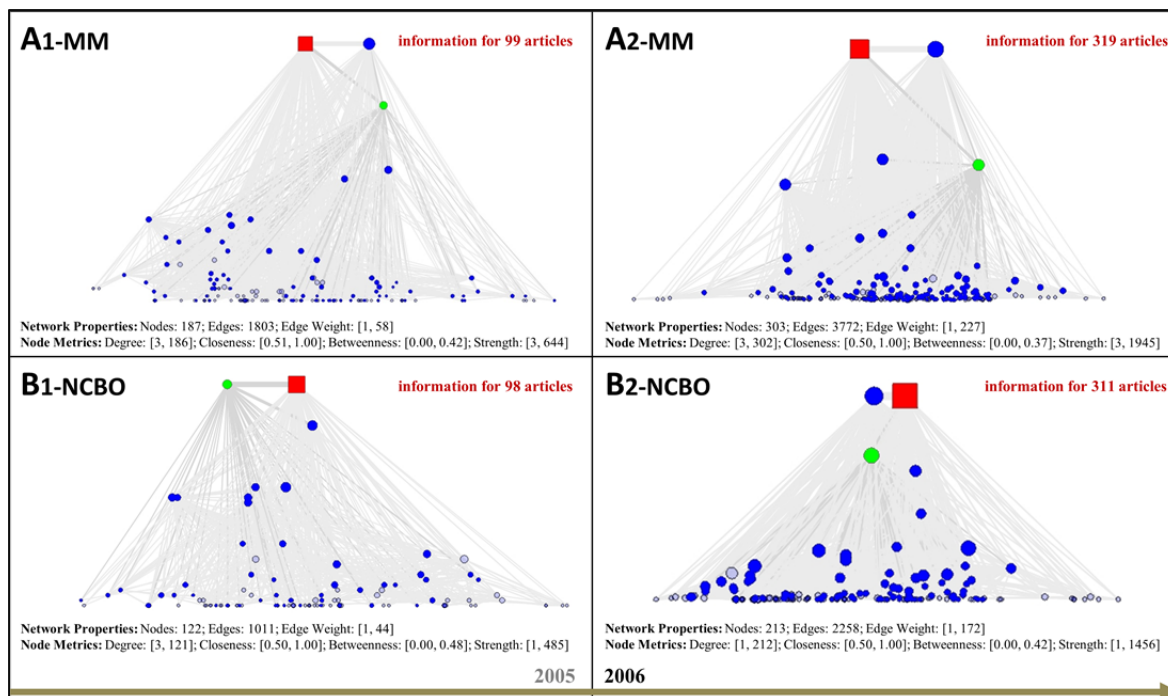


Figure 6. The four subnetworks for each element network that represented the A and B datasets. Red square, blue circle, and green circle nodes represent the Rotavirus Vaccine, MedDRA codes, and Intussusception (IS) code, respectively. Network properties and node metric ranges are provided for informational purposes only.

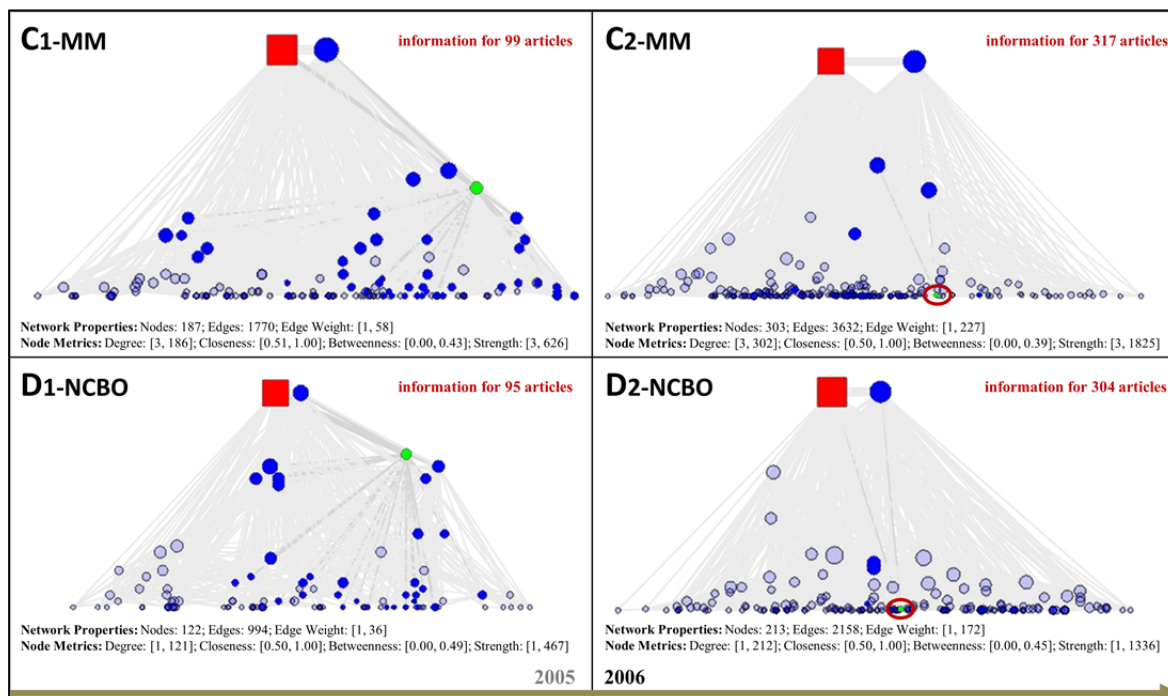


Figure 7. The four subnetworks for each element network that represented the C and D datasets. Red square, blue circle, and green circle nodes represent the Rotavirus Vaccine, MedDRA codes, and Intussusception (IS) code, respectively. IS is one of the top nodes before 2006 but appears at the bottom of the visualization in the second period though (within the magenta eclipse) indicating the lack of association between rotavirus vaccines and intussusception. Network properties and node metric ranges are provided for informational purposes only.

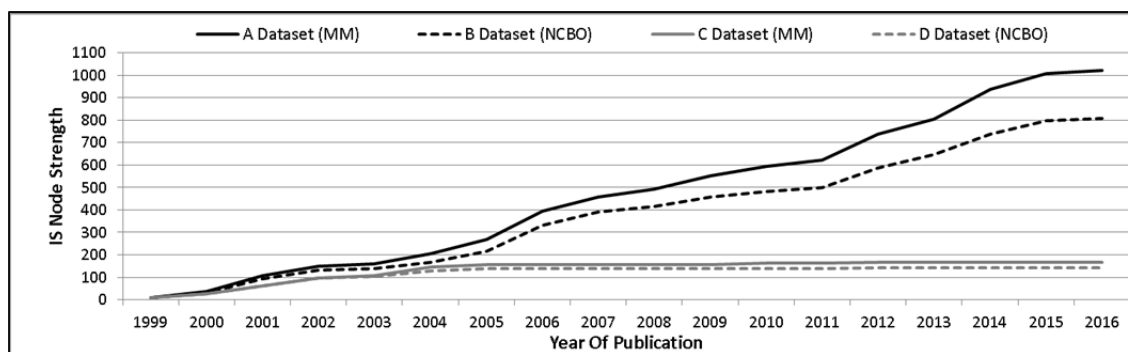


Figure 8. The intussusception (IS) node strength for the four element networks.

Discussion

We evaluated the potential for monitoring information published in the biomedical literature for rotavirus vaccine products and the intussusception adverse event. We initially text-mined the available abstracts with ETHER, a tool developed at the US FDA for the processing of post-market reports, and normalized the output with the NLM MetaMap and the NCBO Annotator, two tools for converting clinical concepts to the codes of particular terminologies, such as MedDRA. The network analysis of the text-based coded information with PANACEA demonstrated that many articles strongly discussed the historical signal of IS until the end of 2005 and the benefits from the vaccinations with the two new risk-free vaccines after 2006. To reach this conclusion we had to perform some manual review and remove the “noise” from the data, i.e. the IS mentions with no clear association with the vaccines.

This study has some limitations. First, we processed the abstracts only and not the full texts that contain more complete descriptions on a topic. The power of the information in abstracts was previously demonstrated and served the purposes of our focused analysis (20). Second, the manual review conducted for the synthesis of the C and D datasets places our approach closer to a semi- rather than a fully-automated approach. While the elimination of the manual steps is often desired by the end users, the partial review of small datasets by humans may be necessary for verification purposes – especially when less effort is required. Finally, we did not conduct a thorough analysis in PANACEA and evaluated certain aspects only, e.g. we did not examine the role of other AEs in the networks and analyzed specific network and node characteristics. The incorporation of more parameters in the network analysis step is definitely one of our goals and we aim at proposing a more complete solution in the future.

The coding of medical information has always been a challenge. The MetaMap tool and NCBO annotator have automated this task. However, there are still some concerns about their efficiency. For example, it was previously discussed that MetaMap demonstrates reduced accuracy in the presence of ambiguity (16). Such weaknesses may or may not be critical depending on the problem that needs to be solved. In this study, we did not validate the MetaMap and NCBO outputs for their completeness and accuracy against a reference standard. Both tools though generated the intussusception code and fully supported the identification of all those abstracts and articles that described a real association between the rotavirus vaccines and IS. The element networks built from the coded data further demonstrated the actual role of the IS code in the datasets, i.e. the corresponding IS node was one of the top ones before the release of the new vaccine and insignificant after 2006. So, without delving into the performance of either the MetaMap tool or the NCBO annotator, we showed that they adequately supported the goals of our study.

One other topic that was mentioned above and should be further discussed is related to the manual review of the abstracts and the subsequent manual curation of the A and B datasets. Although brief and limited, these manual steps showed that ETHER with either MetaMap or NCBO annotator alone could not support context evaluation. On the other hand, abstracts do not generally provide the full (clinical) information collected in a study and included in the full text of the article. For example, we would expect the listing of multiple gastrointestinal symptoms in a paper describing a real association between a rotavirus vaccine and intussusception. If a few articles included such information, the gastrointestinal pattern would emerge in the network visualization. We foresee two ways to handle the generation of datasets for investigation of the kind. First, develop algorithms that will support the full context evaluation in paper abstracts. Second, perform a more complete and focused data collection by processing the paper main texts.

The processing of biomedical literature represents a vast and promising field, but meaningful information retrieval, which is required by the post-market safety reviewers at the FDA among many others, can be very challenging. Our study demonstrated the feasibility for fully- or semi-automated solutions that may assist medical reviewers in the monitoring of biomedical literature. We envision the development of an efficient system that will be incorporated into the regulatory environment and will support end users in their decision making process. As we demonstrated here, the use of advanced approaches with the combined use of multiple tools is the ideal way to move forward.

Acknowledgments

This work was supported in part by the appointments of Abhishek Pandey, Kory Kreimeyer, and Matthew Foster to the Research Participation Program administered by ORISE through an interagency agreement between the US Department of Energy and the US FDA. It should be also noted that this work is at the early stages of development and does not constitute any FDA guidance or official strategy for the processing of biomedical literature.

References

1. Botsis T, Woo EJ, Ball R. Application of information retrieval approaches to case classification in the vaccine adverse event reporting system. *Drug Saf.* 2013;36(7):573-82.
2. Botsis T, Woo EJ, Ball R. The contribution of the vaccine adverse event text mining system to the classification of possible Guillain-Barre syndrome reports. *Appl Clin Inform.* 2013;4(1):88-99.
3. Baer B, Nguyen M, Woo EJ, Winiiecki S, Scott J, Martin D, et al. Can Natural Language Processing Improve the Efficiency of Vaccine Adverse Event Report Review? *Methods Inf Med.* 2016;55(2):144-50.
4. Botsis T, Buttolph T, Nguyen MD, Winiiecki S, Woo EJ, Ball R. Vaccine adverse event text mining system for extracting features from vaccine safety reports. *J Am Med Inform Assoc.* 2012;19(6):1011-8.
5. Botsis T, Nguyen MD, Woo EJ, Markatou M, Ball R. Text mining for the Vaccine Adverse Event Reporting System: medical text classification using informative feature selection. *J Am Med Inform Assoc.* 2011;18(5):631-8.
6. Botsis T, Scott J, Goud R, Toman P, Sutherland A, Ball R. Novel algorithms for improved pattern recognition using the US FDA Adverse Event Network Analyzer. *Stud Health Technol Inform.* 2014;205:1178-82.
7. Ball R, Botsis T. Can network analysis improve pattern recognition among adverse events following immunization reported to VAERS? *Clin Pharmacol Ther.* 2011;90(2):271-8.
8. Botsis T, Ball R. Network analysis of possible anaphylaxis cases reported to the US vaccine adverse event reporting system after H1N1 influenza vaccine. *Stud Health Technol Inform.* 2011;169:564-8.
9. Botsis T, Scott J, Woo EJ, Ball R. Identifying Similar Cases in Document Networks Using Cross-Reference Structures. *IEEE J Biomed Health Inform.* 2015;19(6):1906-17.
10. Botsis T, Jankosky C, Arya D, Kreimeyer K, Foster M, Pandey A, et al. Decision support environment for medical product safety surveillance. *J Biomed Inform.* 2016;64:354-36.
11. Botsis T, Ball R. Automating case definitions using literature-based reasoning. *Appl Clin Inform.* 2013;4(4):515-27.
12. Niu MT, Erwin DE, Braun MM. Data mining in the US Vaccine Adverse Event Reporting System (VAERS): early detection of intussusception and other events after rotavirus vaccination. *Vaccine.* 2001;19(32):4627-34.
13. CDC. Intussusception among recipients of rotavirus vaccine--United States, 1998-1999. *MMWR Morbidity and mortality weekly report.* 1999;48(27):577.
14. FDA. Approval Letter - RotaTeq 2006 [cited 2016 5/10/2016]. Available from: <http://www.fda.gov/BiologicsBloodVaccines/Vaccines/ApprovedProducts/ucm142303.htm>.
15. FDA. Approval Letter - Rotarix 2008 [cited 2016 5/10/2016]. Available from: <http://www.fda.gov/BiologicsBloodVaccines/Vaccines/ApprovedProducts/ucm133541.htm>.
16. Aronson AR, Lang FM. An overview of MetaMap: historical perspective and recent advances. *J Am Med Inform Assoc.* 2010;17(3):229-36.
17. Jonquet C, Shah NH, Musen MA. The open biomedical annotator. *Summit Transl Bioinform.* 2009;2009:56-60.
18. de Nooy W, Mrvar A, Batagelj V. Exploratory Social Network Analysis with Pajek. Second ed. Granovetter M, editor: Cambridge University Press; 2011.
19. Geier DA, King PG, Sykes LK, Geier MR. The temporal relationship between RotaTeq immunization and intussusception adverse events in the Vaccine Adverse Event Reporting System (VAERS). *Medical Science Monitor.* 2012;18(2):PH12-PH7.
20. van Haagen HH, t Hoen PA, Botelho Bovo A, de Morree A, van Mulligen EM, Chichester C, et al. Novel protein-protein interactions inferred from literature context. *PLoS One.* 2009;4(11):e7894.