Discovering Outliers of Potential Drug Toxicities Using a Large-scale Data-driven Approach

Jake Luo[1–3] and Ron A. Cisler[1–5]

[1]Center for Biomedical Data and Language Processing, Department of Health Informatics and Administration, University of Wisconsin–Milwaukee, Milwaukee, WI, USA. [2]Department of Health Informatics and Administration, College of Health Sciences, University of Wisconsin–Milwaukee, Milwaukee, WI, USA. [3]College of Health Sciences, University of Wisconsin–Milwaukee, Milwaukee, WI, USA. [4]Center for Urban Population Health, Milwaukee, WI, USA. [5]Joseph J. Zilber School of Public Health, University of Wisconsin–Milwaukee, Milwaukee, WI, USA.

ABSTRACT: We systematically compared the adverse effects of cancer drugs to detect event outliers across different clinical trials using a data-driven approach. Because many cancer drugs are toxic to patients, better understanding of adverse events of cancer drugs is critical for developing therapies that could minimize the toxic effects. However, due to the large variabilities of adverse events across different cancer drugs, methods to efficiently compare adverse effects across different cancer drugs are lacking. To address this challenge, we present an exploration study that integrates multiple adverse event reports from clinical trials in order to systematically compare adverse events across different cancer drugs. To demonstrate our methods, we first collected data on 186,339 clinical trials from ClinicalTrials.gov and selected 30 common cancer drugs. We identified 1602 cancer trials that studied the selected cancer drugs. Our methods effectively extracted 12,922 distinct adverse events from the clinical trial reports. Using the extracted data, we ranked all 12,922 adverse events based on their prevalence in the clinical trials, such as nausea 82%, fatigue 77%, and vomiting 75.97%. To detect the significant drug outliers that could have a statistically high possibility of causing an event, we used the boxplot method to visualize adverse event outliers across different drugs and applied Grubbs' test to evaluate the significance. Analyses showed that by systematically integrating cross-trial data from multiple clinical trial reports, adverse event outliers associated with cancer drugs can be detected. The method was demonstrated by detecting the following four statistically significant adverse event cases: the association of the drug axitinib with hypertension (Grubbs' test, $P < 0.001$), the association of the drug imatinib with muscle spasm ($P < 0.001$), the association of the drug vorinostat with deep vein thrombosis ($P < 0.001$), and the association of the drug afatinib with paronychia ($P < 0.01$).

KEYWORDS: drug toxicity, cancer informatics, clinical trial, adverse events, outlier discovery, big clinical data

CITATION: Luo and Cisler. Discovering Outliers of Potential Drug Toxicities Using a Large-scale Data-driven Approach. *Cancer Informatics* 2016:15 211–217 doi: 10.4137/CIN.S39549.

TYPE: Original Research

RECEIVED: April 04, 2016. **RESUBMITTED:** September 25, 2016. **ACCEPTED FOR PUBLICATION:** October 03, 2016.

ACADEMIC EDITOR: J. T. Efird, Editor in Chief

PEER REVIEW: Three peer reviewers contributed to the peer review report. Reviewers' reports totaled 726 words, excluding any confidential comments to the academic editor.

FUNDING: UWM College of Health Sciences SEED Fund. The authors confirm that the funder had no influence over the study design, content of the article, or selection of this journal.

COMPETING INTERESTS: Authors disclose no potential conflicts of interest.

CORRESPONDENCE: jakeluo@uwm.edu

Introduction

Anticancer drug treatments normally have a higher risk of adverse events[1] compared with noncancer drug treatments. For example, in chemotherapy, one of the most commonly used methods for treating cancer, agents control cells that divide rapidly,[2] which is the major characteristic of cancer cells. However, many chemotherapy agents can produce severe toxicities,[3–6] such as gastrointestinal toxicity,[7] cardiovascular toxicity,[8,9] and nephrotoxicity (renal toxicity).[10,11] The toxicity effect leads to a wide range of side effects,[12,13] such as decreased production of blood cells, suppression of the normal immune system, hair loss, and bleeding.

To better understand the adverse effects of various cancer drugs, several clinical trial studies have focused on analyzing adverse event patterns. For example, Larrar et al.[14] discovered severe hematological side effects from rituximb treatments in children when the drug is used to treat autoimmune diseases. Norden et al.[15] studied the toxicity of bevacizumab when treating patients with recurrent malignant gliomas, finding common adverse events such as nausea and vomiting and severe events such as hemorrhage, proteinuria, and thromboembolic complications. Andritsos et al.[16] discovered that higher doses of lenalidomide cause life-threatening toxicity in patients with chronic lymphocytic leukemia. Furthermore, several studies have analyzed a specific type of adverse event, such as the intravitreal toxicity of bevacizumab studied by Manzano et al.[17], but most of these clinical studies focused on a specific drug or type of event. Due to the complexity of clinical trials and the challenges in recruiting patients, few studies have been designed to systematically analyze significant adverse events across a large number of cancer therapy agents. In this study, we aim to address this gap by developing a new large-scale, data-driven informatics method to help investigators systematically explore and analyze adverse events in cancer treatments. This systematic analysis complements existing analysis methods for clinical research and offers many potential clinical applications, such as detecting significant side effects of drug for postmarketing monitoring and providing evidence for comparing cancer therapies across different drugs.

In our exploratory study, we developed a method to extract and formalize adverse event data from multiple clinical trial reports for cross-trial analysis. An informatics pipeline was designed and built for clinical investigators to systematically analyze the adverse events using existing clinical study outcomes. To demonstrate our method, we studied the adverse events of 30 cancer therapy agents using data that were automatically extracted from clinical trial reports on ClinicalTrials.gov. The adverse event results were identified and integrated from trial reports. Using the data, we summarized the prevalence of adverse events across different study drugs. We conducted an analysis to compare and rank the adverse event incidences using the extracted data. The results show that the method provided an effective way to discover significant adverse event outliers associated with cancer therapies.

Methods

We first collected data on 186,339 clinical trials from ClinicalTrials.gov. We then conducted a study by using the 1602 cancer clinical trials that targeted 30 common cancer drugs such as bevacizumab, imatinib, lenalidomide, and pemetrexed. The selected cancer drugs included the top eight most commonly used chemotherapies; the complete drug list is summarized in Table 3. A parser was developed to traverse and extract data from the clinical trial reports. From these reports, we extracted the clinical trial title, target condition, recruitment location, and adverse events. To recognize medical concepts and standardize terminologies in the text reports, the extracted data elements were mapped to the Unified Medical Language System (UMLS).[18] For example, *breast cancer* was mapped to *CUI: C0006142* and *ST: Neoplastic Process*, and the drug *vorinostat* was mapped to UMLS concept *CUI: C0672708* and *ST: Pharmacologic Substance*. All the data were stored in a Hadoop-based cloud computing platform[19] for parallel big data retrieval and analysis. The data platform provides a distributed storage of the data that allows us to examine multiple drugs simultaneously. We conducted the following three different exploratory analyses on the extracted data: (1) prevalence and incidence analysis of cross-trial adverse events, (2) ranking analysis of event–drug association, and (3) outlier analysis of event–drug.

Results

Summary of data elements. Table 1 summarizes the study data in the Hadoop data warehouse. The adverse event data table is the main focus of this study, which stores 12,922 distinct adverse events. This data table contained the event name, UMLS concept of the event, number of affected subjects, number of at-risk subjects, and event type (eg, serious event and nonserious event). Other data tables, including 1602 clinical trial descriptions, 30 selected cancer drugs, and 1989 cancer disease conditions, were linked to the adverse event data table through their unique trial reference keys. The Hadoop data warehouse not only stored the adverse event data in a structured format but also provided parallel access to the data elements for data mining analysis.

Table 1. Data extraction and summary.

DATA ELEMENTS	EXAMPLE	DISTINCT DATA SUMMARY
Trials	NCT00403754, NCT00594464, NCT00831701	1,602 (trials)
Cancer Drugs	Afatinib, cyclophosphamide, capecitabine	30 (drugs)
Cancer Conditions	Lung neoplasm, pancreatic cancer, carcinoma	1,989 (conditions)
Adverse Events	Cough, nausea, diarrhea, deep vein thrombosis	12,922 (events)

High prevalence and incidence adverse events. Understanding the prevalence and incidence of adverse events can provide a useful reference for conducting clinical studies and monitoring the postmarketing of toxic drug effects.[20,21] Table 2 shows the top 30 adverse events according to the ranking of trial prevalence. The *trial prevalence rate* of an adverse event in this paper is defined as the proportion of trials that reported the event. It was calculated as the percentage of trials that contained the adverse event among the 1602 trials analyzed. The *subject incidence rate* was calculated by the number of affected patients divided by the number of patients at risk for the event, revealing the probability of occurrence of an adverse event in the trial population.

To the best of our knowledge, there is a lack of large-scale systematic analysis on the prevalence and incidence of the adverse events. Therefore, our study complements existing toxicity research for cancer drugs, providing a fundamental baseline to understand the common events. If an adverse event had a high prevalence, it meant that the event was more common among different drugs. For example, nausea was the top adverse event among trials, with a very high prevalence at 82.77%, followed by fatigue at 77.34%, vomiting at 75.97%, constipation at 72%, and cough at 63%. All the top five adverse events had a prevalence of greater than 60% among trials and an incidence rate of greater than 10% among patients. We also calculated the incidence rates of all the adverse events. High incidence of an adverse event indicated that the risk of observing the event on a patient was high. For example, among the top 30 high-prevalence events, alopecia (hair loss) had the highest incidence rate at 26.43%, which is higher than nausea at 23.17%; even the prevalence of alopecia among trials was about half of nausea. This indicated that if patients are exposed to drugs that cause alopecia, the likelihood of observing the alopecia event was high.

Average adverse event incidence rate per cancer drug. To compare adverse event risks across different cancer drugs, we compared the summarized incident rate of the 30 selected drugs. Table 3 shows the individual event incidence rate for

Table 2. Top 30 ranking of adverse events based on the prevalence analysis.

PREVALENCE RANK	ADVERSE EVENTS	AT-RISK SUBJECTS	AFFECTED SUBJECTS	INCIDENCE RATE %	AFFECTED TRIALS	TOTAL TRIALS	PREVALENCE RATE %
1	Nausea	73339	316530	23.17%	1326	1602	82.77%
2	Fatigue	68739	280228	24.53%	1239	1602	77.34%
3	Vomiting	43863	323263	13.57%	1217	1602	75.97%
4	Constipation	36613	256170	14.29%	1167	1602	72.85%
5	Cough	24355	208757	11.67%	1022	1602	63.80%
6	Insomnia	17435	161872	10.77%	915	1602	57.12%
7	Dizziness	14501	215821	6.72%	905	1602	56.49%
8	Dehydration	9068	207805	4.36%	900	1602	56.18%
9	Headache	22019	213857	10.30%	881	1602	54.99%
10	Anorexia	21452	133662	16.05%	824	1602	51.44%
11	Hypertension	16651	176491	9.43%	782	1602	48.81%
12	Abdominal pain	16065	228908	7.02%	770	1602	48.06%
13	Alopecia	36478	138035	26.43%	739	1602	46.13%
14	Back pain	13608	200385	6.79%	727	1602	45.38%
15	Neutropenia	41805	225771	18.52%	719	1602	44.88%
16	Rash	28990	170759	16.98%	716	1602	44.69%
17	Diarrhea	18121	91457	19.81%	666	1602	41.57%
18	Hypotension	3961	142941	2.77%	659	1602	41.14%
19	Febrile neutropenia	7279	160411	4.54%	654	1602	40.82%
20	Pyrexia	19604	242352	8.09%	637	1602	39.76%
21	Arthralgia	16448	175291	9.38%	629	1602	39.26%
22	Dry skin	9637	105717	9.12%	614	1602	38.33%
23	Dyspepsia	11185	136388	8.20%	614	1602	38.33%
24	Pain	6859	121076	5.67%	590	1602	36.83%
25	Pruritus	9835	118451	8.30%	588	1602	36.70%
26	Pneumonia	3849	143151	2.69%	577	1602	36.02%
27	Leukopenia	18443	164917	11.18%	568	1602	35.46%
28	Myalgia	13196	133999	9.85%	564	1602	35.21%
29	Anaemia	22876	205572	11.13%	558	1602	34.83%
30	Urinary tract infection	5426	154088	3.52%	554	1602	34.58%

the 30 selected drugs, which were ranked by the incidence rate. There was a significant difference among the cancer drugs: the event incidence rates range from vorinostat at 12.41% to lenalidomide at 3.20%. The higher incidence rate of a drug indicated that adverse events were more likely to be observed when a drug was used on a patient. The analysis was a summarized estimation of the total risk of adverse events when administering a drug to a patient. For example, when designing chemotherapy treatment for a patient with breast cancer, a doctor may use a combination of capecitabine and cyclophosphamide. If adverse events were an important factor to consider for the treatment,[22] eg, treating a weak patient, a higher dose of capecitabine could be combined with a lower dose of cyclophosphamide, because the average event incidence rate of capecitabine was more than 40%, which was less than that of cyclophosphamide. Combining high- and low-toxicity drugs to create a therapy could lead to a better-tolerated treatment plan.[23] To further design a better treatment strategy, a clinical investigator may need to compare a specific adverse event across several different drugs. In the next section, we discuss specific adverse events and analyze their potential risks.

Association analysis between drugs and adverse events. To compare the association of an adverse event across different drugs, we used the Apriori[24] association mining method to extract significant drug–event pairs from the clinical trial reports. We excluded low-quality adverse event cases where the at-risk patient count is less than 5 patients and the affected rate is 100%. These trials are usually small and provide little statistical power to the analysis. Figure 1A–D shows the

Table 3. Average incidence rate of adverse event per cancer drug.

INCIDENCE RANK	THERAPY DRUG	EVENT CASES #	AT-RISK CASES #	INCIDENCE RATE %	TRIAL #	INCIDENCE RANK	THERAPY DRUG	EVENT CASES #	AT-RISK CASES #	INCIDENCE RATE %	TRIAL #
1	Vorinostat	12195	98224	12.42%	43	16	Prednisone	90860	1870212	4.86%	71
2	Cyclophosphamide	162748	1899234	8.57%	154	17	Carboplatin	250959	5303753	4.73%	187
3	Ixabepilone	27845	353250	7.88%	29	18	Afatinib	32932	703933	4.68%	27
4	Gemcitabine	116824	1640646	7.12%	156	19	Paclitaxel	248673	5318996	4.68%	169
5	Cisplatin	182902	2726079	6.71%	171	20	Cetuximab	86734	1869277	4.64%	80
6	Bortezomib	31745	482677	6.58%	70	21	Temozolomide	43781	960300	4.56%	62
7	Sorafenib	87235	1384384	6.30%	85	22	Imatinib	55634	1245010	4.47%	55
8	Rituximab	57745	967078	5.97%	105	23	Docetaxel	252105	5707316	4.42%	167
9	Ofatumumab	9691	163157	5.94%	16	24	Sunitinib	89450	2105188	4.25%	88
10	Axitinib	24588	427977	5.75%	22	25	Oxaliplatin	105852	2626162	4.03%	90
11	Bevacizumab	235655	4205831	5.60%	205	26	Degarelix	11395	288061	3.96%	25
12	Topotecan	8493	156510	5.43%	23	27	Erlotinib	78284	2211594	3.54%	91
13	Pemetrexed	76534	1437851	5.32%	104	28	Trastuzumab	75318	2179633	3.46%	57
14	Dasatinib	33876	654005	5.18%	51	29	Lapatinib	82220	2564740	3.21%	54
15	Capecitabine	117388	2287261	5.13%	113	30	Lenalidomide	56585	1766380	3.20%	63

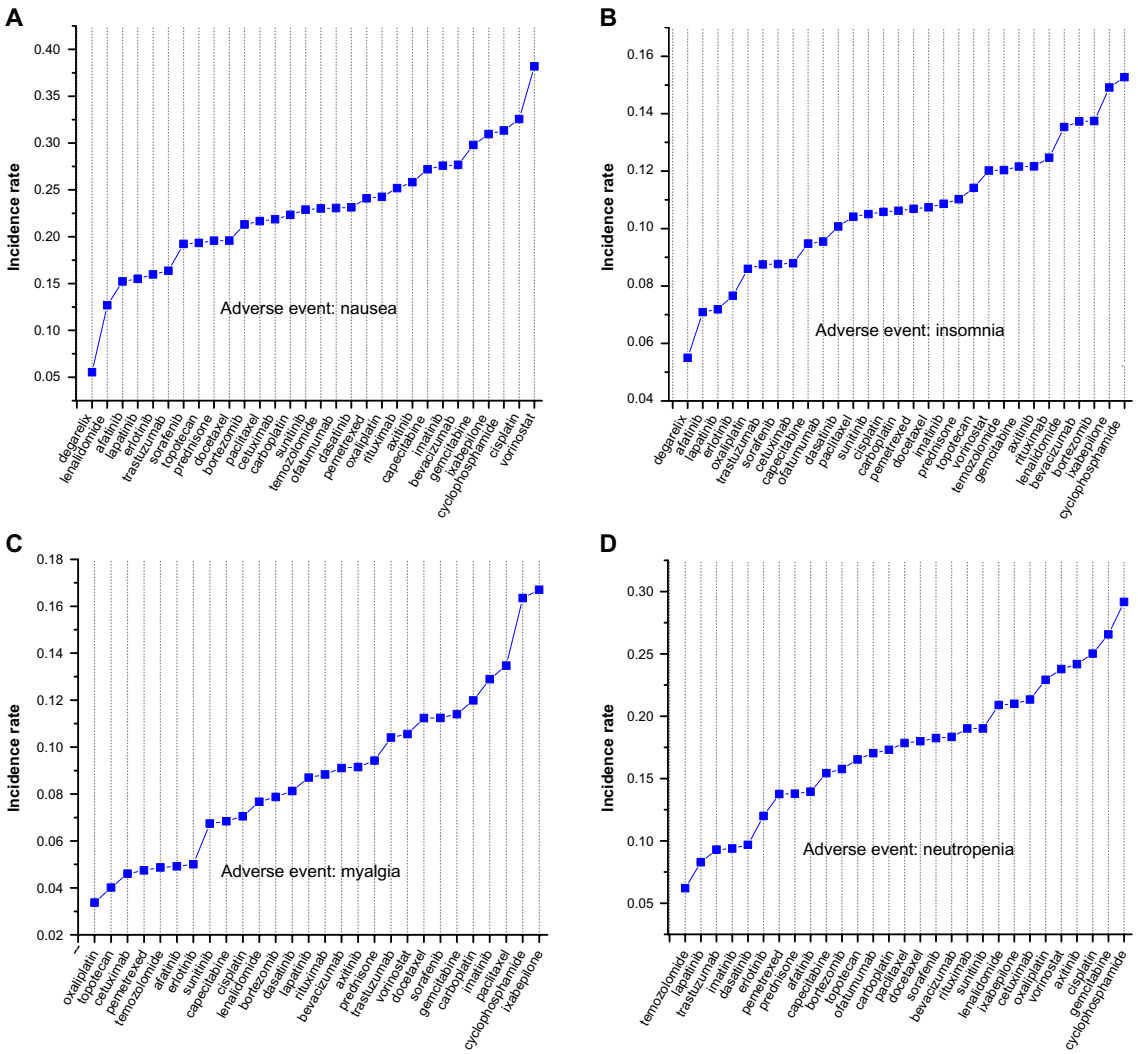

Figure 1. Comparison of adverse event associations across different cancer drugs. (**A**) nausea event; (**B**) insomnia event; (**C**) myalgia; (**D**) neutropenia.

ranking of four adverse events across 30 cancer drugs based on the incidence rate of the events. Given an adverse event, there is a significant difference across the 30 cancer drugs in terms of adverse event incidences. For example, in Figure 1A and B, we can see that degarelix had many fewer nausea and insomnia events than other drugs. Degarelix was a hormonal therapy normally used for prostate cancer treatment, and it is well tolerated. Comparing the nausea events between degarelix (5.53%) and vorinostat (38.19%), vorinostat was significantly more likely (6.9 times) to cause nausea. Some cancer drugs generally have higher incidences of adverse events. For example, cyclophosphamide had high number of insomnia (15.27%) and neutropenia (29.16%) events, and it was also among the top three for nausea (31.34%) and myalgia (16.35%). Comparing the incidence variances of adverse events among cancer drugs is crucial for designing personalized therapy for patients. For example, studies[25,26] found that when a cancer drug causes the neutropenia event, the patient is very likely to develop bacteremia. Therefore, if a cancer drug has high possibility of inducing neutropenia (eg, cyclophosphamide and cisplatin, Fig. 1D), the treatment plan should consider the use of colony-stimulating factors[26] and antibiotics.[25]

The results of our work show that, given an adverse event, the incidence rates across different drugs can be significant. Our method provides an effective way to compare adverse events across multiple drugs by systematically combining evidences from multiple trials.

Detecting significant drug–event association outliers. The previous analysis helped us rank and compare adverse event incidence across different cancer drugs, and we determined that the variance of adverse event incidence could be significant. Based on this observation, we hypothesized that there could be adverse event drugs that are statistically associated with some adverse events when compared with other drugs. Here, we explored a method to visualize and identify significant outliers of drugs that could cause an adverse event. In the data we extracted, we found that for a given drug, there could be many trials conducted to evaluate it. We grouped together trials that tested the same drug and then compared them with other drug groups. To visualize and compare the outlier groups of drugs, the boxplot[27] method was used to examine adverse events across different drugs. The boxplot shows the statistical results of an adverse event among a drug group, including the mean and median; 75th percentile, 25th percentile, 95th percentile, and 5th percentile; and the maximum and minimum values of the incidence. The boxplot provides an effective and intuitive way to estimate the variation and dispersion of drug adverse events. In this study, when the boxplot showed a potential outlier, we further calculated the statistical significance of the outlier using Grubbs' test.

Figure 2 shows four examples of drug–adverse event association outliers. The first example (Fig. 2A) shows that axitinib had a high possibility of inducing hypertension in cancer patients. The event distribution of axitinib was significantly higher than that of other drugs. Previous studies[28,29] have analyzed the impact of axitinib on blood pressure. Using our data, we found that the association significance was $P < 0.001$ using Grubbs' test. Based on these analyses, when using axitinib to treat carcinoma, it is recommended to closely monitor changes in blood pressure. Figure 2B shows the plot of a serious adverse event, deep vein thrombosis (DVT). The results showed that vorinostat has a significant higher degree ($P < 0.001$) of association with DVT. We verified this on the Food and Drug Administration (FDA) drug label, which identifies DVT in the warning and precaution section of adverse event. Figure 2C shows that the adverse event muscle spasms was significantly associated with imatinib. Although the pathophysiology of muscle spasms when using imatinib is not clear, several clinical studies have reported that 20%–40% of patients experienced musculoskeletal effects.[30,31] The last boxplot, Figure 2D, shows that when using afatinib, the possibility of observing the paronychia event was significantly higher than that using other drugs. Lacouture et al.[32] reported that afatinib is frequently associated with dermatologic adverse events. Most of the dermatologic events have a small impact, such as itching and pain; however, paronychia is one of the more serious dermatologic adverse events that can have a significant impact on a patient's well-being. To help lower this risk, patient education and proactive treatment interventions are required.

Discussion

In this study, we proposed a large-scale, systematic approach to analyze and compare the adverse effects of cancer drugs. Using clinical trial reports to extract adverse events from clinical studies, we showed that integrating large amounts of clinical trial data can effectively detect significant adverse events from cancer drugs. Clinical trials are the gold standard for evaluating the safety of drugs, and clinical trial results are valuable resources for clinical research and practice. However, conducting clinical trial studies is expensive and slow: a typical clinical trial could cost millions of dollars within five years.[33,34] Sometimes, even after a trial is completed, clinical investigators still face the challenge of not having enough statistical power to support the analysis of drug toxicity and adverse events. A common way to address this problem is to use meta-analysis[35] to enhance the statistical power.

Meta-analysis aims to aggregate data from multiple clinical trials to test a hypothesis. In a meta-analysis, the investigator combines the results from multiple clinical trials to conduct a statistical analysis, which could provide greater information for evaluating drug toxicity. For example, Silva et al.[36] combined 18 trials to analyze the statin-related adverse events. They manually reviewed the data on 18 trials and applied the Fisher's test to find significant adverse events across trials. Meta-analysis could improve estimation of the effect and reduce the uncertainty of clinical studies. However, meta-analysis is not immune to human bias,[37] and the method can be applied only to a limited number of trials, because it is a labor-intensive process that requires

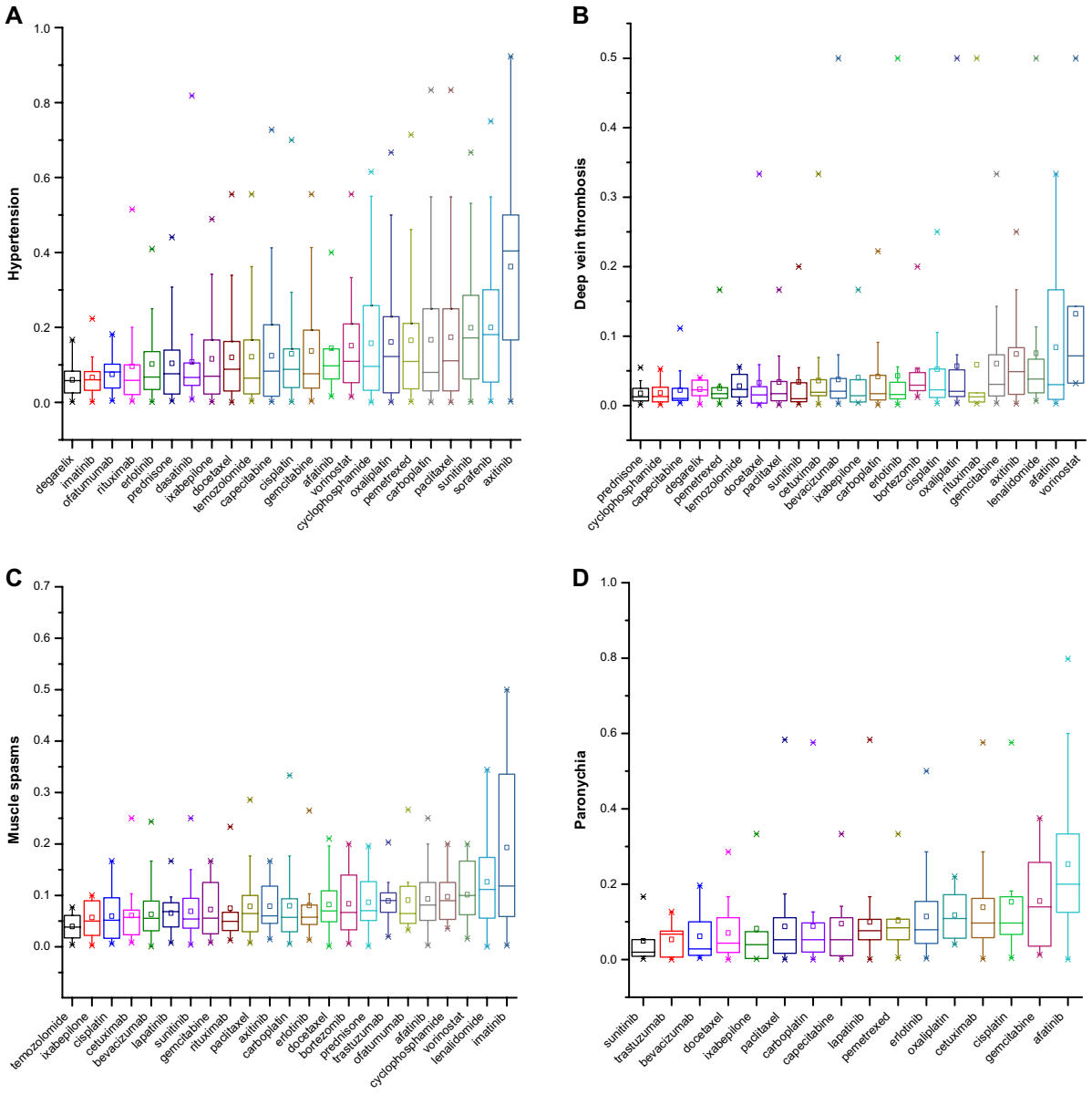

Figure 2. Adverse event outliers when comparing cancer drugs (**A**) hypertension; (**B**) deep vein thrombosis; (**C**) muscle spasms; (**D**) paronychia.

a high level of domain expertise. In this study, we proposed a new, data-driven approach to complement the analysis of adverse events for clinical research. By systematically integrating large numbers of clinical trial reports, we could summarize the prevalence of adverse events across different cancer drugs. We conducted exploratory studies to compare and rank the incidences of adverse events using the extracted data. In the "Results" section, we demonstrated that the method can effectively discover significant adverse event outliers for cancer drugs.

This is an exploratory study, and our method and analysis can be improved in several ways: (1) we did not use the extracted placebo results, which could have been used to establish a baseline standard to discover adverse event outliers. (2) We excluded therapies that use multiple drugs, which helped to reduce the noise of the signal. However, analyzing the combinational effect of drugs is an important topic, and analyzing drug combination therapies can be a future work. (3) About 2% of the adverse event report data contains complex elements that cannot be directly mapped to the UMLS concepts, such as "infection without neutropenia, nasal pharynx" and "late radiotherapy toxicity: subcutaneous tissue (within radiotherapy field)". For these cases, if part of the string can be mapped, we will use the first recognizable string as the adverse event, such as *radiotherapy toxicity* (CUI:C1298616) in the second example. To improve extraction performance, in a future study, we intend to develop a specified parser to extract sophisticated adverse event statements. (4) There could be other factors that associated with adverse events of cancer drugs, such as targeted disease and participant age. We focus on the association between drugs and observed adverse event in this study. In the next step, we will stratify the data to analyze other potential factors associated with adverse events.

Conclusion

We proposed a method to support the outlier detection of adverse events in cancer clinical trials. We used a data-driven approach to synthesize clinical trial results that studied cancer therapy agents. Among the retrieved 186,339 clinical trial data, we focused on 1602 cancer trials that studied 30 cancer drugs. From the trial data, 12,922 distinct adverse events were extracted. We conducted a systematic analysis to rank all the 12,922 adverse events based on their prevalence in trials, such as nausea 82%, fatigue 77%, and vomiting 75.97%. To demonstrate the effect of finding significant adverse events among cancer drugs, we used the boxplot method to visualize adverse event outliers across different drugs. We showed that by systematically integrating clinical trial reports, significant adverse event outliers associated with cancer drugs can be detected. The method is demonstrated by detecting the following four statistically significant adverse event cases: the association of the drug axitinib with hypertension (Grubbs' test, $P < 0.001$), the association of the drug imatinib with muscle spasm ($P < 0.001$), the association of the drug vorinostat with DVT ($P < 0.001$), and the association of the drug afatinib with paronychia ($P < 0.01$). The results show that the outlier detection method is effective for the associations of significant adverse event with cancer drugs.

Author Contributions

Conceived and designed the experiments: JL. Analyzed the data: JL, RAC. Wrote the first draft of the manuscript: JL, RAC. Contributed to the writing of the manuscript: JL, RAC. Agree with manuscript results and conclusions: JL, RAC. Jointly developed the structure and arguments for the paper: JL, RAC. Made critical revisions and approved final version: JL , RAC. Both authors reviewed and approved of the final manuscript.

The corresponding author JL had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

REFERENCES

1. Fernandez-Ortega P, Caloto M, Chirveches E, et al. Chemotherapy-induced nausea and vomiting in clinical practice: impact on patients' quality of life. *Support Care Cancer*. 2012;20(12):3141–8.
2. Malhotra V, Perry MC. Classical chemotherapy: mechanisms, toxicities and the therapeutc window. *Cancer Biol Ther*. 2003;2(suppl 1):1–3.
3. Hurria A, Togawa K, Mohile SG, et al. Predicting chemotherapy toxicity in older adults with cancer: a prospective multicenter study. *J Clin Oncol*. 2011;29(25):3457–65.
4. Repetto L. Greater risks of chemotherapy toxicity in elderly patients with cancer. *J Support Oncol*. 2003;1(4 suppl 2):18–24.
5. Frei E, Canellos GP. Dose: a critical factor in cancer chemotherapy. *Am J Med*. 1980;69(4):585–94.
6. Stam HJ, Challis GB. Ratings of cancer chemotherapy toxicity by oncologists, nurses, and pharmacists. *J Pain Symptom Manage*. 1989;4(1):7–12.
7. Mitchell EP. Gastrointestinal toxicity of chemotherapeutic agents. *Semin Oncol*. 1992;19(5):566–79. WB Saunders Co Independence Square West Curtis Center, Philadelphia, PA.
8. Curigliano G, Cardinale D, Suter T, et al. Cardiovascular toxicity induced by chemotherapy, targeted agents and radiotherapy: ESMO clinical practice guidelines. *Ann Oncol*. 2012;23(suppl 7):vii155–66.
9. Altena R, Perik PJ, Van Veldhuisen DJ, De Vries EG, Gietema JA. Cardiovascular toxicity caused by cancer treatment: strategies for early detection. *Lancet Oncol*. 2009;10(4):391–9.
10. Perazella MA, Moeckel GW. Nephrotoxicity from chemotherapeutic agents: clinical manifestations, pathobiology, and prevention/therapy. *Semin Nephrol*. 2010;30(6):570–81. Elsevier.
11. Skinner R, Pearson AD, Coulthard MG, et al. Assessment of chemotherapy-associated nephrotoxicity in children with cancer. *Cancer Chemother Pharmacol*. 1991;28(2):81–92.
12. Greene D, Nail L, Fieler V, Dudgeon D, Jones L. A comparison of patient-reported side effects among three chemotherapy regimens for breast cancer. *Cancer Pract*. 1993;2(1):57–62.
13. Lindley C, McCune JS, Thomason TE, et al. Perception of chemotherapy side effects cancer versus noncancer patients. *Cancer Pract*. 1999;7(2):59–65.
14. Larrar S, Guitton C, Willems M, Bader-Meunier B. Severe hematological side effects following rituximab therapy in children. *Haematologica*. 2006;91(8 suppl):ECR36.
15. Norden A, Young G, Setayesh K, et al. Bevacizumab for recurrent malignant gliomas efficacy, toxicity, and patterns of recurrence. *Neurology*. 2008;70(10):779–87.
16. Andritsos LA, Johnson AJ, Lozanski G, et al. Higher doses of lenalidomide are associated with unacceptable toxicity including life-threatening tumor flare in patients with chronic lymphocytic leukemia. *J Clin Oncol*. 2008;26(15):2519–25.
17. Manzano RP, Peyman GA, Khan P, Kivilcim M. Testing intravitreal toxicity of bevacizumab (Avastin). *Retina*. 2006;26(3):257–61.
18. Bodenreider O. The unified medical language system (UMLS): integrating biomedical terminology. *Nucleic Acids Res*. 2004;32(suppl 1):D267–70.
19. Shvachko K, Kuang H, Radia S, Chansler R. The hadoop distributed file system. *2010 IEEE 26th Symposium on Mass Storage Systems and Technologies (MSST)*. IEEE, Incline Village, NV, USA; 2010:1–10.
20. Fellay J, Boubaker K, Ledergerber B, et al. Prevalence of adverse events associated with potent antiretroviral treatment: Swiss HIV cohort study. *Lancet*. 2001;358(9290):1322–7.
21. Moore RA, McQuay HJ. Prevalence of opioid adverse events in chronic non-malignant pain: systematic review of randomised trials of oral opioids. *Arthritis Res Ther*. 2005;7(5):R1046.
22. Trotti A, Colevas AD, Setser A, et al. CTCAE v3. 0: development of a comprehensive grading system for the adverse effects of cancer treatment. *Semin Radiat Oncol*. 2003;13(3):176–81. Elsevier.
23. Ohno S, Mitsuyama S, Tamura K, et al. Dosage of capecitabine and cyclophosphamide combination therapy in patients with metastatic breast cancer. *Anticancer Res*. 2007;27(2):1009–13.
24. Agrawal R, Srikant R. Fast algorithms for mining association rules in large databases. *Proceedings of the 20th International Conference on Very Large Data Bases*. Morgan Kaufmann Publishers Inc, San Francisco, CA, USA; 1994:487–99. 672836.
25. Bochud P-Y, Eggiman P, Calandra T, Van Melle G, Saghafi L, Francioli P. Bacteremia due to viridans *Streptococcus* in neutropenic patients with cancer: clinical spectrum and risk factors. *Clin Infect Dis*. 1994;18(1):25–31.
26. Elting LS, Rubenstein EB, Rolston KVI, Bodey GP. Outcomes of bacteremia in patients with cancer and neutropenia: observations from two decades of epidemiological and clinical trials. *Clin Infect Dis*. 1997;25(2):247–59.
27. Williamson DF, Parker RA, Kendrick JS. The box plot: a simple visual method to interpret data. *Ann Intern Med*. 1989;110(11):916–21.
28. Qi WX, He AN, Shen Z, Yao Y. Incidence and risk of hypertension with a novel multi - targeted kinase inhibitor axitinib in cancer patients: a systematic review and meta-analysis. *Br J Clin Pharmacol*. 2013;76(3):348–57.
29. Rini BI, Schiller JH, Fruehauf JP, et al. Diastolic blood pressure as a biomarker of axitinib efficacy in solid tumors. *Clin Cancer Res*. 2011;17(11):3841–9.
30. Zekri JM, Robinson MH, Woll PJ. Relative hypocalcaemia and muscle cramps in patients receiving imatinib for gastrointestinal stromal tumour. *Sarcoma*. 2006;2006(1):48948.
31. Verweij J, Casali PG, Zalcberg J, et al. Progression-free survival in gastrointestinal stromal tumours with high-dose imatinib: randomised trial. *Lancet*. 2004;364(9440):1127–34.
32. Lacouture ME, Schadendorf D, Chu C-Y, et al. Dermatologic adverse events associated with afatinib: an oral ErbB family blocker. *Expert Rev Anticancer Ther*. 2013;13(6):721–8.
33. Glick HA, Doshi JA, Sonnad SS, Polsky D. *Economic Evaluation in Clinical Trials*. Oxford University Press; 2014.
34. Rubin RR, Fujimoto WY, Marrero DG, et al. The Diabetes Prevention Program: recruitment methods and results. *Control Clin Trials*. 2002;23(2):157–71.
35. DerSimonian R, Laird N. Meta-analysis in clinical trials. *Control Clin Trials*. 1986;7(3):177–88.
36. Silva MA, Swanson AC, Gandhi PJ, Tataronis GR. Statin-related adverse events: a meta-analysis. *Clin Ther*. 2006;28(1):26–35.
37. Jüni P, Witschi A, Bloch R, Egger M. The hazards of scoring the quality of clinical trials for meta-analysis. *JAMA*. 1999;282(11):1054–60.