The Oncologist, 2022, **27**, 434–440
https://doi.org/10.1093/oncolo/oyac059
Advance access publication 18 April 2022
Original Article

Signal Detection of Adverse Events Associated with Trastuzumab in a Cohort of Elderly Patients with Breast Cancer

Efstathia Polychronopoulou[1,2], Sharon H. Giordano[3], Lin-Na Chou[1,2], Xiaoying Yu[1,2], Yong-Fang Kuo[1,2,*]

[1]Office of Biostatistics, University of Texas Medical Branch, Galveston, TX, USA
[2]Preventive Medicine and Population Health, University of Texas Medical Branch, Galveston, TX, USA
[3]Department of Health Services Research, University of Texas MD Anderson Cancer Center, Houston, TX, USA
*Corresponding author: Yong-Fang Kuo, PhD, 700 Harborside Drive, Ewing Hall, 1.134, Galveston, TX 77555-1148, USA. Tel: +1 409 772 6355; Email: yokuo@utmb.edu

Abstract

Aim: Utilization of signal detection methods in longitudinal claims data can improve post-marketing drug surveillance, but to date there has been limited application. The aim of this study is to use 3 approaches, the proportional reporting ratio, Gamma Poisson Shrinker, and tree-based scan statistic in detecting adverse drug events (ADEs) attributed to trastuzumab using an administrative claims dataset.

Methods: Using data from the Texas Cancer Registry and SEER linked to Medicare from 2010 to 2013, we conducted 1:2 propensity score matching. Breast cancer HER2+ patients treated with trastuzumab in addition to standard chemotherapy were matched to HER2– patients treated with standard chemotherapy. Inpatient and outpatient encounters up to 6 months from start of therapy were used to identify adverse events.

Results: A total of 4191 patients were included in the study. Across all methods, use of trastuzumab generated signals on 9 distinct body systems. Cardiomyopathy and heart valve disease were the most consistently detected signals. Clinical review determined that most signals represented known ADEs.

Conclusions: We showed that claims data can be used to complement current ADE monitoring using common data mining methods with propensity score matching. Our analysis identified all expected ADEs associated with trastuzumab, and additional signals of valvular heart disorders.

Key words: signal detection; adverse drug reactions; drug safety; pharmacovigilance; breast cancer; trastuzumab.

Implications for Practice

Signal detection of adverse drug events is the first, yet fundamental, step in post-marketing drug surveillance. In this study, we used longitudinal insurance claims data to examine whether current data mining methods can adequately detect adverse event signals in the elderly patient with breast cancer. Based on our findings claims data can complement and expand current drug surveillance systems and provide more timely insight.

Introduction

The introduction of trastuzumab as a targeted therapy for HER2+ breast cancer has contributed to significant improvements in disease progression and survival in the past decade. However, the drug's toxicity profile differs from that of typical chemotherapy agents. Active surveillance of associated adverse reactions is imperative to prevent events that may compromise patients' quality of life.

Post-marketing adverse drug event (ADE) monitoring is an essential component of pharmacovigilance. Despite the rigorous process preceding drug approval, clinical trials have limited generalizability due to selective populations and relatively short follow-up times. Post-marketing surveillance evaluates drugs taken under "real-world" circumstances and is more likely to detect rare ADEs.

Current practice relies mainly on spontaneous reporting systems (SRS); while these datasets are useful tools for surveillance, since they include all marketed drugs and broad patient populations, concerns arise in regard to reporting bias, misattributed drug-event pairs, incomplete or duplicate reporting and lack of a control population. Longitudinal healthcare data, such as insurance claims, are more representative of routine healthcare and could enhance the current pharmacovigilance system. However, use of ADE signaling detection methods in such datasets is limited.[1-3] The present study aims to evaluate 3 data

Received: 21 March 2021; Accepted: 15 December 2021.

mining algorithms (DMAs) in detection of incident ADEs associated with trastuzumab in a cohort of elderly patients with breast cancer using an administrative claims dataset. Specifically, we compare the number of signals detected by each DMA that are considered related to trastuzumab and the number of likely false positives, when using 2 different approaches on outcome definition and 2 different signaling thresholds.

Materials and Methods

Source of Data

The linked Texas Cancer Registry (TCR) Medicare and Surveillance, Epidemiology, and End Results (SEER) Medicare databases were analyzed in this study.[4,5] The SEER program has been collecting information on newly diagnosed cancer cases in SEER registry areas since 1973. Currently SEER covers approximately 34.6% of the US population.[5] The TCR program was initiated in 1976 and follows the same collection and reporting requirements, but does not yet contribute to SEER, assuring that reported cases in each registry are unique. Approximately 120 000 new cases are reported annually in TCR and approximately 17 000 are breast cancer cases.[6] Medicare is the primary health insurance of approximately 98% of the elderly US population.[7] The linkage of the Medicare and cancer registries is based on matching a person's social security number, name, date of birth and sex. Linkage algorithms were developed and applied by the National Cancer Institute.

Study Cohort

We included females aged ≥ 66years diagnosed between 2010 and 2013, with either HER2+ breast cancer who received trastuzumab in addition to standard chemotherapy (exposed group) or females with HER2– cancer who received standard chemotherapy alone (unexposed group). Start of treatment was defined as the first day of trastuzumab use or the first day of chemotherapy depending on group. Patients were followed for ADEs for 6 months from the start of treatment. We required patients to have Medicare Part A and B coverage for the year prior to diagnosis and until the end of follow-up, as well as not to participate in Health Maintenance Organizations. Enrollment in Medicare Part D was additionally required during the follow-up period. Patients were excluded if the diagnosis originated from an autopsy or death certificate or they did not have a confirmed HER2 status at diagnosis. Patients that received pertuzumab or other adjuvant therapies in addition to trastuzumab were excluded. To increase comparability between HER2+ and HER2– patients, a propensity score for HER2+ was generated using a logistic regression model including age, race, cancer stage at diagnosis, presence of diabetes, and radiation therapy. Then, we used greedy nearest neighbor algorithm with a 0.025 caliper of the logit of the propensity score to match each HER2+ patient to 2 HER2– patients. HER2+ patients were not included if they did not match with exactly 2 HER2– patients.

Treatment with trastuzumab and/or chemotherapy was identified from Medicare claims using International Classification of Diseases 9th revision (ICD-9), Current Procedural Terminology and National Drug codes from the Inpatient, Outpatient, Carrier, Durable Medical Equipment, and Part D Prescriber Public Use files as previously described (Supplementary Table S1).[8]

Study Outcomes

All inpatient diagnoses and the primary diagnosis from outpatient claims after the start of treatment that were not present in the prior 6 months, were used to identify incident ADEs for each patient. We excluded secondary outpatient diagnoses from our analysis because these are commonly used to document past health history, chronic conditions, or diagnostic services required for reimbursement purposes. In contrast, secondary inpatient diagnoses were included as they reflect only those diseases that coexist at the time of admission or develop subsequently or affect the treatment received.[9] We excluded ICD-9 codes unlikely to be associated with trastuzumab use, such as radiation, breast cancer surgery, chemotherapy, diagnostic pathology, imaging, injuries, or congenital anomalies (Supplementary Tables S2 and S3). The remaining diagnosis codes were grouped into clinically meaningful categories using the Multi-Level Clinical Classification Software (MLCCS). The MLCCS is a categorization scheme developed by the Agency of Healthcare Research and Quality to organize ICD-9 codes by employing a hierarchical system with 4 levels. The first level consists of 18 body systems and each system has up to 3 sublevels, which consist of progressively more specific diagnosis definitions.

Two approaches were used in outcome definition. In the subject level approach, for each ADE we counted the number of subjects in the exposed and unexposed group with at least one claim for it during follow-up. In the encounter level approach, for each ADE we counted the number of associated encounters in each group during the follow-up period, allowing each subject to contribute more than once with the same diagnosis. Distinct encounters were defined as having either different date or different provider. We used 2 approaches because the same drug-ADE pair for the same patient may be reported multiple times from different sources in SRS. Therefore, the encounter level approach simulates the current SRS more closely than the subject level approach. Additionally, ADEs diagnosed in multiple visits for the same patient may indicate disease persistence or higher severity; the encounter level approach may therefore be more likely to detect rare, but persistent ADE signals. For sensitivity, we also considered a different definition of encounter where the same diagnosis was counted once per month at most.

DMA Methods

DMAs are automated methods to detect drug-event pairs with higher than expected frequencies that may warrant further investigation. We evaluated 3 DMAs to detect signals of ADEs; the proportional reporting ratio (PRR), the Gamma Poisson Shrinker (GPS), and a tree-based scan statistic (TBSS). PRR and GPS are well-validated methods that are easy to implement and interpret, whereas TBSS is a newer method that can additionally account for multiple comparisons.

The PRR is calculated as the ratio of the frequency of an ADE in the exposed group over the frequency of the same event in the unexposed.[10,11] Two commonly used signaling thresholds were assessed; a PRR with a lower 90% 2-sided CI (LCI) ≥2 and at least 3 cases was considered as a strict signal and a PRR ≥ 2, chi-square ≥ 4, and at least 3 cases as a moderate signal.[10,12]

The GPS method is designed to avoid spurious false positives due to small observed frequencies.[13] The method assumes that the observed frequency of a drug-event

combination follows the Poisson distribution. Expected frequencies are calculated from the marginal counts of a drug-event pair assuming independence between event and exposure. The GPS uses the Empiric Bayes Geometric Mean (EBGM) of the posterior distribution in lieu of the observed/expected ratio. Conventionally, a lower bound of the 90% 2-sided EBGM CI ≥2, denoted as EB05, is used as a signaling threshold.[1,12] In this study, we consider this cutoff as strict and a threshold of EB05 >1 with EBGM ≥1.5 as moderate.[1]

The TBSS signal detection method[14] requires ADEs to be classified as a tree-structure, where related diagnoses are closer together. The MLCCS scheme, used to categorize ICD-9 codes in 4 levels of granularity, provides this hierarchical structure. The TBSS simultaneously evaluates whether a group of closely related diagnoses (branch) has a higher than expected risk for the exposed population. The method adjusts for multiple testing due to overlapping diagnoses in the various evaluated branches. Inference is made with Monte-Carlo hypothesis testing. We used the unconditional Bernoulli TBSS, which is the recommended statistic for treatment comparisons among matched cohorts.[15] A *P*-value of <.05 with a relative risk (RR) ≥2 was considered as strict signal, whereas a $P < .05$ and an RR <2 as moderate.

Data management and analysis were conducted using SAS 9.4 (SAS Institute Inc., Cary, NC). For TBSS analysis the Tree Scan Software v1.4 (https://www.treescan.org) was used.[14] The Institutional Review Board at the University of Texas Medical Branch approved this study.

Results

A total of 12 133 patients (9722 in SEER-Medicare and 2411 in TCR-Medicare) met the selection criteria at time of diagnosis (Supplementary Figs. S1 and S2). During the 6 months follow-up period, 6661 of total patients satisfied the continuous enrollment criterion. Of those, 5213 received chemotherapy only (HER2– patients) and 1448 received trastuzumab as an adjuvant therapy (HER2+ patients). The mean age at diagnosis was 73 years old. The 2 cohorts differed in race,

Table 1. Patient characteristics before and after propensity score matching.

Variables	Before matching				After matching			
	HER2+ (*N* = 1448)	HER2– (*N* = 5213)	Absolute SMD	*P*-value	HER2+ (*N* = 1397)	HER2– (*N* = 2794)	Absolute SMD+	*P*-value
	Mean (SD)				Mean (SD)			
Age at diagnosis	73.2 (6.0)	73.2 (6.1)	.0004	.99	73.1 (5.9)	73.1 (5.9)	0.0169	.855
	Count (%)				Count (%)			
Race								
White	1248 (86.2)	4490 (86.1)	.1371	<.001*	1241 (88.8)	2492 (89.2)	0.0044	.895
Black	107 (7.4)	507 (9.7)			102 (7.3)	193 (6.9)		
Other	93 (6.4)	216 (4.2)			54 (3.9)	109 (3.9)		
Radiation								
Yes	569 (39.3)	2340 (44.9)	.1226	<.001*	562 (40.2)	1154 (41.3)	0.0052	.547
No	782 (54.0)	2542 (48.8)			741 (53.0)	1474 (52.8)		
Unknown	97 (6.7)	331 (6.3)			94 (6.7)	166 (5.9)		
Surgery								
Yes	1246 (86.1)	4563 (87.5)	.0595	.157	1208 (86.5)	2431 (87.0)	0.0293	.860
No	>191 (>13.2)	650 (12.5)			>178 (>12.7)	352 (12.6)		
Unknown	<11 (<0.8)	28 (.5)			<11 (<0.8)	11 (0.4)		
Charlson Comorbidity Index								
0	817 (56.4)	2826 (54.2)	.0457	.307	791 (56.6)	1614 (57.8)	0.0718	.572
1	348 (24.0)	1288 (24.7)			334 (23.9)	664 (23.8)		
2	138 (9.5)	575 (11.0)			133 (9.5)	274 (9.8)		
≥3	145 (10.0)	524 (1.1)			139 (10.0)	242 (8.7)		
Diabetes								
Yes	352 (24.3)	1397 (26.8)	.0571	.056	336 (24.1)	653 (23.4)	0.0189	.625
No	1096 (75.7)	3816 (73.2)			1061 (75.9)	2141 (76.6)		
Stage at diagnosis								
In situ/localized	704 (48.6)	2432 (46.6)	.0896	.018*	681 (48.8)	1393 (49.9)	0.0440	.820
Regional	531 (36.7)	2120 (4.7)			518 (37.1)	1025 (36.7)		
Distant	199 (13.7)	621 (11.9)			187 (13.4)	359 (12.9)		
Unstaged	14 (1.0)	40 (0.8)			11 (0.8)	17 (0.6)		

*Accepted significance level: $P < .05$.
SMD: standardized mean difference.

Table 2. Adverse event signals detected from subject- and encounter-level approach at 6 months follow-up.

MLCCS	Diagnosis	Subject level					Encounter level				
		N HER2+	N HER2 *N* =	PRR	EBGM	Tree-scan RR	N HER2+	N HER2 *N* =	PRR	EBGM	Tree-scan RR
3	**Endocrine; nutritional; metabolic diseases; immunity disorders**						1344	2503	.	.	1.08
3.8	Fluid and electrolyte disorders	312	495	.	.	1.18	433	688	.	.	1.17
3.8.2	Hypovolemia						245	358	.	.	1.22
4	**Diseases of the blood and blood-forming organs**										
4.1	Anemia	363	531	.	.	1.17	514	817	.	.	1.14
4.1.3.7	Anemia; unspecified	215	307	.	.	1.24					
5	**Mental illness**										
5.1	Adjustment disorders	15	11	2.55	.	.	16	13	2.22	.	.
6	**Diseases of the nervous system and sense organs**										
6.8	Other otitis media and related conditions						18	15	2.06	.	.
7	**Diseases of the circulatory system**	1053	1736	.	.	1.18	2555	3919	.	.	1.16
7.2	Diseases of the heart	734	937	.	.	1.29	1370	1871	.	.	1.23
7.2.1	Heart valve disorders	327	203	3.01	1.70	1.88	394	233	3.03	1.71	1.86
7.2.1.1	Chronic rheumatic disease of the heart valves	46	30	2.87	.	1.82	49	31	2.72	1.54	1.84
7.2.1.2	Nonrheumatic mitral valve disorders	181	89	3.80	1.71	2.01	191	93	3.53	1.78	2.02
7.2.1.3	Nonrheumatic aortic valve disorders	74	56	2.47	.	1.71	82	66	2.14		1.66
7.2.1.4	Other heart valve disorders	85	54	2.94	1.52	1.83	90	61	2.54	1.55	1.79
7.2.2	Peri-, endo-, and myocarditis; cardiomyopathy	92	48	3.58	1.73	1.99	104	62	3.01	1.60	1.89
7.2.2.1	Cardiomyopathy	77	33	4.36	1.68	2.10	86	41	3.60	1.74	2.03
7.2.2.2	Other peri-, endo-, and myocarditis	17	15	2.12	.	.					
7.2.7	Other and ill-defined heart disease	71	32	4.15	1.78	2.07	72	36	3.58	1.63	2.00
9	**Diseases of the digestive system**										
9.4.2.1	Gastric ulcer	<11	<11	2.80	.	.					
9.4.2.2	Duodenal ulcer	<11	<11	2.80	.	.	≥ 11	< 11	3.15	.	.
9.12	Other gastrointestinal disorders						281	405	.	.	1.22
9.12.3	Other/unspecified gastrointestinal disorders						182	245	.	.	1.28
13	**Diseases of musculoskeletal system and connective tissue**										
13.1	Infective arthritis/osteomyelitis (expect caused by TB or STD)	<11	<11	5.61	.	.	≥11	<11	10.84	1.51	2.57
13.6.2	Other acquired deformities	≥11	<11	2.43	.	.	≥11	<11	2.33	.	.

Table 2. Continued

MLCCS	Diagnosis	Subject level					Encounter level				
		N HER2+	N HER2 N =	PRR	EBGM	Tree-scan RR	N HER2+	N HER2 N =	PRR	EBGM	Tree-scan RR
13.8	Other connective tissue disease						430	663	.	.	1.18
16	**Injury and poisoning**	377	582	.	.	1.18	547	848	.	.	1.18
16.10	Complications						347	522	.	.	1.20
16.11	Poisoning	50	43	2.18	.	1.60	52	44	2.14	.	1.61
16.11.2	Poisoning by other medications and drugs	47	41	2.14	.	1.60	48	41	2.10	.	1.62
17	**Symptoms; signs; and ill-defined conditions and factors influencing health status ‘**										
17.1.6	Nausea and vomiting						161	215	.	1.17	1.28

Pink shading indicates strict-level signal. Yellow shading indicates moderate-level signal. Centered dot (.) indicates no signal detected.
Abbreviations: EBGM, Empiric Bayes Geometric Mean; MLCCS, Multi-level Clinical Classification System; PRR, Proportional Reporting Ratio; Tree-scan RR, relative risk estimated by TBSS.

receipt of radiation therapy, and cancer stage at diagnosis (Table 1). Using propensity score matching, 1397 (96.5%) of the HER2+ patients were matched to 2794 HER2– women consisting the exposed and unexposed groups, respectively. The 2 groups were balanced in demographic and treatment characteristics when evaluated with the absolute standardized mean difference (Table 1). In accordance with SEER-Medicare policies, all cells with a value <11 and cells that allow such values to be derived are either suppressed or indirectly reported.

Subject-Level Approach

At 6 months, we detected a total of 22 signals, across all methods and signaling thresholds (Table 2). Six of the 16 PRR and 3 out of 16 TBSS signals met the stricter threshold. PRR and TBSS signals were identified across 5 and 4 body systems respectively, of which only the circulatory system and poisoning were common. Heart-related signals were the only ones exceeding the stricter threshold. All 6 GPS signals were detected at the moderate signaling threshold and restricted to the circulatory system. The majority of identified signals, indicating higher than expected risk in the group that received trastuzumab, are known ADEs.[16-19] Others, such as poisoning, acquired deformities or adjustment disorders are likely false positives.

Encounter-Level Approach

A total of 27 signals were detected by either method involving 9 body systems (Table 2). Similarly to the subject-level approach, none of the 9 GPS signals met the strict criterion. The 15 PRR and 23 TBSS signals spanned across multiple body systems, with considerable agreement between them (Fig. 1), particularly for the circulatory, poisoning and musculoskeletal systems. Cardiomyopathy and heart disease, heart valve disorder and infective arthritis were the only signals detected at the strict threshold.

Overall, there were no substantial differences in the body systems and conditions signaled by either the subject or the encounter-level approach. Results remained similar when we employed a different definition for encounter (Supplementary Table S4). The use of a propensity score matched cohort substantially reduced the number of spurious false-positive signals detected (Supplementary Table S5).

Signal strengths varied depending on method and approach. For instance, cardiomyopathy signals ranged from 1.7 to 4.4 for the HER2+ group whereas mitral valve disorders ranged from 1.7 to 3.8.

Discussion

In oncology, co-administration of systemic and targeted therapies and patients' underlying health conditions can complicate the timely identification and accurate reporting of ADEs. Administrative claims data can complement the current pharmacovigilance practice by addressing fundamental limitations, such as underreporting, lack of proper control group or "real-world" circumstances. In this study, we used DMAs to identify toxicities associated with trastuzumab using Medicare claims data. The choice of the most suitable method and approach depends largely on study purpose and potential ADE severity (Supplementary Appendix). Overall, our techniques detected known side effects such as cardiotoxicity, but also identified diagnoses which have not been previously associated, such as valve disorders or duodenal ulcers.

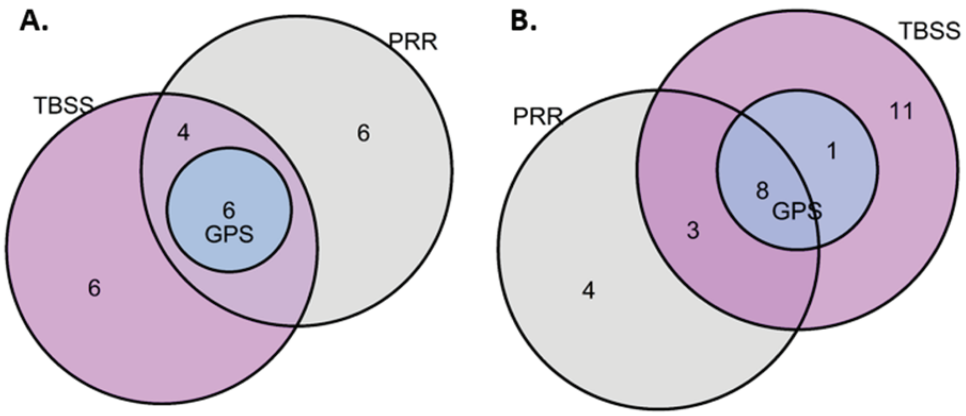

Figure 1. Number of detected signals at the moderate signaling threshold at (**A**) subject-level approach, (**B**) encounter-level approach.

Cardiotoxicity is a well-established ADE in patients who receive trastuzumab that more commonly manifests as cardiomyopathy.[17,20,21] In this study, heart-related toxicity was detected by all methods and approaches, albeit not always at the stricter signaling threshold. Heart valve disorders, particularly mitral, were also consistently signaled. Reports of mitral valve regurgitation potentially attributed to trastuzumab ar[22]e scarce, and no prior clinical trial has established any such association.[23,24] To our knowledge, no population-level study has examined whether trastuzumab increases the risk for valve disease, particularly in elderly patients. Instead, valvular disease in patients with breast cancer has been reported as an adverse effect of radiation therapy.[25] In our study however, the percent of patients that received radiation therapy was similar between the 2 matched groups and therefore cannot justify the detected signal in the trastuzumab group. Further study is needed to evaluate whether this finding is a true ADE of trastuzumab or a reflection of increased cardiac monitoring in the exposed group; regardless, clinicians should screen for and monitor the presence of valvular heart disease in patients receiving trastuzumab, as it can increase the risk for cardiotoxicity in this population.[22]

Several signals were detected in 8 body systems other than the circulatory. The majority of those signals are known ADEs of trastuzumab; anemia, gastrointestinal disorders, arthralgia, infections, or electrolyte disorders have been reported after prolonged trastuzumab use.[16,18,19] Otitis media has not been reported previously as an ADE of trastuzumab, it could, however, occur as a complication of other respiratory infections that are known to increase with this drug. Similarly, low tolerability to bacterial infections could exacerbate duodenal ulcer; duodenal ulcer had a low incidence rate (<1%) in this cohort, which would explain why it may have not been detected during clinical trials. Because targeted therapy is more likely than conventional chemotherapy to cause acute, non-specific, immune reactions,[26] complications such as ulcers or otitis media, while rare, are likely to be associated with trastuzumab. It is important that clinicians are aware of this safety risk to promote timely recognition and intervention.

Some well-established adverse events of trastuzumab, such as infusion reactions or fatigue, were not detected; this is expected however, as the comparator group received chemotherapy which is likely to present with similar reactions. The signals detected in this study therefore represent "excess" adverse events in addition to any chemotherapy toxicity. We also consistently detected signals for poisoning by other medications in the exposed group. This is unlikely to be associated with trastuzumab; it is unclear whether this group of ICD-9 codes "captures" true adverse events only or also includes cases of medication misuse. Overall, use of common data mining methods in administrative data resulted in few false-positive signals, while detecting all commonly anticipated adverse events.

There are several limitations in this study due to the nature of claims data, where billing errors or discrepancies cannot be accounted for; the same disease may be coded differently among providers. In addition, the use of MLCCS to classify related diagnoses may not be the most appropriate for this study. The sample size may have been inadequate for some rare ADEs to be detected.

In contrast to previous studies, we did not limit our analysis to one diagnosis per subject, since some of the trastuzumab-induced ADEs may manifest after a longer period, preceded by other acute events. Further, we did not pre-specify a set of ADEs, but allowed any possible event to be evaluated. In addition, we used propensity score matching to reduce the impact of potential confounding in detected signals. Wang et al have shown through simulations that the performance of propensity score matching with TBSS is comparable to that of full confounding adjustment.[15] Compared with typical pharmacovigilance studies, the use of "real-world" longitudinal data with propensity score matching is the major strength of our study.

Conclusion

Pharmacotherapy in oncology is continuously evolving with the production of new therapeutic drugs that act on specific molecular targets; expansion of current monitoring systems is of paramount importance to improve understanding of new toxicity profiles. In this study, we showed that longitudinal claims data can detect known ADE signals in patients with breast cancer with few or no false positives and could therefore complement and expand current drug surveillance systems. We detected previously unreported signals of heart valve disorders associated with use of trastuzumab, as well as rare cases of otitis media and duodenal ulcers, that require confirmation in future epidemiologic studies. Signal detection of ADEs is the first, yet fundamental, step in post-marketing drug surveillance; clinical review, and further epidemiologic study are warranted.

Acknowledgments

Cancer data have been provided by the Texas Cancer Registry, Cancer Epidemiology and Surveillance Branch, Texas Department of State Health Services, 1100 West 49th Street, Austin, TX 78756, https://www.dshs.texas.gov/tcr/ and the linked SEER-Medicare database. The interpretation and reporting of these data are the sole responsibility of the authors.

Conflict of Interest

The authors indicated no financial relationships.

Author Contributions

Conception/design: E.P., Y.-F.K., S.G. Provision of study material/patients: Y.-F.K., S.G. Collection and/or assembly of data: L.-N.C., E.P. Data analysis and interpretation: E.P., X.X., L.-N.C., Y.-F.K., S.G. Manuscript writing: E.P., X.X., L.-N.C., Y.-F.K., S.G. Final approval of manuscript: All authors.

Data Availability

The data underlying this article will be shared on reasonable request to the corresponding author.

Funding

This research was supported by the Cancer Prevention and Research Institute of Texas (RP160674), the Komen Foundation (SAC150061) and the National Insitutes for Health/National Cancer Institute (P30 CA016672).

References

1. Brown J, Petronis K, Bate A, et al. Drug adverse event detection in health plan data using the Gamma Poisson Shrinker and comparison to the tree-based scan statistic. *Pharmaceutics*. 2013;5(1):179-200. https://doi.org/10.3390/pharmaceutics5010179.
2. Curtis JR, Cheng H, Delzell E, et al. Adaptation of Bayesian data mining algorithms to longitudinal claims data: coxib safety as an example. *Med Care*. 2008;46(9):969-975. https://doi.org/10.1097/MLR.0b013e318179253b.
3. Schachterle SE, Hurley S, Liu Q, Petronis KR, Bate A. An implementation and visualization of the tree-based scan statistic for safety event monitoring in longitudinal electronic health data. *Drug Saf*. 2019;42(6):727-741. https://doi.org/10.1007/s40264-018-00784-0.
4. Texas Cancer Registry. Texas Department of State Health Services. 2019. Accessed October 1, 2019. https://www.dshs.texas.gov/tcr/home.aspx
5. Overview of the SEER Program. National Cancer Institute. 2019. Accessed June 14, 2019. https://seer.cancer.gov/about/overview.html
6. 2018 Texas Cancer Registry Annual Report. Texas Department of State Health Services. 2018. Accessed June 8, 2019. https://www.dshs.texas.gov/tcr/
7. Davis K, Schoen C, Bandeali F. Medicare 50 years of ensuring coverage and care. *The Commonwealth Fund* 2015. https://www.commonwealthfund.org/publications/fund-reports/2015/apr/medicare-50-years-ensuring-coverage-and-care
8. Huo J, Giordano SH, Smith BD, Shaitelman SF, Smith GL. Contemporary toxicity profile of breast brachytherapy versus external beam radiation after lumpectomy for breast cancer. *Int J Radiat Oncol Biol Phys*. 2016;94(4):709-718. https://doi.org/10.1016/j.ijrobp.2015.12.013.
9. *ICD-9-CM Official Guidelines for Coding and Reporting*. 2011. https://www.cdc.gov/nchs/data/icd/icd9cm_guidelines_2011.pdf
10. Evans SJW, Waller PC, Davis S. Use of proportional reporting ratios (PRRs) for signal generation from spontaneous adverse drug reaction reports. *Pharmacoepidemiol Drug Saf*. 2001;10(6):483-486. https://doi.org/10.1002/pds.677.
11. Wilson AM, Thabane L, Holbrook A. Application of data mining techniques in pharmacovigilance. *Br J Clin Pharmacol*. 2004;57(2):127-134. https://doi.org/10.1046/j.1365-2125.2003.01968.x.
12. Candore G, Juhlin K, Manlik K, et al. Comparison of statistical signal detection methods within and across spontaneous reporting databases. *Drug Saf*. 2015;38(6):577-587. https://doi.org/10.1007/s40264-015-0289-5.
13. Dumouchel W. Bayesian data mining in large frequency tables, with an application to the FDA spontaneous reporting system. *Am Stat*. 1999;53(3):177-190.
14. Kulldorff M, Fang Z, Walsh SJ. A tree-based scan statistic for database disease surveillance. *Biometrics* 2003;59(2):323-331. https://doi.org/10.1111/1541-0420.00039.
15. Wang SV, Maro JC, Baro E, et al. Data mining for adverse drug events with a propensity score-matched tree-based scan statistic. *Epidemiology* 2018;29(6):895-903. https://doi.org/10.1097/EDE.0000000000000907.
16. Goldhirsch A, Gelber RD, Piccart-Gebhart MJ, et al. 2 years versus 1 year of adjuvant trastuzumab for HER2-positive breast cancer (HERA): An open-label, randomised controlled trial. *Lancet* 2013;382(9897):1021-1028. https://doi.org/10.1016/S0140-6736(13)61094-6
17. Chavez-MacGregor M, Zhang N, Buchholz TA, et al. Trastuzumab-related cardiotoxicity among older patients with breast cancer. *J Clin Oncol*. 2013;31(33):4222-4228. https://doi.org/10.1200/JCO.2013.48.7884.
18. Cabiddu M, Guarneri P, Petrelli F, Barni S, Lonati V. The risk for anemia with targeted therapies for solid tumors. *Oncologist*. 2012;17(5):715-724. https://doi.org/10.1634/theoncologist.2012-0024
19. Al-Dasooqi N, Bowen JM, Gibson RJ, et al. Trastuzumab induces gastrointestinal side effects in HER2-overexpressing breast cancer patients. *Invest New Drugs*. 2009;27(2):173-178. https://doi.org/10.1007/s10637-008-9152-1.
20. Slamon DJ, Leyland-Jones B, Shak S, et al. Use of chemotherapy plus a monoclonal antibody against HER2 for metastatic breast cancer that overexpresses HER2. *N Engl J Med*. 2001;344(11):783-792. https://doi.org/10.1056/NEJM200103153441101.
21. Nemeth BT, Varga ZV, Wu WJ, Pacher P. Trastuzumab cardiotoxicity: from clinical trials to experimental studies. *Br J Pharmacol*. 2017;174(21):3727-3748. https://doi.org/10.1111/bph.13643
22. Sato A, Yoshihisa A, Miyata-Tatsumi M, et al. Valvular heart disease as a possible predictor of trastuzumab-induced cardiotoxicity in patients with breast cancer. *Mol Clin Oncol*. 2019;10(1):37-42. https://doi.org/10.3892/mco.2018.1764.
23. Günaydin ZY, Bektaş O, Çiltaş A, Karagöz A, Kaya A. Does Trastuzumab cause permanent severe mitral regurgitation?. *Int J Cardiol*. 2015;182(C):8-9. https://doi.org/10.1016/j.ijcard.2014.12.108.
24. Lange SA, Ebner B, Wess A, et al. Echocardiography signs of early cardiac impairment in patients with breast cancer and trastuzumab therapy. *Clin Res Cardiol*. 2012;101(6):415-426. https://doi.org/10.1007/s00392-011-0406-0.
25. Yusuf SW, Sami S, Daher IN. Radiation-induced heart disease: a clinical update. *Cardiol Res Pract*. 2011;2011:1-9. https://doi.org/10.4061/2011/317659.
26. Kroschinsky F, Stölzel F, von Bonin S, et al. New drugs, new toxicities: severe side effects of modern targeted and immunotherapy of cancer and their management. *Crit Care*. 2017;21(89). https://doi.org/10.1186/s13054-017-1678-1.