参考資料1

PHARMACOEPIDEMIOLOGY AND DRUG SAFETY 2007; **16**: 1275–1284
Published online 22 October 2007 in Wiley InterScience (www.interscience.wiley.com) **DOI**: 10.1002/pds.1509

ORIGINAL REPORT

Early detection of adverse drug events within population-based health networks: application of sequential testing methods[†,‡]

Jeffrey S. Brown PhD[1,2*], Martin Kulldorff PhD[1], K. Arnold Chan MD, MPH, ScD[3,4], Robert L. Davis MD, MPH[5], David Graham MD[6], Parker T. Pettus MS[1,2], Susan E. Andrade ScD[2,7], Marsha A. Raebel PharmD[2,8], Lisa Herrinton PhD[2,9], Douglas Roblin PhD[2,10], Denise Boudreau PhD[2,11], David Smith PhD[2,12], Jerry H. Gurwitz MD[2,7], Margaret J. Gunter PhD[2,13] and Richard Platt MD, MSc[1,2]

[1]*Department of Ambulatory Care and Prevention, Harvard Medical School and Harvard Pilgrim Health Care, Boston, MA, USA*
[2]*The HMO Research Network Center for Education and Research in Therapeutics, Boston, MA, USA*
[3]*Harvard School of Public Health, Boston, MA, USA*
[4]*i3 Drug Safety, Waltham, MA, USA*
[5]*Immunization Safety Office, Centers for Disease Control and Prevention, Atlanta, GA, USA*
[6]*Office of Drug Safety, Food and Drug Administration, Rockville, MD, USA*
[7]*Meyers Primary Care Institute (University of Massachusetts Medical School, the Fallon Foundation, and Fallon Community Health Plan), Worcester, MA, USA*
[8]*Kaiser Permanente Colorado, Denver, CO, USA*
[9]*Kaiser Permanente Northern California, Oakland, CA, USA*
[10]*Kaiser Permanente Georgia, Atlanta, GA, USA*
[11]*Center for Health Studies, Group Health Cooperative, Seattle, WA, USA*
[12]*Kaiser Permanente Northwest, Portland, OR, USA*
[13]*Lovelace Clinic Foundation, Albuquerque, NM, USA*

SUMMARY

Purpose Active surveillance of population-based health networks may improve the timeliness of detection of adverse drug events (ADEs). Active monitoring requires sequential analysis methods. Our objectives were to (1) evaluate the utility of automated healthcare claims data for near real-time drug adverse event surveillance and (2) identify key methodological issues related to the use of healthcare claims data for real-time drug safety surveillance.

Methods We assessed the ability to detect ADEs using historical data from nine health plans involved in the HMO Research Network's Center for Education and Research on Therapeutics (CERT). Analyses were performed using a maximized sequential probability ratio test (maxSPRT). Five drug-event pairs representing known associations with an ADE and two pairs representing 'negative controls' were analyzed.

* Correspondence to: Dr J. S. Brown, Department of Ambulatory Care and Prevention, 133 Brookline Ave., 6th floor, Boston, MA 02215, USA. E-mail: jeff_brown@harvardpilgrim.org

[†]Dr Brown, Dr Kulldorff, Dr Davis, Dr Graham, Dr Raebel, Dr Boudreau, Dr Roblin, Dr Gurwitz, Dr Platt and Dr Gunter and Mr Pettus report no conflict of interest. Dr Chan, Dr Andrade, Dr Herrinton and Dr Smith report receiving industry research funding on issues unrelated to the study.

[‡]Dr Platt has included Figure 3 from this manuscript as an example of active surveillance in the following recent presentations: (1) IOM Forum on Drug Safety (12 March 2007), The Future of Drug Safety—Challenges for the FDA (Drug Safety Symposium, The National Academies, IOM), (2) IOM-FDA meeting (24 April, 2007), Emerging Safety Science: A Forum on Drug Discovery, Development and Translation and (3) Keynote address at AMIA workshop on drug safety (13 June 2007), American Medical Informatics Association, Invitational Conference on Drug Safety and Pharmacovigilance.

Results Statistically significant ($p < 0.05$) signals of excess risk were found in four of the five drug-event pairs representing known associations; no signals were found for the negative controls. Signals were detected between 13 and 39 months after the start of surveillance. There was substantial variation in the number of exposed and expected events at signal detection.
Conclusions Prospective, periodic evaluation of routinely collected data can provide population-based estimates of medication-related adverse event rates to support routine, timely post-marketing surveillance for selected ADEs.

KEY WORDS — adverse drug event; methodology; sequential analysis; drug safety surveillance; SPRT

Received 10 November 2006; Revised 29 August 2007; Accepted 7 September 2007

INTRODUCTION

Current prospective post-marketing drug monitoring in the U.S. relies principally on passive surveillance via MedWatch reports to the Adverse Event Report System (AERS).[1–3] Passive monitoring systems have well-recognized drawbacks, including underreporting, reporting bias, incomplete data and limited information on the exposed population and a lack of denominators, thereby making it difficult to know if the spontaneous reports represent an increase in incidence over baseline.[1,4] Additionally, passive surveillance systems cannot provide quantitative information about the frequency or relative risk of reported reactions. These problems are particularly troubling for reported adverse reactions that are also common occurrences in the absence of the drug exposure.

Active surveillance of health plans' populations may improve the timeliness of detection of adverse drug events (ADEs). To realize the full potential of prospective surveillance, the accumulated drug exposure and event experience should be evaluated as it accumulates. Frequent prospective monitoring requires new capacity for extracting information from healthcare data systems, for aggregating information from multiple sources and for analyzing this information in a manner that avoids problems associated with repeated statistical tests on the same data. The CDC sponsored Vaccine Safety Datalink (VSD) has described such a prospective monitoring system for adverse vaccine reactions, using data from eight health plans.[5,6] Applying this methodology to surveillance for ADEs is considerably more complicated than it is for vaccines, in part because many drug exposures are chronic or intermittent, in contrast to vaccines which are usually administered only once or twice. In addition, risk windows for drug exposure may vary considerably by drug type and length of exposure.

This report describes our application of sequential analysis within a well-defined population to detect ADE signals. The goals were to use historical data to (1) evaluate the utility of automated healthcare claims data for near real-time drug adverse event surveillance and (2) identify key methodological issues related to the use of healthcare claims data for real-time drug safety surveillance.

METHODS

Overview

We assessed the ability of sequential analysis to detect ADEs using historical data from nine health plans involved in the HMO Research Network's Center for Education and Research on Therapeutics (CERT).[7] Multiple drug-event pairs were selected for analysis. We define a 'signal' as a statistically significant association between a drug and selected diagnosis codes that requires further attention to determine causality.[8] Key findings regarding the performance of the methodology are reported and key methodological issues are discussed.

Study population and data source

The study cohort was drawn from health plan members who were enrolled at any time from 1 January 2000 to 31 December 2005 in one of the nine health plans involved in the HMO Research Network CERT. The nine health plans are located in different geographic regions across the U.S.

Each of the nine participating health plans created four datasets corresponding to demographic, health plan enrollment, dispensing and diagnosis information based on specifications provided by the study coordinating center for the period 1 January 2000 through 31 December 2005. The demographic file contained date of birth and sex and the enrollment file contained start and stop dates for health plan enrollment and an indicator for whether or not the

DOI: 10.1002/pds

member had drug coverage during the enrollment period. The dispensing file contained all dispensing records for the cohorts of interest and included dispense date, national drug code for the dispensing, units dispensed, days supplied and generic name. The diagnosis file contained records for all ambulatory and inpatient diagnoses recorded on health plan automated claims, including each ICD-9-CM code recorded, date of diagnosis and an indicator for whether the care was provided in an inpatient or outpatient setting.

All health plan members who had at least one membership period with drug coverage of greater than 270 days were included in the analyses. Membership gaps of 60 days or less were bridged to create continuous membership periods. For analytic simplicity, only the first valid membership period was used. The study was approved by the human subjects committees at each health plan.

Drug-event pairs and comparisons

We constructed seven drug-event pairs to assess the stability and performance characteristics of the methodology (Table 1). The drug-event pairs were selected by the authors in 2003 and 2004. Five of the drug-event pairs were selected as known associations between a drug and a serious ADE and two were selected as negative controls for which no association was expected. There were at least two comparison cohorts for each drug-event pair: (a) all health plan members who did not use the drug of interest (non-users) and (b) all health plan members who were incident users (defined below) of a pre-selected comparison drug or drug class. The comparison to non-users was included because active comparators are not commonly available in other safety surveillance and data mining activities (e.g., AERS analyses) and it is expected that future active surveillance studies might involve drugs with no relevant comparator.

Calculating observed and expected counts

We performed analyses that simulated monthly prospective surveillance. For each month the maximized sequential probability ratio test (maxSPRT) requires information about the number of observed adverse events during the month and the expected number under the assumption that the null hypothesis of no excess risk is correct.

Definition of incident exposure. This study focused on incident users of the drug of interest or the comparator agent.[9] Incident use was defined as a dispensing for which there was no other dispensing of the drug of interest or a comparator drug during the prior 181 days (i.e., 181-day exposure-free period). Members who failed to meet the incident dispensing criteria, either due to continuous drug exposure or insufficient membership time, were excluded from the relevant comparison drug analyses. Multiple incident dispensings during the membership period were allowed as long as the requirement of a 181-day exposure-free period was satisfied.

Definition of incident outcome. To be considered an ADE, the diagnosis code of interest was required to be assigned in an inpatient setting; this need not be a criterion for use of this technique, but was adopted for this demonstration to improve the likelihood of detection of serious occurrences. Additionally, these events had to meet our incident ADE criteria that we defined as having no observed inpatient or outpatient

Table 1. Listing of all comparisons included in the evaluation

Drug of interest	Drug comparators	Outcome	Definition of outcome*
Celecoxib	Diclofenac, naproxen	Acute myocardial infarction	Acute myocardial infarction: 410.xx
Rofecoxib	Diclofenac, naproxen	Acute myocardial infarction	Acute myocardial infarction: 410.xx
Valdecoxib	Diclofenac, naproxen	Acute myocardial infarction	Acute myocardial infarction: 410.xx
Lisinopril	ARBs	Angioedema	Angioedema: allergic, any site, with uticaria: 995.1
Cerivastatin	Other statins	Rhabdomyolysis	Rhabdomyolysis: 728.89
Cetirizine**	Fexofenadine and loratadine	Thrombocytopenia	Thrombocytopenia: 287.4 and 287.5
Clemastine**	Loratadine	Stevens–Johnson syndrome, toxic epidermal necrolysis	Erythematous conditions: 695.0 toxic erythema; 695.1 erythema multiforme

*Diagnosis codes based on ICD-9-CM classifications. Based on the study criteria, only inpatient diagnoses were included as adverse events.
**These represent negative controls; no association between the drug and event was expected.
ARBs, angiotensin II antagonists.

DOI: 10.1002/pds

diagnoses during the 181-day period before the incident dispensing.

Eligible person-time. For each comparison eligible person-time began after the first 181-day exposure-free and diagnosis-free period and ended at the first occurrence of any of the following events: end of membership, dispensing of a comparison drug (for analyses using a comparator drug), the first observed inpatient diagnosis of interest (e.g., incident diagnosis of acute myocardial infarction(AMI) or 31 December 2005 (end of the observation period).

Exposed and unexposed person-time. Eligible person-time was classified as exposed or unexposed. Exposed time began on the day after an incident drug dispensing and continued as long as the member was exposed to the drug (based on days supplied in the pharmacy file) plus 14 days.[10,11] Consecutive drug dispensings were combined based on days supplied; exposure gaps of 14 days or less were considered to represent continued medication exposure. Unexposed person-time was defined as all eligible person-time without any drug exposure.

Calculating exposed and unexposed days and diagnoses. The number of exposed and unexposed days was summed by strata defined by health plan, month, sex and age group (5 year groups starting at 0–4 and going through 86+) separately for each drug of interest and comparator. The number of incident diagnoses observed during exposed days and unexposed days also was summed separately by strata defined by health plan, month, sex and age group.

Calculating expected counts: comparison to non-users. We calculated the probability of an unexposed incident ADE within each health plan, sex and age group stratum by dividing the number of unexposed incident ADEs by the number of unexposed days. We then multiplied the probability of an unexposed incident ADE by the number of exposed days for the drug of interest within each health plan, sex, age group and month stratum. This product is the number of incident ADEs expected in each stratum if members exposed to the drug of interest had not been exposed. The number of expected incident ADEs was then summed to the monthly level to generate the number of expected incident ADEs per month. This monthly expected count was compared to the number of observed incident ADEs.

Calculating expected counts: comparison to comparator drug users. We calculated the probability of an incident AE for a person exposed to the comparison drug by dividing the number of ADE during exposure to the comparison drug with the number of exposed days to the comparison drug for each health plan, sex and age group stratum. We then multiplied the probability of an incident ADE among the comparison drug users within each health plan, sex and age group stratum by the number of days of exposure to the drug of interest within each stratum. This product is the number of incident ADEs expected in each stratum if members exposed to the drug of interest had been exposed to the comparator. The number of incident ADEs was then summed across the strata to the monthly level to generate the number of expected incident ADEs per month.

This approach for calculating expected counts is valid if there is (i) a sufficient number of ADEs when exposed to the comparator drug and (ii) if there are considerably more ADE in the comparator group (including historical data).

For this preliminary work we used data from the entire 2000 to 2005 period to calculate expected counts throughout the period. This helped generate stable expected counts for these preliminary analyses. Prospective application of this method might use historical, concurrent or self-controls; all three methods are either being used or considered for vaccine safety surveillance.

Analyses

The maximized sequential probability ratio test. Sequential analysis[12–14] is used when there are repeated looks at data over time, on a continuous, daily, weekly or month basis, adjusting for the multiple testing inherent in the method. We use a maxSPRT, developed by VSD researchers for use in vaccine safety surveillance, in this signal detection study.[15] This is a refinement of the classical sequential probability ratio test[12–14] in that it uses a composite alternative hypothesis of relative risk > 1 rather than a single alternative such as relative risk $= 2$. With the maxSPRT, a drug adverse event signal is generated if and when the log likelihood ratio (LLR) reaches a critical value. The LLR test statistic at time t is calculated as:

$$\mathrm{LLR}(t) = \max_{r>1} \ln\left(\frac{P(c_t|\mathrm{RR}=r)}{P(c_t|\mathrm{RR}=1)}\right)$$

DOI: 10.1002/pds

where c_t is the observed number of adverse events up until and including time t. For this analysis using a large cohort of historical controls we used a Poisson distribution to calculate the LLR.[15]

Critical values. The null hypothesis is rejected the first time the LLR exceeds a critical value, B (i.e., when $LLR(t) > B$). To establish the critical value, it is necessary to specify the alpha level, which we chose to be 0.05, and a pre-specified upper limit on the length of surveillance defined in terms of the expected number of observations (events) under the null hypothesis. For this retrospective analysis of multiple comparisons, a different upper limit was chosen for each drug-event pair in such a way that the length of surveillance would be approximately 72 months, but with a minimum requirement of five expected events under the null. The critical values were generated via simulations and are available from tables provided by Kulldorff *et al.*[15]

RESULTS

The nine participating health plans extracted data from administrative and membership records for over 8 million members over the 6 year study period. The average membership period ranged from approximately 800 to 1500 days across the sites; 6.1 million members had a membership of at least 270 days and therefore qualified for inclusion in the analyses.

Table 2 presents summary data for all study comparisons. A signal of excess risk of AMI among celecoxib users compared to naproxen users was identified in month 25, with 13 observed and about 5 expected AMIs. Excess risk of AMI among rofecoxib users as compared to naproxen users was identified in month 34 with 28 observed and 15.6 expected AMIs (Figure 1). We identified a signal of excess risk of rhabdomyolysis among cerivastatin users compared to users of other statins, but the signal appeared after only 1 observed ADE. As expected, we did not identify a signal of excess risk for the two negative control comparisons (clemastine and cetirizine). Clemastine had 0 observed and less than 1 expected ADEs and cetirizine had 6 observed and about 6 expected ADEs (Figure 2).

Although a signal was detected for the celecoxib and rofecoxib versus diclofenac, and lisinopril versus ARBs comparisons, there were few exposed events among the comparators (diclofenac and ARBs). This is inconsistent with the requirement that the data used to generate expected counts be large enough to generate stable estimates. These results are presented for illustration and to highlight issues related to selection of comparators.

When rofecoxib users were compared to non-users, the signal was detected in month 39, with 39 observed and 23 expected AMIs (Figure 3). This was 5 months later than when rofecoxib was compared to naproxen. The month of signal detection was unchanged for the other drug-event pairs when the comparison was made against non-users. As shown in Table 2, each comparison to non-users was based on hundreds if not thousands of observed outcomes, thereby providing stable estimates for the calculation of expected outcomes.

DISCUSSION

We used health plan automated claims data to conduct a proof of principle evaluation of a prospective safety monitoring system under some of the circumstances that would apply if this method was applied prospectively. Additional work will be required to implement this method for active surveillance. In this dataset, representing approximately 13 million person years of experience, principally in health plans that are relatively slow adopters of new medications, a signal of excess risk was detected in four of the five comparisons of known drug-event associations; we did not observe a signal in the two negative controls. Our findings support the continued investigation of these data as a potentially important contribution to drug safety surveillance using sequential methods.

The intent of sequential analysis is to quickly and efficiently detect signals of excess risk that can then be thoroughly investigated in clinical trials or by other available epidemiological methods. Signal detection using this methodology is not a substitute for confirmatory studies and is not intended to imply a causal relationship. Clearly, sequential analysis using automated healthcare claims data will only be useful if it has reasonable sensitivity and does not generate an unacceptable number of false positives. One way to reduce false positives is to only assess risk for those signals that are flagged in pre-licensure studies, or are of particular biologic relevance. Additional research is needed to investigate the potential for false signaling and the factors associated with false signaling.

Key implementation decisions include the identification of exposed health plan members, selection of comparators, determination and definition of outcomes and the classification of eligible person-time. Decisions related to these specifications affect the number of exposed and unexposed days and events

DOI: 10.1002/pds

Table 2. Summary results for all comparisons

Drug of interest (DOI)	Comparator	Outcome	Months to signal*	Observed events at signal or end of follow-up	Expected events at signal or end of follow-up	Exposed days at signal or end of follow-up (DOI)	Total exposed days for each comparator	Observed events for each comparator	Critical value at $p < 0.05$ (upper limit of expected events)
Celecoxib	Diclofenac†	AMI	25	13	5.01	316 180	2 021 960	18	3.72 (30)
Celecoxib	Naproxen	AMI	25	13	5.19	316 180	15 828 636	124	3.68 (25)
Celecoxib	Non-users	AMI	25	13	5.30	316 180	n/a	21877	3.68 (25)
Rofecoxib	Diclofenac†	AMI	25	17	7.76	623 255	2 021 960	18	3.78 (40)
Rofecoxib	Naproxen	AMI	34	28	15.61	1 078 466	15 828 636	124	3.78 (40)
Rofecoxib	Non-users	AMI	39	39	23.35	1 339 837	n/a	21885	3.83 (50)
Valdecoxib	Diclofenac†	AMI	—	3	2.19	196 867	2 021 960	18	3.30 (5)
Valdecoxib	Naproxen	AMI	—	3	1.80	196 867	15 828 636	124	3.30 (5)
Valdecoxib	Non-users	AMI	—	3	2.15	196 867	n/a	21849	3.30 (5)
Lisinopril	ARBs†	Angioedema	13	3	0.19	828 776	7 682 415	3	3.68 (25)
Lisinopril	Non-users	Angioedema	13	3	0.06	828 776	n/a	282	3.47 (10)
Cerivastatin	Other statins	Rhabdomyolysis	13	1	0.01	15 803	81 481 995	80	3.30 (5)
Cerivastatin	Non-users	Rhabdomyolysis	13	1	0.01	15 803	n/a	1527	3.30 (5)
Cetirizine**	Fexofenadine and Loratadine	Thrombo.	—	6	6.17	5 064 534	31 177 653	59	3.47 (10)
Cetirizine**	Non-users	Thrombo.	—	6	6.07	5 064 534	n/a	9761	3.47 (10)
Clemastine**	Loratadine	SJS/TEN	—	0	0.03	405 676	10 831 935	8	3.30 (5)
Clemastine**	Non-users	SJS/TEN	—	0	0.34	405 676	n/a	319	3.30 (5)

*Months from January 2000 to the first signal at $p < 0.05$ level; emdash indicates no signal was found or not applicable.
**Included as negative controls; no association between the drug and event was expected.
†See text regarding the interpretation of findings related to the limited number of observed events for these comparators.
n/a, not applicable; there were approximately 4.5 billion unexposed days for the non-user comparison groups; AMI, acute myocardial infarction; ARBs, antiotensin II antagonists; SJS/TEN, Stevens–Johnson syndrome, toxic epidermal necrolysis; Thrombo., thrombocytopenia.

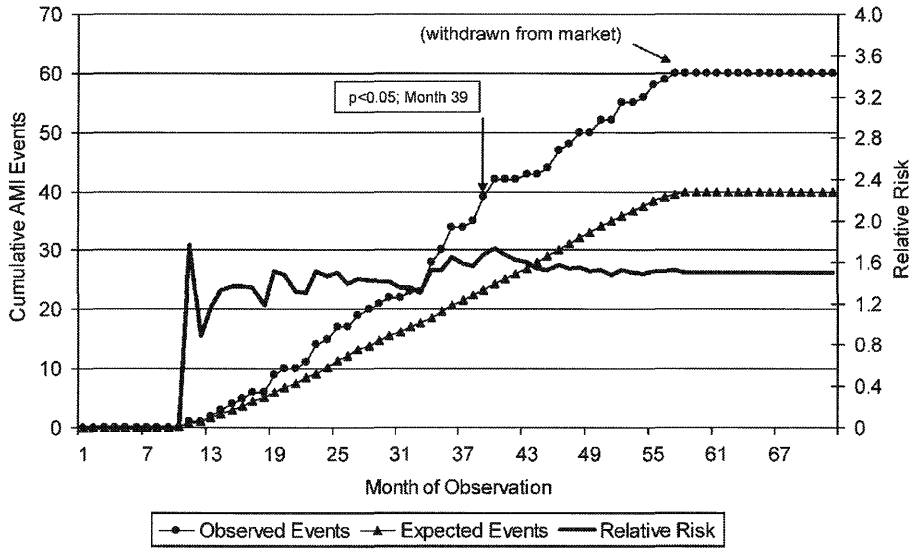

Figure 1. Observed and expected outcomes for rofecoxib users compared to naproxen users: 2000–2005. Outcome: acute myocardial infarction. Adjusted for age, sex and health plan

included in the analyses, thereby influencing the timing of signal detection. For example, shortening the exposure and diagnosis-free interval will include more people in the analyses at the expense of incorporating less of their prior exposure and diagnosis experience.

Signal detection using sequential analysis is closely tied to population size; in general the more people included the faster a signal will be detected or the surveillance will be stopped. For example, all things being equal, doubling the population should halve the time to signal detection. Specification decisions also impact cohort size and those decisions therefore must balance the desire to include as many people in the analyses as possible with the potential for confounding and bias by comparing disparate groups.

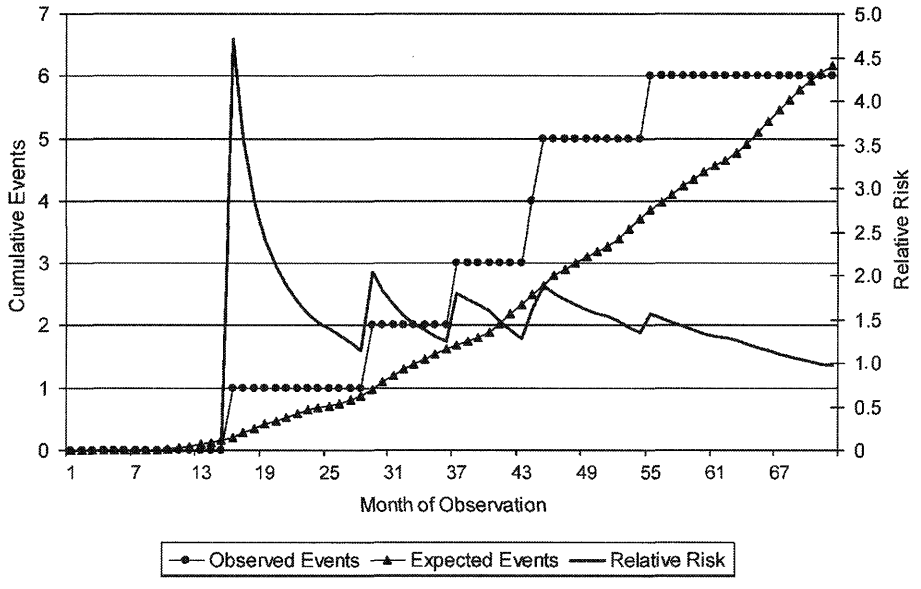

Figure 2. Observed and expected outcomes for cetirizine users compared to non-users: 2000–2005. Outcome: thrombocytopenia. Adjusted for age, sex and health plan

DOI: 10.1002/pds

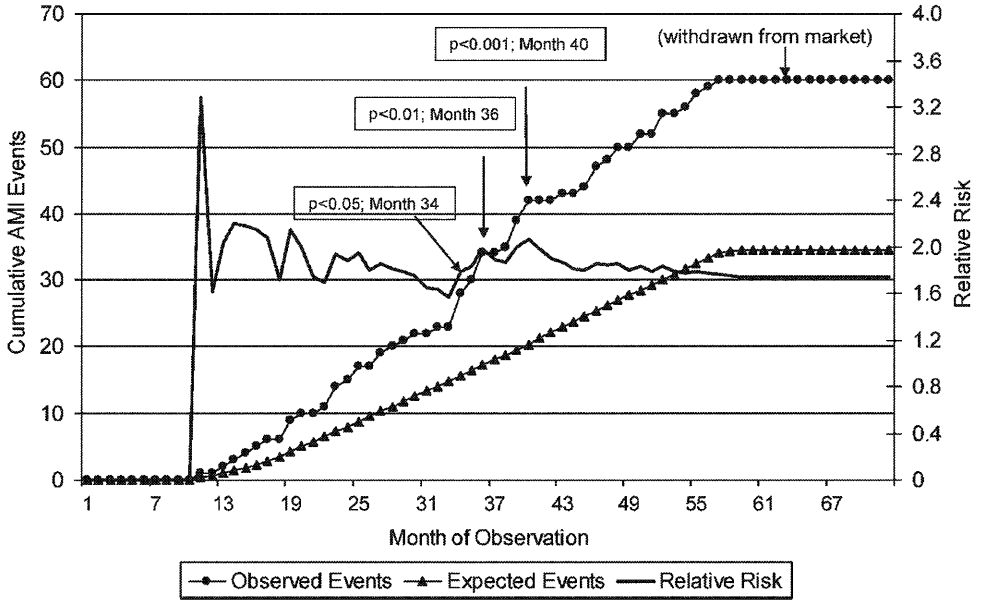

Figure 3. Observed and expected outcomes for rofecoxib users compared to non-users: 2000–2005. Outcome: acute myocardial infarction. Adjusted for age, sex and health plan

Identification of an exposed population

We focused on incident users who had no exposure to the drug of interest or any comparator drugs for the 6 months before the incident dispensing. Our decision to include only incident users was meant to approximate the intended application of the methodology—the prospective monitoring of a newly marketed product. Although excluding prevalent users helped avoid some biases (e.g., survivor bias),[9] it is an open question as to whether prevalent users could also be used for this type of analysis. Our choice of a 6-month exposure-free period is likely conservative and other periods could be reasonable based on drug-specific or other factors. In addition, we did not stratify by dosage; we believe that this level of precision will be important in confirmatory studies but less so in signal detection.

Selection of comparators

We compared the users of each drug of interest to an active comparator cohort of health plan members exposed to a drug or drug class used to treat similar conditions as well as a group that comprises all non-users of the drug of interest. Although we controlled for age, sex and health plan variation, we did not specifically control for treatment selection bias, including confounding by indication, or co-morbidities, incorporating an active comparator was intended to address these issues and should be considered whenever possible. However, the selection of the comparator may introduce other treatment selection biases that must be considered. For this reason, it may be desirable to perform simultaneous evaluation of different comparator groups to help interpret signals. In any event, additional adjustment for confounding is likely to be a useful refinement of this method.

Expected counts of events were generated using a comparison group identified during the same period as users of the drug of interest (i.e., concurrent controls). Prospective application of this methodology may limit the size of a concurrent control group, especially if the control group is based on an infrequently used product. Therefore, it will be important to carefully balance the benefits of a concurrent control group with the benefits of generating stable expected counts using historical exposure and event data. It also may prove desirable to use the self-control case series method for some drugs with brief exposure intervals.[16,17]

Determination and definition of outcomes

Selection of an appropriate outcome is an important aspect of study implementation. Clinically well-defined outcomes are those that can be identified in

DOI: 10.1002/pds

medical claims data with a high level of certainty and little potential for misclassification. To maximize specificity, we selected outcomes that were clinically serious and that required treatment in an inpatient setting.[6] We specified that each ADE was a 'new' event, defined as not having the same event in the 6 months before the incident dispensing. These outcome decisions must be carefully considered to balance the speed of signal detection with the possibility of false signals. For this demonstration work, we did not confirm diagnoses by review of medical records, and there is likely some misclassification because of this. We anticipate that, in practice, it will be necessary to confirm many of the outcomes identified through diagnosis or procedure codes by review of full text medical records or other means.

Eligible person-time

Classification of eligible person-time into exposed and unexposed categories is substantially more complicated for medications than for vaccines.[6] Drug exposure may be continuous or intermittent over long periods, and may include exposure to multiple agents either in sequence or concomitantly. Assigning days as exposed, unexposed or non-contributed for the sequential analyses requires substantial clinical and methodological consideration to balance the inclusion of more patients versus the potential for confounding and bias; additional work is needed to more thoroughly understand these considerations.

Prospective evaluation of accumulating experience of defined cohorts complements the passive safety surveillance because it addresses the main limitations of spontaneous reporting, that is, no denominator. Whereas spontaneous reporting systems often lack information on the exposed population, our system uses a known population with detailed exposure information, thereby allowing calculation of relative risk among various population cohorts. To the extent that the relevant outcomes are reported within claims-based systems, this method avoids the shortfalls associated with both underreporting and reporting bias.

Key benefits of the methodology relate to its use of routinely collected health plan encounter and dispensing data that are commonly used in epidemiological research, minimal data requirements in terms of needed data elements, the ability to simultaneously apply the methodology within numerous data systems and the use of highly summarized data structure for aggregation across systems and analysis. Most public and private health insurers in the U.S. have data that could support sequential analyses. Because only highly summarized data are needed for analysis, concerns over sharing confidential data and patient confidentiality are minimized. Expanding the available population to publicly funded systems like the Veterans Administration, TRICARE, Medicare and Medicaid and to private health insurers would substantially improve the performance of the methodology by increasing the sample size. This would be especially important for monitoring newly marketed products or those that have limited use.

An important study limitation was the limited number of observed events for most comparisons, including the negative controls and the active comparators. Other limitations relate to the relatively complicated set of decisions that need to be made for implementation, pre-specification of the number of expected events to continue surveillance, reliance on the quality and timely availability of the underlying health plan claims data, the need for enough historical or concurrent comparator data to generate stable expected counts and the need for frequent data updates. Additionally, there is limited practical experience with implementation, analysis and reporting of results. Guidelines will be needed to help investigators establish methodological criteria to address issues of exposure, events and setting a minimum number of observations before accepting a signal. Additional methodologic work that will enhance our understanding of the utility and limitations of this method include assessment of the impact of assessing multiple outcomes for each new drug on the likelihood of identifying a signal, reporting confidence intervals for the relative risk, better delineation of sensitivity and specificity through simulation and comparison of maxSPRT to other sequential methods.

The present study supports the use of a more fully developed version of this method for actively monitoring drug safety. Active surveillance is an important complement to passive safety surveillance as it holds

KEY POINTS

- Sequential analysis of near 'real-time' health plan network data may be useful for drug safety surveillance.
- There are a number of methodological issues associated with drug safety surveillance in health plan networks that must be addressed.
- The automated data needed to conduct near real-time drug safety signal detection are routinely collected by health plans.

DOI: 10.1002/pds

the potential to rapidly investigate the drug exposure and safety experience within large, well-defined and diverse populations.

ACKNOWLEDGEMENTS

We are indebted to the statistical programmers at each of the sites for their work in extracting the study data and testing the programming algorithms, to Kimberly Lane and the HMORN CERT Data Coordinating Center for their help in overseeing the study and Meredith Chace for editing and data checking. This study was funded by a grant from AHRQ (2U18HS010391) supporting the HMO Research Network Centers for Education and Research on Therapeutics (CERT).

REFERENCES

1. Committee on the Assessment of the US Drug Safety System. *The Future of Drug Safety: Promoting and Protecting the Health of the Public*. Institute of Medicine of the National Academies: Washington D.C., 2006.
2. Wysowski DK, Swartz L. Adverse drug event surveillance and drug withdrawals in the United States, 1969–2002: the importance of reporting suspected reactions. *Arch Intern Med* 2005; **165**(12): 1363–1369.
3. Piazza-Hepp TD, Kennedy DL. Reporting of adverse events to MedWatch. *Am J Health Syst Pharm* 1995; **52**(13): 1436–1439.
4. Brewer T, Colditz GA. Postmarketing surveillance and adverse drug reactions: current perspectives and future needs. *JAMA* 1999; **281**(9): 824–829.
5. Davis RL, Kolczak M, Lewis E, *et al.* Active surveillance of vaccine safety: a system to detect early signs of adverse events. *Epidemiology* 2005; **16**(3): 336–341.
6. Lieu TA, Kulldorff M, Davis RL, *et al.* Real-time vaccine safety surveillance for the early detection of adverse events. *Medical Care.* 2007; **45**(10):S89–S95.
7. Platt RAS, Davis RL, Destefano F, *et al.* Pharmacovigilance in the HMO Research Network. In *Pharmacovigilance*, Ronald D, Mann EBA (eds). New York, Wiley: 2002; 582.
8. Clark JA, Klincewicz SL, Stang PE. Spontaneous adverse event signaling methods: classification and use with healthcare treatment products. *Epidemiol Rev* 2001; **23**(2): 191–210.
9. Ray WA. Evaluating medication effects outside of clinical trials: new-user designs. *Am J Epidemiol* 2003; **158**(9): 915–920.
10. McMahon AD, Evans JM, McGilchrist MM, McDevitt DG, MacDonald TM. Drug exposure risk windows and unexposed comparator groups for cohort studies in pharmacoepidemiology. *Pharmacoepidemiol Drug Saf* 1998; **7**(4): 275–280.
11. van Staa TP, Abenhaim L, Leufkens H. A study of the effects of exposure misclassification due to the time-window design in pharmacoepidemiologic studies. *J Clin Epidemiol* 1994; **47**(2): 183–189.
12. Spiegelhalter D, Grigg O, Kinsman R, Treasure T. Risk-adjusted sequential probability ratio tests: applications to Bristol, Shipman and adult cardiac surgery. *Int J Qual Health Care* 2003; **15**(1): 7–13.
13. Wald A. Sequential tests of statistical hypotheses. *Ann Math Stat* 1945; **16**: 177–186.
14. Wald A. *Sequential Analysis*. Wiley: New York 1947.
15. Kulldorff M, Davis RL, Kolczak M, Lewis E, Lieu TA, Platt R. *A maximized sequential probability ratio test for drug and vaccine safety surveillance. Working paper, Department of Ambulatory Care and Prevention.* 2007; available at www.dacp.org/faculty_kulldorff.html
16. Tata LJ, Fortun PJ, Hubbard RB, *et al.* Does concurrent prescription of selective serotonin reuptake inhibitors and non-steroidal anti-inflammatory drugs substantially increase the risk of upper gastrointestinal bleeding? *Aliment Pharmacol Ther* 2005; **22**(3): 175–181.
17. Whitaker HJ, Farrington CP, Spiessens B, Musonda P. Tutorial in biostatistics: the self-controlled case series method. *Stat Med* 2006; **25**(10): 1768–1797.

DOI: 10.1002/pds