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.