ORIGINAL RESEARCH ARTICLE

Drug Saf 2012; 35 (5): 395-406
0114-5916/12/0005-0395/$49.95/0

A Pharmacoepidemiological Network Model for Drug Safety Surveillance

Statins and Rhabdomyolysis

Ben Y. Reis,[1] *Karen L. Olson,*[1] *Lu Tian,*[2] *Rhonda L. Bohn,*[3] *John S. Brownstein,*[1] *Peter J. Park,*[1] *Mark J. Cziraky,*[3] *Marcus D. Wilson*[3] and *Kenneth D. Mandl*[1]

1 Children's Hospital Informatics Program, Harvard-MIT Division of Health Sciences and Technology, Children's Hospital, Harvard Medical School, Boston, MA, USA
2 Department of Health Research and Policy, Stanford University, Stanford, CA, USA
3 HealthCore, Inc., Wilmington, DE, USA

Abstract

Background: Recent withdrawals of major drugs have highlighted the critical importance of drug safety surveillance in the postmarketing phase. Limitations of spontaneous report data have led drug safety professionals to pursue alternative postmarketing surveillance approaches based on healthcare administrative claims data. These data are typically analysed by comparing the adverse event rates associated with a drug of interest to those of a single comparable reference drug.

Objective: The aim of this study was to determine whether adverse event detection can be improved by incorporating information from multiple reference drugs. We developed a pharmacological network model that implemented this approach and evaluated its performance.

Methods: We studied whether adverse event detection can be improved by incorporating information from multiple reference drugs, and describe two approaches for doing so. The first, reported previously, combines a set of related drugs into a single reference cohort. The second is a novel pharmacoepidemiological network model, which integrates multiple pair-wise comparisons across an entire set of related drugs into a unified consensus safety score for each drug. We also implemented a single reference drug approach for comparison with both multi-drug approaches. All approaches were applied within a sequential analysis framework, incorporating new information as it became available and addressing the issue of multiple testing over time. We evaluated all these approaches using statin (HMG-CoA reductase inhibitors) safety data from a large healthcare insurer in the US covering April 2000 through March 2005.

Results: We found that both multiple reference drug approaches offer earlier detection (6–13 months) than the single reference drug approach, without triggering additional false positives.

Conclusions: Such combined approaches have the potential to be used with existing healthcare databases to improve the surveillance of therapeutics in the postmarketing phase over single-comparator methods. The proposed network approach also provides an integrated visualization framework enabling decision makers to understand the key high-level safety relationships amongst a group of related drugs.

Withdrawals of major drugs such as Baycol® (cerivastatin)[1-3] and Vioxx® (rofecoxib)[4-7] have highlighted the critical importance of drug safety surveillance in the postmarketing phase.[8,9] Pre-launch clinical trials often fail to detect drug-related adverse events due to limitations in sample size, demographic diversity, and range of co-morbidities and drug interactions represented in the trials.[9] Improved methods for identifying adverse drug effects earlier and more reliably have the potential to prevent serious illness and save lives. The Institute of Medicine has called for the adoption of a 'lifecycle' approach to drug safety surveillance that includes the postmarketing stage,[10,11] and the US Congress has reauthorized the Prescription Drug User Fee Act with support for postmarketing surveillance.[12]

For the past 40 years, postmarketing surveillance has been performed in the US primarily through the Spontaneous Reporting System/MedWatch/Adverse Event Reporting System (AERS), used to monitor spontaneous reports submitted voluntarily by health professionals and consumers. This approach has critical limitations, including under-ascertainment, noise, delay in reporting and the lack of information about the total number of people taking each drug,[13,14] leading some to suggest that spontaneous reporting does not provide a sufficient basis on which to make appropriate drug safety decisions.[15]

Because of the limitations of relying on spontaneous reports, recent efforts have focused on leveraging an alternative data source: electronic administrative healthcare claims databases.[8,15-17] These databases can provide broader population coverage than spontaneous reports or pre-market clinical trials. Studies have shown successful use of these data for detecting a variety of adverse drug effects,[18,19] including cyclooxygenase-2 inhibitors[20] and statins (HMG-CoA reductase inhibitors).[21-23] The use of these administrative databases carries with it potential methodological issues, including selection bias, lack of chart review for case validation in some cases, and the inability to assign causality for adverse events. While grappling with these challenges, the scientific and public health communities are evaluating these data to determine whether they can improve on current methods for early, sensitive and specific detection of adverse drug effects in the postmarketing setting.

In a postmarket surveillance setting, these claims data are typically analysed by comparing adverse event rates among patients prescribed a particular drug versus patients prescribed a comparable reference drug, and a sequential analysis framework may be used to address the issue of multiple testing over time. Brown et al.[24] present a general framework for implementing sequential analysis for postmarketing drug safety surveillance, and Kulldorff et al.[25] describe the maximized sequential probability ratio test (MaxSPRT) methodology in greater detail.

Comparisons are usually performed against a single reference drug, often a drug in the same class[1,23] chosen to maximize comparability with regard to unmeasured factors of the reference group, or against all non-users of the drug.[24] However, there are typically a number of possible reference drug candidates that can be used. For example, cerivastatin is often compared against the most commonly used statin – atorvastatin – but may also be compared against other statins such as pravastatin and simvastatin.[21-23]

In this study, we hypothesize that information contained in the relationships among multiple pairs of drugs can improve drug safety surveillance over the single reference drug methods used

today. There are two major potential benefits to incorporating information from multiple reference drugs. First, combining multiple analyses can improve detection performance. We propose that combining information from multiple reference drugs within a single study has the potential to improve detection performance, whether by combining multiple sub-threshold signals into a super-threshold result, or by exploiting the fact that certain reference drugs enable detection of signals that other reference drugs do not. For a given dataset available for analysis, we hypothesize that combining multiple drug-drug comparisons has the potential to enable earlier detection, and/or detection of signals that would otherwise not be detected at all. By detecting and removing dangerous drugs from the market in a more timely fashion, drug-related morbidity and mortality may be minimized.

Second, integrating multiple comparisons can provide a comprehensive view of the drug safety environment. Choosing a drug as a reference drug implies that the drug is assumed to have a 'normal' level of safety for a given adverse event. However, this assumption may not be valid – the chosen reference drug may actually be more 'dangerous' than assumed and thus make truly 'dangerous' drugs appear 'safe'. Conversely, it may have unusually low levels of an adverse event, making truly 'safe' drugs appear 'dangerous'. By analysing and visualizing multiple relationships amongst a group of drugs, such potential false assumptions are more likely to be detected.

We describe two approaches to combining information from multiple reference drugs – one previously reported and one novel – and compare their performance to that of a single reference drug approach. The first multiple reference drug approach, described herein as the 'combined cohort approach', involves straightforwardly combining a set of related reference drugs into a single reference cohort, an approach similar to the one used by Brown et al.[26] (see the Discussion section for a full comparison with the paper by Brown et al.). For the second approach, we describe a novel class of pharmacoepidemiological networks that analyse and capture information pertaining to multiple drug-drug safety relationships. Building on prior work developing network-based approaches[27] for prospective infectious disease outbreak detection in public health surveillance systems,[28,29] we present a network-based approach for supporting postmarketing surveillance of drug safety. In addition to improved detection performance, the networks are designed to provide the comprehensive high-level perspective that enables decision-makers to understand the safety relationships across a group of drugs, and to comprehend the different sources of evidence that combine to form the final opinion of the system.

We begin by describing the single reference drug approach, then proceed to describe the combined-cohort approach and, finally, the general class of pharmacoepidemiological network models. Using statin safety as a case study, we evaluated the performance of the combined-cohort approach and the pharmacoepidemiological network approach versus the single reference drug approach. For each of the approaches, we performed a retrospective study in which we simulated a prospective analysis based on chronological analysis of the historical data. The goal of this research is not to evaluate the general validity of using healthcare administrative claims databases for postmarketing drug surveillance, nor is it to determine the presence or absence of specific safety issues associated with statin drugs. Rather, the goal of this study is to determine whether, given the same healthcare data, the proposed multiple reference drug approaches can offer improved adverse event detection performance versus the single reference drug approach. We discuss the benefits and limitations of each of the approaches and identify areas for future research.

Methods

Data

We analysed de-identified administrative payer-based healthcare information from the HealthCore Integrated Research Database (HIRD), representing diverse regions in the US, including Western, Midwestern, Mid-Atlantic and Southeastern states,

and covering 9 million fully commercially insured lives at the time of the study. The data consist of health plan enrollment, medical and pharmacy administrative claims files. All study materials were handled in compliance with the Health Insurance Portability and Accountability Act of 1996 (HIPAA) regulations, and the analysis was conducted using a limited dataset. The data include prescription and diagnostic data for all patients prescribed at least one statin drug from April 2000 through March 2005.

Case Study: Statins

We focused on the association between statin use and the subsequent diagnosis of rhabdomyolysis. The family of statin drugs have been shown to effectively reduce blood lipid levels and reduce atherosclerotic coronary artery disease morbidity and mortality, as well as all-cause mortality.[30,31] However, cerivastatin has been found to significantly increase risk for rhabdomyolysis (a condition involving the breakdown of muscle tissue)[22,30] and was withdrawn from the market in August 2001 for this reason.[1,2]

Cohort Study Design

All three detection approaches described below are based on the same retrospective cohort analysis, conducted as follows: for a given pair of target and reference drugs, two cohorts are defined – one for the target drug and one for the reference drug (or combination of reference drugs). Each cohort includes all patients aged 18 years or older who received at least one prescription of the relevant drug during the period April 2000–March 2005. The cohort includes patients with strict statin monotherapy only, meaning that patients prescribed one or more of the other statin drugs at any point in the study period are excluded. The total exposure time is calculated for each person based on the aggregate amount of time covered by the prescriptions. Gaps in exposure of 30 or fewer days are considered as exposed. Dosage information is not considered.

Regarding the case definition, potential claims-based cases of hospitalized rhabdomyolysis were identified by hospital discharge claims associated with International Classification of Diseases, 9th edition (ICD-9) codes, including all secondary diagnoses, for myoglobinuria (791.3x), disorders of the muscle, ligament, and fascia (728.89) and rhabdomyolysis (728.88). These ICD-9 codes were selected on the basis of previous studies, reports, adverse medical events included in the prescribing information of the lipid-lowering drugs (LLD), and the advice of expert panels of clinicians convened by the National Lipid Association (NLA). Available patient medical records were abstracted using a chart abstraction form developed by an expert panel of physicians from the NLA. The chart abstractions were conducted by two physicians who were blinded to the LLD exposure status of the patients. Cases of rhabdomyolysis were confirmed if the medical chart indicated that the attending physician at the time of hospitalization made a clinical diagnosis of rhabdomyolysis or if the patient experienced muscle injury at the time of hospitalization, and their creatine kinase (CK) level during hospitalization was at least 10 times the upper limit of normal with signs of muscle weakness (normal CK: male 24–204 U/L; female 24–173 U/L). Only cases of rhabdomyolysis unrelated to myocardial infarction or other types of injuries were attributed to the use of statins. Furthermore, only cases occurring during the course of treatment with the drug or within 30 days of stopping treatment with the drug were included.

A person enters the cohort upon receiving the first prescription of the drug. A person exits the cohort upon the first occurrence of any the following: (i) receiving a diagnosis matching the case definition; (ii) stopping the course of treatment with the drug; or (iii) termination of the study period. The two cohorts are analysed with logistic regression (SAS software, SAS Institute Inc., Cary, NC, USA), using total exposure time as an offset and adjusting for sex, age and co-morbidity. Each analysis yields a 1-tailed p-value.

Co-morbidity was estimated using the Charlson Comorbidity Index[32] a commonly used and established method for correcting for potential confounders. Other methods can be used, including propensity score analysis.

Single-Comparison Method with Sequential Analysis

We begin by describing the single reference drug approach that serves as a basis for comparison for the two multiple reference drug approaches described below. In the single reference drug approach, the retrospective cohort study described above is used to compare adverse event rates between two cohorts - the drug of interest and the reference drug. The output of the analysis is a 1-sided p-value (the test statistic). In a prospective setting, this analysis takes place periodically – in this study once per month. To address the issue of multiple testing over time, a sequential analysis framework is used to spread the desired level of type I error (chosen to be 0.05) evenly over a total of M periodic (monthly) tests and to calculate a time-dependent decision boundary. For this study, we have chosen a value of $M = 60$, a period of 5 years. The choice of the value of M is up to the user: larger values of M allow for a longer surveillance period, while smaller values of M concentrate the desired type I error over a shorter period of time, with the likely effect of increased sensitivity during the shorter surveillance period. We chose 5 years as a reasonable period of time during which adverse drug events could be expected to be detected, although other values of M are certainly possible.

The sequential analysis procedure proceeds as follows: performing the regression-based comparison once per time period yields M test statistics, $p^*(m)$, where m is the time period (month) and p^* is the one-sided test statistic which has not been adjusted for multiple testing over time. 30 000 permutations (the total number of permutations depends on the desired precision of the final analysis) are performed by randomly reassigning patients to one of the two cohorts while maintaining the relative cohort sizes. A patient is assigned to a cohort consistently for the entire duration of a permutation, i.e. there is no switching between cohorts from month to month within a permutation. A full set of regression analyses are then calculated for each of the permutations, yielding 30 000M test statistics, $p^*_r(m)$, where r is the permutation number.

Once all the permuted p^* values have been calculated, a table is set up with M columns (one for each month) and 30 000 rows (one row for each permutation). Each cell contains the value $p^*_r(m)$ for the specific time period m and specific permutation r. The following steps are then performed iteratively:

Step 1: The table is sorted in ascending order according to the p^* values in the first column. A cutoff value c^*_1 is identified so that $(0.05/2)/M$ of the p^* values in the first column are smaller than c^*_1. Conceptually, this results from setting the monthly false-positive threshold in an empirical fashion to ensure that the false positive rate is spread evenly over each month. In this particular study, the desired type I error rate of 0.05 was divided by 2 because of our interest in one-tail directional results, resulting in a value for c^*_1 of 0.000415.

Step 2: The implication of $p^*_r(m)$ being smaller than c^*_m for a given month m is that, within permutation r, an alert signal has been triggered at month m. Therefore, for each cell in the first column that is smaller than c^*_1, the entire row containing that cell is deleted from the table, thereby removing permutation r from the remaining months of the analysis, as required by the sequential analysis framework.

Step 3: Steps 1 and 2 are performed iteratively, each time analysing the next column in the dataset, until all columns have been analysed, or until the stopping rule, below, has been triggered.

Stopping rule: The result of the sequential analysis is a vector $c^*_1, c^*_2, \ldots, c^*_M$ that defines a time-sensitive decision boundary (alarm thresholds) for the test statistics. This decision boundary is applied to the original un-permuted data $p^*(m)$. The first month for which $p^*(m)$ is lower than c^*_m is considered to be the month where an alert signal has been generated for the target drug. A summary p-value is calculated for month m by determining the rank placement of the unpermuted $p^*(m)$ value within the range of all permuted $p^*(m)$ values in column m.

Combined-Cohort Method with Sequential Analysis

The first multiple-reference drug approach we describe is identical to the single reference drug

approach described above, with one key difference: a set of related reference drugs are combined into one master reference cohort. This master reference cohort is then used in place of the single reference drug cohort above. All other analysis steps are identical.

The Pharmacoepidemiological Network Model

The pharmacoepidemiological network model integrates the results of multiple drug-drug comparisons across a set of related drugs within a sequential analysis framework. In the pharmacoepidemiological network model, the sequential methods described above are modified as follows.

Given a set of N drugs (for example, cerivastatin plus six other control statins, $N=7$), we maintain the assumptions of M periodic measurements and a desired overall type I error of 0.05 over time. Using the same logistic-regression-based comparison as above, we perform a directional comparison between every possible pair of drugs, for a total of $N(N\text{-}1)$ comparisons each month ($7\times6=42$ comparisons in the present example). Performing these comparisons for the M time periods results in $N(N\text{-}1)M$ test statistics, $p*(t,b,m)$, where t is target drug, b is the reference drug and m is the month. As above, these test statistics have not yet been adjusted for the effects of multiple testing.

For each month, we combine the $p*$ values from the $N\text{-}1$ individual comparisons, $p*(t,b,m)$, summing over all values of b. We do this for each target drug t by using a combination function, weighted by the exposure time of the referent drug b:

$$C*_{(t,m)} = 2\sum_{b} \ln(E_b p*_{(t,b,m)})$$

where E_b is the aggregate exposure time across all patients in the cohort for the referent drug b. (This combination approach is used only as a convenient method of combining scores and the results are not interpreted as p-values; no assumptions are made about the validity of combining test statistics using Fisher's data fusion method[33] in this context.) The result is NM combination-test-statistics, $C*(t,m)$.

As above, 30 000 permutations are performed on the data. The regression-based comparisons are performed on each of the permutations, yielding the value $p*_r(t,b,m)$ for each permutation r. The weighted combinations are then applied to these $p*$ values to yield the combination statistic $C*_r(t,m)$ for each permutation r.

Once all the permuted $C*$ values have been calculated, a table is set up with M columns (one for each month) and $30\,000N$ rows (one row for each target drug and permutation). Each cell contains the value $C*_r(t,m)$ for the specific target drug, month and permutation. The three-step iterative procedure described above is then followed.

The result of the sequential analysis is a vector $c_1*, c_2*,\ldots,c_M*$ that defines a time-sensitive decision boundary (alarm thresholds) for the test statistics. This decision boundary is applied to the original un-permuted data $C*(t,m)$, and the first month for which $C*(t,m)$ is less than $c*_m$, an alert signal is generated for that target drug. Summary p-values are calculated by determining the rank placement of the un-permuted $C*(t,m)$ value within the range of all permuted $C*(t,m)$ values in column m.

One of the primary contributions of the pharmacoepidemiological network model is the network visualization (figure 1). In the network

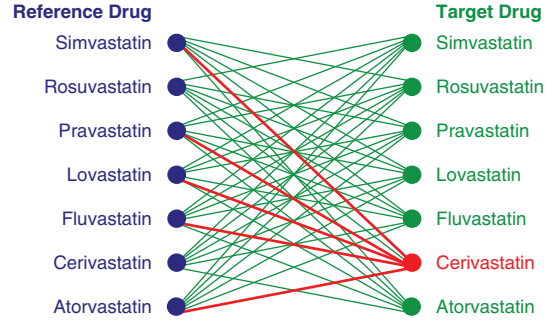

Fig. 1. A pharmacoepidemiological network showing the safety relationships between different statin drugs for rhabdomyolysis. The network shown is a snapshot in time taken in March 2005, the last month of the study period. Reference drugs appear on the left and target drugs appear on the right. Each edge connecting two nodes represents the results of a single pair-wise comparison of a target drug and a reference drug (green = no alert triggered, red = alert triggered). The colour of each target node reflects the consensus score (green = no consensus alert triggered, red = consensus alert triggered), calculated by integrating the opinions of all the other reference drugs pertaining to the safety of that target node.

Table I. Description of inception cohorts for patients receiving statin monotherapy

	Atorvastatin	Cerivastatin	Fluvastatin	Lovastatin	Pravastatin	Rosuvastatin	Simvastatin
N	254 617	9597	16 048	39 092	66 161	15 597	63 182
Female (%)	44.46	48.24	50.70	48.42	47.26	44.74	44.77
Age (average; y)	55.94	57.38	58.85	57.39	56.76	54.50	58.06
Exposure (person-years)	209 866	2456	8657	20 684	49 696	5568	42 188
Mean Charlson comorbidity score[32]	0.590	0.507	0.498	0.528	0.641	0.665	0.780
Rhabdomyolysis cases [% of exposure]	216 [0.10]	10 [0.41]	6 [0.07]	20 [0.10]	50 [0.10]	4 [0.07]	40 [0.09]

visualization, each drug appears twice – once on the left as a *reference node* and once on the right as a *target node*. Each edge connecting two nodes represents the results of the regression-based comparison of the target drug with the reference drug, with red colour and greater thickness of the edge reflecting a greater difference in adverse event levels between the two nodes. In this visualization, we choose an overall type I error level of 0.05 as representing a significant level of alarm, although other thresholds may be chosen depending on the requirements of users of the system. The colour of each target node reflects the consensus score, calculated by integrating the opinions of all the other reference drugs pertaining to the safety of that target node. This visualization provides a comprehensive high-level perspective that enables decision makers to understand the safety relationships across a group of drugs, and to comprehend the different sources of evidence that combine to form the final opinion of the system.

Results

Table I provides descriptions of the seven cohorts used in this study – one for each of the seven commonly used statin drugs. Table II reports the relative risks for rhabdomyolysis for six of the drugs, compared with atorvastatin (with 95% CIs), adjusted for age, sex, Charlson comorbidity score,[32] and exposure time.

Figure 2 shows the results of the single reference drug approach. Empirical p-values resulting from comparisons involving cerivastatin are shown as coloured lines, while all comparisons between other statins are shown as thin grey lines. Using this approach, the earliest detection of the risk of rhabdomyolysis associated with cerivastatin would have occurred in June 2001, assuming an optimal choice had been made ahead of time and simvastatin was used as a reference. However, many studies of cerivastatin-related rhabdomyolysis risk use atorvastatin as a reference, in which case the single reference drug approach would have generated a signal only in October 2001. (As a reminder, cerivastatin was withdrawn from the market in August 2001.) Use of other statins as references would have yielded even later detection, or even no detection at all through the end of the study period, as in the case of rosuvastatin.

Based on the same data, the combined-cohort approach would have detected cerivastatin's rhabdomyolysis risk in September 2000, 9 months earlier than the earliest detection achieved using the single reference drug approach when choosing the most favourable referent drug and 13 months earlier than the single-referent approach when

Table II. Relative risk of drug safety for different adverse events compared with atorvastatin (with 95% CIs), adjusted for age, sex, Charlson comorbidity score[32] and exposure

Drug	Rhabdomyolysis
Cerivastatin	4.08 (1.38, 12.05)
Fluvastatin	0.69 (0.17, 2.78)
Lovastatin	0.96 (0.44, 2.09)
Pravastatin	0.96 (0.57, 1.61)
Rosuvastatin	0.66 (0.12, 3.68)
Simvastatin	0.88 (0.50, 1.56)

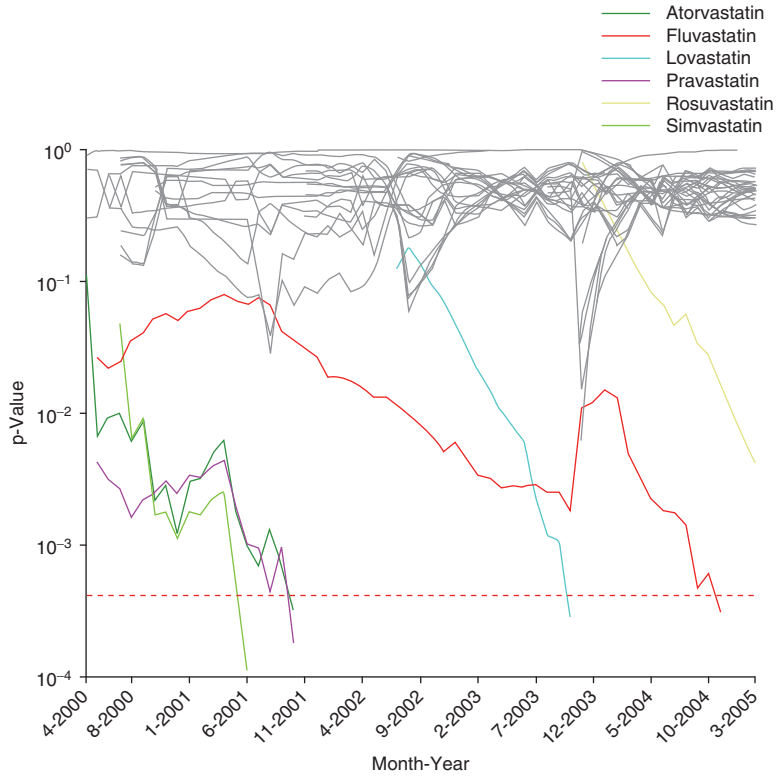

Fig. 2. The results of drug safety analyses using the single reference drug approach. Coloured lines indicate empirical p-values resulting from comparisons between cerivastatin and one of the other statin drugs. Grey lines indicate empirical p-values resulting from comparisons between a pair of statin drugs that do not include cerivastatin. The threshold of 0.000415, indicated by the dashed red line, represents a desired type-I error of 0.05 divided over 60 months within the sequential analysis framework using one-sided t-tests. Comparison against simvastatin would have yielded an alert for cerivastatin in June 2001, while comparison against atorvastatin or pravastatin would have yielded an alert for cerivastatin in October 2001.

using the most common referent drug, atorvastatin. This is accomplished without flagging any other statin drugs as dangerous using the combined-cohort approach. Thus, no additional false positives appear in the present experiment.

Figure 3 shows the results achieved using the pharmacoepidemiological network model. Based on the same available healthcare data, the pharmacoepidemiological network model would have detected cerivastatin's rhabdomyolysis risk in December 2000, 6 months earlier than the earliest detection achieved using the single reference drug approach when choosing the most favourable referent drug and 10 months earlier than the single reference drug approach when using the most common referent drug, atorvastatin. This is also accomplished without flagging any other statin drugs as dangerous using the pharmacoepidemiological network model.

Discussion

Based on the results obtained in the present analysis, the multiple reference drug approaches yielded improved detection timeliness over the single reference drug approach, without generating additional false positives. Each multiple reference drug approach has its own benefits. While the combined-cohort approach achieves slightly earlier detection, the novel pharmacoepidemiological network approach provides a detailed framework to help decision makers understand the high-level relationships between drugs, and comprehend the different components of the

evidence used to generate the final consensus opinion, a level of detail that is not available with the combined-cohort approach.

As mentioned above, such a detailed framework would be useful in cases where it is unclear whether a given comparator drug itself has high levels of a given adverse event. In pharmacoepidemiological analyses of claims data, the 'safety' of a given drug with regard to a particular adverse event is a relative measure that depends on the comparator drug used. A comparator with high rates of a given adverse event could make even a 'dangerous' target drug look 'safe'. Assuming that there are other 'safe' comparators included in the analysis, the pharmacoepidemiological network model can allow investigators to instantly identify the 'dangerous' comparator – it will be revealed as having higher rates of the adverse event than these other comparators.

As mentioned above, the goal of this work is not to evaluate the general validity of using existing healthcare databases for postmarketing drug surveillance, nor is it to determine the presence or absence of specific safety issues associated with statin drugs. Rather, it is to determine whether multiple reference drug approaches for analysing these healthcare databases can offer improved detection performance over the single reference drug approach. Therefore, we compared the detection performance achieved by all the approaches using the same dataset.

While the single reference drug approach requires choosing the best reference drug to use, the network approach allows the use of multiple

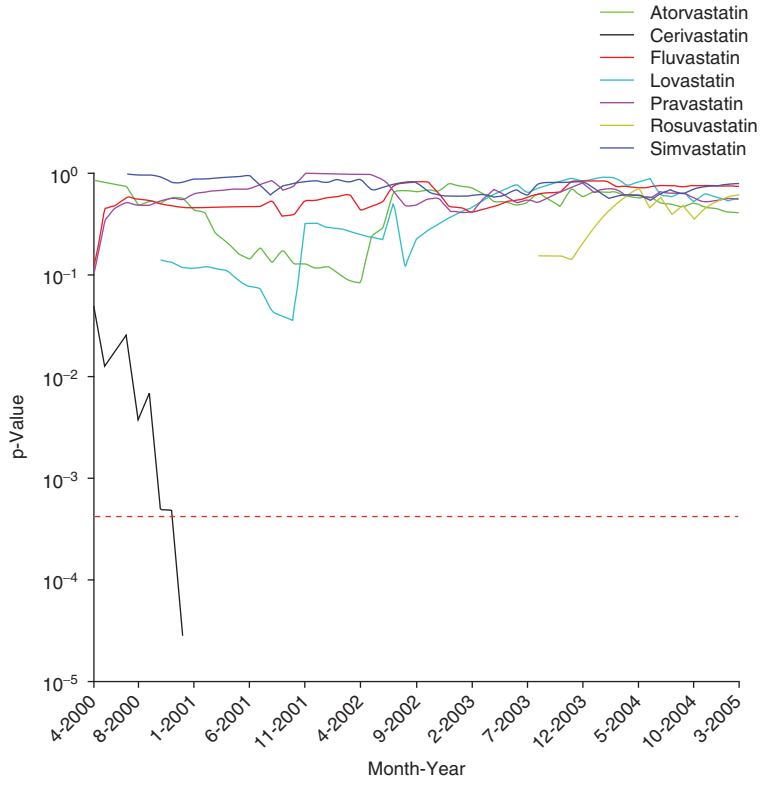

Fig. 3. The results of drug safety analyses using the pharmacoepidemiological network (PEN) model. Each coloured line represents the empirical p-values resulting from the consensus derived from comparing the target statin with all the other statins. The results show that using the PEN model, an alert for cerivastatin would have been generated in December 2000, 6 months earlier than any of the single comparisons in figure 2, and 9 months earlier than the single comparison using atorvastatin as a reference. Importantly, no false alarms would have been generated for any of the other statins.

reference drugs. In the chosen case of statin safety, a well-defined class of seven drugs is readily available for mutual comparisons. In cases where reference drugs from the same class are not available, other drugs with similar indications or the non-exposed population may be used as references. We expect that the quality of the results may vary as less-related drugs are used; the more a reference drug is different from the target drug, the greater likelihood that differences in the treated populations may create confounders that skew the results of the comparison. For this reason it is possible that the use of a more heterogenous group of comparator drugs may lead to poorer performance. In addition to conducting additional evaluative examples, future work will assess both the value of including additional reference drugs in a network, as well as the effects of similarity between target drug and reference drug on network performance. A long-term goal of this future work is to inform the development of a possible set of systematic guidelines for multiple comparator selection, especially in cases where there are many possible comparators to choose from and there is concern that subjective selection of comparator drugs might bias the results of the analysis.

The network approach includes adjustments for multiple comparisons of a target drug with multiple reference drugs. It is possible that these adjustments would decrease the sensitivity of the system to the point where the benefits of combining evidence would be outweighed by the measures taken to combat multiple testing. Based on the results obtained in the present analysis, the benefits of combining evidence outweighed the measures taken to combat multiple testing, leading to earlier detection without causing additional false positives. Additional analyses will enable further study of these performance comparisons.

As mentioned above, Brown et al.[26] also reported using a combined cohort approach within a sequential analysis framework to detect a cerivastatin/rhabdomyolysis signal by comparing case rates for cerivastatin with those of all other statins. Starting with data in January 2000, the single reference drug approach generates a signal for cerivastatin/rhabdomyolysis 13 months later in February 2001. Starting with data in April 2000, the pharmacoepidemiological network method generates a signal 8 months later in December 2000, and the combined cohort approach generates a signal 5 months later in September 2000. While the results presented here are earlier, both in relative and absolute terms, it is important to realize that these two studies analyse different datasets, covering different sized populations, seeking care at different health systems and using different case definitions. Specifically, the signal reported by Brown et al. is based on one observed case of rhabdomyolysis amongst 198 person-years of cerivastatin exposure, while the present signal is based on 10 observed cases of rhabdomyolysis amongst 2456 person-years of cerivastatin exposure. Given the larger sample size used in the present study, it would be expected that a similar method would achieve earlier detection. In general, the main difference between the present study and the study by Brown et al. is the introduction of the pharmacoepidemiological network method, which offers a novel analytical approach to combining information from multiple comparator drugs, and provides a comprehensive high-level perspective that enables decision makers to understand the safety relationships across a group of drugs.

Like many other drug safety surveillance methods, the pharmacoepidemiological network model has been designed to analyse the safety of a specified set of candidate drugs and adverse events, in which the drug safety professional has a specific interest. One of the primary questions in the emerging field of postmarketing drug safety surveillance is whether and how such methods can also be used for broad screening (or 'data mining') of potential safety issues across the entire universe of drugs and adverse events. The primary issue is the numerous false positives that would likely be generated when analysing the full range of drugs and adverse events. While corrective measures can be taken to account for these false positives, a broad analysis of the entire spectrum of drugs and adverse events would likely require a very powerful correction, significantly limiting the sensitivity of the system and causing true cases to be missed or detected late. This is a

fundamental issue in the field that applies to most drug safety surveillance methods, and future work will continue to address this issue.

Differences between the actual safety findings of these analyses and those published in the clinical trial and epidemiological literature may be due to a number of factors, including differences between the populations covered, prescribing practices and co-morbidities, as well as differences in study design. While the database used provides broad coverage of over 9 million lives in diverse regions of the US, as described in Cziraky et al.,[23] it may not be representative of the entire US population. Still, the data used are typical of those that exist in real-world healthcare settings, and would be available for performing similar analyses in other environments. Furthermore, the database analysed reflects a commercially insured population, and results may differ in other populations. The results reported here are subject to the limitations of using existing healthcare databases to perform pharmacoepidemiological analyses in a prospective postmarketing drug safety surveillance setting, including selection bias. Established methods for dealing with selection bias include propensity score analysis.[34,35] Combining such methods with the proposed network approach is an important area for future study. Another area for future study is measuring the safety of combination therapies and identifying drug-drug interactions, such as the interaction between statins and fibrates.[23] While the present study does not explicitly incorporate dosage information, future studies could include dosage information in the model. This could be done, among other approaches, through stratified analyses that separately analyse different ranges of dosages.

While the chosen case study of cerivastatin and rhabdomyolysis illustrates the usefulness of this method in a commonly used group of drugs, rhabdomyolysis has been closely tied to statins, so additional case studies would allow a more complete evaluation of the method's performance. In order to evaluate our methods, we chose a strong, well-established and specific adverse drug event as a case study. Additional studies can be conducted with weaker associations and less specific adverse events in order to test for generalizability in those cases. Our future work will validate and extend the model across additional case studies and data environments.

Conclusions

The pharmacoepidemiological network method provides a promising new approach for drug safety surveillance using healthcare claims data. The method presented here can be used with data available today to improve on current methods for adverse drug event detection in the postmarketing stage.

Acknowledgements

This work was supported by grants R01 GM89731 and R01 GM085421 from the National Institute of General Medical Sciences, and grants R01 LM009879 and R01 LM007677 from the National Library of Medicine, National Institutes of Health. The funding sources had no role in the research.

Although Mark Cziraky, Marcus Wilson and Rhonda Bohn have not received any direct payment, Healthcore, Inc. conducts multiple research projects supported by various pharmaceutical companies in the area of biotech and device and government work. All other authors have no conflicts of interest directly relevant to the content of this study.

References

1. Staffa JA, Chang J, Green L. Cerivastatin and reports of fatal rhabdomyolysis. N Engl J Med 2002 Feb 14; 346 (7): 539-40
2. Charatan F. Bayer decides to withdraw cholesterol lowering drug. BMJ (Clin Res Ed) 2001 Aug 18; 323 (7309): 359
3. Psaty BM, Furberg CD, Ray WA, et al. Potential for conflict of interest in the evaluation of suspected adverse drug reactions: use of cerivastatin and risk of rhabdomyolysis. JAMA 2004 Dec 1; 292 (21): 2622-31
4. Bresalier RS, Sandler RS, Quan H, et al. Cardiovascular events associated with rofecoxib in a colorectal adenoma chemoprevention trial. N Engl J Med 2005 Mar 17; 352 (11): 1092-102
5. Curfman GD, Morrissey S, Drazen JM. Expression of concern reaffirmed [comment]. N Engl J Med 2006 Mar 16; 354 (11): 1193
6. McGettigan P, Henry D. Cardiovascular risk and inhibition of cyclooxygenase: a systematic review of the observational studies of selective and nonselective inhibitors of cyclooxygenase 2. JAMA 2006 Oct 4; 296 (13): 1633-44
7. Juni P, Nartey L, Reichenbach S, et al. Risk of cardiovascular events and rofecoxib: cumulative meta-analysis. Lancet 2004 Dec 4-10; 364 (9450): 2021-9
8. Wadman M. Experts call for active surveillance of drug safety. Nature 2007 Mar 22; 446 (7134): 358-9

9. Strom BL. How the US drug safety system should be changed. JAMA 2006 May 3; 295 (17): 2072-5
10. Baciu A, Stratton K, Burke SP, editors. The future of drug safety: promoting and protecting the health of the public. Institute of Medicine ad hoc Committee on the Assessment of the US Drug Safety System. Washington, DC: National Academies Press, 2006 Sep 22
11. Psaty BM, Burke SP. Protecting the health of the public: Institute of Medicine recommendations on drug safety. N Engl J Med 2006 Oct 26; 355 (17): 1753-5
12. Hennessy S, Strom BL. PDUFA reauthorization: drug safety's golden moment of opportunity? N Engl J Med 2007 Apr 26; 356 (17): 1703-4
13. Fontanarosa PB, Rennie D, DeAngelis CD. Postmarketing surveillance: lack of vigilance, lack of trust. JAMA 2004 Dec 1; 292 (21): 2647-50
14. Hampton T. Postmarket "pharmacovigilance" program on alert for adverse events from drugs. JAMA 2007 Aug 22; 298 (8): 851-2
15. Klein DF. The flawed basis for FDA post-marketing safety decisions: the example of anti-depressants and children. Neuropsychopharmacology 2006 Apr; 31 (4): 689-99
16. Klein DF, O'Brien CP. Improving detection of adverse effects of marketed drugs. JAMA 2007 Jul 18; 298 (3): 333-4
17. Platt R, Madre L, Reynolds R, et al. Active drug safety surveillance: a tool to improve public health. Pharmacoepidemiol Drug Saf 2008 Dec; 17 (12): 1175-82
18. Tiihonen J, Lonnqvist J, Wahlbeck K, et al. Antidepressants and the risk of suicide, attempted suicide, and overall mortality in a nationwide cohort. Arch Gen Psychiatry 2006 Dec; 63 (12): 1358-67
19. Meijer WE, Heerdink ER, Nolen WA, et al. Association of risk of abnormal bleeding with degree of serotonin reuptake inhibition by antidepressants. Arch Intern Med 2004 Nov 22; 164 (21): 2367-70
20. Brownstein JS, Sordo M, Kohane IS, et al. The tell-tale heart: population-based surveillance reveals an association of rofecoxib and celecoxib with myocardial infarction. PLoS ONE 2007; 2 (9): e840
21. Andrade SE, Graham DJ, Staffa JA, et al. Health plan administrative databases can efficiently identify serious myopathy and rhabdomyolysis. J Clin Epidemiol 2005 Feb; 58 (2): 171-4
22. Graham DJ, Staffa JA, Shatin D, et al. Incidence of hospitalized rhabdomyolysis in patients treated with lipid-lowering drugs. JAMA 2004 Dec 1; 292 (21): 2585-90
23. Cziraky MJ, Willey VJ, McKenney JM, et al. Statin safety: an assessment using an administrative claims database. Am J Cardiol 2006 Apr 17; 97 (8A): 61-8C
24. Brown JS, Kulldorff M, Chan KA, et al. Early detection of adverse drug events within population-based health networks: application of sequential testing methods. Pharmacoepidemiol Drug Saf 2007 Dec; 16 (12): 1275-84
25. Kulldorff M, Davis RL, Kolczak M, et al. A maximized sequential probability ratio test for drug and vaccine safety surveillance [online]. Available from URL: http://www.populationmedicine.org/content/docs/maxSPRTsubm.pdf [Accessed 2012 Mar 29]
26. Brown JS, Kulldorff M, Petronis KR, et al. Early adverse drug event signal detection within population-based health networks using sequential methods: key methodologic considerations. Pharmacoepidemiol Drug Saf 2009 Mar; 18 (3): 226-34
27. Reis BY, Kohane IS, Mandl KD. An epidemiological network model for disease outbreak detection. PLoS Medicine 2007 Jun; 4 (6): e210
28. Reis BY, Kirby C, Hadden LE, et al. AEGIS: a robust and scalable real-time public health surveillance system. J Am Med Inform Assoc 2007 Sep-Oct; 14 (5): 581-8
29. Reis BY, Pagano M, Mandl KD. Using temporal context to improve biosurveillance. Proc Natl Acad Sci U S A 2003 Feb 18; 100 (4): 1961-5
30. Law M, Rudnicka AR. Statin safety: a systematic review. Am J Cardiol 2006 Apr 17; 97 (8A): 52-60C
31. Bays H. Statin safety: an overview and assessment of the data – 2005. Am J Cardiol 2006 Apr 17; 97 (8A): 6-26C
32. Charlson ME, Pompei P, Ales KL, et al. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. J Chronic Dis 1987; 40 (5): 373-83
33. Fisher RA. Combining independent tests of significance. Am Stat 1948; 2 (5): 30
34. Wang J, Donnan PT, Steinke D, et al. The multiple propensity score for analysis of dose-response relationships in drug safety studies. Pharmacoepidemiol Drug Saf 2001 Mar-Apr; 10 (2): 105-11
35. Seeger JD, Kurth T, Walker AM. Use of propensity score technique to account for exposure-related covariates: an example and lesson. Med Care 2007 Oct; 45 (10 Suppl. 2): S143-8

Correspondence: Dr *Ben Y. Reis*, Harvard Medical School and Children's Hospital Informatics Program at the Harvard-MIT Division of Health Sciences and Technology, 1 Autumn St, Boston, MA 02115, USA.