SHORT COMMUNICATION

Drug Safety 2009; 32 (2): 137-146
0114-5916/09/0002-0137/$49.95/0

Drug-versus-Drug Adverse Event Rate Comparisons

A Pilot Study Based on Data from the US FDA Adverse Event Reporting System

Alan M. Hochberg, Ronald K. Pearson, Donald J. O'Hara and *Stephanie J. Reisinger*

ProSanos Corporation, Harrisburg, Pennsylvania, USA

Abstract

Background: A number of published studies compare adverse event rates for drugs on the basis of reports in the US FDA Adverse Event Reporting System (AERS). While the AERS data have the advantage of timely availability and a large capture population, the database is subject to many significant biases, and lacks complete patient information that would allow for correction of those biases. The accuracy of comparative AERS-based data mining has been questioned, but has not been systematically studied.

Objective: To determine whether AERS could be used as a data source to accurately compare the adverse event rates for pairs of drugs, using predefined, stringent criteria to dictate whether a given pair of drugs was considered eligible for such a comparison.

Methods: The Fisher's Exact test was utilized to detect differences in adverse event rates between such pairs of drugs. Concordance was determined between statistically significant AERS-based adverse event rate differences, and adverse event rate differences published in the literature from clinical trials and case-control studies. The conditions for validity included (i) data that are free of 'extreme duplication' in AERS reports; (ii) drugs used in similar patient populations; (iii) drugs used for similar indications; (iv) drugs used with the same spectrum of concomitant medications; and (v) drugs not widely disparate in time on the market.

Results: For 19 drugs studied, a total of 36 evaluable adverse event rate comparisons were identified. Comparisons were classified as favouring 'drug A', favouring 'drug B' or detecting no difference. Concordance for the resulting 3×3 table (AERS vs literature) gave a kappa statistic of 0.654, indicating moderately good agreement. In only two cases was there absolute discordance, with AERS designating one drug as having a lower rate, while the published study designated the other drug as having a lower rate, with respect to a given adverse event.

Conclusions: This pilot study encourages further research regarding the use of spontaneous report databases such as AERS, under stringently defined

conditions, to compare adverse event rates for drugs. While not hypothesis proving, such estimates can be used for purposes such as generating hypotheses for controlled studies, and for designing those studies.

Background

Controlled studies that compare adverse event rates for two or more alternative drugs are relatively rare. They are difficult to design and run, expensive and potentially time consuming, particularly when adverse events occur in relatively few patients.

The US FDA Adverse Event Reporting System (AERS) database[1] is a readily available source of safety information for many drugs. Whether or not 'data mining' of the AERS database can be used to compare adverse event rates of related drugs is a matter of active debate in the pharmacovigilance community. An FDA-Pharmaceutical Research and Manufacturers of America (PhRMA) expert working group stated: ''It is tempting to compare signal scores at some level, and it certainly is easy to construct various statistics for this purpose. However, differences between reporting ratios do not imply differences in risk because spontaneous reporting databases are biased in ways that cannot be measured or controlled. It is not legitimate to infer that differences between scores imply differences between treatments without carefully considering the mechanisms that generate reports, including the known and unknown biases.''[2]

The known biases are many. They include differences in the propensity to prescribe different drugs depending on demographic and clinical factors. Differences in drug cost can introduce biases due to socioeconomic factors. Certain drugs are specifically contraindicated for patients at risk for certain adverse events, while other drugs are recommended for those same patients. This phenomenon can drastically bias measured adverse event rates. While some biases affect the actual rates of occurrence for adverse events, others affect the rate at which adverse events are reported to the AERS database. These factors include the age of the drug, as well as medico-legal and publicity considerations, to name just a few.

While the advisability of using AERS for drug comparison is unclear, the fact is that AERS-based comparisons are being conducted and published, largely because of the relative timeliness of the data, its free availability to the public and the relatively large capture population for adverse events. While acknowledging the potentially overwhelming biases, the FDA itself recently used comparative data-mining methodology in a study of 16 antibacterials, in order to elucidate the adverse event profile of a drug of interest, telithromycin. The resulting study report was made available through the FDA website.[3] A number of additional comparative studies based on AERS data have appeared in the peer-reviewed literature, including, for instance, a paper on gatifloxacin and dysglycaemia.[4]

In practice, the bias situation may be somewhat ameliorated by the fact that important comparisons often involve pairs of drugs that are chemically and therapeutically very similar. A subset of these is sometimes referred to as 'me-too' drugs. In some cases, such drugs may be nearly equivalent in efficacy, but may differ in important ways in their safety profiles. For instance, dysglycaemia is seen with gatifloxacin much more frequently than with other fluoroquinolone antibacterials, despite their chemical and microbiological similarities. This is a fascinating example of an idiosyncratic phenomenon, the opposite of a 'class effect'.

Our goal in this study was to determine whether valid, qualitative comparisons of adverse event rates for restrictively defined pairs of similar drugs could be derived from AERS data. We recognized that such comparisons would not ever be definitive, but might be useful in corroborating other studies, in assessing the magnitude of a safety issue and in generating hypotheses and

designing controlled studies. By comparing AERS data for pairs of drugs, and determining the degree of concordance or discordance with published studies, we also set out to confirm our *a priori* definition of the restrictions under which such comparisons may be valid, and to determine what circumstances might cause an AERS-based comparison to be misleading.

Comparative Adverse Event Rates from Adverse Event Reporting System (AERS) Data

This study utilized the public release of AERS data from January 2001 to December 2005. The study was originally a component of a Small Business Innovation Research (SBIR) research project funded by the Department of Defense. So to select drugs of particular interest to the sponsor, drugs were chosen from the Department of Defense Joint Deployment Formulary, in a sample of therapeutic areas that would reflect usage in primary care for defense forces. Those drugs are listed in table I. We chose six American Hospital Formulary Service (AHFS) drug categories that contained multiple therapeutic alternatives within each category, so that comparisons could be made on a pair-wise basis.

Table I. Drugs included in the comparison of adverse event rates, showing the six comparison groups

Drug class	Drug
Antihelminthics	Albendazole
	Mebendazole
Antihistamines	Brompheniramine
	Cetirizine
	Chlorpheniramine
	Diphenhydramine
	Loratadine
	Promethazine
	Pseudoephedrine
	Triprolidine
Antimalarials	Atovaquone
	Chloroquine
	Mefloquine
Fluoroquinolone antibacterials	Ciprofloxacin
	Gatifloxacin
	Levofloxacin
NSAIDs	Diclofenac
	Indometacin
	Ketorolac
	Naproxen
HMG-CoA reductase inhibitors ('statins')	Atorvastatin
	Lovastatin
	Pravastatin
	Simvastatin

For a given pair of drugs 'A' and 'B' within a group, we tallied all MedDRA[1] preferred terms that occurred at least once for either drug. (MedDRA, the Medical Dictionary for Regulatory Activities terminology, is the international medical terminology developed under the auspices of the International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use [ICH]).

For each preferred term, we compared 'drug A' versus 'drug B' as follows: we considered the number of reports mentioning the preferred term of interest for 'drug A' as a proportion of all reports for 'drug A'. We then considered the number of reports mentioning the preferred term of interest for 'drug B' as a proportion of all reports for 'drug B'. Proportions were compared using the Fisher's Exact test. (This calculation is similar in philosophy to the well known 'proportional reporting rate' calculation, but is tailored specifically for the comparison of two drugs.) The p-value threshold for a statistically significant difference in proportions was adjusted by the Bonferroni method. That is, the p-value was compared with 0.05 after it was multiplied by N, where N is the number of different MedDRA preferred terms that occurred at least once for either drug. While the Bonferroni method has been superseded in many clinical research applications, it has been successfully used for adverse

1 MedDRA is a registered trademark of the International Federation of Pharmaceutical Manufacturers and Associations (IFPMA).

event report threshold adjustment[5] where a similarly high degree of multiplicity is present.

Criteria for a Comparison

Based on our prior experience in drug safety data mining, we prospectively defined stringent criteria for the inclusion of drug comparisons in this study. The experience base for these criteria included statistical and epidemiological 'first principles' regarding the validity of comparisons and known confounding factors, as well as previous attempts to compare adverse event rates, by ourselves and others. We benefitted from critiques of those previous efforts, both published and unpublished, including discussions at oral presentations.

We adhered to these criteria, even though doing so severely limited the number of comparable drug pairs. We felt that if we were unable to validate comparisons under these stringent criteria, it was a moot question as to how much the criteria could be loosened to increase inclusiveness. As we noted earlier, there are many pairs of chemically and pharmacologically similar drugs on the market, so even a highly selective approach to drug comparability may have practical utility.

The criteria for comparability are as follows:

1. Data for the drug-event combination under study must be free of 'extreme duplication'. As discussed by Hauben et al.,[6] a small number of adverse event reports in the AERS database appear to be replicated dozens of times, in a way that can skew the statistical distribution of adverse event reports for the affected drug, and materially alter data-mining results.
2. For the event under study, the two drugs must be used in the same patient population with respect to age and sex.
3. For the event under study, the drugs must be used for a similar spectrum of indications.
4. For the event under study, the drugs must be used along with a similar spectrum of concomitant medications.
5. Drugs that have been on the market for more than 2–3 years are not compared with drugs in the peri-approval period.

These criteria were assessed subjectively by the authors, by tabulating the AERS data records relevant to the drug-event combinations by age, sex, indication and concomitant medication. Duplicates were assessed by visually examining the event date, age, sex and weight contained in each AERS report, along with the list of drugs and adverse events reported. While these assessments were subjective, they were also prospective: they were completed before the comparison between the AERS findings and the publications was performed. In the following discussion, we consider how these subjective criteria might be translated into objective counterparts. This, however, would require an entirely separate study for validation.

Qualification of Drug Pairs and Events for Comparison

The criteria described disqualified a number of drug pairs and events for comparison purposes:

- Ketorolac: ketorolac is used as a component of a buprenorphine-ketorolac regimen for the treatment of opioid withdrawal.[7] Ketorolac therefore shows an unusually strong association with morphine and other opiates as concomitant medications, relative to other NSAIDs. As a result, certain events such as hypotension are seen with unusual frequency. Comparisons of ketorolac with other NSAIDs were prospectively disqualified for events where this association with opiates was found.
- Brompheniramine: phenylpropanolamine frequently was seen as a concomitant medication with brompheniramine, while cetirizine and other antihistamines did not share this frequent association. Therefore, comparisons involving this drug were prospectively disqualified (see the Discussion section).
- Promethazine: this drug was eliminated from comparison with other antihistamines, because it was apparent from the AERS data that its primary use is as an anti-nausea agent rather than as an anti-allergy antihistamine. This was detectable, not through the AERS indication field, which was frequently missing for this

drug, but through the association with a number of unusual concomitant medications, including paclitaxel and amphotericin B, not seen as frequently with other antihistamines.
- Indometacin: this drug was eliminated from comparison with other NSAIDs for certain adverse events, such as premature birth, because of its dual use as a tocolytic agent and treatment for patent ductus arteriosus, as well as an NSAID. This was detectable using the above criteria, based on the distribution of patient ages and weights, on a bias toward female sex, and through a unique pattern of concomitant medications for an NSAID, for example, a strong association with dopamine.
- Triprolidine: this drug was not eliminated for any of these reasons; however, no relevant adverse event rate differences were found in published studies, so it does not appear in the table of results.

Concordance with Published Studies and Prescribing Information

We obtained the results of quantitative adverse event rate comparisons by searching the PubMed database through the first quarter of 2006, with no limitation on the earliest publication date, using the generic names for the drug pair of interest as keywords, for example, 'cetirizine AND diphenhydramine'. We selected articles that were specifically designed to study safety, as well as efficacy trials that also compared rates of adverse events, with sufficient patient numbers to establish statistically significant differences. For each adverse event term discussed in each article (e.g. headache), we tabulated stated conclusions from the article that took one of three alternative forms:

- The rate for drug A is significantly higher than for drug B.
- The rate for drug B is significantly higher than for drug A.
- There is no statistically significant difference in the rate between drugs A and B.

Studies that were based on spontaneous report data were eliminated from consideration.

It should also be noted that some but not all of the published literature described adverse events using MedDRA terminology. Where necessary, AERS results were mapped to the terminology of the relevant literature on an *ad hoc* basis, for example, liver function test elevation was counted as concordant to MedDRA preferred terms including 'Aspartate aminotransferase increased', 'Alanine aminotransferase increased' and 'Hepatic function abnormal NOS' (not otherwise specified).

Concordance Results

For 198 drug/drug/adverse event combinations that were screened by the above criteria and considered evaluable, we identified a total of 36 instances in which a published comparative study had appeared. In most cases, the publications described more than one eligible adverse event. Results for significant differences are shown in table II, and the concordance of AERS with published studies is shown in figure 1. In 26 cases (72% of the total), the AERS-based comparison agreed with the published study. In two cases (6%), the AERS-based comparison showed a difference between the drugs, while the published study did not. In six cases (17%), the published study showed a difference between the drugs, but the AERS-based comparison did not. However, in four of those cases, the AERS data showed a trend in the same direction as the published study that did not reach statistical significance. In only two cases (6%), was there absolute discordance: the AERS-based comparison showed one drug had a lower adverse event rate, while the published study stated that the other drug did.

The two cases of absolute discordance between the AERS-based comparison and the published study are sedation for cetirizine compared with diphenhydramine, and rhabdomyolysis for lovastatin versus pravastatin. Diphenhydramine is known to be a sedating antihistamine, yet there was a higher rate of reports of sedation in AERS for cetirizine. This may well be due to the fact that sedation is expected with diphenhydramine[23] and, therefore, it is not reported to AERS. This

Table II. Differences in adverse event (AE) rates, either detected by analysis of the Adverse Event Reporting System (AERS) data as described in the text, or described in published studies or drug labelling

First drug	Second drug	AE MedDRA preferred terms	AERS: drug with lower AE rate; Bonferroni corrected p-value for the Fisher's Exact test[b]	Publication: drug with lower AE rate
Albendazole	Mebendazole	Headache	Mebendazole; $p < 10^{-9}$	Mebendazole[8]
Cetirizine	Chlorpheniramine	Nausea	Chlorpheniramine; $p = 0.006$	Chlorpheniramine[9]
Cetirizine	Chlorpheniramine	Headache	Chlorpheniramine (trend); $p = 0.055$	Chlorpheniramine[9]
Cetirizine	Chlorpheniramine	Upper abdominal pain	Chlorpheniramine (trend); $p = 0.15$	Chlorpheniramine[9]
Cetirizine	Diphenhydramine	Sedation	Diphenhydramine; $p < 10^{-9}$	Cetirizine[9]
Cetirizine	Diphenhydramine	Electrocardiographic abnormalities	Cetirizine (trend); $p = 0.2$	Cetirizine[9]
Cetirizine	Loratadine	Sedation	Loratadine; $p = 10^{-4}$	Loratadine[9]
Cetirizine	Loratadine	Headache	(Tie); p = NS	Loratadine[9]
Cetirizine[a]	Loratadine	Hypospadias	Cetirizine; $p = 10^{-4}$	Cetirizine[10]
Cetirizine[a]	Loratadine	Hypospadias	Cetirizine; $p = 10^{-4}$	(Tie)[11]
Atovaquone	Chloroquine	Neuropsychiatric symptoms	Atovaquone; $p < 10^{-9}$	Atovaquone[12]
Atovaquone	Chloroquine	Fever	(Tie); p = NS	(Tie)[13]
Atovaquone	Chloroquine	Headache	(Tie); p = NS	(Tie)[13]
Atovaquone	Mefloquine	Psychiatric symptoms	Atovaquone; $p = 10^{-5}$	Atovaquone[12]
Ciprofloxacin	Gatifloxacin	Electrocardiographic abnormalities	Ciprofloxacin; $p = 10^{-5}$	Ciprofloxacin[14]
Ciprofloxacin	Gatifloxacin	Hyperglycaemia/hypoglycaemia	Ciprofloxacin; $p < 10^{-9}$	Ciprofloxacin[15]
Ciprofloxacin	Levofloxacin	Torsade de pointes	Ciprofloxacin; $p = 10^{-5}$	Ciprofloxacin[14]
Ciprofloxacin	Levofloxacin	Tendon rupture	Ciprofloxacin; $p < 10^{-9}$	Ciprofloxacin[16]
Gatifloxacin	Levofloxacin	Hyperglycaemia/hypoglycaemia	Levofloxacin; $p < 10^{-9}$	Levofloxacin[15]
Gatifloxacin	Levofloxacin	Liver function test elevation	Levofloxacin; $p < 10^{-9}$	Levofloxacin[17]
Gatifloxacin	Levofloxacin	Tendon disorders	Levofloxacin; $p < 10^{-9}$	Levofloxacin[15]
Diclofenac	Ketorolac	Liver function test elevation	Ketorolac; $p = 10^{-5}$	(Tie)[18]
Diclofenac	Ketorolac	Surgical site bleeding	(Tie); p = NS	(Tie)[18]
Diclofenac	Ketorolac	Acute renal failure	(Tie); p = NS	(Tie)[18]
Diclofenac	Ketorolac	Allergic reaction	(Tie); p = NS	(Tie)[18]
Diclofenac	Naproxen	Liver function test elevation	Naproxen; $p < 10^{-9}$	Naproxen[19]
Diclofenac	Naproxen	Gastrointestinal events	(Tie); p = NS	(Tie)[20]
Atorvastatin	Lovastatin	Rhabdomyolysis	(Tie); p = NS	(Tie)[21]
Atorvastatin	Pravastatin	Rhabdomyolysis	Pravastatin; $p = 0.002$	Pravastatin[21]
Atorvastatin	Pravastatin	Liver function test elevation	(Tie); p = NS	Pravastatin[21]
Atorvastatin	Simvastatin	Liver function test elevation	Simvastatin; $p = 0.004$	Simvastatin[22]
Atorvastatin	Simvastatin	Diarrhoea	Simvastatin (trend); p = NS	Simvastatin[22]
Atorvastatin	Simvastatin	Rhabdomyolysis	(Tie); p = NS	(Tie)[21]
Lovastatin	Pravastatin	Rhabdomyolysis	Lovastatin; $p = 0.01$	Pravastatin[22]
Lovastatin	Simvastatin	Rhabdomyolysis	(Tie); p = NS	(Tie)[21]
Pravastatin	Simvastatin	Rhabdomyolysis	Pravastatin; $p < 10^{-9}$	Pravastatin[21]

a Note that for cetirizine, loratadine and hypospadias, we identified two studies in the literature that differed in their conclusions. Both met the criteria for inclusion, and both are reflected in this table and in the results.

b Bonferroni-corrected p-values <0.001 are listed as order of magnitude; those >0.2 are listed as not siginificant (NS) since the Bonferroni correction is not valid for larger values of p.

difference was not noted when the eligibility criteria were applied, so this drug pair was prospectively considered evaluable with respect to sedation. A reviewer of this manuscript noted that cetirizine is a prescription drug, while diphenhydramine is available over the counter, so a reporting bias in AERS toward prescription drugs may explain this discrepancy as well.

We have no clear explanation with regard to the discrepancy for lovastatin versus pravastatin and rhabdomyolysis, but we note that the p-value for the AERS comparison is much larger (less significant) than most others.

Using three categories, A > B, no difference and B > A, the inter-rater agreement between the literature and data-mining results can be calculated using Cohen's kappa statistic.[24] Within each pair of drugs, the assignment of which drug is A was done on the basis of alphabetical order. For the 36 cases, we obtain kappa = 0.654, which corresponds to a moderately high degree of concordance.

While the original focus of AERS was on serious adverse events, the database contains information on both serious and non-serious events, and the most frequently used MedDRA terms are for non-serious events. We have included both types in this study. Of the 36 events studied, there are 25 serious events, and considering these by themselves yields a kappa of 0.695, slightly higher but essentially unchanged from the value above for all events.

Discussion

Under the conditions used in this study, AERS-based detection of adverse event rate differences between similar drugs were in agreement with published studies 72% of the time. In only two cases was there outright discordance. This indicates at least a moderate degree of agreement between AERS comparative results and published studies.

Restrictive conditions for performing a comparison have been outlined. These include: (i) de-duplication of the data; (ii) both drugs are used in the same patient population; (iii) drugs are used for the same indication; and (iv) drugs are used with a similar spectrum of concomitant medications. As more is learned about the use of AERS for comparisons of adverse event rates, it may be possible to refine these conditions for greater accuracy, or to relax them to include more drug pairs.

We note subjectivity involved in assessing the five criteria for comparability that we have outlined. We are optimistic that objective criteria can be identified and defined. We know from other studies that there are a number of methods for dealing with duplicate records, some of which have been incorporated into commercial data-mining software packages. For example, in one commercial package, duplicate records are identified on the basis of AERS event date, sex, age and weight. The rule for duplicate detection is that three of these four fields must match, and the fourth field must not conflict. Missing data are scored as matching another missing data value, and not conflicting with any data value.

We expect that the other comparability criteria could be made objective by considering the observed and expected frequency distributions of age, sex, indication and concomitant medications for the two drugs, for each event of interest. It should be possible to detect when observed counts for categories based on these variables deviate from expected frequencies beyond a pre-defined threshold. A separate study with

	AERS count rate for Drug A > Drug B	AERS count rate not significantly different	AERS count rate for Drug B > Drug A
Publication states rate for Drug A > Drug B	9	5	1
Publication finds no significant difference	1	9	1
Publication states rate for Drug B > Drug A	1	1	8

Fig. 1. Concordance of Adverse Event Reporting System (AERS) results vs published studies for the comparison of adverse event rates for drug pairs. 'Drug A' was taken to be the first drug alphabetically for each pair of drugs compared. Shaded cells indicate agreement.

independent data would be required to validate such a methodology.

One of the two comparison errors detected by this study leads us to consider that it may be appropriate to add another comparison criterion that prevents a prescription drug from being compared with an over-the-counter drug.

The field of drug safety inherently requires making timely decisions on the basis of imperfect and incomplete information. While they are not definitive by themselves, the AERS-based comparisons described here might provide corroborating evidence for anecdotal information, or they can be used to generate hypotheses for controlled studies. They may be used as well to discourage studies that are likely to be futile: if the AERS data says, ''A has a lower rate than B for adverse event X,'' our results show that the likelihood of proving ''B has a lower rate than A,'' in a controlled study, is fairly low.

Limitations of the Study

This study has a number of significant limitations, among which is the possibility that some of the studies cited here had an influence on the reporting of adverse events through the FDA MedWatch system. Published studies, and publicity of drug safety issues in the professional and lay press, can stimulate AERS reporting. This phenomenon is best known for serious adverse events that become the subject of legal action.

Published studies could also inhibit AERS reporting, since clinicians and the public might then consider a reaction to be 'well known' and not worth reporting. Thus, publications could result in a change in either direction in the rate of AERS reporting. We note that some drug pairs were prospectively eliminated from this study because of strong concomitant use with (and, in some cases, co-formulation with) phenylpropanolamine. The adverse events of phenylpropanolamine were widely publicized at the time when it was removed from the non-prescription market.

Conversely, FDA-AERS data may have influenced the authors of the published studies, even though we eliminated any published study where the AERS data were explicitly cited. The median publication date for the studies was 2004, though some were reviews that cited early work. This time period overlaps the collection period, 2001–5, for the AERS data used here.

An additional characteristic of this study that should be kept in mind is that we did not include recently approved drugs. Using data from 2001–5 avoided the peri-approval period for all of the drugs chosen, including the newest ones. Recently approved drugs are often subject to an initial inrush of adverse event reports, known as the 'Weber effect'.[25] The Weber effect is generally considered to encompass the first 2 years of the life of a drug as a marketed product, though it may extend considerably beyond that time. Thus, a comparison of drugs that had been on the market for substantially different periods of time might produce biased results. The drugs in this study had all been on the market for at least several years, and in many cases, far longer. This was a consequence of our choosing drugs from the Joint Deployment Formulary. The conclusions of this study do not necessarily extend to newer drugs, and, given the literature on the Weber effect, we would caution against comparing drugs that had been on the market for substantially different lengths of time, particularly where one or both of the drugs has been on the market for less than 2 or 3 years.

The criteria for whether or not drugs were considered to be 'different' in this study were statistical ones, not clinical ones. In any study that statistically compares drugs, including this one, it is important to consider whether any detected difference is large enough to be clinically meaningful, not just statistically detectable.

We note that our statistical criteria for detecting differences using AERS data mining were based on the Bonferroni correction, while the criteria used in the peer-reviewed publications varied from paper to paper. Multiplicity corrections are rarely used in published safety analyses for reasons of conservatism. Thus, even if AERS contained identical data to that found in one of the papers, differences in statistical tests might lead us to different conclusions regarding comparisons of drug-event associations.

We also note that this study considered both serious and non-serious adverse events. While the initial purpose of the AERS database was for the identification of serious adverse events, data on non-serious adverse events are collected as well. In fact, the most frequently occurring terms in the database, such as 'nausea', 'headache', etc., relate to these. Rules for the collection of AERS data call for curtailment in the collection of non-serious adverse event reports under certain conditions. But, in practice, these events continue to be reported incidentally to both related and unrelated serious adverse events, resulting in ongoing increases in total report counts for these events.

One of the limitations of the AERS database is the lack of complete and accurate information regarding the indication for which drugs are being prescribed. One must draw inferences from the data on concomitant medications and reported reactions, in order to determine that one is not attempting to compare drugs across different indications. There is also a significant amount of missing age and sex information (7–25%), so it is difficult to establish that the populations being treated are comparable. One possible way to improve these situations is through the mining of electronic health record databases for adverse event rate information.

Conclusion

This pilot study encourages further research into the possibility of using data mining of spontaneous reports as a method of comparing rates of adverse events, at least under certain stringent conditions. While not hypothesis proving, such comparisons might be useful in the preliminary steps of a safety investigation. For example, they might be used for generating hypotheses for testing with additional data sources, particularly electronic health record systems. Electronic health record data permit the use of propensity methods and other statistical techniques, which can correct for some of the confounding factors that were used as exclusion criteria in this study.

Acknowledgement

The authors gratefully acknowledge the helpful advice of LTC Trinka S. Coster, MD, and Kevin Hall, MD, regarding the design and interpretation of this study. This work was supported by the US Army Medical Research and Materiel Command under Contract No. W81XWH-06-C-0059. The views, opinions and/or findings contained in this report are those of the authors and should not be construed as an official Department of the Army position, policy or decision unless so designated by other documentation. The authors also wish to thank the anonymous reviewers of this manuscript for many helpful comments.

Alan Hochberg, Donald O'Hara and Stephanie Reisinger disclose that they are employees of ProSanos Corporation, a vendor of drug safety data-mining products and services. Ronald Pearson was an employee of ProSanos Corporation at the time of this study. The work described here was funded by a SBIR grant from the Department of Defense.

References

1. Rodriguez EM, Staffa JA, Graham DJ. The role of databases in drug postmarketing surveillance. Pharmacoepidemiol Drug Saf 2001 Aug-Sep; 10 (5): 407-10
2. Almenoff J, Tonning JM, Gould AL, et al. Perspectives on the use of data mining in pharmaco-vigilance. Drug Saf 2005; 28: 981-1007
3. Szarfman A, Levine J. Memorandum: NDA 21-144 KETEK (telithromycin) 400MG tablets® (Sanofi-Aventis): data mining analysis of adverse events and outcomes reported for telithromycin and 15 comparator drugs (AERS data) [online]. Available from URL: http://www.fda.gov/ohrms/dockets/AC/06/briefing/2006-4266b1-02-06-FDA-appendic-f.pdf [Accessed 2007 Feb 21]
4. Frothingham R. Glucose homeostasis abnormalities associated with use of gatifloxacin. Clin Infect Dis 2005 Nov 1; 41 (9): 1269-76
5. Hochberg AM, Reisinger SJ, Pearson RK, et al. Using data mining to predict safety actions from FDA Adverse Event Reporting System data. Drug Inf J 2007; 41 (5): 633-43
6. Hauben M, Reich L, DeMicco J, et al. 'Extreme duplication' in the US FDA Adverse Events Reporting System database. Drug Saf 2007; 30 (6): 551-4
7. Telias D, Nir-Hod J. Buprenorphine-ketorolac vs. clonidine-naproxen in the withdrawal from opioids. Int J Psychosocial Rehab 2000; 5: 29-33
8. Gibbon C, Blockman M, Barnes K, et al., editors. The South African medicines formulary. 7th ed. Cape Town: Division of Clinical Pharmacology, University of Cape Town, 2005
9. Helfand M. Drug class review on newer antihistamines. Portland (OR): Oregon Health & Science University, 2006
10. Loratadine, desloratadine and pregnancy: don't use, risk of hypospadias. Prescrire Int 2003 Oct; 12 (67): 183
11. Centers for Disease Control and Prevention (CDC). Evaluation of an association between loratadine and hypospadias: United States, 1997–2001. MMWR Morb Mortal Wkly Rep 2004 Mar 19; 53 (10): 219-21

12. Schlagenhauf P, Steffen R. Neuropsychiatric events and travel: do antimalarials play a role? J Travel Med 2000; 7: 225-6
13. Camus D, Djossou F, Schilthuis HJ, et al. Atovaquone-proguanil versus chloroquine-proguanil for malaria prophylaxis in nonimmune pediatric travelers: results of an international, randomized, open-label study. Clin Infect Dis 2004 Jun 15; 38 (12): 1716-23
14. Frothingham R. Rates of torsades de pointes associated with ciprofloxacin, ofloxacin, levofloxacin, gatifloxacin, and moxifloxacin. Pharmacotherapy 2001 Dec; 21 (12): 1468-72
15. Park-Wyllie LY, Juurlink DN, Kopp A. Outpatient gatifloxacin therapy and dysglycemia in older adults. N Engl J Med 2006 Mar 30; 354 (13): 1352-61
16. Kashida Y, Kato M. Characterization of fluoroquinolone-induced Achilles tendon toxicity in rats: comparison of toxicities of 10 fluoroquinolones and effects of anti-inflammatory compounds. Antimicrob Agents Chemother 1997 Nov; 41 (11): 2389-93
17. Rubenstein E. History of the quinolones and their side effects. Chemotherapy 2001; 47 Suppl. 3: 3-8; discussion 44-8
18. Forrest JB, Camu F, Greer IA, et al. Ketorolac, diclofenac, and ketoprofen are equally safe for pain relief after major surgery. Br J Anaesth 2002 Feb; 88 (2): 227-33
19. Rostom A, Goldkind L, Laine L. Nonsteroidal anti-inflammatory drugs and hepatic toxicity: a systematic review of randomized controlled trials in arthritis patients. Clin Gastroenterol Hepatol 2005 May; 3 (5): 489-98
20. Hippisley-Cox J, Coupland C, Logan R. Risk of adverse gastrointestinal outcomes in patients taking cyclo-oxygenase-2 inhibitors or conventional non-steroidal anti-inflammatory drugs: population based nested case-control analysis. BMJ 2005; 331: 1310-6
21. Law M, Rudnicka AR. Statin safety: a systematic review. Am J Cardiol 2006 Apr 17; 97 (8A): 52C-60C
22. Helfand M, Carson S, Kelley C. Drug class review on HMG-CoA reductase inhibitors (statins). Portland (OR): Oregon Health & Science University, 2004
23. Tu RH, Grewall P, Leung JW, et al. Diphenhydramine as an adjunct to sedation for colonoscopy: a double-blind randomized, placebo-controlled study. Gastrointest Endosc 2006 Jan; 63 (1): 87-94
24. Cohen J. A coefficient of agreement for nominal scales. Educ Psychol Meas 1960; 20: 37-46
25. Hartnell NR, Wilson JP. Replication of the Weber effect using postmarketing adverse event reports voluntarily submitted to the United States Food and Drug Administration. Pharmacotherapy 2004 Jun; 24 (6): 743-9

Correspondence: Mr *Alan M. Hochberg*, ProSanos Corporation, 225 Market Street, Suite 502, Harrisburg, PA 17101, USA.
E-mail: alan.hochberg@prosanos.com