SHORT COMMUNICATION

Drug Saf 2010; 33 (2): 139-146
0114-5916/10/0002-0139/$49.95/0

Prospective Data Mining of Six Products in the US FDA Adverse Event Reporting System

Disposition of Events Identified and Impact on Product Safety Profiles

Steven Bailey, Ajay Singh, Robert Azadian, Peter Huber and *Michael Blum*

Global Safety Surveillance and Epidemiology, Wyeth, Collegeville, Pennsylvania, USA

Abstract

Background: The use of data mining has increased among regulators and pharmaceutical companies. The incremental value of data mining as an adjunct to traditional pharmacovigilance methods has yet to be demonstrated. Specifically, the utility in identifying new safety signals and the resources required to do so have not been elucidated.

Objectives: To analyse the number and types of disproportionately reported product-event combinations (DRPECs), as well as the final disposition of each, in order to understand the potential utility and resource implications of routinely conducting data mining in the US FDA Adverse Event Reporting System (AERS).

Methods: We generated DRPECs from AERS for six of Wyeth's products, prospectively tracked their dispositions and evaluated the appropriate DRPECs in the company's safety database. We chose EB05 (the lower bound of the 90% confidence interval around the Empirical Bayes Geometric Mean) ≥2 as the appropriate metric, employing stratification based on age, sex and year of report.

Results: A total of 861 DRPECs were identified – the average number of DRPECs was 144 per product. The proportion of unique preferred terms (PTs) in AERS for each drug with an EB05 ≥2 was similar across the six products (5.1–8.5%). Overall, 64.0% (551) of the DRPECs were closed after the initial screening (44.8% labelled, 14.3% indication related, 4.9% non-interpretable). An additional 9.9% (85) had been reviewed within the prior year and were not further reviewed. The remaining 26.1% (225) required full case review. After review of all pertinent reports and additional data, it was determined which of the DRPECs necessitated a formal review by the company's ongoing Safety Review Team (SRT) process. In total, 3.6% (31/861) of the DRPECs, yielding 16 medical concepts, were reviewed by the SRT,

leading to seven labelling changes. These labelling changes involved 1.9% of all DRPECs generated. Four of the six compounds reviewed as part of this pilot had an identified labelling change. The workload required for this pilot, which was driven primarily by those DRPECs requiring review, was extensive, averaging 184 hours per product.

Conclusion: The number of DRPECs identified for each drug approximately correlated with the number of unique PTs in the database. Over one-half of DRPECs were either labelled as per the company's reference safety information (RSI) or were under review after identification by traditional pharmacovigilance activities, suggesting that for marketed products these methods do identify adverse events detected by traditional pharmacovigilance methods. Approximately three-quarters of the 861 DRPECs identified were closed without case review after triage. Of the approximately one-quarter of DRPECs that required formal case review, seven resulted in an addition to the RSI for the relevant products. While this pilot does not allow us to comment on the utility of routine data mining for all products, it is significant that several new safety concepts were identified through this prospective exercise.

Background

Data mining tools are increasingly being utilized by regulators, pharmacovigilance centres, and pharmaceutical companies to detect potential new safety signals in postmarketing spontaneous reporting databases. The use of these methods has been extensively reviewed[1-3] and their limitations are well recognized.[4,5] Therefore, data mining in spontaneous reporting databases is considered an adjunct to other pharmacovigilance activities and needs to be integrated into existing pharmacovigilance processes. Although data mining algorithms may have utility in identifying unexpected or 'surprise' reactions, the incremental value remains unclear and needs further elucidation.[6] The analysis of the validity and usefulness of data mining algorithms in large safety databases reported to date has generally been of three types.

The first is the theoretical testing of the various measures of drug-event association performed on simulated databases.[7,8] While instructive, the practical application of such exercises is limited as they are based on models of, rather than on, actual safety databases.

Second, the concordance between the various measures of disproportionality has been examined, measuring the degree of 'agreement' between the results of applying different methods to the same database.[9-12] Such studies have focused on validity between methods rather than whether the 'signals' identified were true safety signals related to an actual product.

This limitation has been addressed by the third form of analysis described in the literature: the sensitivity and specificity of data mining methods to detect 'true signals' have generally been studied by comparing the signals generated retrospectively with terms noted in product labelling. Various authors have taken this approach using the Bayesian Confidence Propagation Neural Network (BCPNN),[13] empirical Bayes[14] and other methodology. Almost all of the published data concerning such studies is based on *retrospective* analyses of either pharmaceutical or regulatory authority databases and existing labels.[15,16]

In order to expand on the previous research, our pilot study evaluated the utility and workload implications of routinely conducting *prospective*

data mining in the US FDA Adverse Event Reporting System (AERS), in which all potential signals (which we refer to as disproportionately reported product-event combinations [DRPECs[1]]) are analyzed and, if necessary, evaluated in Wyeth's safety database. By prospectively performing our analysis, we were able to determine not just the number of terms already labelled, but the status of all DRPECs, especially those that led to additions to the product's label. To do so, we applied the Multi-item Gamma Poisson Shrinker (MGPS) method to AERS for six products, using EB05 (the lower bound of the 90% confidence interval around the Empirical Bayesian Geometric Mean[17]) as our metric. We evaluated the DRPECs identified using the company database, and prospectively tracked their disposition. Initial triage was performed to determine which DRPECs required further database review, which determined any potential association with the product and possible inclusion in the label. The final dispositions of each DRPEC were carefully recorded to determine the impact of this adjunctive process on the existing safety profile.

Objectives

Our aims were to (i) gain further insight into the numbers and types of DRPECs identified by a prospective data mining run on the AERS database across a variety of products; (ii) to analyze all DRPECs identified and evaluate the distribution of their ultimate dispositions; and (iii) understand the potential value of such a data mining programme by evaluating the labelling changes generated during this pilot.

Methods

Disproportionately Reported Product-Event Combinations (DRPEC) Identification Methods

There is no clear consensus on the best statistical approach to determine DRPECs.[18] Numerous data mining algorithms are available, including Proportional Reporting Rate (PRR),[19] BCPNN[20,21] and MGPS.[22] Each has its merits and limits, as noted in the literature.[23,24] The specific methodology used is often based on an assessment of the sensitivity of the method and the attendant workload implications, and will be determined by many situational factors.

After a review of this literature, Wyeth settled on the MGPS methods, with a threshold of EB05 ≥2, as this generally provides relative confidence that the DRPEC is reported at least twice the rate expected than if there were no association with the product.[25] Based on our previous experience and research, the company chose to stratify by age, sex and year of report.[26] Together, these result in a more reasonable workload for the pharmacovigilance physicians and scientists (termed 'pharmacovigilance group' in this paper).

DRPECs were identified by applying the proprietary Qscan® software of DrugLogic® to the AERS database from inception cumulatively through the second quarter of 2006, inclusive. Based on a literature review, the company chose to include only reports in which the drug in question was listed as the suspect drug (either primary or secondary).[16,27] All data mining runs were performed against all other products in the AERS database. The proprietary algorithm employed by DrugLogic® was used for eliminating duplicates in the AERS database, which includes

1 The authors note that a number of terms exist in the literature and in the pharmacovigilance community for these events, and no clear consensus on nomenclature has been reached. For instance, the authors are aware that the CIOMS Working Group VIII is currently discussing this issue and has yet to publish an agreed-upon term. Many in the community, including the authors, feel that a higher reporting rate for a particular product-event combination does not raise it to the level of a true safety signal and therefore feel that they should not be described by any term that includes the word 'signal' itself (such as Signal of Disproportionate Reporting). We have chosen to use the term 'disproportionately reported product-event combination (DRPEC)' here as we believe it accurately describes the true nature of these entities without labelling them signals: disproportionately reported product-event combinations.

using only the latest report, eliminating duplicate reports that were sent to the FDA, both electronically and on paper, as well as eliminating cases with the same manufacturer number and name.

DRPEC Closure Methods

The above DRPEC-generation methods were applied to six of Wyeth's marketed products. As this was a pilot study, all products chosen were mature in their lifecycle, with well established safety profiles. The six products had a median of 12 years and a minimum of 7 years on the market.

All DRPECs were triaged and a full database review of the safety topic was *not* conducted if:

1. it was already noted in the reference safety information (RSI);
2. it was deemed to be indication related;
3. it was deemed non-interpretable from a medical perspective.

In addition, DRPECs reviewed by the pharmacovigilance team within the previous 12 months were not further pursued unless the new information warranted additional evaluation.[2] Following this triage process, any remaining DRPECs were placed in the category of requiring full case review. Table I provides further details regarding these categories.

Table I. Closure codes used in the triage of disproportionately reported product-event combinations

Closure code	Description of MedDRA® preferred term
Verbatim labelled	Corresponds word-for-word to an event listed in the RSI
Synonymously labelled	Has the same meaning as a currently listed term in the RSI (e.g. fever and pyrexia)
Conceptually labelled	Similar in notion or idea to an event listed in the RSI
Indication-related	Directly related to the indication or describing the purpose or reason for the use of the product
Non-interpretable	Not descriptive of an actual medical event
Prior case review	Was subject to a formal safety review that occurred within the previous year
Case review	Requires a postmarketing safety database case review

MedDRA = Medical Dictionary for Regulatory Activities; **RSI** = reference safety information.

The pharmacovigilance group systematically evaluated the DRPECs undergoing case review by querying the company's Safety Surveillance System (S^3) database. All reports with the Medical Dictionary for Regulatory Activities (MedDRA®) preferred term (PT) that had exceeded the threshold, as well as related PTs, were reviewed. If, in the judgement of the pharmacovigilance physician, the data was not suggestive of a possible association, the DRPEC was closed. By contrast, if the data suggested a possible association between the adverse event and the drug, or if such an association could not be clearly excluded, the adverse event was referred for review by the Safety Review Team (SRT). At the company, the SRT consists of members from a variety of disciplines – the core members are comprised of physicians from Pharmacovigilance, Clinical Research and Global Medical Affairs. Adverse events referred to the SRT undergo a comprehensive evaluation, with review of the pertinent literature and data from clinical trials, as well as the postmarketing information. If the totality of the data suggests a reasonable suspicion of association between the drug and the adverse event, the event is referred to the labelling team for addition to the RSI.

Results

A total of 861 DRPECs were identified – the average number of DRPECs was 144 per product (range 74–209). When adjusted for the number of unique PTs in the AERS database for the corresponding drug, the proportion of PTs with EB05 ≥2 was similar across the six products (5.1–8.5%) [see table II for details].

A total of 44.8% of all 861 DRPECs were deemed to already be listed in the RSI. Table III provides further details of the type of labelling correspondences across products (verbatim,

2 The pharmacovigilance groups reviewing the data may have chosen to further review any of the terms that were closed with initial triage based on this new information (signalling score), including those already stated on the label or previously reviewed; however, the details of this activity are not further discussed here.

Table II. Disproportionately reported product-event combinations (DRPECs) identified in terms of unique preferred terms (PTs) in the Adverse Event Reporting System (AERS) database

Drug reviewed	DRPECs identified	Unique PTs in the AERS database	Unique PTs with EB05 ≥2 (%)
A	209	3570	5.9
B	116	1853	6.3
C	106	1315	8.1
D	74	1064	7.0
E	200	2365	8.5
F	156	3041	5.1
Total	**861**	**13 208**	**6.5**

EB05 = the lower bound of the 90% confidence interval around the Empirical Bayes Geometric Mean[17]

synonymous or conceptual [see table I for definitions]). An additional 14.3% of all terms were determined to be indication related, while 4.9% were considered non-interpretable from a medical perspective.

An additional 9.9% (85/861) of the DRPECs had been reviewed within the prior year, and were not further reviewed unless warranted by the new information. Thus, only 26.1% (225) of the DRPECs required full case review (range 11.5–34.4 across products). Table IV summarizes the disposition of all DRPECs.

Although the DRPECs were identified in the AERS database, they were closed via review in our own proprietary database, with interesting results. Thirty-two of the 225 (14.2%) had no reports in the S^3 database that exactly matched the PT identified in AERS. A further 15 had only one report with a matching PT (6.6%).

As approximately 20% of events requiring case review had one or no reports in the S^3 database with an exact match, an exploration of the AERS database for a select number of these DRPECs was undertaken. Wyeth had submitted many of the reports corresponding to the PT that had exceeded the EB05 threshold; however, it appears that during the AERS data entry process, the original PTs submitted by the company had been recoded or new events added.

After review of all pertinent reports in the S^3 database, as well as additional data, the therapeutic area teams determined which DRPECs necessitated a formal review by the company's SRT. A total of 31 of the 861 (3.6%) DRPECs eventually required review at this level. Many of these terms described related medical concepts; consequently, a total of 16 concepts were reviewed at the SRT level. These SRT-level reviews resulted in seven changes to the RSI, involving four of the six compounds reviewed as part of this pilot.

The workload required for this exercise was extensive. The total time spent by pharmacovigilance professionals in reviewing all DRPECs was 1101 hours, ranging from 32 to 550 hours per drug. Notably, this workload was primarily driven by the 225 DRPECs requiring further review. Initial case review of the 225 events averaged 3.8 hours per event, and accounted for 77% of the total time. Further review accounted for an additional 17% of total workload. A complete review of all data collected concerning the workload impact is not within the scope of this article, and will be the subject of another manuscript.

Table III. Number (%) of disproportionately reported product-event combinations (DRPECs) closed during initial triage[a]

Drug	Total DRPECs	Verbatim labelled	Synonymously labelled	Conceptually labelled	Indication-related	Non-interpretable	Total closed without review
A	209	12 (5.7)	2 (1.0)	75 (35.9)	21 (10.0)	2 (1.0)	112 (53.6)
B	116	18 (15.5)	13 (11.2)	30 (25.9)	4 (3.4)	19 (16.4)	84 (72.4)
C	106	18 (17.0)	6 (5.7)	22 (20.8)	19 (17.9)	3 (2.8)	68 (64.2)
D	74	16 (21.6)	15 (20.3)	3 (4.1)	18 (24.3)	1 (1.4)	53 (71.6)
E	200	28 (14.0)	7 (3.5)	64 (32.0)	34 (17.0)	10 (5.0%)	143 (71.5)
F	156	36 (23.1)	14 (9.0%)	7 (4.5%)	27 (17.3%)	7 (4.5%)	91 (58.3)
Total	**861**	**128 (14.9)**	**57 (6.6)**	**201 (23.3)**	**123 (14.3)**	**42 (4.9)**	**551 (64.0)**

a Please see table I for definitions of the closure codes.

Table IV. Distribution of all disproportionately reported product-event combination (DRPEC) dispositions [n (%)]

Drug	DRPECs	Closed without review	Reviewed past year	Case review	Referred to Safety Review Team	Changes to reference safety information
A	209	112 (53.6)	25 (12.0)	72 (34.4)	10 (4.8)	0 (0)
B	116	84 (72.6)	0 (0)	32 (27.6)	6 (5.2)	3 (2.6)
C	106	68 (64.2)	2 (1.9)	36 (34.0)	9 (8.5)	1 (0.9)
D	74	53 (71.6)	0 (0)	21 (28.4)	0 (0)	0 (0)
E	200	143 (71.5)	11 (5.5)	46 (23.0)	4 (2.0)	2 (1.0)
F	156	91 (58.3)	47 (30.1)	18 (11.5)	2 (1.3)	1 (0.6)
Total	**861**	**551 (53.6)**	**85 (9.9)**	**225 (26.1)**	**31 (3.6)**	**7 (0.8)**

Discussion

Given that this was the first data mining exercise we conducted for these six products, it is not surprising that a large number of terms (861) were identified with an EB05 ≥2. There was a wide range in the number of DRPECs identified amongst the six products (74–209).

As noted in other studies,[28,29] a number of the terms identified were in the existing product label: approximately 45% of all terms were considered verbatim, synonymous or conceptually labelled. Furthermore, 9.9% of the terms had already been identified by Wyeth as potential safety signals as part of its routine pharmacovigilance, and had undergone recent review. Thus, 55% of the DRPECs had been previously identified by ongoing pharmacovigilance methods. Additionally, 14.3% of the terms were deemed to be clearly reflective of the co-morbidities commonly seen in the population to which the drug is administered. Taken together, the data indicate that the data mining techniques do identify safety concepts that overlap significantly with traditional pharmacovigilance methods; however, the additional events identified do suggest a potential as an adjunctive tool.

Most of the products studied were relatively mature in their lifecycle, which may partially explain the large number of labelled and otherwise previously reviewed terms identified. The number of terms identified for each drug did not appear to relate to time on market; instead, the number of DRPECs identified appears to be related to the total number of unique PTs for that drug in the underlying database. This finding could be valuable in predicting the workload impact of an initial data mining run prior to initiating a signal detection programme. Given the impact in terms of both resources and safety information, this relationship deserves greater exploration.

While the percentage of labelled DRPECs was relatively constant across all drugs, the distribution of events by relationship to the label was notable. Approximately half of the labelled terms were considered to be represented in the RSI as either labelled verbatim or synonymously, while the other half were considered to be labelled conceptually. This distribution varied greatly for different compounds studied. This difference has important implications in terms of whether DRPECs may be closed in an automated fashion (e.g. by using a list of labelled terms) and in how much medical judgement is required in the initial triage of DRPECs identified via data mining.

Approximately 1 in 20 DRPECs was non-interpretable from a medical review perspective (e.g. 'nonspecific reaction', 'blood creatinine', 'neutrophil count'). Although these terms were closed within this pilot without further review, exploration of such events showed few matching reports within the Wyeth S^3 database. While we could not determine the reason these non-interpretable events were identified in our data mining runs, the number of such terms will have workload and other consequences as data mining in the AERS and other large databases increases. Consequently, further exploration of the underlying source of these non-specific events is warranted.

Using a conservative approach that encouraged review of DRPECs in the S^3 safety database, 225 DRPECs (26.1%) required review. Despite

the variety of drugs studied, when recent reviews were included the DRPECs requiring case review generally fell into a narrow range between one-quarter and one-third of those terms identified across compounds. It was this group that proved the most valuable adjunct to traditional pharmacovigilance activities, as these product-event combinations led to full case review, had the greatest impact in terms of resources required and were most informative in terms of gaining insight into the safety profile of the medication under review. As such, they represent the most critical subset of DRPECs identified.

Of the aforementioned 225 DRPECS, 31 (3.6% of all DRPECs [14% of the subset]), representing 16 broad medical concepts, were identified as representing potentially new safety signals. These topics were fully reviewed by the SRT. Seven of these topics were subsequently added to the RSI – approximately one labelling change for each 125 DRPECs generated. This suggests that these data mining techniques do have value as an adjunct to traditional pharmacovigilance activities.

Limitations

While this data mining exercise was extensive, including six active pharmaceutical products across a broad spectrum of indications and patient populations, there were several limitations to our pilot study. First, all of the drugs selected were marketed for a minimum of seven years, and the data cannot be extrapolated to drugs currently in development or in the first few years of marketing. It is possible that data mining the AERS database for a new pharmaceutical product could identify important product-event combinations earlier than traditional pharmacovigilance methods.

Second, interpretation of the results is limited by the company's choice of methodology for triage and review of product-event combinations identified. Of particular interest is that the DRPECs identified in the AERS database were evaluated using cases in our own proprietary safety database, as the limited individual case information available in the AERS database precluded a thorough evaluation using AERS. Furthermore, the pharmacovigilance group was given a fair amount of latitude with regard to utilizing medical judgement to initially close DRPECs.

Conclusions

Several salient observations were noted in our review of the data from this pilot. We noted that the number of DRPECs identified for each drug approximately correlated with the number of unique PTs in the database. Analysis demonstrated that the majority of DRPECs had been identified by ongoing pharmacovigilance methods, having already been included in the RSI for the product, or having undergone recent or concurrent review. This suggests that the DRPECs identified by these methods do correlate with those noted by ongoing pharmacovigilance methods, but are sufficiently different to serve as an important adjunct to these traditional methods.

Approximately three-quarters of the 861 DRPECs identified were closed without case review. Of the one-quarter of DRPECs that required formal case review, seven resulted in an addition to the RSI for the relevant products. The vast majority of the workload involved was related to the formal case analysis and subsequent SRT review.

This pilot study does not allow us to comment on the utility of routine data mining for all products, especially newer products. Such decisions will have to be individualized, taking into account the signal detection methods employed and workload required. However, it is significant that several new safety concepts were identified through this prospective exercise, which we believe demonstrates that these methods do have value as an adjunct to ongoing pharmacovigilance methods.

Acknowledgements

All authors were employees of Wyeth Pharmaceuticals at the time of this pilot, which funded the activities performed in completing and publishing this research. None of the authors has a potential financial gain from the results described herein.

References

1. Hauben M, Madigan D, Gerrits CM, et al. The role of data-tamining in pharmacovigilance. Expert Opin Drug Saf 2005 Sep; 4 (5): 929-48
2. Almenoff JS, Pattishall EN, Gibbs TG, et al. Novel statistical tools for monitoring the safety of marketed drugs. Clin Pharmacol Ther 2007 Aug; 82 (2): 157-66
3. Bate A, Edwards IR. Data mining in spontaneous reports. Basic Clin Pharmacol Toxicol 2006 Mar; 98 (3): 324-30
4. Stephenson WP, Hauben M. Datamining for signals in spontaneous reporting databases: proceed with caution. Pharmacoepidemiol Drug Saf 2007; 16: 359-65
5. Hauben M. Early postmarketing drug safety surveillance: data mining points to consider. Ann Pharmacother 2004 Oct; 38 (10): 1625-30
6. Hauben M, Horn S, Reich L. Potential use of data-mining algorithms for the detection of 'surprise' adverse drug reactions. Drug Saf 2007; 30 (2): 143-55
7. Roux E, Thiessard F, Fourrier A, et al. Evaluation of statistical association measures for the automatic signal generation in pharmacovigilance. IEEE Trans Inf Technol Biomed 2005 Dec; 9 (4): 518-27
8. Matsushita Y, Kuroda Y, Niwa S, et al. Criteria revision and performance comparison of three methods of signal detection applied to the spontaneous reporting database of a pharmaceutical manufacturer. Drug Saf 2007; 30 (8): 715-26
9. Almenoff JS, LaCroix KK, Yuen NA, et al. Comparative performance of two quantitative safety signalling methods: implications for use in a pharmacovigilance department. Drug Saf 2006; 29 (10): 875-87
10. Kubota K, Koide D, Hirai T. Comparison of data mining methodologies using Japanese spontaneous reports. Pharmacoepidemiol Drug Saf 2004 Jun; 13 (6): 387-94
11. Gould AL. Practical pharmacovigilance analysis strategies. Pharmacoepidemiol Drug Saf 2003 Oct-Nov; 12 (7): 559-74
12. van Puijenbroek EP, Bate A, Leufkens HG, et al. A comparison of measures of disproportionality for signal detection in spontaneous reporting systems for adverse drug reactions. Pharmacoepidemiol Drug Saf 2002 Jan-Feb; 11 (1): 3-10
13. Lindquist M, Stahl M, Bate A, et al. A retrospective evaluation of a data mining approach to aid finding new adverse drug reaction signals in the WHO international database. Drug Saf 2000 Dec; 23 (6): 533-42
14. Lehman HP, Chen J, Gould AL, et al. An evaluation of computer-aided disproportionality analysis for post-marketing signal detection. Clin Pharmacol Ther 2007 Aug; 82 (2): 173-80
15. van Puijenbroek EP, van Grootheest K, Diemont WL, et al. Determinants of signal selection in a spontaneous reporting system for adverse drug reactions. Br J Clin Pharmacol 2001 Nov; 52 (5): 579-86
16. Hauben M, Reich L. Potential utility of data-mining algorithms for early detection of potentially fatal/disabling adverse drug reactions: a retrospective evaluation. J Clin Pharmacol 2005 Apr; 45: 378-84
17. DuMouchel W, Smith E, Beasley R, et al. Association of asthma therapy and churg-strauss syndrome: an analysis of postmarketing surveillance data. Clin Ther 2004; 26 (7): 1092-104
18. Almenoff J, Tonning JM, Gould AL, et al. Perspective on the use of datamining in pharmacovigilance. Drug Saf 2005; 28 (11): 981-1007
19. Evans SJ, Waller PC, Davis S. Use of proportional reporting ratios (PRRs) for signal generation from spontaneous adverse drug reaction reports. Pharmacoepidemiol Drug Saf 2001 Oct-Nov; 10 (6): 483-6
20. Bate A, Lindquist M, Edwards IR, et al. A Bayesian neural network method for adverse drug reaction signal generation. Eur J Clin Pharmacol 1998 Jun; 54 (4): 315-21
21. Noren GN, Bate A, Orre R, et al. Extending the methods used to screen the WHO drug safety database towards analysis of complex associations and improved accuracy for rare events. Stat Med 2006 Nov 15; 25 (21): 3740-57
22. DuMouchel W. Bayesian data mining in large frequency tables, with an application to the FDA spontaneous reporting system. Am Stat 1999; 53 (3): 177-90
23. Hauben M. A brief primer on automated signal detection. Ann Pharmacother 2003 Jul-Aug; 37 (7-8): 1117-23
24. Hauben M, Zhou X. Quantitative methods in pharmacovigilance: focus on signal detection. Drug Saf 2003; 26 (3): 159-86
25. Szarfman A, Machado S, O'Neill R. Use of screening algorithms and computer systems to efficiently signal higher-than-expected combinations of drugs and events in the us FDA's spontaneous reports database. Drug Saf 2002; 25 (6): 381-92
26. Hauben M, Reich L, Chung S. Postmarketing surveillance of potentially fatal reactions to oncology drugs: potential utility of two signal-detection algorithms. Eur J Clin Pharmacol 2004; 60: 747-50
27. Hauben M, Reich L, Gerrits CM, et al. Illusions of objectivity and a recommendation for reporting data mining results. Eur J Clin Pharmacol 2007 May; 63 (5): 517-21
28. Hauben M, Reich L. Safety related drug-labelling changes: findings from two data mining algorithms. Drug Saf 2004; 27 (10): 735-44
29. Hochberg AM, Reisinger SJ, Pearson RK, et al. Using data mining to predict safety actions from FDA adverse event reporting system data. Drug Info J 2007; 41 (5): 633-43

Correspondence: Dr *Steven R. Bailey*, Senior Director – Vaccines and Technology, Medical Pharmacovigilance, Wyeth Research, 500 Arcola Road, Collegeville, PA 19426, USA.
E-mail: baileysr@wyeth.com