Drug Saf (2013) 36:371–388
DOI 10.1007/s40264-013-0053-7

ORIGINAL RESEARCH ARTICLE

The Development and Evaluation of Triage Algorithms for Early Discovery of Adverse Drug Interactions

Johanna Strandell · Ola Caster · Johan Hopstadius · I. Ralph Edwards · G. Niklas Norén

Published online: 3 May 2013

Abstract
Background Around 20 % of all adverse drug reactions (ADRs) are due to drug interactions. Some of these will only be detected in the postmarketing setting. Effective screening in large collections of individual case safety reports (ICSRs) requires automated triages to identify signals of adverse drug interactions. Research so far has focused on statistical measures, but clinical information and pharmacological characteristics are essential in the clinical assessment and may be of great value in first-pass filtering of potential adverse drug interaction signals.
Objective The aim of this study was to develop triages for adverse drug interaction surveillance, and to evaluate these prospectively relative to clinical assessment.
Methods A broad set of variables were considered for inclusion in the triages, including cytochrome P450 (CYP) activity, explicit suspicions of drug interactions as noted by the reporter, dose and treatment overlap, and a measure of interaction disproportionality. Their unique contributions in predicting signals of adverse drug interactions were determined through logistic regression. This was based on the reporting in the WHO global ICSR database, VigiBase™, for a set of known adverse drug interactions and corresponding negative controls. Three triages were developed, each producing an estimated probability that a given drug–drug–ADR triplet constitutes an adverse drug interaction signal. The triages were evaluated against two separate benchmarks derived from expert clinical assessment: adverse drug interactions known in the literature and prospective adverse drug interaction signals. For reference, the triages were compared with disproportionality analysis alone using the same benchmarks.
Results The following were identified as valuable predictors of adverse drug interaction signals: plausible CYP metabolism; notes of suspected interaction by the reporter; and reports of unexpected therapeutic response, altered therapeutic effect with dose information and altered therapeutic effect when only two drugs had been used. The new triages identified reporting patterns corresponding to both prospective signals of adverse drug interactions and already established ones. They perform better than disproportionality analysis alone relative to both benchmarks.
Conclusions A range of predictors for adverse drug interaction signals have been identified. They substantially improve signal detection capacity compared with disproportionality analysis alone. The value of incorporating clinical and pharmacological information in first-pass screening is clear.

J. Strandell · O. Caster (✉) · J. Hopstadius · I. R. Edwards · G. N. Norén
Uppsala Monitoring Centre, WHO Collaborating Centre for International Drug Monitoring, Box 1051, 751 40 Uppsala, Sweden
e-mail: ola.caster@who-umc.org

J. Strandell
Department of Drug Research/Clinical Pharmacology, Linköping University, Linköping, Sweden

O. Caster
Department of Computer and Systems Sciences, Stockholm University, Stockholm, Sweden

G. N. Norén
Department of Mathematics, Stockholm University, Stockholm, Sweden

1 Background

Adverse drug reactions (ADRs) are a major health problem with dire consequences for patients as well as for society

[1–3]. One of several factors that increase the risk of ADRs are drug–drug interactions [4], which are reported to be responsible for approximately 20 % of all ADRs [3, 5]. Whereas large collections of individual case safety reports (ICSRs) are one of the key sources in detection of novel ADRs related to single drugs, these reports are still an underutilized source to detect adverse drug interactions systematically, i.e. drug interactions resulting in ADRs. So far most of the methodological research on automatic screening for drug interactions in large report repositories has focused on purely quantitative measures of disproportionality [6–8], although the individual usefulness of these measures in detection of novel adverse drug interactions has not been fully demonstrated. Our previous results indicate that reported clinical and pharmacological information are other potentially useful components when screening for adverse drug interactions [9].

Triage algorithms are selection strategies with predefined criteria which aid to focus the analysis of ICSRs on important associations, i.e. associations that are most likely to lead to signals [10]. The development of interaction triages is possibly an important step to facilitate the systematic detection of adverse drug interaction signals in large collections of reports. Our previous study which identified variables with potential to highlight adverse drug interactions early was an important step in this direction [9]. The aims of this study are (i) to develop triages for adverse drug interaction surveillance; and (ii) to evaluate those triages prospectively relative to expert clinical assessment, including a performance comparison to disproportionality screening.

2 Development of Triage Algorithms

2.1 Methods

Three candidate triage algorithms were developed in a process combining regression modelling with expert clinical judgement. An overview is provided in the upper panel of Fig. 1.

2.1.1 *Known Adverse Drug Interactions and Drug–Drug–Adverse Drug Reactions Not Known to Interact*

A set of drug–drug–ADR (DDA) triplets was constructed based on Stockley's Interaction Alerts [11], a comprehensive international source of drug interaction information. The set contains 324 DDAs representing *known adverse drug interactions* and 6480 *DDAs not known to interact*, where the categorization refers to the DDAs' occurrence in Stockley's Interaction Alerts as of the third quarter of 2009 [9]. This set of DDAs provides a basis to identify what variables are predictive of subsequent inclusion into Stockley's Interaction Alerts, and can therefore be used to develop triages for prospective identification of adverse drug interaction signals. For more details on the extraction of the known adverse drug interactions and the DDAs not known to interact, see Appendix 1.

2.1.2 *Potential Triage Variables*

Twenty variables providing clinical, pharmacological or statistical support of adverse drug interactions were considered as potential components of the triages [7, 9, 12]. The majority of variables correspond to qualitative clinical information provided on ICSRs, such as a reported explicit suspicion of a drug–drug interaction; two variables are pharmacological: both drugs belonging to the same Anatomical Therapeutic Chemical (ATC) group, and both drugs displaying activity on a common cytochrome P450 (CYP) enzyme in a way that can lead to an interaction; and one variable (Omega_{025} (Ω_{025}) > 0) is purely quantitative, highlighting higher than expected reporting of the DDA within the ICSR database [7].

Table 1 describes the 20 potential triage variables, which are all defined on the level of the DDA. While the quantitative and pharmacological variables described above (18–20 in Table 1) are binary by construction, the clinical variables (1–17 in Table 1) are defined in terms of the number of reports on the DDA fulfilling certain criteria. Of the latter, variables 1–11 require fulfilment of a single criterion, whereas variables 12–17 require multiple criteria. For example, variable 17 (see Table 1) is defined as the number of reports fulfilling each of the following three criteria: solely two drugs listed, overlapping treatment duration for these two drugs, and a positive dechallenge for either of the two drugs with the ADR. The multiple-criteria variables were constructed to identify simultaneous reporting of several clinical aspects suggestive of an adverse drug interaction, corresponding to unusually strong reports.

2.1.3 *Dataset for Derivation of Triages*

To construct a dataset suitable for derivation of adverse drug interaction triages, data on the 20 potential triage variables was extracted for all included known adverse drug interactions and DDAs not known to interact. The WHO global ICSR database, VigiBase™, was used as the data source for the clinical and quantitative variables; ATC information was gathered from the WHO Collaborating Centre for Drug Statistics Methodology, and the CYP Standardized Drug Grouping available in WHO Drug Dictionary Enhanced was used for the CYP variable (for more information, see Table 1). The reports pooled in VigiBase™ are collected within the WHO Programme for International Drug Monitoring [13] and

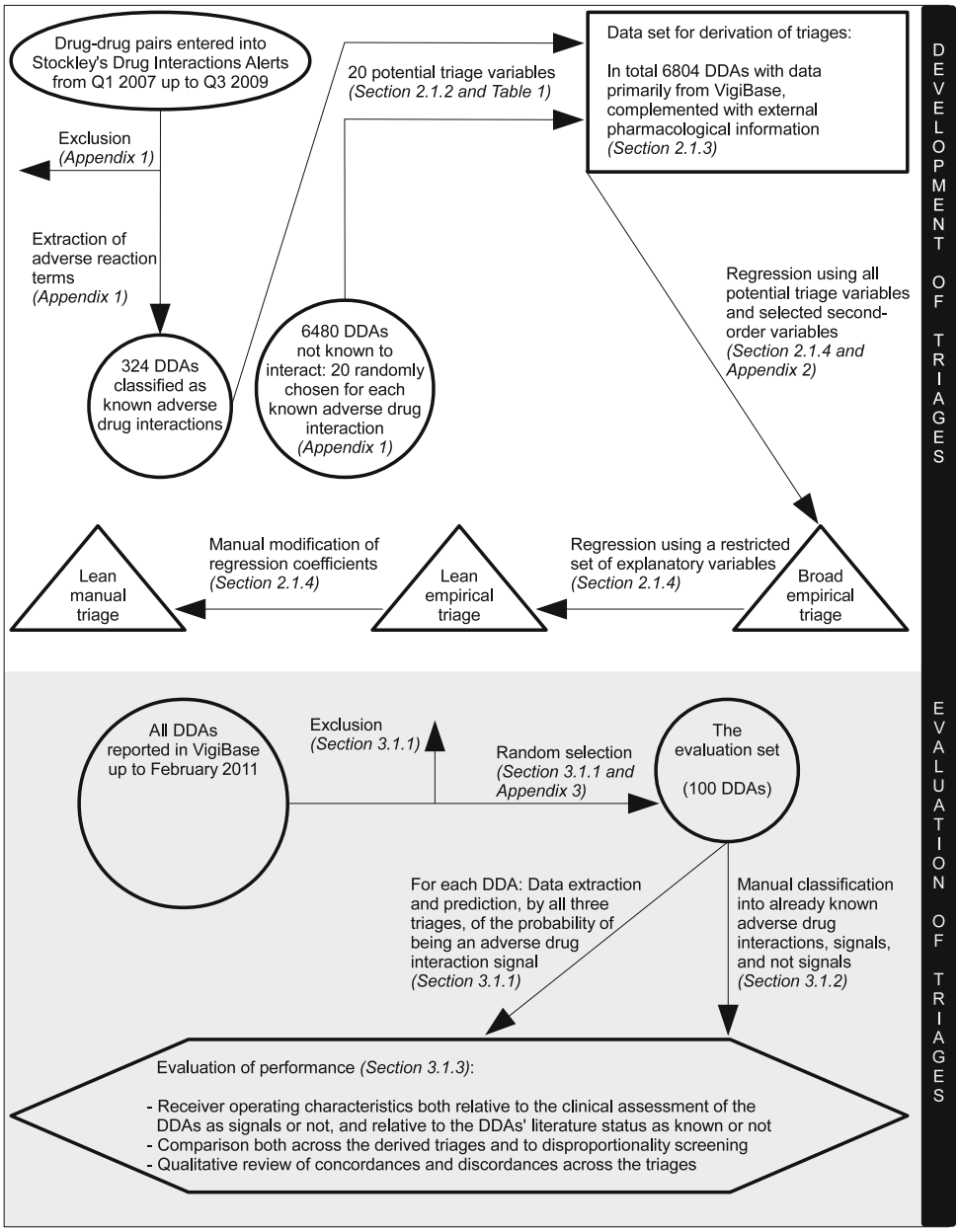


Fig. 1 Overview of the processes for development and evaluation of triages for adverse drug interaction surveillance presented in this study. Note that while we have previously described how the known adverse drug interactions and the DDAs not known to interact were extracted [9], a summary of this process is included in Appendix 1 for completeness. *DDAs* drug–drug–adverse drug reactions

amount to a vast resource of safety information: as of February 2013, VigiBase™ contained almost 8 million reports forwarded from 111 countries worldwide. In this study, suspected duplicates were removed [14].

For each known adverse drug interaction, the only VigiBase™ reports considered were those entered between 1 January 1990 and the quarter prior to inclusion of that DDA into Stockley's Interaction Alerts. The purpose was to mimic the setting for early prospective detection of adverse drug interactions, by approximating the situation when currently known adverse drug interactions were still unknown. Each DDA not known to interact had a corresponding data extraction endpoint, inherited from the known adverse drug interaction to which it had been matched (see Appendix 1). The second quarter of 2009 was the last quarter from which data was extracted from VigiBase™ for any DDA.

Drugs can be listed as either suspected (S), interacting (I) or concomitant (C) on reports. When studying adverse drug interactions, it is not self-evident whether or not the evaluation of potential triage variables should include concomitant drugs [12]. Therefore, for variables 1–2 and 4–18 in Table 1, data was extracted from VigiBase™ based both on reports where the drug pair was reported as suspected or interacting (SI) and as suspected, interacting or concomitant (SIC). Consequently, the dataset contained 37 variables for each included known adverse drug interaction and DDA not known to interact.

Table 1 The 20 variables considered for inclusion into triage algorithms for adverse drug interaction signal detection

	Short description	Definition	Motivation	Type	SI and SIC
Clinical variables defined by the number of ICSRs fulfilling a single criterion					
1	Effect decreased	Number of reports where the DDA was co-reported with any of the following WHO-ART terms: 'Therapeutic response decreased', 'Drug level below therapeutic', 'Drug level decreased', 'Medicine ineffective', 'Medicine ineffective unexpected'	The co-reported term indicates that the therapeutic effect was lower than expected. This can be suggestive of a drug interaction [33]	Numerical	Yes
2	Effect increased	Number of reports where the DDA was co-reported with any of the following WHO-ART terms: 'Therapeutic response increased', 'Drug level increased'	The co-reported term indicates that the therapeutic effect was greater than expected. This can be suggestive of a drug interaction [33]	Numerical	Yes
3	Interacting	Number of reports where both drugs within the DDA were recorded as interacting	The reporter(s) suspect(s) an interaction between the drugs as involved in causing the ADR	Numerical	No
4	MedDRA® [a] interaction	Number of reports where the DDA was co-reported with any of the following MedDRA® terms: 'Drug interaction', 'Labelled drug–drug interaction medication error', 'Inhibitory drug interaction', 'Potentiating drug interaction'	The reporter(s) suspect(s) the ADR to be a result of a possible interaction. At the time of study WHO-ART did not include interaction-related terms and therefore MedDRA® terms were used	Numerical	Yes
5	Narrative information	Number of reports on the DDA with a case narrative including the word fragments 'interact' or 'interakt'	The reporter(s) describe(s) a possible interaction in the case narrative	Numerical	Yes
6	Unexpected therapeutic response	Number of reports where the DDA was co-reported with the WHO-ART term 'Unexpected therapeutic effect'	The co-reported term indicates that the expected therapeutic effect has been altered. This can be suggestive of a drug interaction [33]	Numerical	Yes
7	Dose information	Number of reports where dose information was provided for both drugs within the DDA	The reporter(s) might suspect and report doses to greater extent if the ADR appears when normal doses are given	Numerical	Yes
8	Solely two drugs	Number of reports where the drug pair within the DDA were the only reported drugs	To exclude influence from other possible interacting agents	Numerical	Yes
9	Overlapping treatment	Number of reports where the treatment of both drugs within the DDA was definitely overlapping	The drugs have definitely been concomitantly used	Numerical	Yes
10	Positive dechallenge	Number of reports where a positive dechallenge outcome is recorded for at least one of the drugs with the ADR	A positive dechallenge can strengthen a potential relationship between one of the drugs and the ADR	Numerical	Yes
11	Positive rechallenge	Number of reports where a positive rechallenge outcome is recorded for at least one of the drugs with the ADR	The re-occurrence of an ADR on re-introduction of one of the drugs can strengthen a potential relationship between the two	Numerical	Yes

Table 1 continued

	Short description	Definition	Motivation	Type	SI and SIC
Clinical variables defined by the number of ICSRs fulfilling multiple criteria					
12	i) Effect decreased or Effect increased ii) Dose information	Number of reports which include dose information for both drugs within the DDA, and the DDA was co-reported with any of the WHO-ART terms listed above indicative of effect increased or effect decreased	The reporter(s) might suspect and report doses to greater extent if the ADR appears when normal doses are given. In addition, the expected therapeutic effect was reported to be altered	Numerical	Yes
13	i) Effect decreased or Effect increased ii) Solely two drugs	Number of reports on which the drug pair within the DDA were the only reported drugs, and the DDA was co-reported with any of the WHO-ART terms listed above indicative of effect increased or effect decreased	The influence from other possible interacting agents is limited, and the expected therapeutic effect was reported to be altered	Numerical	Yes
14	i) Effect decreased or Effect increased ii) Overlapping treatment	Number of reports on which the treatment of both drugs within the DDA was definitely overlapping, and the DDA was co-reported with any of the WHO-ART terms listed above indicative of effect increased or effect decreased	The concurrent use of the two drugs within the DDA may have led to changes of the expected therapeutic effect	Numerical	Yes
15	i) Effect decreased or Effect increased ii) Positive dechallenge	Number of reports on which the DDA was co-reported with any of the WHO-ART terms listed above indicative of effect increased or effect decreased, and the ADR was reported to abate upon withdrawal of one of the drugs within the DDA	The expected therapeutic effect was reported to be altered. Furthermore, the ADR was reported to abate upon drug withdrawal, which strengthens a potential relationship between one of the drugs and the ADR	Numerical	Yes
16	i) Effect decreased or Effect increased ii) Positive rechallenge	Number of reports on which the DDA was co-reported with any of the WHO-ART terms listed above indicative of effect increased or effect decreased, and the ADR was reported to re-occur upon re-introduction of one of the drugs within the DDA	The expected therapeutic effect has been changed. Re-occurrence of an ADR on re-introduction of one of the drugs further strengthens a potential relationship between the two	Numerical	Yes
17	i) Solely two drugs ii) Positive dechallenge iii) Overlapping treatment	Number of reports on which the treatment of both drugs within the DDA was definitely overlapping; the drugs were the only drugs reported; and the ADR was reported to abate upon withdrawal of one of the drugs	The drug pair has definitely been concomitantly used and the influence from other possible interacting agents is limited. In addition, the ADR was reported to abate upon drug withdrawal, which strengthens a potential relationship between one of the drugs and the ADR	Numerical	Yes
Quantitative variable computed within the ICSR database					
18	$\text{Omega}_{025} > 0$	Ω is a shrinkage observed-to-expected ratio for the number of reports of the ADR with the two drugs together. Ω_{025} is the lower limit of a 95% credibility interval for Ω	When Ω_{025} exceeds zero the DDA is reported reliably more often than expected if the attributable risks of the ADR from each drug would add together [7]	Binary	Yes

Table 1 continued

	Short description	Definition	Motivation	Type	SI and SIC
Pharmacological variables not dependent on an ICSR database					
19	ATC	The drug pair within the DDA is classified with the same chemical group (third-level ATC code)[b]	Could indicate an additive effect of the two drugs	Binary	NA
20	CYP	The drug pair within the DDA are drugs that may induce, inhibit or be substrates in the phase I metabolism via the same CYP enzyme(s). CYP information is gathered from countries' drug reference sources [34, 35], original articles referred to on Indiana University's homepage (the so-called Flockhart table) [36], and Stockley's Drug Interactions [16]. The substances were required to have activity that either competed for the same enzyme, or inhibited or induced each other's metabolism[c]	Indicates a potential pharmacokinetic mechanism Furthermore, drugs that undergo hepatic metabolism through the CYP enzymes are more likely to cause changes in other drugs' concentrations and therefore result in dose–related ADRs	Binary	NA

Variables 1–18 are defined in relation to an ICSR database. Variables 12–17 count the number of reports fulfilling all criteria listed with small Roman numerals in the short description. The rightmost column indicates whether the variable is dependent on the choice of only including drugs listed as suspected (S) and interacting (I) on the reports, or whether also concomitant (C) drugs are included. All variables are defined for drug–drug–ADR (DDA) triplets

[a] MedDRA® terminology is the medical terminology developed under the auspices of the International Conference on Harmonization of technical requirements for Registration of Pharmaceuticals for Human Use (ICH)

[b] This definition deviates from the one used in a previously reported preliminary analysis, where the fourth level ATC code was used [9]

[c] This definition is stricter than the one used previously, where any effect on the same CYP enzyme was sufficient [9]

ADR adverse drug reaction, *ATC* Anatomical Therapeutic Chemical, *CYP* cytochrome P450, *ICSR* individual case safety report, *MedDRA®* Medical Dictionary for Regulatory Activities, *NA* not applicable, *WHO-ART* WHO Adverse Reactions Terminology

2.1.4 *Construction of Triages by Regression Modelling and Manual Modification*

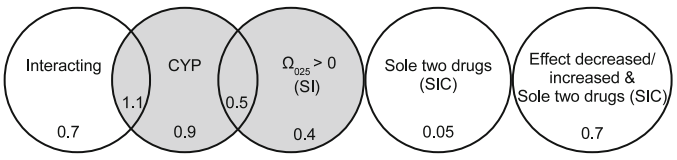
Whereas we have previously assessed the potential value of the variables by considering them one at a time [9], this study used shrinkage logistic regression to evaluate all variables simultaneously. In this context, the estimated regression coefficients determine the variables' unique contributions in predicting signals of adverse drug interactions, with positive values strengthening the likelihood of a signal. The shrinkage employed results in variable selection, as some coefficients are set to zero. For each DDA, the model yields a predicted probability that this DDA represents an adverse drug interaction. The predicted probabilities can be used to triage a given list of DDAs (see Fig. 2).

Our triage algorithms could include so-called second-order variables, with the purpose of capturing synergistic effects. For example, even if each report where the drug pair is listed as interacting (variable 3) is attributed a certain reward, and if a common CYP pathway for the two drugs (variable 20) is attributed another reward, a second-order variable would allow an additional reward to be given if the two variables were fulfilled for the same DDA. Based on clinical considerations on possible synergies in practice, each of $\Omega_{025} > 0$, ATC and CYP were combined with variables 1–6 and 12–17, respectively, to form second-order variables. Furthermore, $\Omega_{025} > 0$ was combined with ATC and CYP, to yield 96 second-order variables in total.

Three different triages were constructed. The first was derived by fitting a shrinkage logistic regression model based on the dataset described in Sect. 2.1.3, using all 37 first-order variables and 96 second-order variables as potential predictors of adverse drug interactions. The details of the regression modelling are given in Appendix 2. This *broad empirical* triage was employed primarily as a basis for identifying a smaller and more coherent set of promising variables to be considered for the other two triages, the *lean empirical* and the *lean manual* triages. The variable selection process for the latter two took into account the estimated coefficients of the broad empirical triage, as well as subjective clinical considerations. For example, coherence was sought with respect to including variables either on SI or SIC level. The lean empirical triage was then derived through regression modelling in the exact same way as for the broad empirical triage (see Appendix 2), but based on this restricted set of potential predictors. In contrast, the coefficients of the lean manual triage were manually set, under the restriction that their absolute values have the same sum as the coefficients of the lean empirical triage. The elicitation process took into account the estimated coefficients of the two empirical triages, but aimed to design a triage as general as possible, to avoid undue customization to the adverse drug interactions included into Stockley's Interaction Alerts. While this source is largely dominated by pharmacokinetic interactions, it has been empirically demonstrated that those interactions amount to fewer than half of all interactions reported to VigiBase™ [15].

2.2 Results

The broad empirical triage is presented in Table 2. In the regression analysis, 67 variables in total were retained with non-zero coefficients. The first-order variables that obtained



DDA 1	Intercept	0 reports	Yes		Yes	5 reports	2 reports	= 0.45 → Pr = 0.61
	-3.0	+0	+0.9	+0.5	+0.4	+0.25	+1.4	
DDA 2	Intercept	1 report	No		Yes	1 report	0 reports	= -1.85 → Pr = 0.14
	-3.0	+0.7	+0		+0.4	+0.05	+0	
DDA 3	Intercept	2 reports		Yes	No	0 reports	0 reports	= 1.50 → Pr = 0.82
	-3.0	+1.4	+2.2	+0.9	+0	+0	+0	

Fig. 2 A schematic description of how an already-fitted logistic regression model can be used to triage DDAs. This imaginary model contains three integer-valued and two binary predictor variables depicted with white and grey circles, respectively. The numbers at the bottom of the circles indicate the variables' coefficients, and the numbers in the intersecting areas indicate the coefficients of the corresponding second-order variables. For each DDA, the coefficients are multiplied by the values specific to the DDA, and the results are added to the intercept to yield a sum total, e.g. 0.45 for DDA 1 here. Each sum total is then transformed to a predicted probability that the DDA corresponds to an adverse drug interaction signal, using the formula $Pr = e^x/(1 + e^x)$, where x is the sum. The probabilities can be used to order the DDAs for prioritization, so that in this example DDA 3 would be assessed first, followed by DDA 1 and DDA 2. Alternatively, a threshold could be used, and only DDAs with predicted probabilities above that threshold would be assessed. *CYP* cytochrome P450, *DDAs* drug–drug–adverse drug reactions, *Pr* probability, *S* suspected, *I* interacting, *C* concomitant

Table 2 Estimated coefficients in the lasso logistic regression model representing the broad empirical triage

Variable	Own coefficient	Coefficient with $\Omega_{025} > 0$ (SI)	Coefficient with $\Omega_{025} > 0$ (SIC)	Coefficient with ATC	Coefficient with CYP
Intercept	−3.372	-	-	-	-
Effect decreased SI			−0.391		−0.634
Effect decreased SIC			−0.087	−0.046	
Effect increased SI				0.016	
Effect increased SIC	0.101	−0.164		0.033	
Interacting	0.928		−0.073	−0.347	0.793
MedDRA® interaction SI	0.114		−0.024	0.169	0.099
MedDRA® interaction SIC	0.004			0.258	
Narrative information SI		0.384			0.286
Narrative information SIC	0.454	−0.481	−0.421		0.464
Unexpected therapeutic response SI					
Unexpected therapeutic response SIC	0.636			−1.313	1.060
Dose information SI		-	-	-	-
Dose information SIC		-	-	-	-
Solely two drugs SI		-	-	-	-
Solely two drugs SIC	0.030	-	-	-	-
Overlapping treatment SI	−0.004	-	-	-	-
Overlapping treatment SIC	−0.071	-	-	-	-
Positive dechallenge SI	0.272	-	-	-	-
Positive dechallenge SIC	0.031	-	-	-	-
Positive rechallenge SI	−0.545	-	-	-	-
Positive rechallenge SIC		-	-	-	-
Effect decreased or Effect increased + Dose information SI	0.490		1.167	0.407	
Effect decreased or Effect increased + Dose information SIC	0.257	−0.976	−0.084	−0.073	−0.249
Effect decreased or Effect increased + Solely two drugs SI	−0.747			−0.306	
Effect decreased or Effect increased + Solely two drugs SIC	1.052				
Effect decreased or Effect increased + Overlapping treatment SI	−0.554	1.859			
Effect decreased or Effect increased + Overlapping treatment SIC	−0.358				−1.079
Effect decreased or Effect increased + Positive dechallenge SI		1.573		0.116	−0.841
Effect decreased or Effect increased + Positive dechallenge SIC			−0.171		
Effect decreased or Effect increased + Positive rechallenge SI	3.201				
Effect decreased or Effect increased + Positive rechallenge SIC				−1.580	−1.370
Solely two drugs + Positive dechallenge + Overlapping treatment SI		−0.132		−1.157	1.291
Solely two drugs + Positive dechallenge + Overlapping treatment SIC		−0.523		−0.847	
$\Omega_{025} > 0$ SI	0.433	-	-	−0.043	0.081
$\Omega_{025} > 0$ SIC	−0.071	-	-		0.342
ATC	−0.283	-	-	-	-
CYP	0.841	-	-	-	-

Blank cells indicate zero estimates, and dashes correspond to variables not considered in the analysis

ATC Anatomical Therapeutic Chemical, *CYP* cytochrome P450, *MedDRA®* Medical Dictionary for Regulatory Activities, *S* suspected, *I* interacting, *C* concomitant

the highest coefficients were Effect decreased or Effect increased + Positive rechallenge (SI), Effect decreased or Effect increased + Solely two drugs (SIC), Interacting, and CYP.

24 variables were selected for further consideration. First-order variables were essentially retained unless they were negative in the full model. SI was considered the generally preferred level, with two exceptions: Unexpected therapeutic response and Solely two drugs. The former is very rarely reported in VigiBase™ and did not occur on SI level for any of the included DDAs; the latter is conceptually stronger on SIC level, as that definition implies that no additional drugs at all are co-reported, not even drugs listed as concomitant. Consequently, the multiple-criteria variable Effect decreased or Effect increased + Solely two drugs was also kept on SIC level.

Table 3 Estimated and elicited coefficients for the variables considered for inclusion in the lean empirical and lean manual triages

Variable	Coefficient lean empirical	Coefficient lean manual	Coefficient broad empirical
Intercept	−3.219	−3.2	−3.372
Effect increased SI			
Interacting	0.312	0.6	0.928
MedDRA® interaction SI	0.187	0.3	0.114
Narrative information SI	0.028	0.4	
Unexpected therapeutic response SIC	0.144	0.2	0.636
Dose information SI			
Solely two drugs SIC			0.030
Positive dechallenge SI			0.272
Effect decreased or Effect increased + Dose information SI	0.492	0.2	0.490
Effect decreased or Effect increased + Solely two drugs SIC	0.110	0.3	1.052
Effect decreased or Effect increased + Positive dechallenge SI		0.2	
Effect decreased or Effect increased + Positive rechallenge SI		0.2	3.201
Solely two drugs + Positive dechallenge + Overlapping treatment SI		0.4	
$\Omega_{025} > 0$ (SI)		0.6	0.433
CYP	0.607	0.4	0.841
CYP with Effect increased SI			
CYP with Interacting	1.187		0.793
CYP with MedDRA® interaction SI			0.099
CYP with Narrative information SI	0.736		0.286
CYP with Unexpected therapeutic response SIC			1.060
CYP with Effect decreased or effect increased + Dose information SI			
CYP with Effect decreased or effect increased + Solely two drugs SIC			
CYP with Solely two drugs + Positive dechallenge + Overlapping treatment SI			1.291
CYP with $\Omega_{025} > 0$ SI			0.081

The coefficients of the broad empirical triage are given as reference (see also Table 2). Blank cells indicate zero estimates

CYP cytochrome P450, *MedDRA®* Medical Dictionary for Regulatory Activities, *S* suspected, *I* interacting, *C* concomitant

The estimated coefficients of the lean empirical triage and the elicited coefficients of the lean manual triage for these 24 variables are displayed in Table 3. The lean empirical triage places much emphasis on a plausible pharmacokinetic mechanism through documented effects on a shared CYP enzyme. This emphasis is reduced in the lean manual triage, bearing in mind that important tasks in pharmacovigilance are to detect the unexpected and previously unknown, as well as to detect problems related to drug usage. Furthermore, the lean manual triage is designed to reward well-described reports to a greater extent than the lean empirical triage.

3 Evaluation of Triage Algorithms

3.1 Methods

After completed development, the candidate triages' respective performance in predicting signals of adverse drug interactions was evaluated. Thus, the development and evaluation phases were securely separated. Two separate benchmarks were used in the evaluation: whether or not DDAs were adverse drug interactions known in the literature, and whether or not DDAs corresponded to prospective adverse drug interaction signals according to expert clinical assessment. Performance was also compared with that of a triage based only on disproportionality screening, which represents the current state-of-the-art. An overview of the evaluation process is provided in the lower panel of Fig. 1.

3.1.1 The Evaluation Set

The performance evaluation was based on a set of 100 randomly sampled DDAs. To be considered for inclusion, DDAs had to fulfil the same requirements as the known adverse drug interactions and DDAs not known to interact previously extracted from Stockley's Interaction Alerts (see Appendix 1). In addition, they were required to have at

least one report on SIC level entered into VigiBase™ during the most recent year, to focus on DDAs of potential current interest. For all DDAs in this evaluation set, data was extracted on the same variables that were previously used to derive the triages. However, this extraction included VigiBase™ reports entered up to February 2011. The three triages were then used to predict the probability of each DDA being an adverse drug interaction signal. For more details on the construction of the evaluation set, see Appendix 3.

3.1.2 Benchmark Classification

All 100 DDAs of the evaluation set were manually classified as already known in the literature, signals or not signals. This classification formed the basis for the two benchmarks in the subsequent evaluation of the triages. In this process, an adverse drug interaction was defined as a drug–drug pair that increases the risk of an ADR more than can be expected based on the independent effects of the two drugs. Relative to the common definition that requires the effects of one of the drugs to be altered [16], our focus was broader and included not only pharmacokinetic and synergistic pharmacodynamic interactions, but also additive pharmacodynamic interactions.

To begin with, one domain expert (JS) reviewed the scientific literature to identify and exclude from further assessment adverse drug interactions established as such in the literature. For each DDA, drug interactions reference literature [11, 16, 17] and PubMed [18] were reviewed to examine whether the drug pair is known to interact. Furthermore, the scientific literature [18–20] was reviewed to determine whether the ADR or a related ADR is listed for the affected drug, and/or for the drug suspected of inducing the interaction.

The DDAs not classified as already known in the literature were then independently assessed by two domain experts (JS and IRE) and classified as either signals or not. In this regard, adverse drug interaction signals were defined in analogy with the CIOMS signal definition for drug–ADR pairs [21]. For each DDA, the experts assessed the available reports in VigiBase™, as well as complementary quantitative and pharmacological information (cf. variables 18–20), and other background information. Discordant classifications were discussed and consensus was reached. Throughout the entire process the triages' predicted probabilities of the DDAs corresponding to adverse drug interactions were unavailable to the assessors.

3.1.3 Evaluation of Performance

The performance of the triages was analysed by relating their predicted probabilities (see Fig. 2) to the outcome of the benchmark classification, for the DDAs in the evaluation set. The analysis primarily relied on so-called receiver operating characteristics (ROC) curves, which display the relation between sensitivity and specificity, for varying thresholds on the triages' predicted probabilities. Because we are mainly interested in predicting adverse drug interaction signals, analysis was first restricted to the DDAs not classified as known in the literature, and the benchmark was defined by the classification of those DDAs into signals or not signals according to clinical assessment. As a complementary analysis, the benchmark was defined by the classification of all 100 DDAs into already known or not known in the literature. Table 4 explains the calculation of sensitivity and specificity using these two complementary benchmarks.

Table 4 Calculation of sensitivity and specificity for generation of receiver operating characteristics (ROC) curves

	Already known in the literature	Not known in the literature	
		Signal	Not signal
Predicted probability $\geq t$	a	c	e
Predicted probability $< t$	b	d	f

For a given threshold t, a triage's sensitivity with respect to the manual classification into signals and not signals is calculated by $c/(c + d)$ and its specificity is calculated by $f/(e + f)$. With respect to the classification of the literature status of the drug–drug–adverse drug reactions, sensitivity is calculated by $a/(a + b)$ and specificity is calculated by $(d + f)/(c + e + d + f)$. By varying t, two complete ROC curves for the triage are constructed. Good prediction corresponds to high sensitivity and high specificity. Note that $a + b + c + d + e + f = 100$

To date, first-pass screening for adverse drug interactions in ICSR databases is almost exclusively based on disproportionality analysis. Therefore our triages were compared with pure disproportionality analysis in an additional ROC analysis. This was done by considering Ω_{025} as a triage algorithm of its own, and varying its threshold to generate ROC curves.

Each triage provides its own ranking of the DDAs in the evaluation set, and discordant results among the various triages were identified and reviewed. This was done to complement the high-level ROC analyses. For a given triage, ranks 1 and 100 correspond to the DDAs most and least likely, respectively, of being adverse drug interactions. When predicted probabilities were tied among several DDAs, all were assigned the best available rank.

3.2 Results

3.2.1 Characteristics of the Evaluation Dataset

Just over 1 million reported DDAs were eligible for inclusion into the evaluation dataset. More than 70 % of

these DDAs did not fulfil any of the variables included in the two lean triages.

Of the 100 clinically evaluated DDAs, 20 were classified as already known adverse drug interactions in the literature. For example, itraconazole – simvastatin – rhabdomyolysis was classified as already known, since concurrent use of itraconazole and simvastatin can raise simvastatin levels, which in turn may increase the risk of myopathy and rhabdomyolysis [22]. Of the remaining 80 DDAs, 30 were classified as signals and 50 as not signals. For example, one of the DDAs classified as signals was warfarin – nystatin – prothrombin decreased. VigiBase™ contained 37 reports on this DDA, for which an interaction was suspected in four, and in two cases these two drugs were the only reported. The association was also supported by a recently published study showing the need to adjust warfarin dosage under concurrent use with nystatin [23]. Among the DDAs classified as not signals was ibuprofen – clindamycin – suicide attempt, where it was clear from the reports that the drugs had been concurrently used in an attempt to commit suicide.

Among DDAs assessed as signals, the majority resulted in increased therapeutic effect, although there were examples of decreased therapeutic effect too. For instance fluconazole – azithromycin – sinusitis had a total of 27 reports, including two strong cases of drug ineffectiveness.

3.2.2 *Evaluation of Performance*

Figure 3a displays the performance of the three triages relative to the clinical assessment of the 80 DDAs in the evaluation set not found in the literature. The lean manual and the lean empirical triages dominate the broad empirical triage, i.e. their curves are consistently above or equal to that of the broad empirical triage. This implies that at any given sensitivity, the lean triages perform equal to or better than the broad triage in terms of specificity. Conversely, for any given specificity, they yield equal or better sensitivity. Neither of the two lean triages dominates the other.

The area under the curve (AUC) is a common summary metric to compare algorithms over all possible thresholds in ROC analyses. In this study, with clinical assessment as the benchmark, AUC corresponds to the probability of ranking a random adverse drug interaction signal higher than a random non-signal. With this benchmark, the lean manual, lean empirical and broad empirical triages obtained AUC values of 0.71, 0.69 and 0.59, respectively. However, these values offer only rough guidance, since in routine screening for potential adverse drug interactions, the higher thresholds are the most relevant: the total list of DDAs is typically too comprehensive to undergo manual assessment, and it is reasonable to start with the highest ranked DDAs. For VigiBase™, thresholds corresponding to specificity below 0.7 are not practically relevant considering their implied workloads, which would be in the order of 150,000 DDAs. We have refrained from computing partial AUCs, corresponding to thresholds above a certain limit, to maintain generalizability to databases that do not match the size of VigiBase™.

Figure 3b displays the performance of the three triages relative to the literature status of all 100 DDAs in the evaluation dataset. The broad empirical triage performs slightly better relative to literature status than to clinical

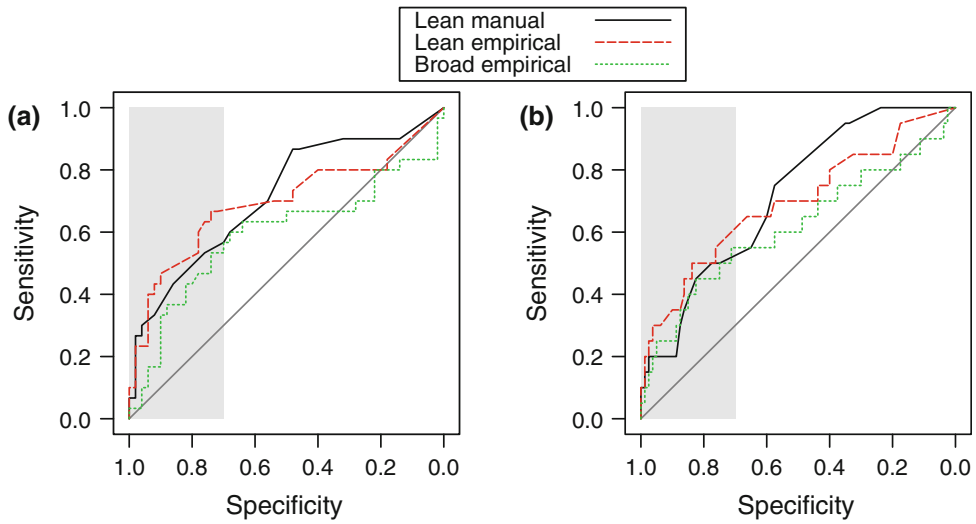


Fig. 3 Receiver operating characteristics curves for the three triages relative to (**a**) the clinical assessment of the 80 DDAs in the evaluation dataset not found in the literature; and (**b**) the literature status of all 100 DDAs in the evaluation dataset. In (**a**), the area under the curve is 0.71, 0.69 and 0.59 for the lean manual, lean empirical and broad empirical triages, respectively. In (**b**), the corresponding values are 0.72, 0.69 and 0.62, respectively. The shaded regions are the ones considered practically relevant for VigiBase™. The 45° lines correspond to an algorithm based on random guessing. *DDAs* drug–drug–adverse drug reactions

assessment. However, it still has clearly lower AUC than both of the lean triages, and is dominated by the lean empirical. The broad empirical triage is not considered further in this study.

3.2.3 Characterization of Lean Triages

The two lean triages are quite similar in their design (see Table 3), which is also reflected in the results. For example, among the ten DDAs ranked highest by the lean manual triage, six are ranked in the top ten also by the lean empirical triage. However, for some DDAs they differ considerably, as reflected by the two examples presented in Tables 5 and 6. The signal fluconazole – venlafaxine – pneumonia is ranked 12th by the empirical triage but only 31st by the manual triage. This discrepancy is due to the higher reward given to the CYP variable in the empirical triage, in particular when combined with the report where the drugs are listed as interacting. For the known additive adverse drug interaction glyceryl nitrate – diltiazem – fall [24], which is ranked 14th by the manual triage and merely 60th by the empirical triage, the explanation is a combination of three factors: the absence of a common CYP pathway, which affects the empirical triage more; the reward given to disproportionality ($\Omega_{025} > 0$) by the manual triage; and the higher reward for reported clinical information directly suggestive of an interaction by the manual triage.

3.2.4 Comparison with Disproportionality Analysis

Figure 4 presents ROC comparisons between the lean triages and disproportionality analysis. Both relative to clinical assessment (Fig. 4a) and literature status (Fig. 4b), the lean triages perform better than Ω_{025}, the particular disproportionality metric used for the purpose of this comparison. This is true in particular around the region of the natural threshold for Ω_{025}, zero. At the specificity obtained by $\Omega_{025} > 0$ with clinical assessment as the benchmark (0.88), the lean manual and lean empirical triages yield a sensitivity of 0.40 and 0.48, respectively, compared with 0.30 for Ω_{025}. Analogously, at the specificity obtained by $\Omega_{025} > 0$ with literature status as the benchmark (0.81), the lean manual and lean empirical triages reach a sensitivity of 0.46 and 0.50, respectively, compared with 0.25 for Ω_{025}. The lean triages also outperform Ω_{025} overall, as reflected by their higher AUC values.

To better understand the differences in practice between the lean triages and disproportionality analysis alone, it is instructive to study their respective top-ranked DDAs. Figure 5 displays such a comparison, based on a probability threshold of 0.15 for the lean triages. This threshold appears to be practically relevant for VigiBase™, since it yields a manageable number of about 14,000 DDAs to assess during the foreseeable period of a decade. Both triages highlight 16 DDAs in the evaluation set at this threshold, so the top 16 DDAs for Ω_{025} are shown for comparison. (Note, however, that Ω_{025} highlights 20 DDAs at its natural threshold zero.) The lean triages display excellent performance: For the lean manual triage, six of the 16 DDAs are known in the literature, and eight of the remaining 10 were classified as signals in the clinical assessment. The lean empirical triage highlights seven

Table 5 Discordant rankings by the lean triages of the DDA fluconazole-venlafaxine-pneumonia

Triage	Intercept	Supporting information			Sum	Probability	Rank
		CYP	Interacting	CYP + Interacting			
Lean empirical	−3.219	0.607	1 report × 0.312	1 report × 1.187	−1.11	0.25	12
Lean manual	−3.2	0.4	1 report × 0.6	1 report × 0.0	−2.2	0.10	31

This DDA was assessed as a signal with two strong reports. Fluconazole is a CYP3A4 inhibitor and is likely to have precipitated the reaction since venlafaxine is a CYP3A4 substrate. There are several case reports in the literature supporting a link between venlafaxine and pneumonia [37, 38]

DDA drug-drug-adverse drug reaction, *CYP* cytochrome P450

Table 6 Discordant rankings by the lean triages of the DDA glyceryl nitrate-diltiazem-fall

Triage	Intercept	Supporting information			Sum	Probability	Rank
		$\Omega_{025} > 0$ (SI)	Interacting	MedDRA® Interaction (SI)			
Lean empirical	−3.219	0.0	1 report × 0.312	1 report × 0.187	−2.72	0.06	60
Lean manual	−3.2	0.6	1 report × 0.6	1 report × 0.3	−1.7	0.15	14

This is listed as an interaction in diltiazem's summary of product characteristics [24]: 'Increased hypotensive effects and faintness (additive vasodilating effects)'

DDA drug-drug-adverse drug reaction, *MedDRA®* Medical Dictionary for Regulatory Activities, *SI* suspected, interacting

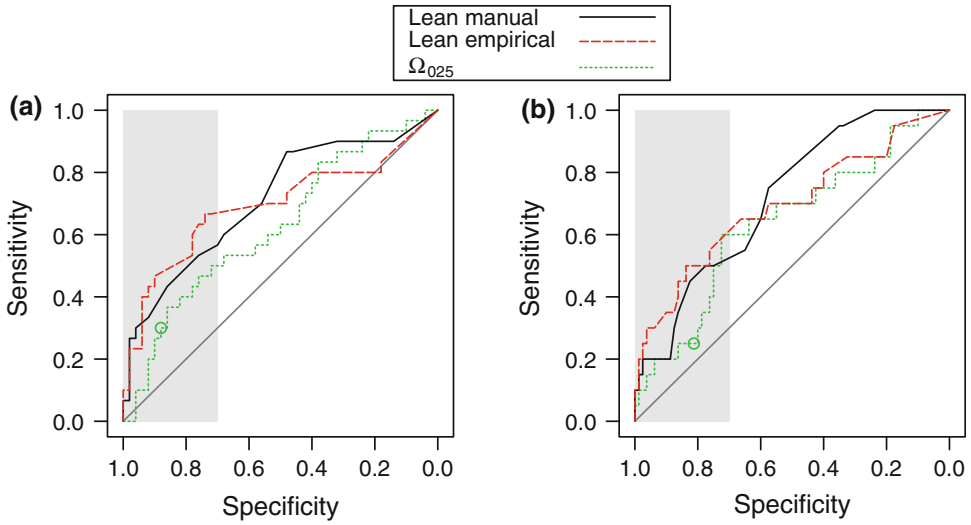


Fig. 4 Receiver operating characteristics curves for the two lean triages and disproportionality analysis alone (Ω_{025}) relative to (**a**) the clinical assessment of the 80 DDAs in the evaluation dataset not found in the literature; and (**b**) the literature status of all 100 DDAs in the evaluation dataset. In (**a**), the area under the curve is 0.71, 0.69 and 0.63 for the manual triage, empirical triage and Ω_{025}, respectively. In (**b**), the corresponding values are 0.72, 0.69 and 0.64, respectively. The shaded regions are the ones considered practically relevant for VigiBase™. The circles correspond to the natural threshold for Ω_{025}, zero. The 45° lines correspond to an algorithm based on random guessing. *DDAs* drug–drug–adverse drug reactions

DDAs reported as adverse drug interactions in the literature, and seven of the remaining nine were classified as signals. For Ω_{025}, 5 of the 16 DDAs are known in the literature, and 6 were classified as signals. Consequently, Ω_{025} highlighted five non-signals, compared with two each for the lean triages.

Several adverse drug interactions known in the literature or classified as signals in the clinical assessment have negative Ω_{025} values but are found among the 16 top-ranked DDAs by the lean manual and lean empirical triages. Two of these DDAs, the signal omeprazole – ritonavir/lopinavir – vomiting and the known adverse drug interaction amiodarone – simvastatin – hepatic function abnormal, are presented in detail in Tables 7 and 8. It is clear that the triages have benefited from the strong clinical and pharmacological information available for these DDAs.

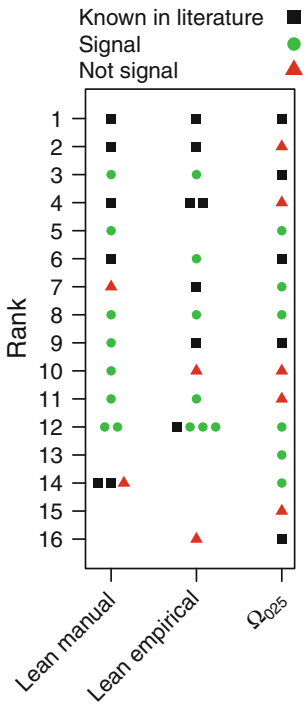


Fig. 5 Outcome from literature review and/or clinical assessment of the top-ranked DDAs by the lean triages and by disproportionality analysis alone (Ω_{025}). A probability threshold of 0.15 has been used for the lean triages, and the list for Ω_{025} has been truncated to include the same number of DDAs. Using the natural threshold of zero for Ω_{025} would yield four additional DDAs, three of which are signals and one not a signal. *DDAs* drug–drug–adverse drug reactions

4 Discussion

We have developed two effective triages for adverse drug interaction surveillance in large collections of ICSRs. The proposed triages automate some of the considerations applied in detailed clinical assessment, and as such use a much broader range of information than what is customary in first-pass screening. For example, the triages take into account whether the drugs are metabolized via the same

Table 7 Rankings by the lean triages of the DDA omeprazole-ritonavir/lopinavir-vomiting not highlighted by disproportionality alone ($\Omega_{025} > 0$)

Triage	Intercept	Supporting information						Sum	Probability	Rank
		CYP	Interacting	MedDRA® Interaction (SI)	Narrative information (SI)	CYP + Interacting	CYP + Narrative information (SI)			
Lean empirical	−3.219	0.607	1 report × 0.312	1 report × 0.187	1 × 0.028	1 report × 1.187	1 report × 0.736	−0.16	0.46	6
Lean manual	−3.2	0.4	1 report × 0.6	1 report × 0.3	1 × 0.4	1 report × 0.0	1 report × 0.0	−1.5	0.18	11

This DDA was assessed as a signal with one very strong report. The mechanism is likely to be primarily pharmacokinetic: ritonavir inhibits CYP3A4, which is one of omperazole's metabolizing enzymes. However, an additive effect cannot be excluded since both ritonavir/lopinavir and omeprazole can cause vomiting [39, 40]

DDA drug-drug-adverse drug reaction, *CYP* cytochrome P450, *MedDRA®* Medical Dictionary for Regulatory Activities, *SI* suspected, interacting

Table 8 Rankings by the lean triages of the DDA amiodarone-simvastatin-hepatic function abnormal not highlighted by disproportionality alone ($\Omega_{025} > 0$)

Triage	Intercept	Supporting information				Sum	Probability	Rank
		CYP	Interacting	MedDRA® Interaction (SI)	CYP + Interacting			
Lean empirical	−3.219	0.607	3 reports × 0.312	1 report × 0.187	3 reports × 1.187	2.07	0.89	4
Lean manual	−3.2	0.4	3 reports × 0.6	1 report × 0.3	3 reports × 0.0	−0.7	0.33	6

It is recognized that simvastatin and amiodarone interact to increase the risk of myopathy [41]. Amiodarone is a CYP inhibitor, although it is not known whether it inhibits simvastatin's metabolism [16]. Simvastatin can elevate the levels of certain hepatic enzymes [41], and hence this was considered a known adverse drug interaction

DDA drug-drug-adverse drug reaction, *CYP* cytochrome P450, *MedDRA®* Medical Dictionary for Regulatory Activities, *S* suspected, interacting

CYP enzyme, and if so whether their activity may affect the metabolism of the other drug involved. They also reward reporting with strong clinical support of an adverse drug interaction. In addition to plausible CYP metabolism, the following were identified as valuable predictors of adverse drug interaction signals: notes of suspected interaction by the reporter, reports of unexpected therapeutic response, reports of altered therapeutic effect with dose information and reports of altered therapeutic effect when only two drugs had been used. In our study, the new triages identified reporting patterns corresponding to both prospective signals of adverse drug interactions and issues established as such.

The proposed triages produce better predictions in our study than disproportionality analysis alone, both relative to what is already known in the literature and relative to prospective signal detection. The value of the triages is well illustrated by their high rankings for the signal of vomiting under concomitant use of omeprazole and ritonavir/lopinavir, and for the known interaction between amiodarone and simvastatin leading to abnormal hepatic function. These are DDAs with strong individual reports that do not stand out in terms of the total numbers, and thereby are not highlighted with disproportionality analysis. The main limitation of pure disproportionality analysis is its exclusive reliance on raw numbers of reports. We did not compare different measures of disproportionality, but previous studies have indicated that Omega compares favourably to standard alternatives such as logistic regression [7], and we expect the added value of our triages will hold for other measures of disproportionality.

On the whole, the two lean triages yield similar performance for both benchmarks. The ROC analysis appears to marginally favour the lean empirical triage, although this depends on what regions of the curves are emphasized. Conceptually, the main advantage of the lean empirical triage is that it is data-driven and unlikely to have over-fitted to the preferences of JS and IRE who were involved in the definition of the lean manual triage and responsible for the clinical assessment in the evaluation phase. The main advantage of the lean manual triage is that it uses a broader range of qualitative information on reports, incorporates disproportionality analysis and relies less heavily on a known CYP pathway. It is therefore less likely to have over-fitted to the dominance of pharmacokinetic interactions in Stockley's Interaction Alerts, and may be at an advantage to detect pharmacodynamic interactions. This is valuable since from our perspective the scope of pharmacovigilance should be to detect any drug combination with increased risk of ADRs. We made the assumption that the broad empirical triage may over-fit to the reference set used in the development phase. This is strongly supported by our evaluation results.

A strength of our evaluation set is that it provides two independent benchmarks: one based on what is known in the literature and one based on clinical assessment. The analysis relative to clinical assessment is our primary interest since it reflects whether the data could support a safety signal, prospectively, and covers any issues with potential implication to patient safety or public health. In addition, whereas the literature benchmark will treat any new signals as false positives, the benchmark based on clinical assessment focuses exclusively on the unchartered territory that is the reality of day-to-day pharmacovigilance. On the other hand, clinical assessment is subjective and a separate set of clinical assessors might have selected a different set of interaction signals from the same 100 DDAs. In this respect, it is reassuring that the evaluation against the literature benchmark yields similar results in our study. As a precaution, our triages were finalized before anyone had access to the 100 DDAs in the evaluation set. As such, the development and evaluation phases were clearly separated. Furthermore, the predicted probabilities were blinded and the order of DDAs was randomized to the assessors throughout the complete evaluation phase.

Our triages produce rankings rather than binary dichotomizations. This is unlike most, if not all, algorithms in routine use for pair-wise ADR surveillance. One advantage is that this allows clinical assessors to start with the DDAs most likely to represent true adverse drug interactions. Should a binary dichotomization be preferred, it is possible to threshold the interaction triage score at an appropriate level. As an illustration, at a threshold of 0.15 predicted probability (that the suspected adverse drug interaction would be reported in the literature in the near future), both lean triages would highlight around 14,000 DDAs in VigiBase™ as suspected adverse drug interactions. In our evaluation set, 16 DDAs met this threshold for each triage. Six or seven (38–44 %) of those were identified as adverse drug interactions known in the literature and seven or eight of the remaining DDAs (78–80 %) were classified as signals worthy of further follow-up. In pair-wise drug-ADR surveillance in VigiBase™, around half of the associations in first-pass screening relate to already known ADRs, and among the other half, around 20 % are sent out for clinical assessment. The high proportion of signals in our study suggests excellent performance, but it should be regarded in the light of being based on quite a small sample of DDAs, and also that it may be favoured by the use of a fairly conservative threshold. Furthermore, the more limited understanding and documentation of clinically relevant interactions can be expected to yield a lower proportion of already unknown causal associations in interaction surveillance.

Drug interactions of clinical importance are likely to increase the risk for a spectrum of adverse reactions. We have evaluated adverse drug interaction surveillance at the level of specific ADRs rather than for pairs of drugs in general. The latter is an interesting topic for future research. It may be particularly valuable for rare drugs and in smaller databases. Most of the triage variables that we have identified are directly applicable to an analysis at the level of drug pairs, without a link to a specific ADR. The main challenge would be to adapt disproportionality analysis to this setting.

As far as we are aware, this is the first application of predictive regression models for first-pass screening of large collections of ICSRs. We opted for lasso shrinkage logistic regression, which was of significant help in the process of reducing the number of potential predictor variables. During the course of this work we have been made aware of interesting recent developments within this methodology. These should be considered in future work and include the possibility to constrain the regression modelling to non-negative coefficients only, and to determine their confidence intervals [25].

Our triages are not a perfect safety net to detect all signals of suspected adverse drug interactions. In our evaluation set, three DDAs clinically assessed as signals received the lowest possible score by both lean triages, i.e. these DDAs did not fulfil any of the variables included into these two triages. Still, for some of these DDAs there were reports with supportive information such as positive dechallenge or dose information. This emphasizes the importance of future work in this area, in particular as regards the ascertainment of additional predictors of suspected adverse drug interactions at the level of individual reports.

Collections of ICSRs suffer from well-known limitations, e.g. extensive and variable underreporting [12]. This may distort aggregated analyses such as our triages, and potential findings from automated screening should always be followed by detailed clinical assessment of the reports.

5 Conclusions

The value of incorporating clinical and pharmacological information in triages for first-pass screening for signals of adverse drug interactions is clear. The proposed triages clearly outperform pure disproportionality analysis. Plausible CYP metabolism, reporter notes of suspected interaction, and unexpected therapeutic response or altered therapeutic effect with dose information or solely two reported drugs were identified as valuable predictors of adverse drug interaction signals.

Acknowledgments The authors thank the Royal Pharmaceutical Society for access to Stockley's Interactions Alerts. Furthermore, the authors are indebted to the National Centres that contribute data to the

WHO Programme for International Drug Monitoring. The opinions and conclusions in this study are not necessarily those of the various centres, nor of the WHO. As of February 2013, Johanna Strandell is a full-time employee of TFS Trial Form Support; however, all her contributions to this work were made prior to that time, during the course of her employment at the Uppsala Monitoring Centre.

No sources of funding were used to conduct this study or prepare this manuscript. The authors have no conflicts of interest that are directly relevant to the content of this study. MedDRA® trademark is owned by the International Federation of Pharmaceutical Manufacturers & Associations (IFPMA) on behalf of the ICH.

Appendix 1: Extraction of Known Adverse Drug Interactions and Drug–Drug–Adverse Drug Reactions (DDAs) not Known to Interact

We have previously described in detail how a reference set of known adverse drug interactions and DDAs not known to interact was constructed based on Stockley's Interaction Alerts [9]. The main steps are repeated here.

Stockley's Interaction Alerts [11] is an electronic quick ready-reference with information on more than 40,000 drug–drug, drug–alcohol and drug–food pairs. It is one component of Stockley's Drug Interactions [16], a well-renowned international drug interaction reference. Among other things, its textual descriptions list adverse drug reactions (ADRs) that might result from the drug interactions.

For drug pairs entered between the first quarter of 2007 and the third quarter of 2009, the drug names were matched to substances in the WHO Drug Dictionary Enhanced [13], and pairs where either drug was ethanol or nicotine were excluded. From each free text description, WHO Adverse Reaction Terminology (WHO-ART) preferred terms [13] were extracted using a customized algorithm [9, 26], and were then combined with the drug pair to form DDAs. For each DDA, a note was made of the quarter when it was first included into Stockley's Interaction Alerts.

Any DDA was excluded if its ADR was part of a variable to be tested for inclusion into the triages (see the definitions of variables 1, 2, and 6 in Table 1). Furthermore, any DDA was excluded that did not fulfil basic reporting requirements into the WHO global individual case safety report database, VigiBase™ [13], which is the main data source in this study: After removal of suspected duplicates [14], at least three reports were required where the two drugs were listed as suspected, interacting, or concomitant, entered between 1 January 1990 and the quarter prior to inclusion of the DDA into Stockley's Interaction Alerts; and at least one report was required with both drugs listed as suspected or interacting, in that same time period. Finally, for the remaining DDAs, the extracted ADRs were manually validated against the source text, and false hits were removed.

This yielded 324 known adverse drug interactions. The previously reported number, 322 [9], excludes two DDAs that fulfilled the reporting requirements in the quarter prior to inclusion into Stockley's Interaction Alerts, but not at the database end date.

The reference set was completed by including a comparison group of DDAs not known to interact. All drugs in Stockley's Interaction Alerts were paired, and those pairs that were actually listed were excluded. Thereafter the remaining pairs were combined with all ADR terms extracted from Stockley's Interaction Alerts in the identification of known adverse drug interactions, thus forming a new group of DDAs. From this group, DDAs with certain drugs or ADRs were excluded, as described above for the known adverse drug interactions. For the remaining DDAs, it was identified for what quarters they fulfilled the VigiBase™ reporting requirements used for the known adverse drug interactions. Finally, 20 DDAs not known to interact were randomly chosen for each known adverse drug interaction, matched on quarter of data extraction. (The same DDA could not be selected for multiple quarters.)

Appendix 2: Regression Modelling

For the purpose of empirically deriving triages, lasso logistic regression was used. The lasso shrinkage restricts the coefficients so that the sum of their absolute values cannot exceed a predefined limit [27]. This reduces the risk of over-fitting to the available dataset, which in turn should yield better predictive performance on new data. At the same time, the lasso performs model selection by setting some coefficients to exactly zero, i.e. it allows only the best predictive variables into the model. Shrinkage regression of this type has grown in usage over recent years, and has been repeatedly and successfully applied in the medical domain [28, 29]. We used the implementation of lasso logistic regression provided by Friedman et al. [30].

Prior to model fitting, all variables were standardized to have unit variance. However, all coefficients presented (see Tables 2, 3) have been transformed back, and should be applied to the variables' respective original scales.

The standard approach of five-fold cross-validation was used to select the appropriate level of shrinkage [31]. Predictive performance was measured in terms of mean squared error [32]. Once the optimal amount of shrinkage had been determined, the entire data set was used to fit a single model.

Appendix 3: Construction of the Evaluation Set

Because clinical assessment is tedious, the size of the evaluation set was limited to 100 DDAs. To make the

evaluation as informative as possible, it was important to ensure coverage of likely interaction signals as well as unlikely ones. To this end, the random selection of DDAs into the evaluation set was stratified according to the predicted probabilities of the two lean triages. Creating strata by simply dividing the 0–1 interval into equally spaced pieces would yield disproportionately many DDAs with high probabilities, since the distribution of DDAs is heavily skewed towards zero. At the same time, to stratify according to the cumulative distribution of predicted probabilities would yield unsatisfactorily many DDAs with the lowest possible probability, since those DDAs amount to more than 70 % of all included DDAs. We used an intermediate alternative, whereby each stratum was created by incrementally increasing its upper limit. A stratum was considered complete as soon as it was filled up by more DDAs than its specific target size, which was computed as the ratio between the remaining number of DDAs and the remaining number of strata.

References

1. Mjörndal T, Boman MD, Hägg S, Bäckstrom M, Wiholm BE, Wahlin A, et al. Adverse drug reactions as a cause for admissions to a department of internal medicine. Pharmacoepidemiol Drug Saf. 2002;11(1):65–72.
2. Lazarou J, Pomeranz BH, Corey PN. Incidence of adverse drug reactions in hospitalized patients: a meta-analysis of prospective studies. JAMA. 1998;279(15):1200–5.
3. Pirmohamed M, James S, Meakin S, Green C, Scott AK, Walley TJ, et al. Adverse drug reactions as cause of admission to hospital: prospective analysis of 18 820 patients. BMJ. 2004;329(7456):15–9.
4. Pirmohamed M, Breckenridge AM, Kitteringham NR, Park BK. Adverse drug reactions. BMJ. 1998;316(7140):1295–8.
5. Leone R, Magro L, Moretti U, Cutroneo P, Moschini M, Motola D, et al. Identifying adverse drug reactions associated with drug–drug interactions: data mining of a spontaneous reporting database in Italy. Drug Safety. 2010;33(8):667–75.
6. Almenoff JS, DuMouchel W, Kindman LA, Yang X, Fram D. Disproportionality analysis using empirical Bayes data mining: a tool for the evaluation of drug interactions in the post-marketing setting. Pharmacoepidemiol Drug Saf. 2003;12(6):517–21.
7. Norén GN, Sundberg R, Bate A, Edwards IR. A statistical methodology for drug–drug interaction surveillance. Stats Med. 2008;27(16):3057–70.
8. Thakrar BT, Grundschober SB, Doessegger L. Detecting signals of drug–drug interactions in a spontaneous reports database. Br J Clin Pharmacol. 2007;64(4):489–95.
9. Strandell J, Caster O, Bate A, Norén GN, Edwards IR. Reporting patterns indicative of adverse drug interactions: a systematic evaluation in VigiBase. Drug Saf. 2011;34(3):253–66.
10. Edwards IR, Biriell C. Harmonisation in pharmacovigilance. Drug Saf. 1994;10(2):93–102.
11. Baxter K, editor. Stockley's interaction alerts (including data up to April 2011). London: Pharmaceutical Press.
12. Strandell J, Bate A, Hägg S, Edwards IR. Rhabdomyolysis a result of azithromycin and statins: an unrecognized interaction. Br J Clin Pharmacol. 2009;68(3):427–34.
13. Lindquist M. VigiBase, the WHO global ICSR database system: basic facts. Drug Inf J. 2008;42(5):409–19.
14. Norén GN, Orre R, Bate A, Edwards IR. Duplicate detection in adverse drug reaction surveillance. Data Min Knowl Discov. 2007;14(3):305–28.
15. Strandell J, Wahlin S. Pharmacodynamic and pharmacokinetic drug interactions reported to VigiBase, the WHO global individual case safety report database. Eur J Clin Pharmacol. 2011;67(6):633–41.
16. Baxter K, editor. Stockley's drug interactions. 7th ed. London: Pharmaceutical Press; 2006.
17. DrugDex. Drug interactions. Thomson Reuters Healthcare MICROMEDEX 2.0. Available from URL: http://www.thomsonhc.com/home/. Accessed 15 Jun 2011.
18. PubMed, National Centre for Biotechnology Information. Available from URL: http://www.ncbi.nlm.nih.gov/pubmed/. Accessed 15 Jun 2011.
19. Drugdex. Thompson Micromedex database. Available from URL: http://www.thomsonhc.com. Accessed 15 Jun 2011.
20. Datapharm Communications Ltd. Electronic medicines compendium (electronic version). Available from URL: http://www.medicines.org.uk/. Accessed 15 June 2011.
21. CIOMS Working Group VIII. Practical aspects of signal detection in pharmacovigilance. Geneva: CIOMS; 2010.
22. Neuvonen PJ, Niemi M, Backman JT. Drug interactions with lipid-lowering drugs: mechanisms and clinical relevance. Clin Pharmacol Ther. 2006;80(6):565–81.
23. Kovac M, Mitic G, Kovac Z. Miconazole and nystatin used as topical antifungal drugs interact equally strongly with warfarin. J Clin Pharm Ther. 2012;37(1):45–8.
24. Datapharm Communications Ltd. Electronic medicines compendium (electronic version). Product information for Tildiem. Available from: http://www.medicines.org.uk/. Accessed 15 April 2012.
25. Chatterjee A, Lahiri SN. Bootstrapping lasso estimators. J Am Stat Assoc. 2011;106(494):608–25.
26. Caster O. Automatic extraction of adverse drug reaction terms from medical free text. Annual conference of the International Society of Clinical Biostatistics (ISCB'09); 2009 Aug 26: Prague. Available from URL: http://www.iscb2009.info/RSystem/Soubory/Prez%20Wednesday/S33.1%20Caster.pdf. Accessed 15 Apr 2012.
27. Tibshirani R. Regression shrinkage and selection via the lasso. J R Stat Soc Series B Stat Methodol. 1996;58(1):267–88.
28. Fleisher AS, Sowell BB, Taylor C, Gamst AC, Petersen RC, Thal LJ. Clinical predictors of progression to Alzheimer disease in amnestic mild cognitive impairment. Neurology. 2007;68(19):1588–95.
29. Laaksonen R, Katajamaa M, Paiva H, Sysi-Aho M, Saarinen L, Junni P, et al. A systems biology strategy reveals biological pathways and plasma biomarker candidates for potentially toxic statin-induced changes in muscle. PloS one. 2006;1:e97.
30. Friedman J, Hastie T, Tibshirani R. Regularization paths for generalized linear models via coordinate descent. J Statist Softw. 2010;33(1):1–22.
31. Hastie T, Tibshirani R, Friedman J. The elements of statistical learning: data mining, interference and prediction. Canada: Springer Science; 2001.
32. Wallach D, Goffinet B. Mean squared error of prediction as a criterion for evaluating and comparing system models. Ecol Modell. 1989;44:299–306.
33. Meyboom RH, Lindquist M, Flygare AK, Biriell C, Edwards IR. The value of reporting therapeutic ineffectiveness as an adverse drug reaction. Drug Saf. 2000;23(2):95–9.
34. MIMS annual: 30th edn. Sydney: MIMS Australia; 2006.
35. FASS. Interaktioner. Available from URL: http://www.fass.se/LIF/home/index.jsp. Accessed 14 Oct 2011.

36. Flockhart table. Indiana University (updated 19 Aug 19 2009). Available from URL: http://medicine.iupui.edu/clinpharm/ddis/table.aspx. Accessed 13 Oct 2011.
37. Drent M, Singh S, Gorgels AP, Hansell DM, Bekers O, Nicholson AG, et al. Drug-induced pneumonitis and heart failure simultaneously associated with venlafaxine. Am J Respir Crit Care Med. 2003;167(7):958–61.
38. Fleisch MC, Blauer F, Gubler JG, Kuhn M, Scherer TA. Eosinophilic pneumonia and respiratory failure associated with venlafaxine treatment. Eur Respir J. 2000;15(1):205–8.
39. Datapharm Communications Ltd. Electronic medicines compendium (electronic version). Product information for Losec. Available from: http://www.medicines.org.uk/. Accessed 15 Apr 2012.
40. Datapharm Communications Ltd. Electronic medicines compendium (electronic version) Product information for Kaletra. Available from: http://www.medicines.org.uk/. Accessed 15 Apr 2012.
41. Datapharm Communications Ltd. Electronic medicines compendium (electronic version). Product information for Zocor. Available from: http://www.medicines.org.uk/. Accessed 15 April 2012.