Drug Saf (2013) 36 (Suppl 1):S195–S204
DOI 10.1007/s40264-013-0112-0

ORIGINAL RESEARCH ARTICLE

# The Impact of Drug and Outcome Prevalence on the Feasibility and Performance of Analytical Methods for a Risk Identification and Analysis System

**Christian G. Reich · Patrick B. Ryan · Marc A. Suchard**



**Abstract**

*Background* A systematic risk identification system has the potential to study all marketed drugs. However, the rates of drug exposure and outcome occurrences in observational databases, the database size and the desired risk detection threshold determine the power and therefore limit the feasibility of the application of appropriate analytical methods. Drugs vary dramatically for these parameters because of their prevalence of indication, cost, time on the market, payer formularies, market pressures and clinical guidelines.

*Objectives* Evaluate (i) the feasibility of a risk identification system based on commercially available observational databases, (ii) the range of drugs that can be studied for certain outcomes, (iii) the influence of underpowered drug-outcome pairs on the performance of analytical methods estimating the strength of their association and (iv) the time required from the introduction of a new drug to accumulate sufficient data for signal detection.

*Methods* As part of the Observational Medical Outcomes Partnership experiment, we used data from commercially available observational databases and calculated the minimal detectable relative risk of all pairs of marketed drugs and eight health outcomes of interest. We then studied an array of analytical methods for their ability to distinguish between pre-determined positive and negative drug-outcome test pairs. The positive controls contained active ingredients with evidence of a positive association with the outcome, and the negative controls had no such evidence. As a performance measure we used the area under the receiver operator characteristics curve (AUC). We compared the AUC of methods using all test pairs or only pairs sufficiently powered for detection of a relative risk of 1.25. Finally, we studied all drugs introduced to the market in 2003–2008 and determined the time required to achieve the same minimal detectable relative risk threshold.

The OMOP research used data from Truven Health Analytics (formerly the Health Business of Thomson Reuters), and includes MarketScan® Research Databases, represented with MarketScan Lab Supplemental (MSLR, 1.2 m persons), MarketScan Medicare Supplemental Beneficiaries (MDCR, 4.6 m persons), MarketScan Multi-State Medicaid (MDCD, 10.8 m persons), MarketScan Commercial Claims and Encounters (CCAE, 46.5 m persons). Data also provided by Quintiles® Practice Research Database (formerly General Electric's Electronic Health Record, 11.2 m persons) database. GE is an electronic health record database while the other four databases contain administrative claims data.

C. G. Reich (✉)
AstraZeneca PLC, 35 Gatehouse Drive, Waltham, MA 02451, USA
e-mail: reich@omop.org

P. B. Ryan
Janssen Research and Development LLC, Titusville, NJ, USA

M. A. Suchard
Departments of Biomathematics, and of Human Genetics, David Geffen School of Medicine at UCLA, University of California, Los Angeles, CA, USA

M. A. Suchard
Department of Biostatistics, UCLA Fielding School of Public Health, University of California, Los Angeles, CA, USA

C. G. Reich · P. B. Ryan · M. A. Suchard
Observational Medical Outcomes Partnership, Foundation for the National Institutes of Health, Bethesda, MD, USA

*Results* The performance of methods improved after restricting them to fully powered drug-outcome pairs. The availability of drug-outcome pairs with sufficient power to detect a relative risk of 1.25 varies enormously among outcomes. Depending on the market uptake, drugs can generate relevant signals in the first month after approval, or never reach sufficient power.
*Conclusion* The incidence of drugs and important outcomes determines sample size and method performance in estimating drug-outcome associations. Careful consideration is therefore necessary to choose databases and outcome definitions, particularly for newly introduced drugs.

## 1 Introduction

A risk identification system aims at the detection and evaluation of adverse effects of all marketed medical products. Although randomized controlled clinical trials are the gold standard for studying drug effects, their use is constrained by their lack of generalizability to real-world clinical settings, their limited study duration and their tightly restricted patient population. These limitations can be overcome by performing retrospective observational studies in longitudinal data of large clinical databases that have emerged over the last two decades [1]. Drug effects can now be studied over broad, diverse patient populations, clinical settings, different health plans and physician specialties [2].

However, in order for a risk identification system to detect an association between a particular drug and a particular outcome, sufficient numbers of cases of drug exposure and outcome occurrences need to be available to detect a certain effect size (relative risk) with sufficient power. Few of these parameters are under the control of the researcher; the number of cases depends on the time since market introduction of the drug, the drug use rate, the incidence rate of the outcome and the size of the database. Larger databases with more complete data capture contain more data that can be studied, but even the best database is eventually limited by the size of the population within a healthcare system. Many authors have suggested that an observational study that yields a larger relative risk is more likely to indicate a true association [3] with some advocating that relative risk (RR) estimates should be larger than 2 to warrant publication [4, 5]. In the context of drug safety, however, many recent high profile issues, such as the myocardial infarction risks of rofecoxib and rosiglitazone, have involved relative risks less than two (although these conclusions were based largely on meta-analyses of randomized trials, or large individual randomized trials).

Coloma et al. [6], have studied the use of electronic health databases for active drug safety surveillance in the Exploring and Understanding Adverse Drug Reactions (EU-ADR) project (encompassing $\sim$20 million lives). The authors determined whether such a healthcare database-based signal detection system has sufficient power to complement spontaneous reporting systems currently used for post-marketing drug surveillance. Using a simple power estimator, they found that for a relative risk of at least 2, depending on the frequency of the outcome, 1 to 23 percent of the 2,289 drugs in the studied databases have sufficient power to be investigated. However, the paper did not reveal the names of the drugs and did not study the performance of different analytical methods used in drug surveillance when testing underpowered drugs.

As part of the Observational Medical Outcomes Partnership (OMOP), we evaluated the feasibility of detecting a signal of association between drugs and outcomes in real-world data from commercially available administrative claims databases across a library of analytical methods with different characteristics and designs [7]. In an experimental setting that applied a set of test cases of drugs with known association to four select health outcomes of interest (HOIs), as well unrelated control drugs, we tested the ability of methods to discriminate between positive and negative controls. We aimed at testing method performance with all test cases or only those with sufficient power. This experiment allowed us to understand the effect of drugs and their usage frequency that are specific to certain outcomes.

Systematic drug surveillance requires the ability to test all drugs for any outcome. To determine the range of drugs sufficiently powered for systematic surveillance, we calculated the minimal detectable relative risk (MDRR) for marketed drugs in the US in a few example outcomes relevant for drug research and development. However, the drugs for which such surveillance is most crucial are those newly introduced to the market and for which a limited safety profile is available from clinical studies. To this end, using drugs introduced during the data capture of the studied databases ranging between 2003 and 2008, we also asked the question of how long it typically takes from the introduction to accumulate sufficient cases for signal detection.

## 2 Methods

### 2.1 Analytical Methods

A variety of analytical methods is under consideration for risk identification of medical products [8]. As part of the OMOP experimental framework, we established a method

library (http://omop.org/MethodsLibrary) for the purpose of empirical testing for their performance. We implemented seven different analytic methods for potential deployment in a risk identification system. These methods are a new user cohort design (CM), case control design (CC), the self-controlled case series (SCCS), a self-controlled cohort design (SCC), temporal pattern discovery (ICTPD), a disproportionality analysis (DP) and a longitudinal gamma Poisson shrinker (LGPS). All specific analysis choices in adapting these methods for use with longitudinal observational healthcare database, such as the pre-exposure control period or the length of the post-exposure time-at-risk, were parameterized. Complete descriptions, references and source code for each method are described elsewhere [8] and available at http://omop.org/MethodsLibrary.

### 2.2 Power Calculation and Minimal Detectable Relative Risk

Exact power calculations differ between analytic methods. To facilitate development, we considered a general approach to power calculations necessary for a risk identification system. In this system, we computed the MDRR across our longitudinal healthcare databases of varying population sizes and exposures to different drugs. Our power calculation generalization relies on the observation that (i) an analytic method reports estimates that can translate into a relative risk (*RR*) and its standard error of developing an adverse event (AE) while exposed to a given drug, (ii) that many of the seven analytic methods take a cohort design approach and (iii) the approximation has found successful previous use in similar studies [6, 9]. The approximation assumes that observed AEs resulting from drug exposure arise from a simple Poisson process with intensity $RR * E$, where $E$ is the expected number of events in an unexposed, background population. Let $C$ count the observed number of events from this process in the cohort. For modestly large values of $RR * E$, the square-root of $C$ is approximately normally distributed, with a mean equaling the square-root of $RR * E$ and a fixed variance of ¼ [10]. Because the variance is known, we consider the construction of a normally distributed $Z$-statistic test involving the null hypothesis that the observed statistic mean $\mu = \sqrt{(RR * E)}$ is fixed at $\mu_0 = \sqrt{E}$, i.e. $RR = 1$. For a given power $(1 - \beta)$, or probability of rejecting the null hypothesis assuming that the alternative is true, and for a nominal $\alpha$-level for Type I errors,

$$Z_{(1-\beta)} = Z_\alpha - 2(\sqrt{(RR)} - 1)(\sqrt{E}), \quad (1)$$

where $Z_x$ is the $x$-th quantile of a standard normal random variable. Within this paper, we take the usual $1 - \beta = 0.80$ and $\alpha = 0.05$. To determine the MDRR for a drug and outcome in a given database, we estimate $E$ and then back-calculate the *RR* from Eq. (1) within 20 age decile-by-gender strata.

### 2.3 Testing of Methods

In the overall OMOP experiment we used five different observational databases, but here we will restrict ourselves to the following three representative databases to allow evaluation of performance across different populations and data capture processes during the years 2003–2008: MarketScan® Lab Supplement (MSLR, 1.2M patients); MarketScan® Medicare Supplemental Beneficiaries (MDCR, 4.6M patients), and MarketScan® Commercial Claims and Encounters (CCAE, 46.5M patients). All of these databases contain administrative claims data. The data are described in more detail elsewhere [11].

During method evaluation, we executed different analysis choice combinations against a set of drug-outcome pairs to generate an effect estimate and standard error for each pair and parameter combination. These test cases included positive controls—active ingredients with evidence to suspect a positive association with the outcome—and negative controls—active ingredients with no evidence to expect a causal effect with the outcome. The full set of test cases and its construction is described elsewhere [12]. A list of HOIs is provided in Appendix.

### 2.4 Discrimination Metrics

The methods were applied to four HOIs—acute liver injury, acute kidney injury, acute myocardial infarction, and upper gastrointestinal bleeding- to evaluate discrimination, i.e. the ability of methods to distinguish between positive and negative controls. The relative risk estimates were used to draw the receiver operator characteristics (ROC) curve, and compute the area under this curve (AUC) as a single metric of performance [13]. An AUC of 1 indicates a perfect distinction of positive test cases from negative ones. An AUC of 0.5 is equivalent to random guessing.

## 3 Results

### 3.1 Method Performance

Figure 1 shows the ROC curve and AUC values of the method-settings combinations that achieved the highest AUCs for the four HOIs. For clarity, only data from the MDCR database are shown. The best performing method was the self-controlled cohort method. The performance utilizing all control cases, whether sufficiently powered or not, is compared to that applying only cases that have sufficient power. In every situation shown, the

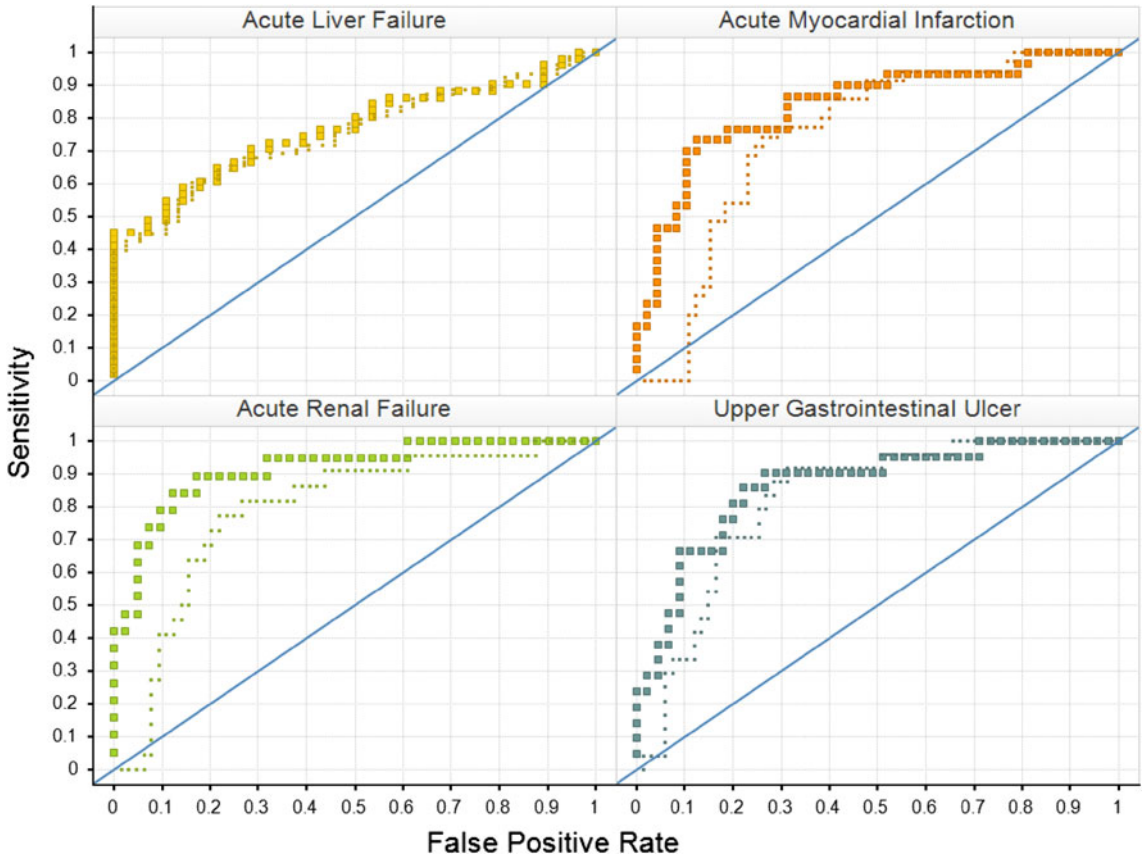


**Fig. 1** Receiver-operator curves (ROC) and area under the curve (AUC) values for methods discriminating between positive and negative drug controls for four select health outcomes of interest (HOIs) in the MDCR database. *Large boxes* indicate the ROC when only cases with sufficient data to detect a relative risk (RR) of 1.25 are used. *Small boxes* show the ROC when all controls are used. The resulting AUC scores and the number of applicable test cases are provided in the table. The method and parameter combination optimal for each health outcome of interest are shown

discriminating power of the method improves after exclusion of the underpowered pairs.

In Fig. 2, we show the AUCs for all method-analytical choice combinations that were evaluated, again focusing on the MDCR database, although other databases showed equivalent results (data not shown). On the *x*-axis the AUC is plotted when using all controls, the *y*-axis shows the AUC when restricting to those drug-HOI pairs where there is enough power to detect a RR of 1.25. Each dot represents an analytical method and a specific set of analytical choices. We see that for most analyses, the AUC improves when restricting to powered drug-HOI pairs, although for upper GI bleeding this trend is not very strong. The reason for this is most likely that upper GI bleeding is a frequent condition, and only few controls had to be eliminated when restricting on MDRR.

### 3.2 Drug-Outcome Pairs with Enough Power

We examined active drug ingredients approved for marketing in the United States and captured in the databases. We found 1,538 active drug ingredients. For each of these drugs, we calculated the MDRR for five different HOIs that represent different magnitudes of prevalence: progressive multifocal encephalopathy (PML), aplastic anemia, mortality after myocardial infarction (MI), venous thromboembolism, and upper GI bleeding. Table 1 lists the number of cases, incidence rate and number of drugs with MDRR <1.25 per HOI in the smallest (MSLR) and largest (CCAE) databases.

Figure 3 shows each drug's rate plotted against the MDRR. The figure demonstrates that for the most frequent HOIs (e.g. upper GI bleeding), a large number of drug ingredients can be tested using observational databases of even modest size (MSLR, 1.2M patients) for a relative risk as low as 1.25. However, no database is large enough to cover all drugs, and some of the rare pharmacologically relevant conditions such as PML cannot be studied as none of the databases examined provides sufficient data.

Drug surveillance is particularly important for newly approved drugs, as they have not had a widespread application to different populations in real clinical settings, and

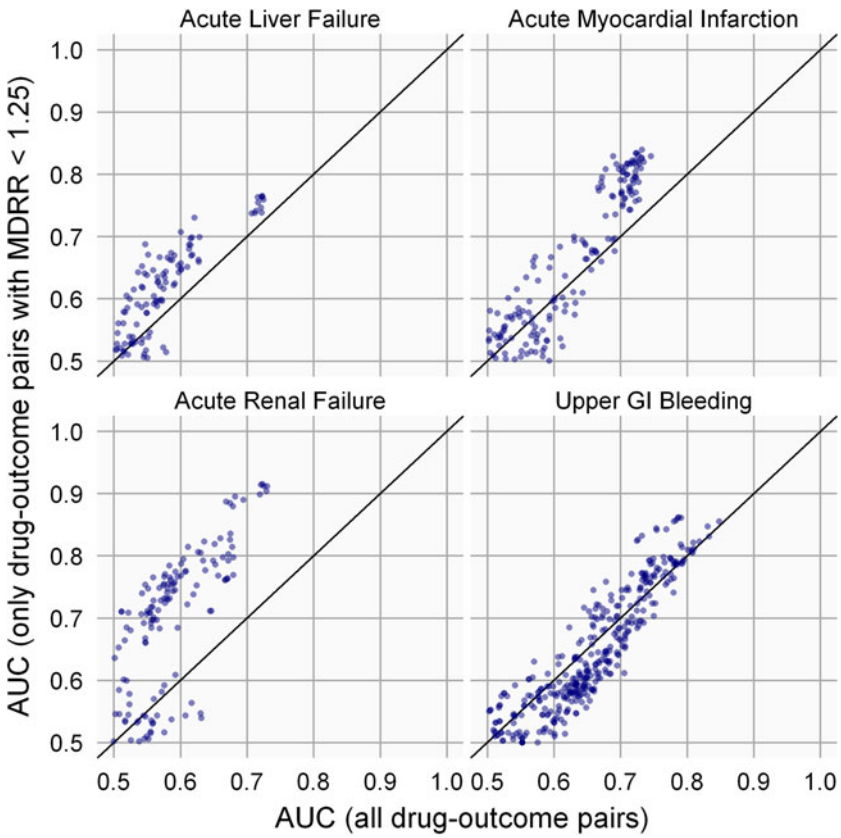


**Fig. 2** AUCs when using all drug-outcome pairs versus using only pairs with enough power to detect a RR of 1.25. Every *dot* represents the performance of a combination of a method and a specific set of analysis choices in the MDCR database

**Table 1** HOIs, the number of cases in the database (patients taking the drug at least once), the rate of cases per 100 thousand patients and the percentage of drugs that can be studied for an effect size (relative risk) as low as 1.25

| Health outcome of interest | MSLR | | | CCAE | | |
|---|---|---|---|---|---|---|
| | # Cases | Cases/100k | % of drugs MDRR ≤1.25 | # Cases | Cases/100k | % of drugs MDRR ≤1.25 |
| PML | 17 | 1.38 | 0.0 | 275 | 0.592 | 0.0 |
| Aplastic anemia | 329 | 26.8 | 0.0 | 6,651 | 14.3 | 7.7 |
| Mortality after MI | 1,909 | 155 | 1.4 | 37,598 | 80.9 | 21.3 |
| Venous thromboembolism | 18,704 | 1,521 | 16.2 | 359,798 | 774.5 | 46.4 |
| Upper GI bleeding | 204,832 | 16,662 | 41.5 | 4,578,767 | 9,856 | 65.8 |

*MSLR*, MarketScan Lab Supplement, *CCAE* MarketScan Commercial Claims and Encounters, *PML* progressive multifocal leukoencephalopathy, *MI* myocardial infarction, *GI* gastrointestinal

vary tremendously in their dynamics of adoption. We therefore calculated the calendar time that it takes for drugs to reach sufficient power to be tested. We chose drugs introduced during the coverage of the CCAE and MSLR databases (2003–2008) and focused on their known side effects as examples of drug-outcome pairs for which surveillance would be expected to generate a signal. Figure 4 shows the MDRR for these drug-outcome pairs over time.

The result varies greatly for each drug-outcome pair. First, market adoption does not necessarily follow the approval date closely. For some drugs such as tipranavir (approved 22-Jun-05), data emerge shortly afterwards in the database. For others such as gemifloxacin (approved 04-Apr-03) the product appears only 18 months later, while for a third group such as bortezomib (approved 13-May-03) the drug had already achieved market penetration before general marketing approval by the FDA, perhaps for reasons of compassionate use in life-threatening cases of its indication. Second, depending on the frequency of the outcome and market adoption of the drug, the speed by which a drug-outcome pair reaches an MDDR of 1.25 varies greatly as well. While bortezomib had achieved full power for detecting acute liver failure already at approval date, gemifloxacin took about 2 years to reach that threshold. Posaconazole (approved 15-Sep-06) was still accumulating acute liver failure cases at the end of the

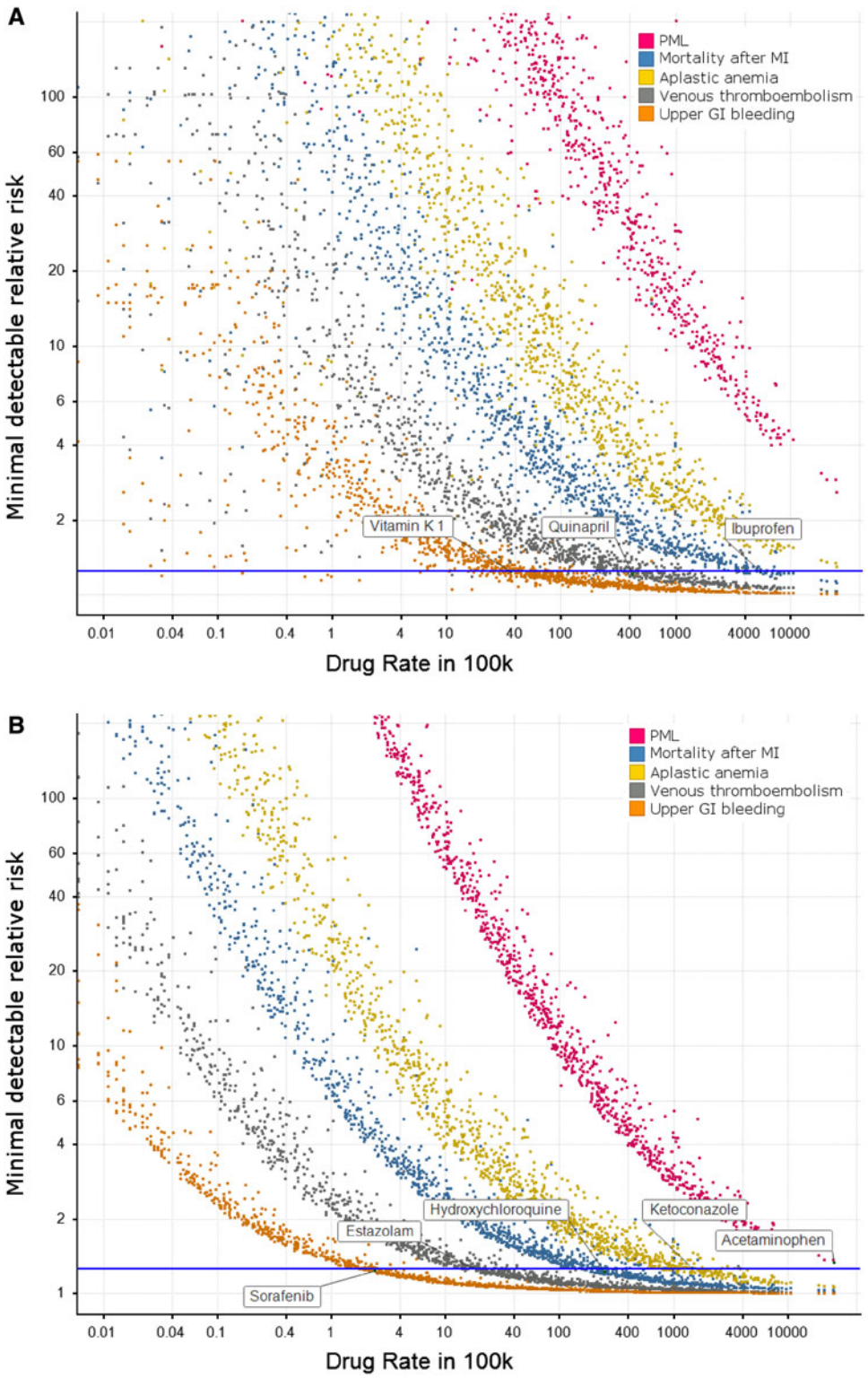


**Fig. 3** Minimal detectable relative risk of drugs for different HOIs in two select databases: **a** MSLR (1.2M patients) and **b** CCAE database (47M patients). The *x*-axis contains the number of patients per 100 thousand having been prescribed or administered the drug at least once. The *horizontal line* denotes the relative risk of 1.25, the *boxes* highlight examples of typical drugs that are close to the detection threshold for each of the HOIs

database coverage period without reaching the threshold. Tipranavir and natalizumab (approved 23-Nov-04) have effectively been discontinued in the MSLR database, reaching a plateau without any chance to test their association with liver failure for a very long time. The reason for this de-facto discontinuation of tipranavir in MSLR is unclear, while the marketing of natalizumab is limited due to its unfavorable safety profile [14].

Table 2 summarizes the time required to reach sufficient sample size among the newly marketed products from

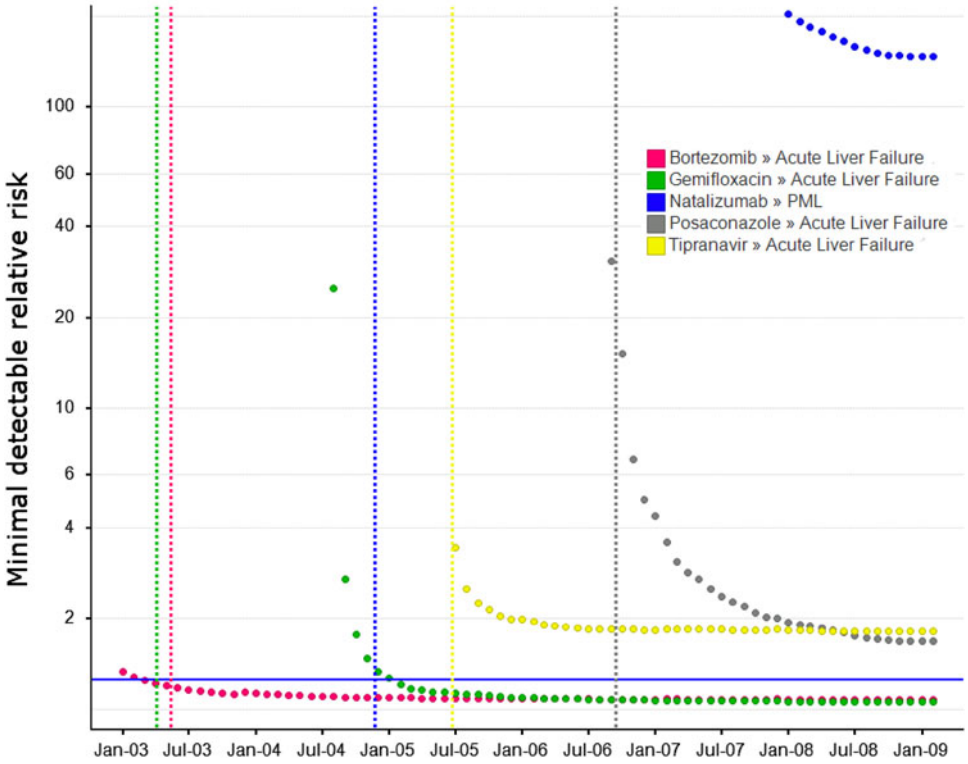


**Fig. 4** MDRR for drugs newly approved by the FDA between 2003 and 2008 in the MSLR database for a selection of HOIs and their known associated outcomes. The *vertical lines* depict the date of FDA approval and the circles the minimal detectable relative risk (MDRR) for the corresponding drug-outcome pair for each month in the same color-coding. The *blue horizontal line* indicates the MDRR of 1.25. See http://omop.org/mdrr for a list of all drugs approved in the period 2003–2007 and the duration to reach an MDRR of 1.25 or less for a selection of relevant HOIs

**Table 2** The time required to reach an MDRR of 1.25 or less in the CCAE database for each HOI, among active compounds approved by the FDA between 2003 and 2008

| | % of drugs meeting MDRR ≤1.25 in: | | | | | Time to MDRR ≤1.25 (months) | |
|---|---|---|---|---|---|---|---|
| | <1 year | <2 year | <3 year | <4 year | <5 year | Avg | Stdev |
| Number of newly marketed drugs with sufficient follow-up | 108 | 89 | 68 | 51 | 24 | | |
| Acute liver injury | 5 % | 8 % | 6 % | 33 % | 67 % | 10.3 | 10.8 |
| Acute myocardial infarction | 5 % | 6 % | 6 % | 31 % | 63 % | 11.8 | 12.1 |
| Acute kidney injury | 1 % | 3 % | 4 % | 16 % | 46 % | 18.4 | 13.4 |
| Aplastic anemia | 0 % | 0 % | 0 % | 0 % | 4 % | 49.0 | |
| Bleeding | 6 % | 10 % | 13 % | 35 % | 75 % | 6.2 | 12.1 |
| Mortality after MI | 2 % | 3 % | 4 % | 18 % | 46 % | 17.7 | 13.3 |
| PML | 0 % | 0 % | 0 % | 0 % | 0 % | | |
| Upper GI bleeding | 5 % | 3 % | 6 % | 24 % | 54 % | 12.7 | 12.7 |

*PML* progressive multifocal leukoencephalopathy, *MI* myocardial infarction, *GI* gastrointestinal, *FDA* Food and Drug Administration, *MDRR* minimal detectable relative risk, *CCAE* MarketScan Commercial Claims and Encounters

2003–2008 for eight health outcomes of interest. Among the 108 drugs with at least 1 year of observation since approval, less than 6 % reached an MDRR of 1.25 for any outcome. Less than 35 % of drugs achieve sufficient sample size 4 years after approval. After 5 years of observation, 75 % of drugs achieved an MDRR of 1.25 for bleeding and 63 % of drugs have sufficient cases to study acute myocardial infarction. Among those drugs that do eventually achieve sufficient sample size to be powered for an RR of 1.25, the average time-to-power is more than 12 months for upper GI bleeding, and more than 18 months for acute kidney injury. No drugs achieved sufficient sample size to study PML.

## 4 Discussion

A drug safety surveillance system is expected to have the ability to identify risks of marketed therapies from large-scale observational data fast and reliably for the majority of

products. In this paper, we studied whether typical commercially available large-scale observational databases can support such a system and contain sufficient data necessary for reliability and short turn-around of the generated signals.

We show that analytical methods have a better discriminatory performance for drug-outcome pairs abundant enough to provide sufficient power. This is dependent on the frequency of drug and outcome and the size of the database. We conclude that a substantial fraction of drugs cannot be tested for rare pharmacologically relevant outcomes, such as PML, aplastic anemia or acute liver injury, even if the database included the entire US or even world population.

For the purpose of this paper we used a crude power calculation assuming a basic cohort design. For example, the calculation makes no adjustment for analytical methods that target relevant subpopulations in the databases or for methods that control for confounding patient characteristics or for concomitant drug exposure. Methods that adjust for these multiple covariates are likely to furnish greater power for similarly sized study populations and true effect size. For such methods, our MDRR estimates here may be upwardly biased, although the size of this bias is unknown and expected to be small. Further, more refined sample size calculations may be available for the SCCS [15]. The SCCS likelihood function parallels a Cox proportional hazards model for which limited sample size calculations exist [16] and their functional form follows closely our approximation in Eq. (1). However, as a cases-only design, power depends first on the total number of patients in database who have both exposure to the drug and experienced at least one HOI event, and then on the proportion of time observed in the database during which the patient is exposed. Power increases as this proportion moves away from 0 or 100 %, suggesting that balancing the inclusion of exposed and unexposed periods is important. The power of analytic methods using matched designs can be improved through the inclusion of increasing numbers of controls matched for possible confounding covariates. Fortunately, the availability of such controls is not limiting with massive healthcare databases.

In spite of these modest differences between analytical methods, one general way of improving performance is to select outcome definitions that produce the highest frequencies in the databases. We studied the effect of those alternative definitions for the same outcome on the performance of methods, and found that the more inclusive definitions are comparable to narrower, more precise definitions in their ability to support discriminating signal from negative control, but are superior in its ability to study less frequent drugs [17].

## 5 Conclusion

The performance of methods estimating the strength of drug-outcome associations in observational databases decays with low sample size. Therefore, careful consideration is necessary to chose databases and outcome definitions to maximize the frequency of drug and outcome cases. This is particularly important for the study of newly introduced drugs.

**Acknowledgements** The Observational Medical Outcomes Partnership is funded by the Foundation for the National Institutes of Health (FNIH) through generous contributions from the following: Abbott, Amgen, AstraZeneca, Bayer Healthcare Pharmaceuticals, Biogen Idec, Bristol-Myers Squibb, Eli Lilly & Company, GlaxoSmithKline, Janssen Research and Development, Lundbeck, Inc., Merck & Co., Novartis Pharmaceuticals, Pfizer, Pharmaceutical Research Manufacturers of America (PhRMA), Roche, Sanofi-Aventis, Schering-Plough, and Takeda. Dr. Reich is an employee of AstraZeneca. Dr. Ryan is an employee of Janssen Research and Development. Dr. Suchard received a grant previously from the FNIH.

This article was published in a supplement sponsored by the Foundation for the National Institutes of Health (FNIH). The supplement was guest edited by Stephen J.W. Evans. It was peer reviewed by Olaf H. Klungel who received a small honorarium to cover out-of-pocket expenses. S.J.W.E has received travel funding from the FNIH to travel to the OMOP symposium and received a fee from FNIH for the review of a protocol for OMOP. O.H.K has received funding for the IMI-PROTECT project.from the Innovative Medicines Initiative Joint Undertaking (http://www.imi.europa.eu) under Grant Agreement no 115004, resources of which are composed of financial contribution from the European Union's Seventh Framework Programme (FP7/2007-2013) and EFPIA companies' in kind contribution.

## Appendix: Definitions of the Health Outcomes of Interest Studied

| HOI | # | Definition |
|---|---|---|
| Aplastic anemia | 2 | Occurrence of at least one diagnostic code ICD-9-CM:<br>• 284.0* Constitutional aplastic anemia[a]<br>• 284.8* Other specified aplastic anemias<br>• 284.9 Aplastic anemia, unspecified AND<br>Occurrence of at least one diagnostic procedure code for bone marrow aspiration or biopsy within 60 days prior to the diagnostic code |
| Acute kidney injury | 1 | Occurrence of at least one diagnostic code ICD-9-CM:<br>• 584* Acute renal failure[a] |
| Acute liver injury | 1 | Occurrence of at least one diagnostic code ICD-9-CM:<br>• 277.4 Disorders of bilirubin excretion<br>• 570* Acute and subacute necrosis of the liver[a]<br>• 572.2 Hepatic coma (hepatorenal syndrome)<br>• 572.4* Hepatorenal syndrome[a]<br>• 573* Other disorders of the liver, including chemical or drug induced[a]<br>• 576.8 Other specified disorders of biliary tract<br>• 782.4 Jaundice, unspecified, not of newborn<br>• 789.1* Hepatomegaly[a]<br>• 790.4* Nonspecific elevation of transaminase or lactic dehydrogenase levels[a]<br>• 794.8* Abnormal liver function test results[a] |
| Acute myocardial infarction | 1 | Occurrence of at least one broad diagnostic code ICD-9-CM:<br>• 410* Acute myocardial infarction[a]<br>• 411.1 Intermediate coronary syndrome<br>• 411.8 Other acute coronary occlusion<br>• 413.9 Other and unspecified angina pectoris on or during hospitalization |
| Bleeding | 3 | Occurrence of at least one diagnostic code[b] |
| Mortality after myocardial infarction | 3 | Occurrence of at least one narrow diagnostic code ICD-9-CM:<br>• 410* Acute myocardial infarction[a] AND<br>Occurrence of at least one diagnostic procedure code within 30 days prior to diagnostic code[c] OR<br>Occurrence of at least one therapeutic procedure code within 60 days after the diagnostic code[c] AND<br>Occurrence of death after the diagnostic code as one of the following:<br>• OBSERVATION_PERIOD_END_DATE where PERSON_STATUS = Death<br>• Occurrence of one condition code indicating death ICD-9-CM:<br>• 798.0 Sudden death, cause unknown<br>• 798.1 Instantaneous death<br>• 798.2 Death occurring in less than 24 h from onset of symptoms, not otherwise explained<br>• 798.9 Unattended death<br>• Occurrence of a diagnostic code ICD-9-CM:<br>• 427.5 Cardiac arrest AND OBSERVATION_PERIOD_END_DATE at the date of the diagnostic code |
| Progressive multifocal leukoencephalopathy | 1 | Occurrence of at least one diagnostic code ICD-9-CM:<br>• 046.3 Progressive multifocal leukoencephalopathy |
| Upper GI Ulcer | 1 | Occurrence of at least one diagnostic code[d] AND hospitalization at date of diagnostic code |

[a] An asterisk indicates a wildcard, i.e. any code with or without additional digits is included in the definition

[b] A detailed list of all codes are available at http://omop.org/Bleeding

[c] A detailed list of all codes are available at http://omop.org/AcuteMyocardialInfarction

[d] A detailed list of all codes are available at http://omop.org/GIulcerHospitalization

## References


1. Takahashi Y, Nishida Y, Asai S. Utilization of health care databases for pharmacoepidemiology. Eur J Clin Pharmacol. 2012;68(2):123–9.
2. Silverman SL. From randomized controlled trials to observational studies. Am J Med. 2009;122(2):114–20.
3. Vandenbroucke JP. When are observational studies as credible as randomised trials? Lancet. 2004;363(9422):1728–31.
4. Temple R. Meta-analysis and epidemiologic studies in drug development and postmarketing surveillance. JAMA. 1999;281 (9):841–4.
5. Taubes G. Epidemiology faces its limits. Science. 1995;269 (5221):164–9.
6. Coloma PM, Trifirò G, Schuemie MJ, Gini R, Herings R, Hippisley-Cox J, et al. Electronic healthcare databases for active drug safety surveillance: is there enough leverage? Pharmacoepidemiol Drug Saf. 2012;21(6):611–21.
7. Ryan PB, Madigan D, Stang PE, Marc Overhage J, Racoosin JA, Hartzema AG. Empirical assessment of methods for risk identification in healthcare data: results from the experiments of the Observational Medical Outcomes Partnership. Stat Med. 2012;31 (30):4401–15.
8. Ryan PB, Stang PE, Overhage JM, Suchard MA, Hartzema AG, DuMouchel W, et al. A comparison of the empirical performance of methods for a risk identification system. Drug Saf (in this supplement issue). doi:10.1007/s40264-013-0108-9.
9. Armstrong B. A simple estimator of minimum detectable relative risk, sample size, or power in cohort studies. Am J Epidemiol. 1987;126(2):356–8.
10. Beaumont JJ, Breslow NE. Power considerations in epidemiologic studies of vinyl chloride workers. Am J Epidemiol. 1981;114 (5):725–34.
11. Overhage JM, Ryan PB, Schuemie MJ, Stang PE. Desideratum for Evidence Based Epidemiology. Drug Saf (in this supplement issue). doi:10.1007/s40264-013-0102-2.
12. Ryan PB, Schuemie MJ, Welebob E, Duke J, Valentine S, Hartzema AG. Defining a reference set to support methodological research in drug safety. Drug Saf (in this supplement issue). doi:10.1007/s40264-013-0097-8.
13. Cantor SB, Kattan MW. Determining the area under the ROC curve for a binary diagnostic test. Med Decis Making. 2000;20 (4):468–70.
14. FDA Drug Safety Communication: risk of progressive multifocal leukoencephalopathy (PML) with the use of Tysabri (natalizumab). 02-05-2010 [cited 2013 January 28]. http://www.fda.gov/drugs/drugsafety/postmarketdrugsafetyinformationforpatientsandproviders/ucm199872.htm.
15. Musonda P, Farrington CP, Whitaker HJ. Sample sizes for self-controlled case series studies. Stat Med. 2006;25:2618–31.
16. Schoenfeld DA. Sample-size formula for the proportional-hazards regression model. Biometrics. 1983;39(2):499–503.
17. Reich CG, Ryan PB, Schuemie MJ. Alternative outcome definitions and their effect on the performance of methods for observational outcome studies. Drug Saf (in this supplement issue). doi:10.1007/s40264-013-0111-1.