

The risk of acute liver injury associated with the use of antibiotics—evaluating robustness of results in the pharmacoepidemiological research on outcomes of therapeutics by a European consortium (PROTECT) project.

Renate Udo¹, Stéphanie Tcherny-Lessenot², Ruth Brauer³, Paul Dolin⁴, David Irvine⁴, Yunxun Wang², Laurent Auclert², Juhaeri Juhaeri², Xavier Kurz⁵, Lucien Abenheim⁶, Lamiae Grimaldi⁶ and Marie L. De Bruin^{1*}

¹Utrecht Institute for Pharmaceutical Sciences, Utrecht University, The Netherlands

²Global Pharmacovigilance and Epidemiology, Sanofi, Chilly-Mazarin, France

³London School of Hygiene and Tropical Medicine, Faculty of Epidemiology and Population Health, London, United Kingdom

⁴Takeda Development Centre Europe, London, United Kingdom

⁵European Medicines Agency (EMA), London, United Kingdom

⁶LA-SER Europe Limited, London, United Kingdom

ABSTRACT

Purpose To examine the robustness of findings of case–control studies on the association between acute liver injury (ALI) and antibiotic use in the following different situations: (i) Replication of a protocol in different databases, with different data types, as well as replication in the same database, but performed by a different research team. (ii) Varying algorithms to identify cases, with and without manual case validation. (iii) Different exposure windows for time at risk.

Methods Five case–control studies in four different databases were performed with a common study protocol as starting point to harmonize study outcome definitions, exposure definitions and statistical analyses.

Results All five studies showed an increased risk of ALI associated with antibiotic use ranging from OR 2.6 (95% CI 1.3–5.4) to 7.7 (95% CI 2.0–29.3). Comparable trends could be observed in the five studies: (i) without manual validation the use of the narrowest definition for ALI showed higher risk estimates, (ii) narrow and broad algorithm definitions followed by manual validation of cases resulted in similar risk estimates, and (iii) the use of a larger window (30 days vs 14 days) to define time at risk led to a decrease in risk estimates.

Conclusions Reproduction of a study using a predefined protocol in different database settings is feasible, although assumptions had to be made and amendments in the protocol were inevitable. Despite differences, the strength of association was comparable between the studies. In addition, the impact of varying outcome definitions and time windows showed similar trends within the data sources. Copyright © 2015 John Wiley & Sons, Ltd.

KEY WORDS—pharmacoepidemiology; case–control studies; case definition; time window at risk; acute liver injury; antibiotics

Received 8 October 2014; Revised 24 June 2015; Accepted 25 June 2015

INTRODUCTION

Although drug-induced liver injury is a well-recognized health problem and an important reason for drug withdrawal, it is a challenge to conduct population-based epidemiological studies and quantify the adverse event.¹ The complexity of diagnosis and

its rarity are key problems. Designing a valid observational study can be further complicated when the exposure of interest is used intermittently and time at risk does not overlap with the time exposed.² Acute liver injury (ALI) associated with antibiotic use was therefore selected as a study topic in the Pharmacoepidemiological Research on Outcomes of Therapeutics by a European Consortium (PROTECT) project, which aims to develop, test and disseminate methodological standards for the design, conduct and

* Correspondence to: M. L. De Bruin, Division of Pharmacoepidemiology and Clinical Pharmacology, Utrecht Institute for Pharmaceutical Sciences, PO-box 80082, 3508 TB Utrecht, The Netherlands. Email: m.l.debruin@uu.nl

analysis of pharmacoepidemiology studies, applicable to different safety issues using different data sources.³

The objective of the present study was to examine the robustness of the estimated risk of acute liver injury (ALI) with the use of antibiotics in three distinct situations. These included (i) the replication of a common protocol in different databases, which comprised different data types, and replication of the study in the same database, but performed by a different research team. (ii) The use of varying case definitions of ALI, namely a narrow definition algorithm was compared with a broader definition algorithm. Also, the impact of manual validation of cases was examined. (iii) The influence of different time windows at risk was evaluated.

METHODS

Within the PROTECT project, five case-control studies were performed to quantify the risk of ALI associated with antibiotic exposure in the primary care setting. To examine the robustness of results the study was repeated in different databases by different research teams. The protocol developed for the case-control studies in BIFAP and CPRD I (initial studies) served as the basis for the protocols and statistical analysis plans of the other studies (follow-up studies) to harmonize outcome and exposure definitions, and statistical analyses. Protocols of the studies are registered in the European Network of Centres for Pharmacoepidemiology and Pharmacovigilance (ENCePP).⁴ Results of the initial case-control studies were submitted centrally to the PROTECT Research Manager at the coordinating center (Utrecht) and blinded to the researchers performing the follow-up studies until finalization of their statistical analysis plans.

Data sources

The studies were conducted in primary care databases from the United Kingdom and Spain, respectively, the Clinical Practice Research Datalink (CPRD) and the Base de datos para la Investigación Farmacoepidemiologica en Atención Primaria (BIFAP), the US health insurance (claims) database Clinformatics Data Mart (ClinFormatics) and the Dutch hospital-based database Utrecht Patient Oriented Database (UPOD), which was linked to the Mondriaan primary care databases Julius General Practitioners Network (JHN) and Almere Health Care (ZGA). Databases BIFAP and CPRD contain nationwide primary care data provided by general practitioners (GPs). ClinFormatics contains

claims from providers, outpatient pharmacy dispensing records, laboratory tests and demographic data. The database UPOD comprises information on hospital discharge diagnoses, in-hospital pharmacy records and laboratory tests for all patients treated at the University Medical Center Utrecht (UMCU) in the Netherlands (inpatients as well as outpatients). JHN and ZGA contain routine health care data provided by general practices in the cities of Almere and Utrecht and their vicinity. All databases are described in detail elsewhere.⁵⁻¹¹

Study design

In all databases a case-control study was performed. In the CPRD database two case-control studies were conducted by two different research teams. The initial CPRD study is referred to as CPRD I, the replication study as CPRD II. All cases with a first recorded occurrence of acute liver injury were identified by the algorithm described below from January 2004 to December 2009, except for UPOD during the period January 2008 to December 2010. Cases were matched to up to five controls by age (within one year), sex, calendar date (month and year) and in CPRD and UPOD also by practice. Controls were sampled from the patients at risk (without a diagnostic code related to liver injury) at the time of occurrence of the event from the same source population that gave rise to the cases. Index date for the control was the index date of the assigned case (date of diagnosis). For cases identified in UPOD, record linkage to the primary care databases JHN and ZGA was performed first, and only linked cases were matched to controls. In the Netherlands, GPs are considered to have preference for one particular hospital to refer patients to. Because cases were also matched to controls by practice, controls were regarded sampled from the same population that gave rise to the cases. The case-control studies in CPRD I and BIFAP are presented in more detail elsewhere.¹²

The study was also performed in the hospital database UPOD to evaluate the transferability and feasibility of the developed methods with hospital-based data and to be able to compare results in different data source types.

Outcome definition

BIFAP and CPRD I. Patients with ALI were identified using an algorithm developed within PROTECT, which was based on diagnostic codes indicative of idiopathic ALI, liver test results and referral for liver injury to any specialist or hospital within two weeks

of the recorded diagnosis. Two distinct case definitions were used. For the primary analyses a narrow algorithm for ALI was applied, and the identified cases are referred to as definite cases. These cases fulfill all three conditions of the algorithm. This meant a case with a liver disease related diagnostic code together with a referral and liver test abnormality. This abnormality could be either an increase of more than two times the upper limit of the normal range (ULN) in alanine aminotransferase (ALT) or a combined increase in aspartate aminotransferase (AST), alkaline phosphatase (AP) and total bilirubin provided one of them is twice their respective ULN.

A secondary outcome definition, including definite plus probable cases of ALI, required the occurrence of a diagnostic code (specific or non specific) together with elevated liver-related enzyme values with or without a referral or hospitalization. This is considered the broad algorithm. Cases with a diagnostic code also indicating other known causes of liver injury were excluded in both algorithms (viral hepatitis, cancer, alcohol related problems, gallbladder disease, pancreatic disease and other chronic liver diseases). Cases with concomitant drug use were not excluded from the analyses as all cases with ALI potentially caused by antibiotic use should be taken into account. In the main regression model we adjusted for concomitant use (prescription until or 30 days prior to the index date) of potentially hepatotoxic agents.

In BIFAP only, all identified definite and probable cases were manually reviewed in medical records. Only validated ALI cases were used in the statistical analyses. Details about the algorithms and utilized diagnostic codes (READ codes in CPRD and International Classification of Primary Care (ICPC) codes in BIFAP) are described in the PROTECT case ascertainment study in CPRD and BIFAP.¹³

ClinFormatics, CPRD II and UPOD. In ClinFormatics, CPRD II, and UPOD the same primary and secondary outcome definitions were applied as for BIFAP and CPRD I with the following adjustments. For UPOD and ClinFormatics diagnostic codes were converted to the International Classification of Diseases, 9th revision, Clinical Modification (ICD-9-CM) codes. The codes applied in UPOD and ClinFormatics are described in the PROTECT validation study performed in UPOD.¹⁴ In CPRD II no time window for referrals or hospitalization was applied and any referral sufficed, not only referrals specific for liver injury as was in the BIFAP and CPRD I studies. In UPOD, the original algorithm definitions

were applied to identify hospital cases with ALI (inpatients as well as outpatients), apart from the following two amendments. (i) Although diagnoses were made in hospital, patients were not necessarily referred to the hospital for liver problems. As a referral is therefore not included in the algorithm, the UPOD cases follow the broad definition. (ii) A second algorithm based on only liver test results (ALT > 10ULN and Hy's law criteria (ALT > 3ULN and bilirubin > 2ULN and absence of AP elevation)) was used. After identification of cases by the algorithms in UPOD, medical hospital records of all identified cases were manually reviewed to validate outcome.¹⁴

In ClinFormatics, CPRD II and UPOD, alcohol related diagnostic codes were not applied to exclude cases. However, in UPOD these patients were excluded following manual review, because in all cases, alcohol was assumed to be the cause of the ALI. Another exception was that in the ClinFormatics and UPOD studies only hepatic cancer/metastases were excluded and not all cancer codes, as was in the other studies.

Exposure definition and time at risk

Exposure to antibiotics was assessed as a recorded prescription for an antibiotic drug for systemic use. Patients were defined as currently at risk if a prescription for an antibiotic drug lasted until or after the index date or ended within 14 days prior to the index date. Time not at risk was defined when supply of the most recent antibiotic drug prescription ended more than 365 days before the index date or when there was no recorded prescription at any time prior to the index date.

Statistical analyses

We computed odds ratios (OR) and 95% confidence intervals of first occurrence of ALI associated with current use of antibiotics (14-day time window) as compared to non-use with multivariable conditional logistic regression. Primary analyses were performed with definite cases. For UPOD this entails validated ALI cases after manual review. Statistical analyses were conducted using Stata, version 11 (CPRD, BIFAP), SAS, version 9.1 (ClinFormatics) and SPSS, version 20 (UPOD).

Sensitivity analyses. To assess the impact of varying ALI definitions, sensitivity analyses were performed. Additional ORs were therefore calculated for ALI cases according to the broad definition in

ClinFormatics, CPRD II and UPOD. In the BIFAP study the broad identified cases were manually validated. In UPOD, risk estimates for manually reviewed true cases were compared to estimates for ALI cases solely identified by the algorithm.

A second sensitivity analysis was performed with a 30-day time window at risk for current use of antibiotics, to assess the impact on risk estimates for validated true cases in UPOD and definite cases in the other databases.

RESULTS

In CPRD I, CPRD II and ClinFormatics, where cases were identified according to the predefined algorithms only, the number of cases were 989, 4064 and 5898 (broad algorithm) and 263, 1367 and 753 (narrow algorithm), respectively. In the BIFAP study, where all identified definite and broad definition cases were manually validated, final numbers were 436 cases according to the broad algorithm, and 124 cases according to the narrow algorithm. In UPOD, approximately one out of ten cases identified in the UPOD hospital database could be linked to primary care databases and the final numbers of 80 cases identified by the broad algorithm and 20 validated true cases were suitable for the case-control analysis. Age and the percentage of women were higher in both linked groups compared to the non-linked groups (Figure 1 online). For the main characteristics of all five studies see Table 1.

The results of the primary analyses showed an increased risk of ALI (according to the narrow definition) with antibiotic use up to 14 days, namely from OR 2.6 (95% CI 1.3–5.4) in BIFAP to 7.7 (95% CI 2.0–29.3) in UPOD (Figure 1, squares). The results of the sensitivity analyses for broader definitions of ALI are also displayed in Figure 1 (triangles). In the studies that performed manual validation (Figure 1, left hand side), defining ALI by a less strict algorithm to broad definition cases in BIFAP or to identified (non-validated) cases in UPOD led to a small increase in estimates and to more precision in risk estimates; from OR 2.6 (95% CI 1.3–5.4) to OR 3.1 (95% CI 2.1–4.6) in BIFAP and from OR 7.7 (95% CI 2.0–29.3) to OR 8.0 (95% CI 2.8–23.2) in UPOD. When analyses were extended to broad definition cases in the studies without manual case validation (Figure 1, right hand side), the risk estimates of outcome association decreased in all three studies, from OR 5.7 (95% CI 3.5–9.4) to OR 3.6 (95% CI 2.8–4.6) in CPRD I, from OR 3.8 (95% CI 3.0–4.8) to OR 3.3 (95% CI 2.8 to 3.8) in CPRD II, and from OR 4.4 (95% CI 3.2–6.1) to OR 2.6 (95% CI 2.3–2.9) in ClinFormatics.

The results of the case-control analyses across databases with a different definition of current use of antibiotics are displayed in Figure 2. The use of a larger time window to define current use (from 14 days to 30 days) led to decrease or no change in risk estimates. The OR decreased from 5.7 (95% CI 3.5–9.4) to 4.4 (95% CI 2.9–6.7) in CPRD I, from 3.8 (95% CI 3.0–4.8) to 3.1 (95% CI 2.5–3.8) in CPRD II, and from 4.4 (95% CI 3.2–6.1) to 2.8 (95% CI 2.2–3.7) in ClinFormatics, remained stable in BIFAP from 2.6 (95% CI 1.3–5.4) to 2.5 (95% CI 1.3–4.9) and was the same in UPOD, being 7.7 (95% CI 2.0–29.3). In all five studies, the use of a larger risk window led mainly to increase of background exposure rate in controls. Increase was about 50% across all databases in controls and no more than 30% in cases (data not shown).

Risks varied for the different antibiotic agents (Table 1 online). Although similar trends could be identified across studies, the numbers of cases in each antibiotic class was low, and therefore confidence intervals were wide.

DISCUSSION

By conducting five case-control studies estimating the association between antibiotic use and ALI in different settings, we were able to evaluate the robustness of the obtained results and to test the reproducibility of the protocol. It appeared not to be possible to exactly replicate the main study protocol in all the performed situations, because of differences in data sources with their own characteristics and variables, differences in research teams and perspectives, and amendments made to the protocol during the course of the project. Despite other data sources and protocol amendments, comparable overall results on risk estimates were found.

However, several subtle differences between the study results were observed, which could be explained as follows. In UPOD, the highest risk estimates were observed. As this is a hospital-based database, cases might be different regarding health status and received health care. In a hospital setting frequent liver tests are being performed as part of routine investigations and elevated liver-related enzyme values will commonly be detected, whereas in general practice only patients with liver injury who experience symptoms are likely to be identified. Therefore, non-symptomatic liver injury is more likely to be detected in a hospital setting. What also might be of influence is that in the Netherlands antibiotics are prescribed relatively conservatively and may be channeled towards more ill patients.¹⁵ Interesting is the small difference in risk estimate observed in the two studies performed in CPRD that is likely to be

Table 1. General covariates assessed as risk factors for acute liver injury (ALI)

	BIFAP		CPRD I		UPOD		ClimFormatics		CPRD II	
	Controls N = 620	Cases N = 124	Controls N = 1284	Cases N = 263	Controls N = 100	Cases N = 20	Controls N = 3765	Cases N = 753	Controls N = 6575	Cases N = 1367
Sex, n (%) Male	310 (50)	62 (50)	638 (50)	130 (50)	35 (35)	7 (35)	1939 (51.5)	388 (51.5)	3259 (50)	671 (49)
Age, mean (SD)	46.6(21.3)	46.6(21.4)	57.2 (20.7)	57.8 (20.9)	56.3 (14.4)	56.4 (14.7)	48.2 (13.3)	48.3 (13.4)	53.9 (17.4)	54.3 (17.5)
Exposure any antibiotics, n (%)										
Non-use	354 (57.1)	49 (39.5)	719 (56.0)	94 (35.7)	79 (79.0)	10 (50.0)	2142 (56.9)	317 (42.1)	4625 (71.8)	676 (49.4)
Current use	40 (6.5)	20 (16.1)	66 (5.1)	59 (22.4)	5 (5.0)	6 (30.0)	275 (7.3)	125 (16.6)	248 (3.8)	229 (16.8)
Recent/past use	226 (36.5)	55 (44.3)	499 (38.9)	110 (41.8)	16 (16.0)	4 (20.0)	1465 (38.9)	369 (49.0)	1602 (24.4)	462 (33.8)
Comorbidities, n (%)										
Heart failure	12 (1.9)	2 (1.6)	35 (2.7)	13 (4.9)	1 (1.0)	1 (5.0)	11 (0.3)	17 (2.3)	105 (1.6)	35 (2.6)
Rheumatoid arthritis	4 (0.65)	1 (0.81)	27 (2.1)	9 (3.4)	0	1 (5.0)	23(0.6)	5 (0.7)	27 (0.4)	15 (1.1)
Hemochromatosis	—	—	0	1 (0.4)	—	—	0	0	2 (0.0)	2 (0.2)
Alpha-antitrypsin-deficiency	—	—	1 (0.1)	0	—	—	0	0	1 (0.0)	3 (0.2)
Diabetes	61 (9.8)	9 (7.3)	97 (7.6)	30 (11.4)	7 (7.0)	2 (10.0)	294 (7.8)	101 (13.4)	363 (5.5)	144 (10.5)
Treatment*, n (%)										
NSAIDs	65 (10.5)	25 (20.2)	83 (6.5)	25 (9.5)	6 (6.0)	4 (20.0)	425 (11.3)	100 (13.3)	589 (9.0)	262 (19.2)
Other analgesics/antipyretics	73 (11.8)	25 (20.2)	291 (22.7)	80 (30.4)	0	8 (40.0)	30 (0.8)	12 (1.6)	—	—
Statins	67 (10.8)	13 (10.5)	213 (16.6)	53 (20.2)	17 (17.0)	7 (35.0)	599 (15.9)	82 (10.9)	706 (10.7)	220 (16.1)
Antidepressants	37 (6.0)	5 (4.0)	97 (7.6)	36 (13.7)	2 (2.0)	7 (35.0)	561 (14.9)	166 (22)	385 (5.9)	165 (12.0)
Oral contraceptives	—	—	39 (6.0)	8 (6.0)	3 (3.0)	2 (10.0)	151 (4)	20 (2.7)	80 (1.2)	17 (1.2)
Oral preparation for acne	—	—	3 (0.2)	1 (0.4)	2 (2.0)	0	8 (0.2)	3 (0.4)	43 (0.7)	21 (1.5)
DMARD	5 (0.8)	0	10 (0.8)	5 (1.9)	0	2 (10.0)	4 (0.1)	0	53 (0.8)	38 (2.8)
Oral corticosteroids	9 (1.5)	5 (4.0)	19 (1.5)	13 (4.9)	2 (2.0)	5 (15.0)	260 (6.9)	54 (7.2)	0	0
Antidiabetic drugs	33(5.3)	6 (4.8)	54 (4.2)	21 (8.0)	4 (4.0)	6 (30.0)	256 (6.8)	96 (12.7)	197 (3.0)	88 (6.4)
Other hepatotoxic drugs (list FDA)	31(5.0)	2 (1.6)	45 (3.5)	28 (10.6)	0	3 (15.0)	—	—	144 (2.2)	77 (5.6)

*Treatment as current use: use of the drug at the index date or within 30 days prior to the index date.

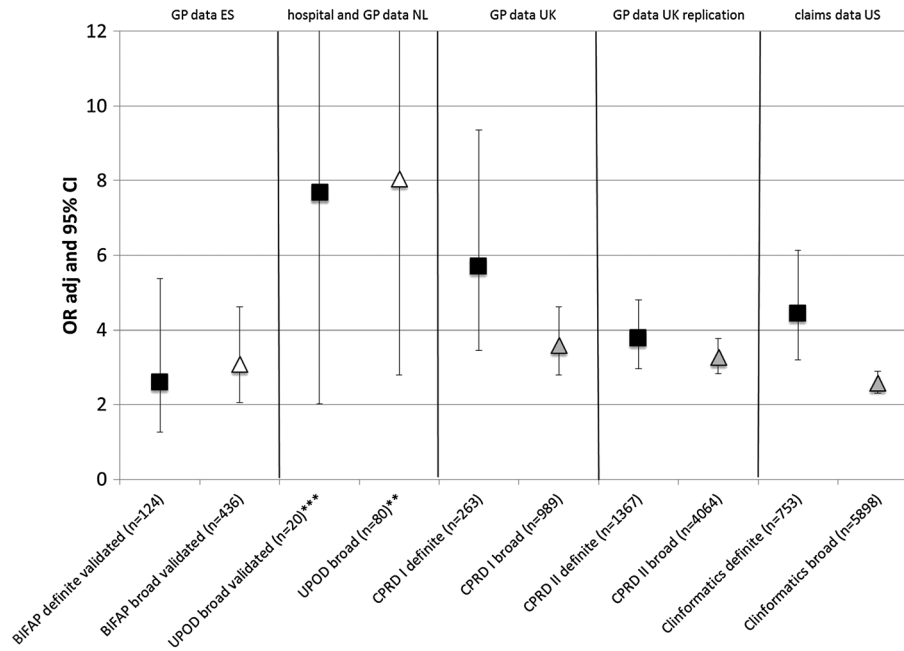


Figure 1. Adjusted odds ratios* of ALI associated with current antibiotic use (14-day time window) for varying ALI definitions in primary care data, hospital data and claims data. Left hand panel: studies that performed manual case validation (squares: primary analyses, triangles: white sensitivity analyses, outcome definition validated); right hand panel: studies that solely used algorithms to identify cases (squares: primary analyses, triangles: grey sensitivity analyses, broad outcome definition) * Odds ratios adjusted on general confounders (BMI, smoking and visits to GP in previous year), comorbidity variables (heart failure, diabetes and rheumatoid arthritis) and comedications (NSAIDs, other analgesics, statins, oral preparations for acne, disease modifying antirheumatism disease drugs, oral corticosteroids, antidiabetics, antidepressants and other hepatotoxic drugs); ** Odds ratio from best fitted model using forward selection of covariates from all potential confounders (general, comorbidity and comedications); *** Odds ratio adjusted on general confounders (BMI, smoking and visits to GP in previous year)

attributable to the difference in number of identified ALI cases. Five times as many definite cases were retained in the CPRD II study when compared to the CPRD I study. Executing the definitions in less strict conditions (i.e. using any referral with no time restriction) as was done in CPRD II resulted in dilution of the risk outcome.

In the ClinFormatics database, the ratio between broad definition cases and definite cases is almost twice the ratio observed in the CPRD I and CPRD II study. This health insurance database in the US reflects potentially different patient management practices. Although the less frequent records of referral to specialist or hospital resulted in a substantial decrease in retained definite cases, the risk estimates remained in the same range as for the CPRD studies.

Sensitivity analyses: outcome definition. The impact of degree of outcome certainty on the estimates of association between ALI and use of antibiotics was explored by using a broader definition of ALI and by comparing results for validated true cases and identified (non-validated) cases. Little is known about the possible error that is introduced when potential ALI cases identified by search algorithms in health care databases are used in estimating associations. The

expectation that associations would be diluted by a less strict definition of cases was confirmed in CPRD I, CPRD II and ClinFormatics, which were all purely algorithm based. The ORs for broad definition ALI cases were lower in comparison to definite cases, but resulted in smaller confidence intervals because of larger numbers of cases. A similar pattern was observed in the cohort studies in BIFAP and CPRD: higher incidence rate ratios were found for definite cases in comparison to broad definition cases.¹² These results could be explained by the inclusion of more false-positive cases following the less restrictive case definition. The validity of the algorithms that were used in the PROTECT case-control studies was assessed in the BIFAP and UPOD databases. In BIFAP the narrow definition resulted in a higher proportion of valid ALI cases compared to the broader definition (35% versus 20%).¹³ In UPOD the stricter criteria of inclusion (diagnostic codes plus laboratory abnormalities) showed a positive predictive value (PPV) of 47%, whereas the PPV in the broader algorithm (based on laboratory results only) was only 26%.¹⁴ The findings in the present study are therefore influenced by the non-differential misclassification of the outcome, which leads to estimates biased towards the null.¹⁶

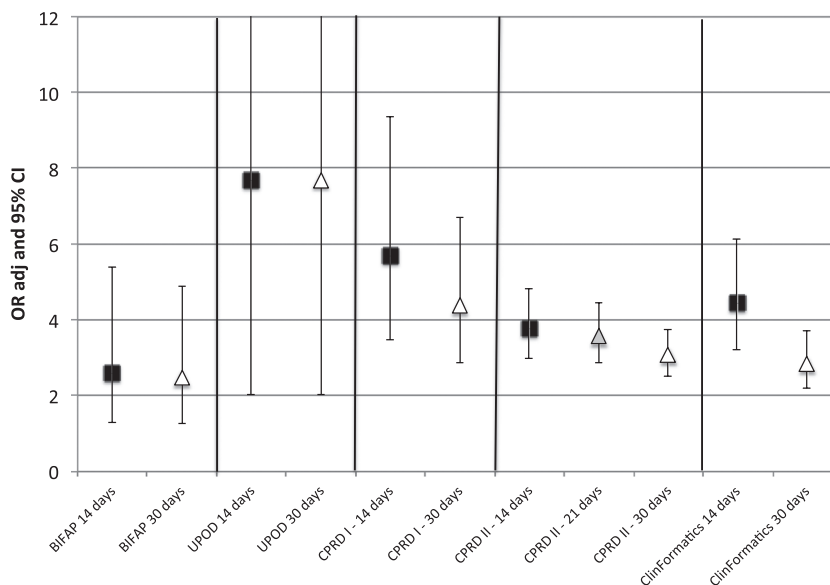


Figure 2. Adjusted odds ratios* of definite cases** of ALI associated with antibiotic use across databases and different definitions of time window at risk. Squares: primary analyses (14-days time window at risk); Triangles: sensitivity analyses (30-days time window at risk). * Odds ratios adjusted on general confounders (BMI, smoking and visits to GP in previous year), comorbidity variables (heart failure, diabetes and rheumatoid arthritis) and comedications (NSAIDs, other analgesics, statins, oral preparations for acne, Disease modifying rheumatic disease drugs, oral corticosteroids, antidiabetics, antidepressants and other hepatotoxic drugs); ** UPOD and BIFAP used validated cases (definite validated cases in BIFAP and broad validated cases in UPOD)

Our assumption that manual validation of outcome to identify true cases would also have a diluting influence on the risk association was not shown in UPOD. However, the limited number of cases and the broad overlap of confidence intervals in UPOD make the interpretation of results difficult. Based on this single hospital-based study no firm conclusions could be drawn on whether case validation should be undertaken when associations are studied between ALI and antibiotic use. Furthermore, validation of the cases identified by the narrow and broad definition in BIFAP resulted in comparable relative risk estimates of association for both definitions. It appears that manual validation overruled the impact of different levels of certainty of ALI represented by the algorithm definitions.

In drug safety issues where risk estimates for ALI are requested and validation is possible and time not a limiting factor, the broad definition could be used to identify potential cases for validation which will result in higher number of cases as this algorithm is more sensitive and therefore result in smaller confidence intervals. However, manual validation is a time-consuming process, and when power is sufficient it might be considered to start with the definite definition as risk outcomes are comparable. If manual validation is not feasible, the use of a narrow definition is recommended as identified cases will yield relatively more true cases. Preferably, a definition including combinations of diagnostic codes, laboratory measurements and procedures should be

used.^{17–19} Unfortunately, as was shown in the previous PROTECT validation studies, the algorithm that identified ALI cases more accurately also selected a lower absolute number of cases.^{13,14} This resulted in less precision of the risk estimate shown in the present study.

Sensitivity analysis: time window at risk. Another objective of the study was to evaluate the effect of different lengths of time window at risk. In selecting an appropriate risk window, the time interval between the beginning or the end of use of the drug and the onset of the adverse event should be considered. In addition, the induction period of the reaction should be taken into account.^{2,20} Liver injury can be classified into distinct types of injury, each with different maximum lengths to onset of the reaction. Causality assessment methods such as the liver specific Council for International Organizations of Medical Sciences (CIOMS) scale presumes that for hepatocellular liver injury causal relationship is less probable after 90 days of initial treatment or more than 15 days after drug cessation. Onset of cholestatic and mixed type liver injury can even be till 30 days after drug cessation.²¹ This makes it appropriate for short-term use such as antibiotics that the time at risk should also include a period of risk after cessation of the drug. The 14-day time window showed an increase in risk estimates compared to the larger risk window (30 days). For antibiotics associated with ALI, the selection of a short

exposure window was efficient to limit the rate of background exposure in the controls. However, peak risk may vary with different individual antibiotics that induce different types of liver injury. Sensitivity analyses should be conducted to estimate this. In case of an outcome associated with a chronic exposure, the change of exposure window may have less impact on the strength of the association.²²

A limitation of this study is the absence of results of analyses on non-validated cases in BIFAP and validated true cases in CPRD to further explore and quantify the impact of validation of cases on effect estimates. Also, because no separate analyses were performed for individual agents (see online annex) to study appropriate time window at risk the finding of a peak risk estimate with a 14-day time window might not be a general rule for antibiotic-associated ALI. Furthermore, a limitation in studies using health care record databases is that information may be incomplete or unavailable. Information especially related to life style factors was not always completely recorded in the used databases.

In conclusion, replication of the study protocol on risk of ALI associated with the use of antibiotics is feasible, although amendments were needed to be able to perform the study in a different setting. Differences in databases and amendments in the main protocol appear not to be of considerable influence on the risk estimates, and clinical conclusions remain generally the same.

There are different ways to get more valid cases for case-control studies using electronic health care record databases. Algorithms using a narrow definition of ALI result in higher proportions of true cases and as a result higher risk estimates, albeit at the expense of lower absolute numbers of cases and hence less precision. As expected, manual validation appears to overrule the impact that different definitions have on risk estimates. In drug safety research, when validation of cases is possible, our study suggests that either definition could be used. In situations where manual validation would not be feasible, the use of a narrow definition should be considered as a method to reach more accurate effect estimates. Time windows at risk for ALI may be partly predicted on beforehand by the clinical manifestation described in case reports (hepatocellular, cholestatic and mixed), but it may be useful to assess various time windows using sensitivity analyses.

CONFLICT OF INTEREST

The research leading to these results was conducted as part of the PROTECT consortium (Pharmacoepidemiology Research0020on Outcomes of Therapeutics

by a European Consortium, www.imi-protect.eu) which is a public-private partnership coordinated by the European Medicines Agency.

KEY POINTS

- Five case-control studies conducted within the PROTECT project show an increased risk of ALI associated with antibiotic use with ORs ranging from 2.6 (95% CI 1.3–5.4) to 7.7 (95% CI 2.0–29.3).
- Risk estimates and the trends observed within the different studies align, supporting the robustness of the results.
- A more restrictive search algorithm to identify ALI patients resulted in higher risk estimates.
- Manual review of algorithmically identified ALI cases appears to neutralize the observed differences in estimates resulting from broader versus narrower algorithms.
- Time at risk in short term treatments, such as in the case of antibiotics, lasts after drug use has been terminated. A time window of 14 days after the last prescription has ended show higher risk estimates compared to a 30-day time window.

ETHICS STATEMENT

The BIFAP protocol was approved by the Spanish Agency for Medical products (AEMPS) as part of the PROTECT project. CPRD ISAC approval was obtained in October 2011 (11_019A_2). The London School of Hygiene and Tropical Medicine ethics committee approved the study in May 2011 (reference number 5973). Approval for the UPOD study was obtained by the Institutional Review Board of the UMC Utrecht. For Clinformatics, no ethics approval was required. Patients' information is only available anonymized in the database.

ACKNOWLEDGEMENTS

This research has been performed on behalf of work package 2 and 6 of the PROTECT project. The authors thank the excellent collaboration of physicians in the participating countries, whose contribution in recording their professional practice with high quality standards enables the availability of databases used in this research. Authors would like to acknowledge Ana Ruigomez (Ceife, Spain) and Consuelo Huerta (BIFAP, Spain) for coordinating the local analyses in Spain and their valuable comments on the manuscript.

FUNDING

The PROTECT project has received support from the Innovative Medicine Initiative Joint Undertaking (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. The views expressed are those of the authors only and not of their respective institution or company.

REFERENCES

- Bell LN, Chalasani N. Epidemiology of idiosyncratic drug-induced liver injury. *Semin Liver Dis* 2009; **29**(4): 337–347.
- McMahon AD, Evans JM, McGilchrist MM, McDevitt DG, MacDonald TM. Drug exposure risk windows and unexposed comparator groups for cohort studies in pharmacoepidemiology. *Pharmacoepidemiol Drug Saf* 1998; **7**(4): 275–280.
- Abbing-Karahagopian V, Kurz X, de Vries F, et al. Bridging differences in outcomes of pharmacoepidemiological studies: design and first results of the PROTECT project. *Curr Clin Pharmacol* 2014; **9**(2): 130–138.
- ENCePP. <http://www.encepp.eu/encepp/studies> (accessed June 2014)
- CPRD. <http://www.cprd.com> (accessed June 2014)
- Garcia Rodriguez LA, Perez GS. Use of the UK general practice research database for pharmacoepidemiology. *Br J Clin Pharmacol* 1998; **45**(5): 419–425.
- BIFAP. <http://www.bifap.org/summary.php> (accessed June 2014)
- Salvador Rosa A, Moreno Perez JC, Sonogo D, Garcia Rodriguez LA, de Abajo Iglesias FJ. The BIFAP project: database for pharmaco-epidemiological research in primary care. *Aten Primaria* 2002; **30**(10): 655–661.
- Clinformatics data mart. <http://www.optuminsight.com/life-sciences/solutions/value-strategy/marketing-analytics/clinformatics-data-mart/overview> (accessed June 2014)
- ten Berg MJ, Huisman A, van den Bemt PM, Schobben AF, Egberts AC, van Solinge WW. Linking laboratory and medication data: new opportunities for pharmacoepidemiological research. *Clin Chem Lab Med* 2007; **45**(1): 13–19.
- Grobbee DE, Hoes AW, Verheij TJ, Schrijvers AJ, van Ameijden EJ, Numans ME. The Utrecht health project: optimization of routine healthcare data for research. *Eur J Epidemiol* 2005; **20**(3): 285–287.
- Brauer R, Douglas I, García Rodríguez L, et al. Risk of acute liver injury associated with use of antibiotics. Comparative cohort and nested case-control studies using two primary care databases in Europe. *Pharmacoepidemiol Drug Saf* 2015.
- Ruigomez A, Brauer R, Rodriguez LA, et al. Ascertainment of acute liver injury in two European primary care databases. *Eur J Clin Pharmacol* 2014; **70**(10): 1227–1235.
- Udo R, Maitland-van der Zee A, Egberts A, et al. Validity of diagnostic codes and laboratory measurements to identify patients with idiopathic acute liver injury. *Pharmacoepidemiol Drug Saf* 2015.
- Brauer R, Ruigómez A, Downey G, et al. Prevalence of antibiotic use: a methodological comparison across various European health care data sources. *Pharmacoepidemiol Drug Saf* 2015.
- Copeland KT, Checkoway H, McMichael AJ, Holbrook RH. Bias due to misclassification in the estimation of relative risk. *Am J Epidemiol* 1977; **105**(5): 488–495.
- Hansen RA, Gray MD, Fox BI, Hollingsworth JC, Gao J, Zeng P. How well do various health outcome definitions identify appropriate cases in observational studies? *Drug Saf* 2013; **36**(Suppl 1): S27–32.
- Katz AJ, Ryan PB, Racoosin JA, Stang PE. Assessment of case definitions for identifying acute liver injury in large observational databases. *Drug Saf* 2013; **36**(8): 651–661.
- Reich CG, Ryan PB, Schuemie MJ. Alternative outcome definitions and their effect on the performance of methods for observational outcome studies. *Drug Saf* 2013; **36**(Suppl 1): S181–93.
- Schneeweiss S. A basic study design for expedited safety signal evaluation based on electronic healthcare data. *Pharmacoepidemiol Drug Saf* 2010; **19**(8): 858–868.
- Danan G, Benichou C. Causality assessment of adverse reactions to drugs—I. A novel method based on the conclusions of international consensus meetings: application to drug-induced liver injuries. *J Clin Epidemiol* 1993; **46**(11): 1323–1330.
- Mokhles MM, Trifiro G, Dieleman JP, et al. The risk of new onset heart failure associated with dopamine agonist use in Parkinson's disease. *Pharmacol Res* 2012; **65**(3): 358–364.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of this article at the publisher's website.